



# 멀티모달 대규모 언어모델과 기계학습을 활용한 도시 가로 경관 쇠퇴 영향요인 분석\*

## Analysis of Influencing Factors of Urban Landscape Decline Using Multi-Modal Large Language Model and Machine Learning

김이정\*\* · 이수기\*\*\*  
Kim, Yijeong · Lee, Sugie

### Abstract

Cities are constantly evolving, with growth, vitality, decline, and shrinkage occurring as interrelated phenomena. While urban vitality and revitalization have been extensively studied, research on urban decline and shrinkage remains comparatively limited. Existing studies on urban decline have primarily focused on diagnostic indicators, such as physical aging, population decline, and reduction in the number of businesses, to assess patterns of decline. However, studies on the subjective perception of urban decline, particularly in relation to urban landscapes, remain limited. Given this gap, it is crucial to identify the causes of urban decline by analyzing the factors that influence the public perception of declining landscapes through subjective evaluations of urban scenery. This study quantitatively analyzes how people perceive declining urban landscapes and identifies the key factors that influence these perceptions, using street view images of Seoul. A survey was first conducted to assess urban landscape decline based on streetscape images. The Trueskill algorithm was applied to quantify perceived level of decline. Subsequently, machine learning was used to analyze the primary factors influencing these perceptions. The results of the analysis are as follows. First, perception of decline decreased as the proportion of physical environmental elements such as roads, green spaces, sidewalks, and cars increased. In contrast, an increased presence of elements such as buildings, bicycles, walls, and fences was associated with a heightened perception of urban decline. Second, an analysis of the importance of contributing factors indicated that roads, sidewalks, green spaces, and cars were the most influential in shaping perception, in that order. Third, the relationship between the proportion of physical environmental elements in urban landscape images and perceptions of decline was found to be non-linear. This study presents a methodology for evaluating urban landscape decline based on people's subjective perceptions and provides policy implications by identifying the streetscape features that substantially influence perceptions of urban decline.

**주제어** 도시 쇠퇴경관, 가로경관 이미지, 멀티모달 대규모 언어모델, 기계학습, 주관적 인식  
**Keywords** Urban Landscape Decline, Street View Image, Multi-Modal Large Language Model, Machine Learning, Subjective Perception

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

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

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

# 1. 서론

## 1. 연구의 배경 및 목적

도시 경관은 인간의 삶의 질에 직접적인 영향을 미친다(Xu et al., 2023). 시간의 흐름에 따라 물리적 활력과 쇠퇴가 반복되며 도시 이미지를 형성하고 인간의 인식에 영향을 미친다(Cividino et al., 2020). 특히, 도시 경관은 도시 활력 혹은 쇠퇴를 직관적으로 보여주는 요인 중 하나로 인간이 도시환경을 인식하는 데에 중요한 역할을 한다. 도시 경관의 변화는 인간의 심리적 반응을 유발하며 도시를 긍정적 또는 부정적으로 인식하는 기준이 된다(Li et al., 2022). 도시 경관에 대한 인간의 인식은 도시 활력과 쇠퇴의 중요한 지표로 작용할 수 있다. 이러한 주관적 인식은 인간이 체감하는 도시 활력 혹은 쇠퇴 수준을 드러낸다(Yang et al., 2023). 따라서 도시 경관 분석을 통해 도시 활력 혹은 쇠퇴에 대한 인간의 인식을 파악하는 것은 변화에 대한 인간의 관점을 이해하고 도시를 개선하기 위한 중요한 단서를 제공할 수 있다.

도시 활력은 경제적, 사회적, 물리적 등의 측면에서 도시가 활성화되는 것을 의미하며 경제성장, 일자리 창출, 인프라 발전 등과 같은 긍정적인 영향을 준다(Yang et al., 2024). 유사한 맥락에서 도시쇠퇴는 여러 측면에서 침체되는 것을 의미하며 경제하락, 인구감소 등의 방향으로 도시에 부정적인 영향을 미친다(Zhang et al., 2018). 도시는 이러한 활력과 쇠퇴의 양상을 나타낼 수 있지만 모든 지역이 동시에 활력과 쇠퇴를 보이는 것은 아니다. 예시로 특정 지역이 상업적으로 번성하여 경제적 활력을 보일 수 있지만 환경적 문제, 사회적 불평등을 유발하여 도시쇠퇴의 요인으로 작용할 수 있다. 반대로 쇠퇴지역도 특정 사업이나 예술 활동을 통해 활력의 요인으로 작용할 수 있다. 이처럼 도시는 복합적인 구조로 되어 있고 활력과 쇠퇴는 단순히 반비례 관계로 설명되지 않는다(Alberti et al., 2018). 그러나 도시 활력에 대한 연구는 정성적 연구, 정량적 연구 등 다양하게 진행되었지만, 도시쇠퇴에 대한 연구는 상대적으로 부족한 실정이다.

도시쇠퇴는 일반적으로 물리적 노후도, 인구감소, 사업체 수 감소 등의 객관적인 지표를 통해 판단되며, 도시활력이 저하되는 상태로 정의된다(도시재생사업단, 2010). 이러한 정의는 통계 기반의 쇠퇴 진단에 활용되어왔으나, 정량적 평가만으로는 도시쇠퇴에 대한 인간의 주관적 인식을 충분히 설명하기 어렵다(von Wirth et al., 2015). 또한, 서울시 2030 도시재생전략계획의 주된 추진전략 중 하나는 “시민수요대응”이지만(서울특별시, 2023), 현재의 쇠퇴진단지표로는 시민 수요를 충분히 반영하기 어렵다. 이에 따라 쇠퇴경관에 대한 인간의 주관적 인식을 파악할 필요가 있다. 본 연구는 이러한 문제의식을 바탕으로 도시쇠퇴를 단순히 ‘활력이 줄어든 상태’로 간주하지 않고, 도시 가로경관에서 물리적 노후, 유지관리 미흡, 공실 등과 같은 부정적 환경

요소들이 직관적으로 인식되는 상태로 개념화하였다. 이러한 개념적 정의는 도시쇠퇴에 대한 인간의 주관적 인식을 실질적으로 반영하는 것으로 정량적 지표 중심의 기존 진단 체계를 보완할 수 있는 새로운 시각을 제시한다.

따라서 본 연구에서는 가로경관 이미지를 기반한 주관적 쇠퇴경관 설문조사 자료를 활용하여 쇠퇴경관에 대한 인간의 인식과 이에 영향을 미치는 물리적 환경요인을 분석하고자 한다. 또한, 본 연구를 통해 서울시 쇠퇴경관 저감과 스마트 도시계획 및 관리에 기여할 수 있는 정책적 시사점을 도출하고자 한다. 이를 위해 본 연구의 흐름은 <그림 1>과 같이 진행되며, 세 가지의 분석 내용을 설정하였다. 먼저, 쇠퇴경관에 대한 인간의 주관적 인식을 정량적으로 추출하기 위해 가로경관 이미지를 활용하여 이미지 쌍별 비교 설문조사를 진행하고 Trueskill 알고리즘을 적용하였다. 다음으로 의미론적 분할(Semantic Segmentation)과 멀티모달 대규모 언어모델(Multi-Modal Large Language Model, MLLM)을 통해 물리적 환경요인의 비율(Quantity)과 상태(Quality) 점수를 도출하였다. 마지막으로 기계학습과 해석가능한 기계학습을 사용

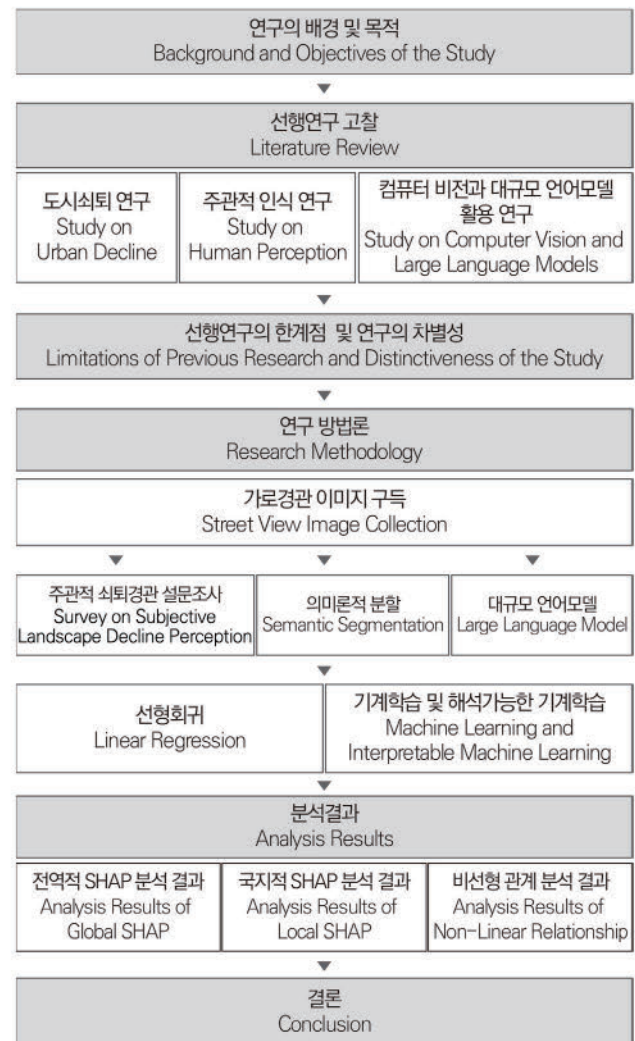


그림 1. 연구의 흐름도  
Figure 1. Research flow chart

하여 쇠퇴경관에 대한 인간의 주관적 인식과 물리적 환경요인 간의 비선형적 관계를 분석하여 정책적 시사점을 도출하였다.

이를 위한 본 연구의 첫 번째 목적은 가로경관 이미지를 활용하여 서울특별시의 쇠퇴경관에 대한 인간의 인식을 분석하는 것이다. 다음, 두 번째 목적은 쇠퇴경관에 대한 인간의 주관적인 인식에 영향을 미치는 도시 내 물리적 환경요인을 분석하고자 한다. 이에 따른 본 연구의 질문은 다음과 같다: 첫째, 인간이 인식하는 쇠퇴경관에 영향을 미치는 물리적 환경요인은 무엇인가? 둘째, 물리적 환경요인과 쇠퇴인식 간에 비선형 관계가 존재하는가? 해당 연구 질문을 해결하기 위해 가로경관 이미지를 활용하여 도출한 쇠퇴경관에 대한 인간의 주관적 인식을 정량화하여 종속변수로 설정하였다. 또한, 쇠퇴경관에 대한 인간의 주관적 인식에 영향을 미치는 물리적 환경요인을 독립변수로써 구축하기 위해 미분적 분할 기법과 멀티모달 대규모 언어모델을 활용하였다.

## II. 선행연구 고찰

### 1. 도시쇠퇴 관련 연구

한국은 도시쇠퇴를 판단하기 위해 도시재생 종합정보체계에서 제공하는 쇠퇴진단지표(인구감소, 사업체 수 감소, 물리적 노후도)를 사용하고 있다(도시재생사업단, 2010). 쇠퇴진단지표에서 판단하는 물리적인 노후도는 건축물의 건축연도, 신규 건축물의 개수, 노후 건축물의 개수 등으로 좌우된다. 도시쇠퇴를 판단하고자 하는 지역의 건축연도가 오래되고, 노후 건축물이 많으며 신규 건축물이 적으면 물리적인 노후도가 높다고 판단되어 쇠퇴 지역이라고 판단한다(도시재생사업단, 2010). 이와 같이 객관적 지표만을 사용하여 물리적 환경의 쇠퇴를 진단한다. 그러나, 재건축, 재개발 등의 변수로 인해 건축물 개수 혹은 건축연도만으로 물리적 환경에 대해 판단하는 것은 다소 무리가 있을 수 있다.

조운애(2014)는 선행연구의 쇠퇴진단지표 연구에서 지표 선정 과정과 산출 과정에는 상이한 측면이 존재하며, 이에 따라 지역마다 상충하는 결과가 도출될 수 있다고 말했다. 따라서 개별지표에 따른 쇠퇴지역 선정의 한계점에 대응하는 '쇠퇴지수D'를 제안하였으며, 이를 위해 인구수와 도시 면적의 증가율을 활용하였다. 쇠퇴지수D는 단위가 없기 때문에 표준화할 필요가 없고, 가중치를 계산하지 않아도 되기 때문에 편리한 활용이 가능하다는 이점이 있지만, 사용되는 지표가 두 가지이기 때문에 초기 연구의 성격이 강하다는 한계점을 가지고 있다. 이정은·박윤미(2021)는 각 개별지표를 이용하여 쇠퇴를 결정하는 것은 지역마다 상이한 결과를 도출한다고 주장한다. 이에 개별지표들을 통합하여 산출한 복합쇠퇴지수를 사용하여 도시쇠퇴를 판단하였다. 이와 같이 도시쇠퇴 관련 연구들은 쇠퇴진단지표의 활용, 새로운 쇠퇴지표의 제안 등의 목적을 가지며 정성적인 연구는 부족한 실정이

다. 더하여 서울시 2030 도시재생전략계획의 새로운 추진전략 중 하나는 사람의 수요에 대응하는 것인 만큼(서울특별시, 2023) 도시쇠퇴에 대한 인간의 주관적인 인식을 수용하고 지표로써 사용해야 할 필요가 있다.

### 2. 주관적 인식 관련 연구

도시 경관에 대한 주관적 인식은 경관에 대한 인간의 인지적 경험을 말하며(Zhang et al., 2023), 물리적 환경에 대한 개인의 주관적 이해, 경험을 의미한다(Sun, 2022). 건물, 차도 등 가로경관 내 물리적 환경요인은 인간의 인식에 큰 영향을 미치며, 도시 이미지를 형성하는 데에도 상당한 영향을 미친다(Dubey et al., 2016).

전통적인 조사는 인터뷰 형식으로 진행되거나 간략한 질문에 대한 척도형 설문조사로 진행되었다. 그러나, 최근 컴퓨터 비전의 발달로 다양한 형태의 설문조사를 진행할 수 있게 되었다(Zhang et al., 2023). 이에 따라 도시 경관에 대한 인간의 인식을 분석하는 연구는 급속하게 발달하였다. 이러한 발전은 인간의 생각을 파악하여 도시계획 및 의사 결정에 명확한 근거를 제공할 수 있게 되었다. Zhang et al.(2018)은 가로경관 이미지를 활용하여 설문조사를 진행하였다. "안전함", "아름다움", "우울함", "활기찬", "부유함", "지루함" 등의 여섯 가지 감정에 대한 응답자의 인식을 수집하였다. 물리적 환경요인과 인식 간의 상관관계를 파악하기 위해 다변량 회귀 분석을 진행하였으나 해당 연구에서는 물리적 환경요인이 여섯 가지 인식에 영향을 미치는 기여도까지만 파악되었으며 비선형 관계에 대해서는 고려되지 않았다.

Xu et al.(2023)은 도시 협곡에서 인식되는 물리적 환경요인과 여섯 가지 인식 간의 비선형 관계를 파악하였다. 그러나 해당 연구에서는 한 가지의 기계학습을 사용하였기 때문에 타 모형과의 성능 비교가 이루어지지 않았다. 이외에도 도시활력에 관한 주관적 인식을 조사한 연구도 존재한다(Wedyan and Saeidi-Rizi, 2024). 인간이 인식하는 도시 안전에 대한 연구가 진행되었으며 인식된 안전은 인간의 정신건강과 직결되었음을 밝혀냈다(van Dinter et al., 2022). 이렇게 가로경관 이미지 기반의 설문조사를 통해 수집한 주관적 인식은 향후 도시계획에 중요한 근거가 될 수 있다(Xie et al., 2024).

### 3. 컴퓨터 비전과 대규모 언어모델 활용 연구

도시 연구에서 가로경관 이미지는 점점 더 많이 활용되고 있으며 컴퓨터 비전의 발전과 관련이 있다(Park et al., 2023). 연구자들은 가로경관 이미지를 제공하는 초기부터 도시 환경을 평가할 수 있는 잠재력을 인식해왔다(Kelly et al., 2013; 박준상, 2023). Google Street View(GSV) 및 Naver Street View(NSV) 등과

같은 서비스는 가로경관 이미지를 제공하여 도시의 통찰력을 제공한다(Helbich et al., 2021). 따라서 가로경관 이미지는 도시환경을 분석하기 위한 도구로써 잠재력이 높다고 판단된다. 기동환 외(2021)의 연구에서는 서울시 녹지변수를 구축하고 녹지의 형평성을 분석하였다. 가로경관 이미지를 통해 보행자 관점의 녹지를 측정하고 가로녹지율을 산출한 후 녹지분포의 불평등을 완화하기 위한 시사점을 제시하였다. He et al.(2023)의 연구에서는 가로경관 이미지를 통해 인간의 감성적 인식을 추출하여 도시 재개발을 평가함으로써 종합지표를 구축했다. 기존 지표와 새로운 지표를 비교분석한 결과, 제안된 지표가 기존 지표보다 더 높은 평가 정확도를 나타냈다. 해당 연구에서는 Street View Image(SVI)를 사용함으로써 건축환경의 공백을 보완하였다.

한편, 최근 MLLM이 도시 데이터를 분석하는 데 중요한 역할을 하고 있다(Li et al., 2024). 컴퓨터 비전이 이미지를 기반으로 도시환경을 분석하는 데 중점을 둔다면 MLLM은 이미지와 텍스트를 결합하여 풍부한 해석을 할 수 있다(Li et al., 2024). 이는 기존 이미지 분석만으로는 해석하기 어려운 새로운 통찰을 파악할 수 있다는 것을 의미한다(Zhanga et al., 2024). 도시계획 분야에서도 점차 LLM을 활용하여 도시 데이터를 증강하여 더욱 정교한 분석에 도출하고 있다. 특히 MLLM의 경우, 이미지와 텍스트 데이터를 함께 분석하여 통합적인 관점을 제시할 수 있다. Driessen et al.(2024)의 연구에서는 GPT-4o를 사용하여 교통 이미지에서 인간이 인식하는 위험도를 예측하였다. 교통 상황에서 인간의 위험도 인식을 예측하는 데 있어서 GPT-4o가 얼마큼 유용한지 분석하였다. GPT-4o가 평가한 위험도는 인간이 평가한 위험도와 높은 상관관계를 나타냈다. 이는 GPT-4o가 복잡한 이미지 데이터를 기반으로 도시를 평가하는 데 유용하게 작용한다는 것을 의미한다. 해당 연구 결과는 MLLM이 도시 연구에 있어서 유용하게 활용될 수 있음을 시사한다.

이처럼 SVI나 MLLM 등의 활용은 도시 공간의 데이터와 컴퓨터 비전 기법 발전 사이의 간극을 메우며 더 나아가 도시 공간 분석에서 자원을 절약하고 더욱 세밀한 분석을 가능하게 하는 이점이 있다. 컴퓨터 비전 기법과 더불어 MLLM을 함께 활용하는 것은 도시 분야에서 중요한 연구 도구로 자리 잡고 있다.

#### 4. 연구의 차별성

한국은 쇠퇴진단지표를 사용하여 도시쇠퇴를 판단하지만, 인간의 주관적 인식은 반영하기에는 한계가 있다. 이는 인간의 수요에 대응하고자 하는 서울시 2030 도시재생전략계획의 추진전략과는 다른 방향을 가리키기 때문에 정성적 지표를 활용할 필요가 있다. 많은 연구에서 가로경관 이미지를 활용하여 도시의 다양한 측면에 대한 인간의 주관적 인식을 추출하고 있다. 대체로 MIT Place Pulse에서 정의한 인식을 지표로 활용하지만, 인간의

인식을 여섯 가지 범주로 정의하기에는 한계가 있다. 따라서 각각의 연구 목적에 부합하는 주관적 인식지표를 사용할 필요가 있다. 마지막으로 대부분의 컴퓨터 비전 활용 연구는 의미론적 분할, 객체 탐지와 같은 딥러닝 기법을 활용하여 가로경관 이미지 내 인식할 수 있는 물리적 환경요인을 정량화한다. 그러나, 이러한 딥러닝 기법은 인식되는 물리적 객체의 비율이나 개수만을 판단할 수 있고 물리적 환경요인의 상태는 판단하지 못한다.

이에 따른 본 연구의 차별성은 다음과 같다. 첫째, 도시쇠퇴 연구에서 주관적 인식을 고려한 연구가 미비하다는 점에서 본 연구에서는 인간이 인식하는 쇠퇴경관을 분석하고 어떠한 환경적 요인이 쇠퇴경관에 영향을 미치는지 분석한다. 둘째, Place Pulse에는 도시쇠퇴 인식에 직접적으로 대응하는 지표가 없으며, 도시쇠퇴는 단순히 활력의 저하로 설명할 수 없는 복합적인 현상이기 때문에 활기참의 역값을 활용하여 도시쇠퇴 인식을 추정하는 방식은 개념적으로 타당하지 않다고 판단하였다. 본 연구의 목적에 부합하는 인식지표인 쇠퇴경관과 관련된 인간의 주관적인 인식을 활용한다. 이에 따라 본 연구의 도시쇠퇴는 “도시 가로경관에서 부정적인 환경요소들이 직관적으로 인식된 상태”로 개념화하였다. 셋째, MLLM을 활용하여 물리적 환경요인의 비율뿐만 아니라 각 객체의 상태(Quality)까지 고려하여 기존 딥러닝 기법의 한계점을 보완한다. 결과적으로 본 연구는 기존 딥러닝 기법과 더불어 MLLM을 활용하여 인간의 주관적 인식을 기반으로 도시쇠퇴를 평가하는 새로운 접근법을 제시함으로써 기존 연구와 차별성을 가진다.

### III. 연구 방법론

#### 1. 연구의 범위

본 연구는 서울특별시의 쇠퇴에 대한 인간의 주관적 인식을 분석하며, 공간적 범위는 서울특별시, 시간적 범위는 2018년으로 설정하였다. COVID-19 팬데믹은 자영업 매출의 급감, 임대료 부담 상승, 상가 공실 증가 등으로 이어졌으며(조영식 외, 2021), 이러한 변화는 도시쇠퇴에 대한 주관적 인식에도 영향을 미칠 수 있다. 건물 외관에 “임대”, “철거” 등의 문구는 쇠퇴한 가로경관으로 인식될 수 있음이 확인되었다(남지현 외, 2019). 이에 따라 본 연구에서는 팬데믹의 영향을 배제하고 일반적인 도시 경관에 대한 인식을 분석하기 위해, GSV에서 제공하는 팬데믹 이전의 최근 연도인 2018년을 시간적 범위로 설정하였다. 또한, 근린 환경을 분석하는 선행연구에서는 대부분 500m 내외의 그리드를 사용하고 있으며 대표적으로 500m를 적정 보행거리로 언급한다. 본 연구의 첫 번째 목적은 도시쇠퇴에 대한 인간의 인식을 분석하는 것이다. 선행연구에 따라 보편적으로 사용하고 있는 보행거리인 500m 그리드를 가로경관 이미지 수집 단위로 설정하였다(그림 2). 본 연구

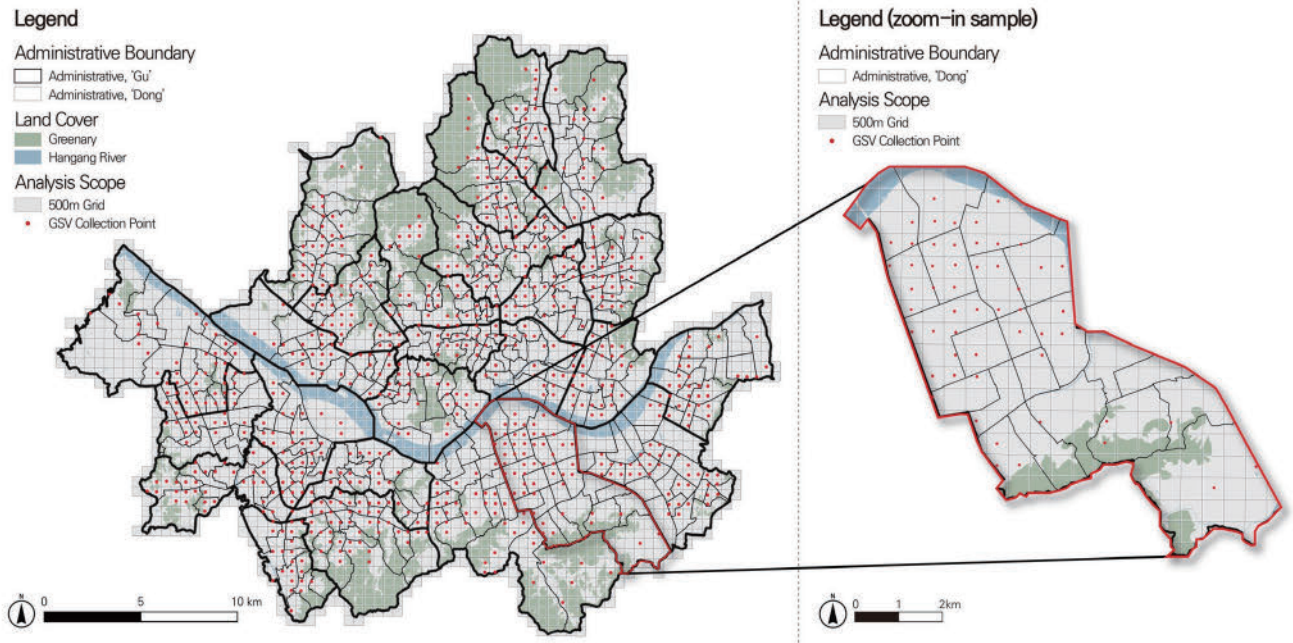


그림 2. 연구의 범위  
Figure 2. Case study area

에서 사용한 데이터는 가로경관 이미지와 주관적 쇠퇴경관 설문 조사이다. 이를 위해 GSV를 활용하여 가로경관 이미지를 수집하였고 이를 기반으로 주관적 쇠퇴경관 설문조사를 진행하였다.

## 2. 분석 자료 및 과정

### 1) 가로경관 이미지 구득

도시 가로경관을 직접적으로 묘사하는 하나의 방법은 SVI를 사용하는 것이다. SVI는 녹지, 가로환경 등의 연구에서 활용되고 있다(Biljecki et al., 2023). SVI의 활용은 물리적 환경요인을 분석함에 따라 도시 현상을 이해하는 데 기여한다(Xu et al., 2023).

본 연구에서는 도시 쇠퇴경관에 대한 인간의 주관적 인식을 분석하기 위해 GSV를 활용하였다. GSV 구득 방법은 다음과 같다. 보행자 관점의 이미지를 수집하고자 12m의 생활도로에서 포인트가 생성되도록 파라미터를 설정한다. 다음으로 인간의 평균 보행거리가 500m임을 고려하여 서울시를 500m 그리드로 분할하여 그리드의 중심점과 50m 간격의 포인트 간의 교차지점을 도출한 후, 이미지의 수집지점으로 설정한다. 해당 지점들의 위도와 경도를 수집하고 GSV API를 통해 이미지를 수집한다. 가로경관 이미지 구득 후에 주관적 쇠퇴경관 설문조사에 사용할 수 있는 이미지 정제 작업을 진행하였다. 일차적으로 수집한 가로경관 이미지 중에서 이미지가 없는 지점에서 촬영된 이미지, 터널 내부에서 촬영된 이미지 등 설문조사에 적합하지 않은 이미지를 제거하였다. 그 결과, 전체 약 1,032장의 가로경관 이미지 중에서 전체 리 진행 후 740장의 이미지를 연구에 활용하였다.

### 2) 주관적 쇠퇴경관 설문조사

본 연구에서 진행한 설문조사 내용은 다음과 같다. “두 이미지 중 더 쇠퇴한 경관을 고르시오.”라는 지문이 제시되며, 응답자는 두 개의 이미지 중에서 더 쇠퇴하였다고 판단되는 이미지를 선택하게 된다. 선행연구에 따르면, 평균 약 20분 내외의 지속시간을 가지는 설문조사는 결과의 질에 유의미한 영향을 미치지 않는다(김권현 외, 2015). 이에 따라 총 10분이 넘지 않도록 사전 실험을 통하여 설문조사를 설계하였으며, 해당 시간 안에 완료될 수 있는 서른 쌍의 설문조사 이미지 세트를 구성하였다. 응답자의 개인 특성으로는 성별과 연령대 항목을 포함하였으며, 한 세트의 설문지는 총 30문항으로 구성되었다. 총 1,256명이 설문조사에 참여하였으며, 응답자의 성별과 연령대 특성은 <표 1>과 같다.

표 1. 설문조사 응답자 특성(단위: 명)

Table 1. Characteristics of survey respondents (unit: person)

| Division           | Male            | Female          | Total             |
|--------------------|-----------------|-----------------|-------------------|
| 10대<br>Ages 10-19  | 92<br>(15.5%)   | 4<br>(0.6%)     | 96<br>(7.64%)     |
| 20대<br>Ages 20-29  | 210<br>(35.5%)  | 449<br>(67.4%)  | 659<br>(52.4%)    |
| 30대<br>Ages 30-39  | 124<br>(21%)    | 93<br>(13.9%)   | 217<br>(17.2%)    |
| 40대<br>Ages 40-49  | 84<br>(14.2%)   | 51<br>(7.65%)   | 135<br>(10.7%)    |
| 50대 이상<br>Ages 50+ | 80<br>(13.5%)   | 69<br>(10.3%)   | 149<br>(11.8%)    |
| 합계<br>Total        | 590<br>(100.0%) | 666<br>(100.0%) | 1,256<br>(100.0%) |

본 연구에서 주관적 쇠퇴경관 설문조사의 결과를 쇠퇴점수라고 정의하였다. 쇠퇴점수는 설문조사 결과를 정량화한 값이다. 설문조사 결과를 정량화하기 위하여 Trueskill 알고리즘을 사용하였다. Trueskill 알고리즘은 베이지안 확률 추론 방법을 기반으로 게임 플레이어의 순위를 결정하는 Microsoft에서 개발한 랭킹 시스템이다. 게임 플레이어의 실력을 평균값(점수)과 표준편차(불확실성)로 나타낸다. 기본적으로 설정된 평균값과 표준편차는 25와 8.3이며 게임이 순차적으로 진행될 때마다 평균값과 표준편차가 실시간으로 업데이트된다. 단순히 점수에 따라서 순위가 결정되는 것이 아닌 불확실성이 순위 결정에 포함되기 때문에 결과의 신뢰성을 높일 수 있는 것이 장점이다. 본 연구에서는 설문조사 결과를 정량화하기 위하여 사용하였으며 불확실성은 사람들이 이미지를 선택할 불확실한 확률을 말한다. 응답자가 특정 경관 이미지를 선택했을 때, 이미지가 선택되는 불확실성을 설문조사 결과에 적용하여 선택의 신뢰성을 높이고 설문조사의 결과를 정량화하여 쇠퇴점수로 도출하기 위해 Trueskill 알고리즘을 사용하였다.

### 3) 의미론적 분할과 멀티모달 대규모 언어모델

도시 내 물리적 환경 객체의 비율과 상태를 파악하기 위해 의미론적 분할과 MLLM을 활용하였다. 의미론적 분할이란 이미지 내 각 픽셀을 사전에 정의된 객체로 분할하는 것을 의미한다(유승재 외, 2021). 본 연구에서 의미론적 분할을 적용하기 위해 사용된 딥러닝 모델은 Seg-former이다. 이는 멀티 스케일 기능을 출력하는 디코더로 구성되어 있고 다층 퍼셉트론 방식을 사용하기 때문에 효율적으로 진행되어 분할 시간을 줄일 수 있다는 장점이 있다. Seg-former에 사용된 데이터 세트는 ADE20K로 다양한 이미지를 포함하고 있는 대규모 이미지 데이터이다. 본 연구에서는 선행연구 고찰에 따라 하늘, 녹지, 잔디, 차량, 빌딩, 벽, 차도, 보도의 여덟 가지 변수를 사용하였다. 물리적 환경 객체의 비율은 값이 클수록 인간이 인식할 수 있는 요소의 비율이 높다는 것으로 해석된다(기동환 외, 2021).

다음, 물리적 환경 객체의 상태를 파악하기 위해 MLLM을 사용하였다. 최근 SVI를 분석하기 위한 방법 중 하나로, MLLM의 활용이 대두되고 있다. MLLM은 기존 LLM에서 발전된 모델로 이미지와 텍스트 데이터 두 가지 유형의 처리가 가능하여 정교한 데이터 분석이 가능하다(OpenAI, 2023). 본 연구에서 사용된 GPT-4o는 복잡한 이미지 분석과 언어적 추론이 필요한 작업에 적합하다(Jin et al., 2024). 이미지를 분석하여 건물, 도로 등의 상태를 해석할 수 있으며 기존의 비전 모델로는 측정하기 어려운 정성 분석이 가능하다는 강력한 장점이 있다(OpenAI, 2023). 물리적 환경의 주관적인 평가 지표를 도출하여 정량화할 수 있어 이러한 특성은 도시 연구에서 유용하게 적용될 수 있다(Zhang et al., 2024).

의미론적 분할은 물리적 환경을 비율로 인식할 수 있지만 객체의 상태는 알 수 없는 한계점이 존재한다. <표 2>와 같이 유사한 녹지 비율을 가지는 두 가로경관 이미지에서 MLLM을 통해 평가한 녹지 상태의 점수는 큰 차이를 나타내고 있다. 본 연구는 GPT-4o를 활용하여 가로경관 이미지 내 인식되는 물리적 환경 객체의 상태를 점수로 추출하였다. 이를 위해 실험한 프롬프트는 총 11가지이다(부록의 <표 1>). 최종적으로 선택된 프롬프트는 가로경관 이미지 내 인식되는 물리적 환경 객체의 상태를 점수로 평가하도록 설계되었다. 실험을 진행했던 11가지 프롬프트 중 모든 명령에 일관성 있게 결과를 도출하는 프롬프트는 9번째 프롬프트이기 때문에 최종적으로 연구에 활용하였다(표 2).

### 4) 기계학습 및 해석가능한 기계학습

본 연구는 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하기 위해 기계학습과 해석가능한 기계학습을 사용하였다. 기계학습을 활용하면 전통적인 회귀분석에서 파악하기 어려웠던 새로운 관계를 확인할 수 있으며 비선형으로 파악되는 변수 간의 관계를 확인할 수 있다. 비선형 관계를 확인하면 특정 변수의 임계점을 함께 확인할 수 있으며(Park et al., 2023), 이를 통해 이전 연구와는 다른 결과가 확인될 수 있다. 본 연구에서는 선형회귀 모델과 세 가지 기계학습을 사용하여 평가하고 성능이 가장 좋은 모델을 선택하여 쇠퇴경관 영향요인 분석 모델로 활용하였다. 본 연구에서 사용한 기계학습 모델은 XGBoost, GBM, RF 이다. 이는 모두 여러 모델을 단일 모델로 결합하는 앙상블 기법이다. XGBoost는 트리 기반의 앙상블 학습에서 주목받는 알고리즘 중 하나이다. GBM의 단점인 느린 수행 시간, 과적합 등의 문제를 극복한 모델이다. GBM은 다양한 약한 학습기를 순차적으로 학습하면서 예측이 잘못된 데이터에 가중치를 부여하여 오류를 개선해 나간다. RF는 같은 알고리즘으로 여러 개의 분류기를 만들어서 최종결정하는 알고리즘이다. 앙상블 기법 중에서 빠른 속도를 가지고 있으며 높은 예측 성능을 가지고 있다. 위 세 가지 기계학습의 공통점은 여러 개의 결정 트리 모델을 결합하여 하나의 결정 트리 모델보다 더 좋은 성능을 도출한다는 것이다. 그러나 기계학습은 결과의 도출 과정을 파악할 수 없는 블랙박스라는 특징이 있다.

따라서 본 연구에서는 블랙박스를 해석하기 위해 해석가능한 기계학습을 사용한다. 해석가능한 기계학습의 종류는 SHAP, LIME 등이 있다. 이 중 SHAP는 쉽게 사용될 수 있도록 배포되어 있기 때문에 SHAP를 사용하여 주관적 인식과 물리적 환경객체 간의 비선형 관계를 분석하였다. SHAP에는 Global SHAP와 Local SHAP가 있다. Global SHAP는 변수 간의 전역적 연관성을 나타내며 변수의 영향력이 높은 순서대로 위에서부터 배열된다. 막대그래프 색깔의 차이로 변수의 영향력(+/-)을 설명한다. 빨간색은 변수 간의 긍정적 연관성이 있음을 나타내고 파란색은

표 2. GPT-4o 최종 명령 프롬프트와 의미론적 분할 비율과 GPT-4o 점수 비교

Table 2. Final command prompt for GPT-4o and comparison of semantic segmentation ratio and GPT-4o score

| Final command prompt   |          |        |       |
|--|----------|--------|-------|
| This street landscape image contains objects such as car, wall, road, building, grass, sidewalk, tree, and sky.<br>I want to evaluate the urban landscape using this image.<br>Please analyze this image and quantify the condition of each object on a scale from 1 (very poor) to 5 (very good).<br>If an object is not present in the image, assign it a score of 0 (not present).<br>The output format must be in JSON as shown below, and each score should be in float format.<br>Do not include any additional explanations output only the JSON. |          |        |       |
| ##Output format<br>{"car":[car score], "wall":[wall score], "road":[road score], "building":[building score], "grass":[grass score],<br>"sidewalk":[sidewalk score], "tree":[tree score], "sky":[sky score]}   |          |        |       |
| (A)  |          | (B)    |       |
| Street view image  | Object   | Ratio  | Score |
|  | Building | 13.91% | 3     |
|  | Road     | 30.95% | 2     |
|  | Sky      | 10.63% | 2.5   |
|  | Wall     | 18.28% | 2.5   |
|  | Sidewalk | 0.95%  | 1     |
|  | Car      | 0.65%  | 1     |
|  | Tree     | 12.92% | 2     |
|  | Grass    | 2.18%  | 1.5   |
| Street view image  | Object   | Ratio  | Score |
|  | Building | 31.15% | 4.5   |
|  | Road     | 34.05% | 4.5   |
|  | Sky      | 12.40% | 4.5   |
|  | Wall     | 0%     | 0     |
|  | Sidewalk | 3.92%  | 4     |
|  | Car      | 0.88%  | 4     |
|  | Tree     | 11.52% | 4     |
|  | Grass    | 1.18%  | 4     |

변수 간의 부정적 연관성이 있음을 나타낸다. Local SHAP는 변수 간의 관계를 산점도를 이용하여 나타낸다. Y축을 기준으로 오른쪽에 산점도가 몰려있다면 양의 관계를 나타낸다. 반면, 왼쪽에 몰려있다면 음의 관계를 나타내며 파란색으로 나타난다. 또한 빨간색과 파란색의 점이 섞여 있는 패턴을 보인다면 비선형 관계가 있음을 나타낸다.

## IV. 분석결과

### 1. 주관적 쇠퇴경관 설문조사

본 연구에서는 쇠퇴경관에 대한 인간의 주관적 인식을 분석하기 위해 Trueskill 알고리즘을 사용하였다. 주관적 쇠퇴경관 설문조사 결과에 Trueskill 알고리즘을 적용한 결과는 <표 3>과 같다. 알고리즘은 기본 평균값 25와 표준편차 8.3에서 시작한다. 평균값은 이미지가 많이 선택될수록 높아지며 표준편차는 낮아진다. 본 연구에서는 평균값이 높아질수록 많은 사람이 쇠퇴했다고 선택한 경관이다. 본 연구에서는 평균값을 종속변수로 사용하였으며 쇠퇴점수로 정의하였다.

우선, 설문조사 결과의 신뢰성 검토를 위해 Split-Half 방식의 신뢰도 분석을 수행하였다. 설문조사 결과에서 랜덤의 두 집단을 추출한 후, 각 그룹의 쇠퇴점수 간의 상관관계를 피어슨 상관분

표 3. Trueskill 알고리즘을 적용한 설문조사 결과

Table 3. Survey results using the TrueSkill

| Division | File name | Mean   | Std.  |
|----------|-----------|--------|-------|
| 1        | 144068    | 36.846 | 1.197 |
| 2        | 141702    | 35.733 | 1.193 |
| 3        | 37775     | 35.576 | 1.315 |
| 4        | 91495     | 35.194 | 1.131 |
| 5        | 45980     | 35.091 | 1.536 |
| :        | :         | :      | :     |
| 736      | 64495     | 13.580 | 1.427 |
| 737      | 107232    | 11.470 | 1.548 |
| 738      | 115338    | 9.433  | 1.508 |
| 739      | 117028    | 8.377  | 1.482 |
| 740      | 167395    | 4.904  | 1.243 |

석으로 분석하였다. 상관분석 결과, 두 그룹 간의 상관계수는 0.996( $p < 0.01$ )으로 유의미한 양의 상관관계가 있음이 확인되었다. 또한, 전체 신뢰도를 추정하기 위해 Spearman-Brown 방법을 활용하였다. Split-Half 방식은 전체 결과의 절반에 대한 신뢰도이기 때문에 전체에 대한 신뢰도를 추정하기 위해서는 보정을 적용해야 한다(Pronk et al., 2022). 따라서 Spearman-Brown

방법을 적용하였으며, 0.998( $p < 0.01$ )의 신뢰도가 나타났다. 이는 설문조사 결과가 내적으로 매우 일관된 평가 결과를 나타낸다는 것을 시사한다.

다음 <그림 3>은 정량화된 설문조사 결과의 공간적 분포이다. 쇠퇴점수가 높아질수록 쇠퇴하였다고 인식한 경관임을 의미한다. 빨간색일수록 쇠퇴점수가 높다는 것을 의미하고 쇠퇴한 경관이라고 인식했다는 것을 말한다. 반면, 파란색일수록 쇠퇴점수가 낮은 것을 의미하며 쇠퇴하지 않은 경관이라고 인식했다는 것을 말한다. 결과에 따르면, 쇠퇴점수가 높은 지역은 양천구, 강북구 등 강북과 강서지역에 집중되어 있는 것을 알 수 있다. 해당 지역의 가로경관 이미지를 확인한 결과, 낡고 관리가 부족한 인프라의 모습이 확인되었다. 이는 해당 지역의 물리적 노후도가 높고 쇠퇴한 경관이라고 인식한다는 것을 시사한다. 반면에 쇠퇴점수가 낮은 지역은 중구, 강남구, 서초구, 송파구 등의 지역에 집중되어 있는 것을 확인할 수 있다. 해당 지역의 가로경관 이미지를 확인한 결과, 자연환경과의 조화로운 경관이 확인되었다. 자연환경과 인공환경이 적절하게 조화로운 경관을 쇠퇴하지 않은 경관이라고 인식하며 긍정적 인식을 유도한다는 것을 시사한다.

## 2. 의미론적 분할과 멀티모달 대규모 언어모델

### 1) GPT-4o 점수 검증

본 연구에서 GPT-4o는 물리적 환경 객체의 상태(Quality)를 평가하기 위해 사용되었다. GPT-4o를 사용하여 도출한 점수의 활용 가능성을 입증하기 위해 쇠퇴점수와와의 피어슨 상관분석을 진행하였다. 우선, <그림 4>의 왼쪽은 GPT-4o 점수와 쇠퇴점수의 이동평균 추세를 확인한 결과이다. 빨간색은 쇠퇴점수이며, 파란색은 GPT-4o 점수를 가리킨다. 두 점수의 추세를 확인한 결과, 두 점수 모두 전반적으로 우상향하는 추세를 보이며 유사한 패턴을 나타내고 있다. 두 점수의 추세선이 일치한다는 것은 GPT-4o가 각 물리적 환경 객체들의 상태를 인간의 주관적인 인식과 유사한 방식으로 평가하고 있음을 시사한다. 다음으로 <그림 4>의 오른쪽 그림은 GPT-4o 점수와 쇠퇴점수의 피어슨 상관분석 결과이다. 그 결과, 쇠퇴점수와 GPT-4o 점수는 0.5875( $p < 0.01$ )의 유의미한 양의 상관관계가 있는 것으로 확인됨으로써 이는 GPT-4o의 활용 가능성을 입증할 수 있음을 시사한다.

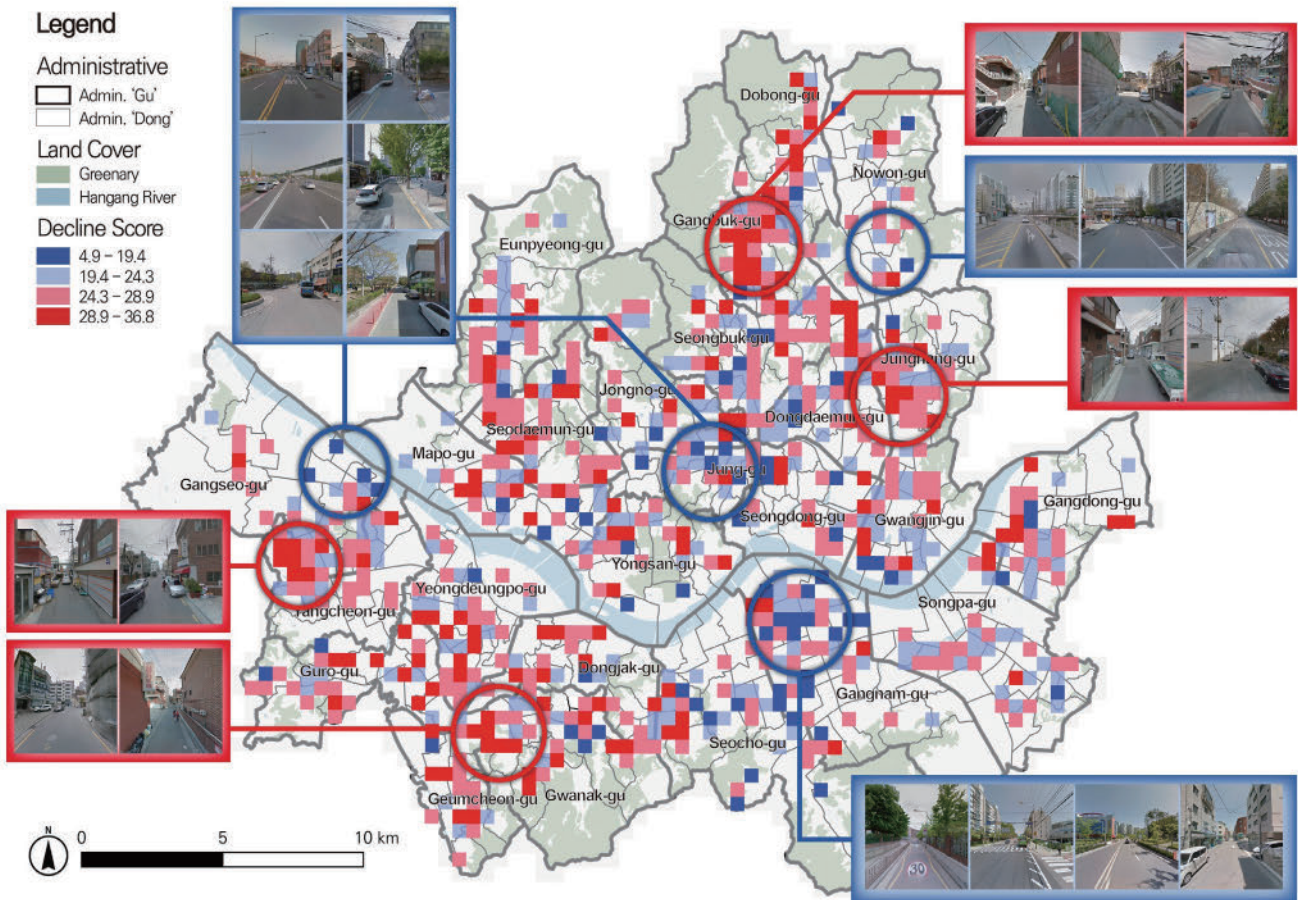


그림 3. 주관적 쇠퇴경관 설문조사 시각화  
 Figure 3. Visualization of subjective urban landscape decline survey

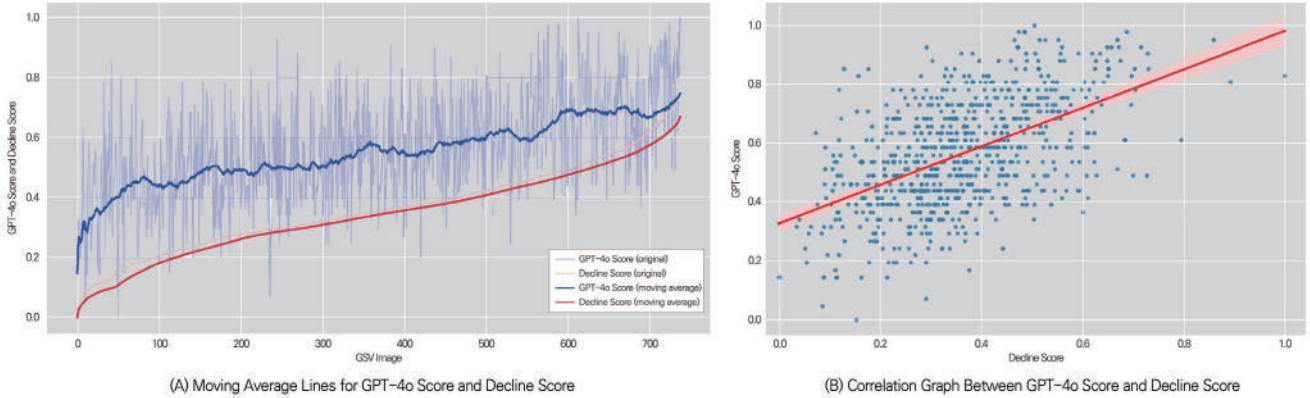


그림 4. GPT-4o 점수와 쇠퇴점수 간의 상관관계 분석

Figure 4. Correlation analysis between GPT-4o score and decline score

2) 의미론적 분할과 GPT-4o 점수

다음 <표 4>는 인간이 가장 쇠퇴하였다고 인식한 경관(<표 4>의 (A))과 가장 쇠퇴하지 않았다고 인식한 경관(<표 4>의 (B))이다. 두 가로경관 이미지를 비교한 결과, 가장 쇠퇴하였다고 인식

한 경관은 건물, 차도, 하늘, 벽 등의 순서로 비율이 높았으며 자연환경 요인은 검출되지 않았다. 또한, 가장 쇠퇴하지 않았다고 선택한 경관은 보도, 나무, 하늘, 건물 등의 순서로 비율이 높았다. 의미론적 분할 결과에 따르면, 인간은 자연적 환경과 인공적

표 4. 의미론적 분할 결과와 GPT-4o 점수 결과

Table 4. Semantic segmentation ratio and GPT-4o score

| Street View Image |  | After semantic segmentation |  | Decline score |              |              |     |
|-------------------|--|-----------------------------|--|---------------|--------------|--------------|-----|
|                   |  |                             |  | 36.846        |              |              |     |
|                   |  |                             |  | <b>Object</b> | <b>Ratio</b> | <b>Score</b> |     |
|                   |  |                             |  | 차량            | Car          | 0.26%        | 2   |
|                   |  |                             |  | 벽             | Wall         | 2.04%        | 3   |
|                   |  |                             |  | 차도            | Road         | 27.77%       | 2   |
|                   |  |                             |  | 건물            | Building     | 44.49%       | 3   |
|                   |  |                             |  | 잔디            | Grass        | 0.00%        | 1   |
|                   |  |                             |  | 보도            | Sidewalk     | 0.44%        | 2   |
|                   |  |                             |  | 나무            | Tree         | 0.00%        | 1   |
|                   |  |                             |  | 하늘            | Sky          | 18.61%       | 2   |
|                   |  |                             |  | 4.904         |              |              |     |
|                   |  |                             |  | <b>Object</b> | <b>Ratio</b> | <b>Score</b> |     |
|                   |  |                             |  | 차량            | Car          | 0%           | 0   |
|                   |  |                             |  | 벽             | Wall         | 5.68%        | 4.5 |
|                   |  |                             |  | 차도            | Road         | 5.70%        | 4   |
|                   |  |                             |  | 건물            | Building     | 5.99%        | 4   |
|                   |  |                             |  | 잔디            | Grass        | 0.68%        | 4.5 |
|                   |  |                             |  | 보도            | Sidewalk     | 37.75%       | 4.5 |
|                   |  |                             |  | 나무            | Tree         | 36.46%       | 4.5 |
|                   |  |                             |  | 하늘            | Sky          | 7.94%        | 4   |

환경이 적절한 비율로 이루어진 경관을 쇠퇴하지 않았다고 인식하는 것으로 나타났다. 또한, GPT-4o 점수에 따르면 인간이 쇠퇴하지 않았다고 판단한 경관의 GPT-4o 점수가 월등히 높게 평가되었다.

### 3. 선형회귀와 기계학습 모형 평가

설문조사, 의미론적 분할, GPT-4o를 통하여 정의된 본 연구의 변수는 <표 5>와 같다. 본 연구에서 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하기 위해서 선형회귀와 기계학습을 사용하였다. 선형회귀 모형과 세 가지 기계학습의 성능을 비교하고, 설명력이 높은 모형을 선택하여 분석에 활용하였다. 각 모형의 설명력을 높이고, 과적합을 최소화하기 위해 파라미터를 조정하였다. 파라미터는 기계학습을 학습할 때, 설명력을 높이기 위하여 조정하는 모형의 변수이다. 해당 변수의 조정에 따라서 모형은 다른 설명력을 가질 수 있다. 본 연구에서 최적의 파라미터를 설정하기 위해 Grid Search를 사용하였다. 이는 조절하고자 하는 변수의 그리드를 설정하고 그리드를 줄여나가면서 기계학습에 알맞은 최적의 파라미터를 찾아가는 방법이다. 본 연구에서는 Grid Search를 사용하여 각 기계학습 별로 가장 높은 설명력을 보이는 하이퍼 파라미터를 설정한 후, 기계학습 모형의 학습을 진행하였으며 본 연구에서 사용한 하이퍼 파라미터는 <표 6>과 같다.

본 연구에서는 기계학습 모형에 대한 추론을 진행하기 위하여

학습과 평가 데이터로 구분하였다. 학습과 평가 데이터의 비율은 선행연구 고찰에 따라 8:2의 비율로 구분하였다. 총 데이터의 80%로 모형에 학습한 후 모형이 학습된 데이터를 잘 설명하는지 확인하기 위해 20%의 평가 데이터로 추론을 진행하였다. 따라서 본 연구에서는 학습 데이터에 대한 설명력으로 기계학습 모형의 설명력을 비교하였다.

기계학습 모형의 성능을 비교하기 위한 지표는  $R^2$ , MSE, MAE를 사용하였다.  $R^2$ 는 독립변수가 종속변수를 얼마나 설명하는지를 평가하는 지표이다. MSE는 오차들의 평균을 의미하며 MAE는 오차 평균의 절댓값을 의미한다. 본 연구는 영향요인을 분석하기 위한 모형을 구축하기 때문에  $R^2$ 가 가장 높은 기계학습 모형을 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하기 위한 최종 모형으로 선정하였다. 또한, 본 연구에서는 GPT-4o 점수를 가중치로 적용한 변수를 독립변수로 활용하였으며 이를 심층적으로 평가하기 위해 의미론적 분할 비율만을 독립변수로 활용한 모형을 추가하여 두 모형의 설명력을 비교하였다. 그 결과, 두 모형 모두 XGBoost가 LR, GBM, RF보다 우수한 설명력을 가진 것으로 나타났으며 의미론적 분할 비율만을 독립변수로 적용한 모형보다 GPT-4o 점수를 가중치로 적용했을 때 모형의 설명력이 더 높아진 것을 확인하였다. 이는 물리적 환경 객체의 비율만 사용하여 쇠퇴경관을 분석하는 것보다 물리적 환경 객체의 비율과 상태를 모두 활용하여 쇠퇴경관을 분석하는 것이 정확한 분석이 가능함을 시사한다. 따라서 의미론적 분할 비율에

표 5. 변수 정의 및 출처

Table 5. Definition of variables and data source

| Variables                     | Definition            |  | Data source                                    |
|-------------------------------|-----------------------|--|--|
| 종속변수<br>Dependent variable    | 쇠퇴점수<br>Decline score | 도시쇠퇴에 대한 인간의 주관적 인식<br>Human subjective perception of urban decline  | 연구자 설계 설문조사<br>Survey designed by a researcher |
| 독립변수<br>Independent variables | 차량<br>Car             | 차량 픽셀 비율 × 차량 상태 점수<br>Car pixel ratio × Car quality score           | Google street view (2018)<br>GPT-4o            |
|                               | 벽<br>Wall             | 벽 픽셀 비율 × 벽 상태 점수<br>Wall pixel ratio × Wall quality score           | Google street view (2018)<br>GPT-4o            |
|                               | 차도<br>Road            | 차도 픽셀 비율 × 차도 상태 점수<br>Road pixel ratio × Wall quality score         | Google street view (2018)<br>GPT-4o            |
|                               | 건물<br>Building        | 건물 픽셀 비율 × 건물 상태 점수<br>Building pixel ratio × Building quality score | Google street view (2018)<br>GPT-4o            |
|                               | 잔디<br>Grass           | 잔디 픽셀 비율 × 잔디 상태 점수<br>Grass pixel ratio × Grass quality score       | Google street view (2018)<br>GPT-4o            |
|                               | 보도<br>Sidewalk        | 보도 픽셀 비율 × 보도 상태 점수<br>Sidewalk pixel ratio × Sidewalk quality score | Google street view (2018)<br>GPT-4o            |
|                               | 나무<br>Tree            | 나무 픽셀 비율 × 나무 상태 점수<br>Tree pixel ratio × Tree quality score         | Google street view (2018)<br>GPT-4o            |
|                               | 하늘<br>Sky             | 하늘 픽셀 비율 × 하늘 상태 점수<br>Sky pixel ratio × Sky quality score           | Google street view (2018)<br>GPT-4V            |

표 6. 기계학습 모형 평가 결과 및 조정 매개변수

Table 6. Results of machine learning model evaluation and adjusted parameters

| Model      | Semantic segmentation ratio only |       |      | GPT-4o score weighted |       |      |
|------------|----------------------------------|-------|------|-----------------------|-------|------|
|            | R <sup>2</sup>                   | MSE   | MAE  | R <sup>2</sup>        | MSE   | MAE  |
| LR         | 0.32                             | 15.11 | 3.11 | 0.43                  | 12.59 | 2.81 |
| XGBoost    | 0.71                             | 6.38  | 1.98 | 0.84                  | 3.48  | 1.44 |
| GBM        | 0.63                             | 7.34  | 2.14 | 0.67                  | 7.02  | 2.05 |
| RF         | 0.43                             | 12.27 | 2.76 | 0.51                  | 10.70 | 2.55 |
| Parameters | Max_depth                        |       |      | 4                     |       |      |
|            | N_estimators                     |       |      | 40                    |       |      |
|            | Learning_rate                    |       |      | 0.2                   |       |      |
|            | Subsample                        |       |      | 0.7                   |       |      |
|            | Colsample_bytree                 |       |      | 0.4                   |       |      |
|            | Gamma                            |       |      | 0.01                  |       |      |
|            | Min_child_weight                 |       |      | 1                     |       |      |

GPT-4o 점수를 가중치로 반영한 변수를 독립변수로 설정한 모형을 최종 모형으로 선정하였다.

#### 4. 쇠퇴경관 영향요인 분석

##### 1) 기술통계분석

선형회귀 분석에 앞서, 각 변수의 평균, 표준편차, 최솟값, 최댓값을 확인하였고, 결과는 <표 7>과 같다. 종속변수는 주관적 쇠퇴경관 설문조사 결과에 Trueskill 알고리즘을 적용한 쇠퇴점수

이다. 독립변수는 가로경관 이미지에 의미론적 분할을 적용하여 도출한 물리적 환경요인의 비율에 GPT-4o를 통해 도출한 물리적 환경요인의 상태 점수를 가중치로써 곱한 값이다. 기술통계 분석 결과, 종속변수인 쇠퇴점수의 평균값은 25.28로 나타났으며 표준편차는 4.67로 비교적 낮기때문에 지역 간 쇠퇴점수의 분포가 균등한 것이 확인되었다. 최솟값은 4.91, 최댓값은 36.85로 나타나 쇠퇴점수의 분포가 극단적이지 않고 안정적인임을 알 수 있다. 이는 인간이 인식하는 서울시의 쇠퇴 수준 차이는 크지 않음을 시사한다.

표 7. 기술통계분석 결과

Table 7. Result of descriptive analysis

| Variables                     |                       | Obs.  | Mean   | S.D.  | Min.   | Max.   | VIF  |
|-------------------------------|-----------------------|-------|--------|-------|--------|--------|------|
| 종속변수<br>Dependent variable    | 쇠퇴점수<br>Decline score | 740   | 25.28  | 4.67  | 4.91   | 36.85  | -    |
| 독립변수<br>Independent variables | 차량<br>Car             | 740   | 14.86  | 15.98 | 0      | 86.88  | 4.57 |
|                               | 벽<br>Wall             | 740   | 5.14   | 13.39 | 0      | 194.95 | 2.64 |
|                               | 차도<br>Road            | 740   | 104.86 | 39.26 | 0      | 211.96 | 2.31 |
|                               | 건물<br>Building        | 740   | 134.59 | 64.14 | 0      | 280.41 | 1.65 |
|                               | 잔디<br>Grass           | 740   | 1.76   | 5.57  | 0      | 62.39  | 1.33 |
|                               | 보도<br>Sidewalk        | 740   | 15.53  | 18.09 | 0      | 156.36 | 1.18 |
|                               | 나무<br>Tree            | 740   | 20.39  | 34.21 | 0      | 202.77 | 1.15 |
| 하늘<br>Sky                     | 740                   | 64.34 | 32.41  | 0     | 201.13 | 1.06   |      |

독립변수의 기술통계 분석 결과, 차량의 표준편차는 15.98로 높은 변동성을 보인다. 일부 경관에서 차량의 비중이 매우 높거나 거의 없는 경관이 있음을 의미한다. 벽은 평균값이 5.14이고 표준편차가 13.39로 나타나 상대적으로 낮은 비중을 나타내었다. 건물은 평균값이 134.59로 가장 높은 값을 나타내며 이는 건물이 차지하는 비중이 크다는 것을 의미한다. 반면에 잔디는 평균값이 1.76으로 매우 낮아 잔디의 비율이 다른 요인에 비해 현저히 낮음을 의미한다. 보도는 평균값이 15.53이고 표준편차가 18.09로 상대적으로 안정적인 분포를 나타내며, 경관에서 일정한 비율을 유지하고 있음을 시사한다. 나무의 평균값은 20.39, 표준편차는 34.21로 나타났으며 하늘은 평균값이 64.34, 표준편차는 32.41로 나타났다. 하늘은 비교적 높은 평균값을 나타내었으며 최댓값은 201.13으로 나타났기 때문에 일부는 폐쇄적인 경관을 보이는 반면, 다른 경관은 개방적인 환경임을 시사한다. 다음으로 다중공선성을 검토하기 위해서 분산팽창지수를 확인한 결과, 5를 초과하지 않아 다중공선성 문제가 없음을 확인하였다.

**2) 선형회귀모형 분석**

다음 <표 8>은 선형회귀분석 결과이다. 먼저 기계학습 모형을 사용하기 전에 선형회귀모형을 사용하여 분석하였다. 선형회귀 결과에 따르면, 전체 변수 중에서 차량, 벽, 차도, 건물, 잔디, 보도, 나무의 변수가 쇠퇴경관과 유의미한 관계를 보이는 것으로 나타났으며, 모두 음의 관계인 것으로 나타났다. 이는 차량, 벽, 차도, 건물, 잔디, 보도, 나무의 비율이 높아질수록 쇠퇴점수는 낮아진다는 것을 의미한다. 이는 인간이 경관을 인식할 때, 차량, 벽, 차도, 건물, 잔디, 보도, 나무의 비율이 높을수록 쇠퇴하지 않은 경관이라고 인식한다는 것을 시사한다. 그러나, 선형회귀 분석에서 유의미한 변수는 많이 도출되었지만, 하늘을 제외한 모든 변수가 유의미한 결과를 보이기 선형적인 관계만으로 쇠퇴경관에 대한 영향요인을 분석하는 것은 심층적인 분석이 어려울 수 있다. 따라서 비선형적인 관계를 분석할 필요가 있다고 판단된다.

**3) 전역적 SHAP와 국지적 SHAP 분석 결과**

다음 <그림 5>는 최종적으로 선정한 XGBoost 기계학습 모형에 SHAP를 적용하여 분석한 결과이다. <그림 5>의 왼쪽은 전역적 SHAP의 결과이다. 전역적 SHAP의 경우, 위에 배치된 막대 그래프부터 변수의 영향력이 높은 순서대로 배치된다. 또한, 막대 그래프의 색깔로 양/음의 관계를 확인할 수 있다. 막대 그래프의 색깔이 빨간색이면 쇠퇴점수와 양의 관계로 설명되고 본 연구에서는 쇠퇴점수가 높아지는 것에 영향을 미치는 것으로 설명된다. 이는 곧 인간이 쇠퇴했다고 인식하는 것을 의미한다. 반대로 막대 그래프의 색깔이 파란색이면 쇠퇴점수와 음의 관계로 설명되며, 쇠퇴점수가 낮아지는 것에 영향을 미치는 것으로 설명된다. 이는 곧 인간이 쇠퇴하지 않았다고 인식되는 것을 의미한다.

**표 8. 쇠퇴점수에 대한 선형회귀분석 결과**

**Table 8. Results of linear regression analysis on decline score**

| Variables               | Decline score |        |      |       |
|-------------------------|---------------|--------|------|-------|
|                         | Coef.         | t      | p> t | β     |
| 차량<br>Car               | -19.327 ***   | -4.39  | 0.00 | -0.13 |
| 벽<br>Wall               | -19.267 ***   | -3.41  | 0.00 | -0.11 |
| 차도<br>Road              | -37.426 ***   | -17.41 | 0.00 | -0.63 |
| 건물<br>Building          | -10.903 ***   | -4.99  | 0.00 | -0.30 |
| 잔디<br>Grass             | -38.261 ***   | -3.16  | 0.00 | -0.09 |
| 보도<br>Sidewalk          | -43.629 ***   | -11.07 | 0.00 | -0.34 |
| 나무<br>Tree              | -26.988 ***   | -9.25  | 0.00 | -0.40 |
| 하늘<br>Sky               | -4.332        | -1.32  | 0.18 | -0.06 |
| const.                  | 39.18         | 30.78  | 0.00 | 39.18 |
| Obs.                    |               |        |      | 740   |
| R <sup>2</sup>          |               |        |      | 0.43  |
| Adjusted R <sup>2</sup> |               |        |      | 0.41  |

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

우선, 선형회귀분석 결과와 비교한 결과, 변수의 중요도가 다르게 나타났다. 선형회귀 결과에서는 하늘은 통계적으로 유의하지 않았으며, 낮은 중요도를 나타내었다. 반면, SHAP 분석에서는 하늘의 중요도가 높게 나타났다. 이러한 차이는 두 분석 비법의 변수 중요도 산정 방식의 차이에서 비롯된다. 선형회귀는 독립변수와 종속변수 간의 선형적 관계를 반영하기 때문에 하늘의 영향력이 단순화될 가능성이 있다. 반면에 SHAP 분석은 변수의 비선형적 영향까지 함께 고려하여 예측에 대한 기여도를 평가하기 때문에 하늘의 기여도가 상대적으로 높게 나타난 것으로 판단된다. 이러한 분석 방법의 차이로 인해 본 연구에서는 SHAP 분석을 통해 선형회귀만으로는 확인하기 어려운 인지적 판단의 복잡성을 정밀하게 설명할 수 있었다. 따라서, 본 연구에서 SHAP 분석을 통해 도출된 변수 중요도는 선형회귀 기반 해석에서 간과되기 쉬운 비선형성을 반영한 결과로 해석할 수 있다. 이는 SHAP 분석이 기계학습 모델의 예측값 변화에 따른 비선형성을 기반으로 해석을 제공함으로써, 전통적인 선형회귀보다 더 풍부하고 정교한 통찰을 제공할 수 있다는 선행연구와 일치한다 (Lundberg and Lee, 2017; Molnar, 2022).

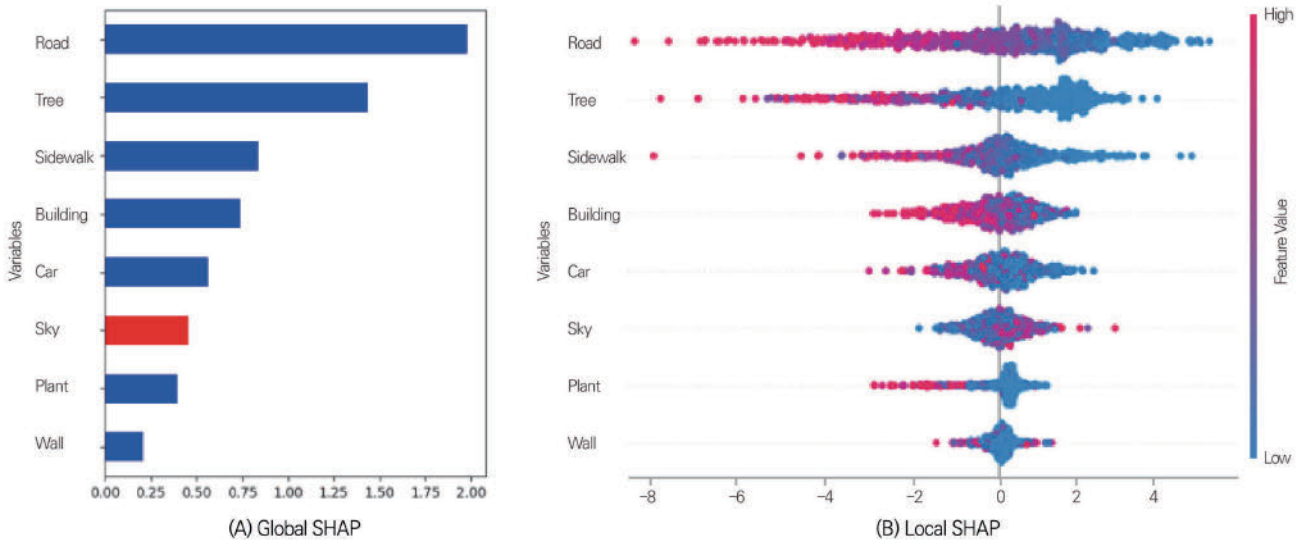


그림 5. 주관적 쇠퇴경관에 영향을 미치는 물리적 환경 객체의 SHAP 분석 결과

Figure 5. SHAP results of physical environmental objects influencing subjective urban landscape decline

주관적 쇠퇴경관에 영향을 미치는 도시 내 물리적 환경 객체는 차도, 나무, 보도, 건물, 차량, 하늘, 잔디, 벽의 순서로 영향력이 높은 것으로 나타났다. 쇠퇴경관에 양의 영향을 미치고 있는 변수는 하늘로 나타났으며 음의 영향을 미치고 있는 변수는 차도, 나무, 보도, 건물, 차량, 잔디, 벽으로 나타났다. 하늘의 비율이 증가하면 쇠퇴점수가 높아지며 이는 인간이 해당 경관은 쇠퇴하였다고 인식한다는 것을 의미한다. 반대로 차도, 나무, 보도, 건물, 차량, 잔디, 벽의 비율이 증가하면 쇠퇴점수가 낮아지며 해당 경우에는 쇠퇴하지 않았다고 인식한다는 것을 의미한다. 국지적 SHAP의 경우에도 가장 위쪽부터 배치된 변수부터 영향력이 높다고 설명된다. 다만, 전역적 SHAP와는 다르게 변수의 영향력을 산점도로 나타내는 것이 특징이며 빨간색 산점도와 파란색 산점도가 섞여 있는 양상이 나타날 경우에 단순하게 양과 음의 관계를 판단하기 어렵다. <그림 5>의 오른쪽을 확인하면 차도, 나무, 보도, 건물, 차량, 하늘의 경우에는 빨간색 산점도와 파란색 산점도가 함께 섞여 있는 양상이 나타나기 때문에 해당 변수들은 비선형적 관계가 있음을 예상할 수 있으며 비선형 관계를 해석하기 위한 추가적인 분석이 필요하다고 판단하였다.

#### 4) 비선형 관계 분석 결과

다음 <그림 6>은 앞선 국지적 SHAP 결과에서 비선형적 결과가 있을 것으로 예상되는 차도, 나무, 보도, 건물, 차량, 하늘 변수를 Scatter Plot으로 시각화한 결과이다. 차도와 건물의 경우에는 약 20-30% 정도까지는 쇠퇴점수 증가에 기여하다가 약 30%를 넘어가면 쇠퇴점수 감소에 기여하는 것으로 나타났다. 이는 인간이 경관을 인식할 때, 차도 및 건물의 비율이 약 20-30% 미만인 경우에는 해당 경관을 쇠퇴하였다고 인식한다는 것을 의미한다. 차도와 건물의 비율이 30% 이상인 가로경관 이미지를 확인

한 결과, 대부분 최소 양복 이차선 도로와 상가가 있는 지역에서 촬영된 가로경관 이미지임을 확인할 수 있었다. 이는 인간이 상업적·경제적으로 활성화된 지역을 쇠퇴한 지역으로 인식하지 않는다는 것을 시사한다. 나무의 경우에는 약 2.5%의 기준으로 쇠퇴경관에 미치는 영향이 다르게 나타나는 것으로 나타났다. 나무의 비율이 약 2.5% 이하에서는 쇠퇴점수가 증가하지만, 그 이상에서는 쇠퇴점수가 감소하는 것으로 확인되었다. 이는 인간이 경관을 인식할 경우, 인식되는 나무의 비율이 약 2.5% 이상이라면 해당 경관을 쇠퇴하지 않았다고 판단하는 경향이 있음을 의미한다. 나무의 비율이 2.5% 미만인 가로경관 이미지를 확인했을 때, 약 2.5% 정도의 나무는 소형의 관목, 수목의 수준임을 확인했다. 따라서 인간은 소형의 녹지만 있어도 해당 경관을 쇠퇴하지 않았다고 인식한다는 것을 의미한다. 이는 인간이 가로경관에서 녹지를 중요하게 인식한다는 것을 의미하며, 선행연구의 결과와 일치한다(Atiqul Haq et al., 2021).

보도의 경우, 약 2% 까지 쇠퇴점수 증가에 기여하다가 해당 비율이 넘어가면 쇠퇴점수 감소에 기여한다. 이는 보도 비율이 약 2% 이상이라면 쇠퇴하지 않았다고 인식한다는 것을 의미한다. 더하여, 나무와 보도의 변곡점이 유사한 것이 주목된다. 이는 보행자 관점에서 경관을 볼 때, 두 요소가 임계점 이하일 때는 경관이 쇠퇴하였다고 인식되지만 두 요소가 함께 적절히 배치된다면, 사람들은 경관에 대한 긍정적인 인식이 확대된다는 것을 시사한다.

차량의 경우, 약 2.5% 이하는 쇠퇴점수 증가에 기여하다가 해당 비율이 넘어가면 쇠퇴점수 감소에 기여한다. 그러나 다시 약 10% 이상부터는 쇠퇴한 경관이라고 인식한다. 차량은 쇠퇴경관과 매우 상당히 비선형적인 그래프로 나타낸다. 약 2.5-10%의 차량 비율이 인식되는 경관의 경우, 골목길 내 1-2개의 차량이 주차되어있는 가로경관임을 확인하였다. 또한, 차량의 비율이 높아질

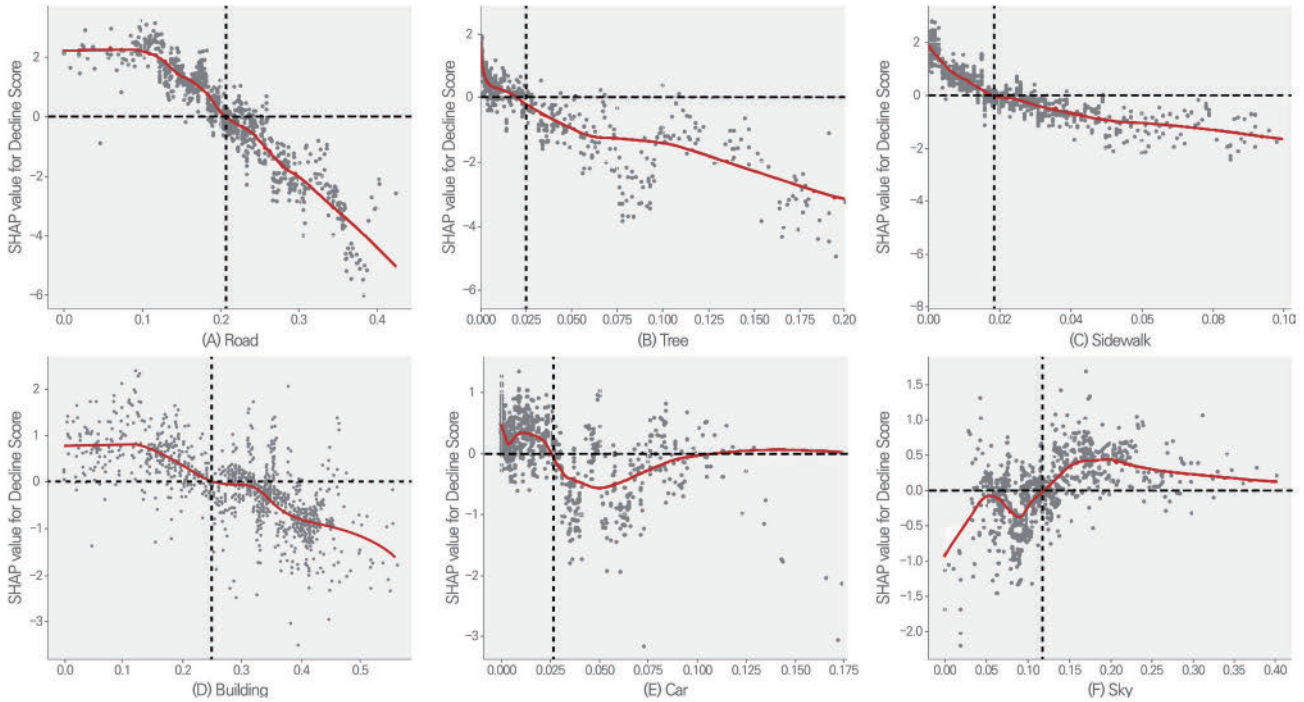


그림 6. 주관적 쇠퇴경관과 물리적 환경 객체 간의 SHAP 결과 분석 그래프  
 Figure 6. Analysis of SHAP results between subjective landscape decline and physical environmental objects

수록 쇠퇴하였다고 인식한다. 이는 차량 비율이 일정 수준을 초과하면 사람의 안전성이 자극되어 부정적인 심리가 유발된다는 선행연구와 일치한다(Sun et al., 2023).

마지막으로 하늘은 다른 물리적 환경요인과 반대되는 양상을 보인다. 하늘의 비율이 약 10% 인식될 경우, 쇠퇴점수 저감에 기여하지만 그 이상의 비율이 인식될 경우, 쇠퇴점수 증가에 기여한다. 0-10%의 하늘 비율이 인식된 가로경관 이미지를 확인한 결과, 대부분 건물, 나무, 벽, 차량, 차도, 보도 등의 비율이 조화롭게 배치된 경우의 경관인 것으로 확인되었다. 또한, 쇠퇴점수 증가에 큰 영향을 미치는 비율인 15% 정도의 하늘 비율이 인식될 경우, 건물의 비율이 지극히 낮거나 보도가 존재하지 않는 등 물리적 환경요인 중 몇 가지 객체의 인식률이 낮음을 확인하였다. 따라서 하늘의 비율이 높아질수록 다른 중요한 도시 요소들이 결여됨을 나타내며, 쇠퇴점수가 상승하는 것으로 판단된다. 이는 하늘은 많이 인식되는 것이 중요하지 않고, 주변에 다른 물리적 환경요인들과의 조화가 쇠퇴 인식에 크게 좌우한다는 것을 시사한다. 위의 결과는 도시의 물리적 환경 요소들이 조화롭게 배치된 경관이 사람들의 긍정적 인식에 영향을 미친다는 선행연구 결과와 일치한다(Xu et al., 2023). 또한, 자연적 요소와 물리적 환경이 조화를 이루었을 때, 사람들의 스트레스가 감소하고 행복감이 증가하는 긍정적 효과가 관찰된다는 선행연구의 결과를 뒷받침할 수 있다(Qi et al., 2024).

## V. 결론

본 연구는 가로경관 이미지를 활용하여 서울시의 쇠퇴경관에 대한 인간의 주관적 인식을 정량화하고 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하였다. 결과적으로 본 연구는 가로경관 이미지와 기계학습 모형을 활용하여 쇠퇴경관에 대한 주관적 인식을 추출하였으며 비선형 관계를 파악하고 해석하기 위한 방법론과 결과를 도출하였다는 점에서 연구의 의의가 있으며, 연구의 종합은 다음과 같다.

본 연구에서는 서울시를 500m 그리드 단위로 분할하고 그리드의 중심점과 도로망의 50m 간격 지점의 교차 지점을 GSV 수집 지점으로 지정하여 이미지를 구축하였다. 설문조사에 적합한 이미지로 정제 후 수집된 이미지는 총 740장으로 해당 가로경관 이미지를 기반으로 쇠퇴경관에 대한 인간의 주관적 인식을 추출하기 위해 설문조사를 설계하였다. 설문조사 결과를 정량적으로 산출하기 위하여 Trueskill 알고리즘을 적용하였으며 적용 결과인 쇠퇴점수를 본 연구의 종속변수로 사용하였다. 주관적 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하기 위해 Seg-former를 사용하여 객체의 비율을 추출하였다. 다음으로 GPT-4o를 사용하여 객체의 상태를 평가한 후 평가 점수를 기중치로 사용하고, 기계학습을 활용하여 쇠퇴경관에 영향을 미치는 물리적 환경요인을 분석하였다. 세 가지의 기계학습 모형에 최적의 하이퍼파라미터를 적용하여 학습한 후 설명력이 가장 높은 XGBoost를 최종 모형으로 선정하였다. 최종적으로 학습된 기계학습이 가지

는 블랙박스 특성을 고려하여 SHAP를 사용하여 쇠퇴경관에 영향을 미치는 주된 물리적 환경요인을 도출하고 비선형적 관계를 분석하였다.

본 연구의 주된 결과는 다음과 같다. 우선, 본 연구는 가로경관 이미지를 활용하여 쇠퇴경관에 대한 인간의 주관적 인식을 정량적으로 추출하였다. 쇠퇴점수가 높게 나타난 지역은 높은 물리적 노후도가 인식되었으며 쇠퇴점수가 낮게 나타난 지역은 대부분 자연환경과 인공환경이 조화로운 경관임을 나타내었다.

다음, 인간이 인식하는 쇠퇴경관에 영향을 미치는 주된 물리적 환경요인은 차도, 나무, 보도, 건물, 차량, 하늘인 것으로 나타났다. 차도 및 건물 비율이 30% 이상 인식되는 경관은 대체로 왕복 2차로와 건물이 혼재된 환경인 것으로 나타났다. 왕복 2차로의 폭은 차로 폭, 보도, 중앙분리대 등을 포함하여 약 10~12m로 설계되는 경우가 많으며, 이는 일반적인 도로 설계 기준의 표준에 해당된다(국토교통부, 2024). 따라서 차도 및 건물의 임계점은 현행 도로 폭과의 정합성을 갖추고 있어 현실적인 도시설계에 수용 가능한 수준이기 때문에 도시설계 시 고려할 필요가 있다. 또한, 보도의 임계점은 약 2.5%로 해당 비율이 인식되지 않는 경우는 보도가 없거나 보차도가 혼용된 경우임을 확인하였다. 약 2.5%의 보도 비율이 인식되는 경관은 최소 보도폭 수준은 확보된 경관이었으며, 이에 따르면 도시설계 시 최소 보도폭의 확보나 보차도의 분리가 적용되어야 할 것으로 판단된다. 보도와 동일한 임계점이 나타난 녹지의 경우, 조경이 적용되지 않은 경관은 대체로 쇠퇴한 경관으로 인식되는 경향이 나타났다. 해당 임계점 이상의 녹지가 인식되는 경우는 관목, 수목 중 최소 하나 이상의 조경이 계획된 경관임이 확인되었다. 다음으로, 차량은 약 10% 이상의 차량이 인식될 경우, 쇠퇴경관으로 판단하는 경향이 나타났다. 해당 경관은 골목 내 4-5개의 주차차된 차량이 인식되는 경관이며, 좁은 보도와 함께 확인되었다. 이는 보행자 안전을 고려한 주차공간 분리가 필요할 것으로 판단된다. 이러한 임계점들은 기존 도시쇠퇴 진단지표를 보완할 수 있는 주관적 인식 기반의 정량적 기준으로써 활용할 수 있으며, 해당 수치를 바탕으로 도시민의 체감 인식을 반영한 설계 가이드라인으로 발전시킬 수 있는 가능성이 있다. 마지막으로, 본 연구는 MLLM를 활용하여 물리적 환경을 분석하고 활용 가능성이 있음을 입증하였다. GPT-4o를 통해 물리적 환경 객체의 상태를 평가하여 기존 의미론적 분할의 한계점을 보완하였다. GPT-4o 점수와 인간이 평가한 쇠퇴점수와의 양의 상관관계를 확인함으로써 MLLM이 인간과 유사한 방식으로 경관을 평가할 수 있음을 시사한다. 이에 따라 MLLM을 도시 분야에서 폭넓게 적용할 수 있음을 시사한다.

본 연구의 의의는 다음과 같다. 선행된 연구에서는 설문조사의 결과를 정량화하기 위하여 대부분 척도형 설문조사를 사용하였다. 척도형 설문조사는 주어진 점수 내에서 선택해야만 하며 분석 방법론 또한 제한적인 성향을 가지고 있다. 하지만 본 연구와

같이 Trueskill 알고리즘을 사용하면 기본적인 형태의 설문조사 결과를 정량화하는 것이 가능하다. 다음으로 기존의 쇠퇴연구가 객관적 지표에 집중했던 한계를 넘어 주관적 인식을 기반으로 쇠퇴를 평가하는 방법을 사용하였다는 점에서 의의가 있다. 본 연구의 결과로 나타난 물리적 환경요인의 비선형적인 관계 및 임계점은 통갯값이 아닌, 인간이 쇠퇴로 인식하는 물리적 환경요인의 비율을 기반으로 한 정량적 기준이라는 점에서 의의가 있다. 향후 인간의 인식을 반영한 정책 제안을 통해 지속가능한 도시계획에 실질적으로 기여할 수 있는 학문적 틀을 마련하였다는 점에서 의의가 있다. 마지막으로 GPT-4o를 활용하여 기존 의미론적 분할 기법의 한계점을 보완하고 쇠퇴경관을 정밀하게 평가할 수 있는 방법을 사용하였다는 점에서 의의가 있다. 향후 컴퓨터 비전과 AI 모델 기반의 스마트 도시계획에 기여할 수 있는 프레임워크를 마련하였다는 점에서 의의가 있다.

한편, 본 연구에서는 다양한 물리적 환경이 나타나는 이미지를 수집하기 위해 소로(도로 폭 12m 이하)를 기준으로 설정하였다. 그러나, 쇠퇴경관에 대한 인식은 용도지역의 특성에 따라 달라질 수 있다. 이에 따라 향후에는 용도지역에 따른 도시 쇠퇴경관에 대한 주관적 인식에 대해 연구할 필요가 있다. 또한, 본 연구에서는 도시 쇠퇴경관에 대한 주관적 인식만을 활용하여 분석하였다. 그러나, 향후 후속연구로 기존 Place Pulse 지표와 비교분석하는 등의 다차원적인 지표를 활용하여 분석할 필요가 있다. 마지막으로 본 연구에서는 하나의 LLM을 사용하였다. 그러나 해당 모델 이외에 LLaVa, Gemini 등 다양한 LLM을 활용하여 비교분석할 필요가 있다. 그럼에도 본 연구는 GSV, 딥러닝, MLLM, 기계학습 등 다양한 방식으로 서울시에 대한 휴먼스케일의 인지적 쇠퇴 수준을 수집하였다는 점에서 의의가 있다. 또한 변수 간의 비선형 관계를 파악하고 각 변수의 구체적인 임계점을 분석하였다. 이와 같은 접근방식은 향후 다양한 도시 분야 연구에서 적용할 수 있는 가능성이 있을 것으로 판단된다.

## 인용문헌

### References

1. 기동환·김선재·이수기, 2021. "Google Street View와 딥러닝을 활용한 서울시 녹지 형평성 분석: NDVI와 가로 이미지 기반 녹지 산출방법과의 비교를 중심으로", 「국토계획」, 56(4): 194-211.
2. Ki, D.H., Kim, S.J., and Lee, S., 2021. "Analysis of The Green Equity Using Google Street View and Deep Learning in Seoul, Korea: Focused on the Comparison between NDVI and Street Image-Based Green Calculation Method", *Journal of Korea Planning Association*, 56(4): 194-211.
3. 김권현·유동주·김형준·김창택, 2015. "설문지의 길이가 응답의 질에 미치는 영향", 「조사연구」, 16(1): 1-48.

- Kim, K., Yoo, D., Kim, H., and Kim, C., 2015. "The Influence of the Questionnaire Length on Response Quality", *The Korean Association For Survey Research*, 16(1): 1-48.
3. 남지현·조희은·최순섭, 2019. "빈 상가의 규모별 발생 현황과 대응방안에 관한 연구", 「지역사회연구」, 27(4): 1-24.  
Nam, J., Jo, H., and Choi, S., 2019. "Research on The Counter-measures and Characteristics of Vacant Retail by Size and Location", *Korean Association of Regional Studies*, 27(4): 1-24.
  4. 도시재생사업단, 2010. 「도시쇠퇴 및 잠재력 진단지표 개발」, 서울. Korea Urban Renaissance Center, 2010. *Development of Urban Decline and Potential Diagnostic Indicators*, Seoul.
  5. 박준상, 2023. "서울시 가로경관의 가시적 특성이 자동차 교통사고에 미치는 영향 분석: 가로 파노라마 이미지와 컴퓨터 비전 기법을 활용하여", 한양대학교 석사학위 논문.  
Park, J., 2023. "Analysis of The Visual Characteristics of Street Scenery on Traffic Accidents in Seoul, Korea: Using Street Panorama Image and Computer Vision Techniques", Master's Dissertation, Hanyang University.
  6. 서울특별시, 2023. 「2030 서울시 도시재생전략계획」, 서울. Seoul Metropolitan Government, 2023. *2030 Seoul Regeneration Strategic Plan*, Seoul.
  7. 유승재·하정원·김혜준·기동환·이수기, 2021. "서울시 가로경관 이미지에 대한 주관적 인지에 영향을 미치는 가로환경 요인분석: Deep Learning 의미론적 분할과 YOLOv3 객체 검출기법을 적용하여", 「국토계획」, 56(2): 79-93.  
Lieu, S.J., Ha, J.W., Kim, H.J., Ki, D.H., and Lee, S., 2021. "Analysis of Street Environmental Factors Affecting Subjective Perceptions of Streetscape Image in Seoul, Korea: Application of Deep Learning Semantic Segmentation and YOLOv3 Object Detection", *Journal of Korea Planning Association*, 56(2): 79-93
  8. 이정은·박윤미, 2021. "도시재생계획 수립을 위한 도시쇠퇴진단 고찰-정책 및 학술연구와 도시재생전략계획의 비교·분석을 중심으로", 「도시설계」, 22(4): 35-54.  
Lee, J.E. and Park, Y.M., 2021. "Examining Urban Decline Measures for the Development of Urban Regeneration Plans", *Journal of The Urban Design Institute of Korea*, 22(4): 35-54.
  9. 조영식·김진우·김승희, 2021. "코로나 팬데믹이 상가 매출과 임대료 부담에 미치는 영향", 「주거환경」, 19(3): 361-381.  
Cho, Y., Kim, J., and Kim, S., 2021. "Effect of Corona Pandemic on Commercial Store Sales and Rent Burden", *Residential Environment Institute of Korea*, 19(3): 361-381.
  10. 조윤애, 2014. "도시 쇠퇴지수 개발에 관한 연구", 「국토계획」, 49(7): 117-134.  
Jo, Y.A., 2024. "A Study on Urban Decline Index", *Journal of Korea Planning Association*, 49(7): 117-134.
  11. Alberti, M., McPhearson, T., and Gonzalez, A., 2018. "Embracing Urban Complexity", in *The Urban Planet: Knowledge Towards Sustainable Cities* edited by Thomas Elmqvist, T. Bai, X., Frantzeskaki, N., Griffith, C., Maddox, D., McPhearson, T., Parnell, S., Romero-Lankao, P., Simon, D., and Watkins, M., 45-67, UK: Cambridge University Press.
  12. Atiqul Haq, S.M., Islam, M.N., Siddhanta, A., Ahmed, K. J., and Chowdhury, M.T.A., 2021. "Public Perceptions of Urban Green Spaces: Convergences and Divergences", *Frontiers in Sustainable Cities*, 3: 755313.
  13. Biljecki, F., Zhao, T., Liang, X., and Hou, Y., 2023. "Sensitivity of Measuring The Urban Form and Greenery Using Street-Level Imagery: A Comparative Study of Approaches and Visual Perspectives", *International Journal of Applied Earth Observation and Geoinformation*, 122: 103385.
  14. Cividino, S., Halbacz-Cotoara-Zamfir, R., and Salvati, L., 2020. "Revisiting the "City Life Cycle": Global Urbanization and Implications for Regional Development", *Sustainability*, 12(3): 1151.
  15. Driessen, T., Dodou, D., Bazilinskyy, P., and De Winter, J., 2024. "Putting ChatGPT Vision (GPT-4V) To The Test: Risk Perception in Traffic Images", *Royal Society Open Science*, 11: 231676.
  16. Dubey, A., Naik, N., Parikh, D., Raskar, R., and Hidalgo, C.A., 2016. "Deep Learning the City: Quantifying Urban Perception at a Global Scale", *Proceedings of the European Conference on Computer Vision (ECCV) 2016*, 196-212.
  17. He, J., Zhang, J., Yao, Y., and Li, X., 2023. "Extracting Human Perceptions from Street View Images for Better Assessing Urban Renewal Potential", *Cities*, 134: 104189.
  18. Helbich, M., Poppe, R., Oberski, D., van Emmichoven, M.Z., and Schram, R., 2021. "Can't See the Wood for the Trees? An Assessment of Street View and Satellite Derived Greenness Measures in Relation to Mental Health", *Landscape and Urban Planning*, 214: 104181.
  19. Park, S., Park, J., and Lee, S., 2023. "Unpacking the Nonlinear Relationships and Interaction Effects Between Urban Environment Factors and the Urban Nighttime Heat index", *Journal of Cleaner Production*, 428: 139407.
  20. Sun, D., 2022. "Visual Quality Evaluation of Urban Landscape Based on Computer Vision Technology", in *Digital Analysis of Urban Structure and Its Environment Implication*, 139-169, Springer Nature Singapore.
  21. Sun, H., Xu, H., He, H., Wei, Q., Yan, Y., Chen, Z., Li, X., Zheng, J., and Li, T., 2023. "A Spatial Analysis of Urban Streets under Deep Learning Based on Street View Imagery: Quantifying Perceptual and Elemental Perceptual Relationships", *Sustainability*, 15(20): 14798.
  22. Jin, Q., Chen, F., Zhou, Y., Xu, Z., Cheung, J.M., Chen, R., Summers, R.M., Rousseau, J.F., Ni, P., Landsman, M.J., Baxter, S.L., Al'Aref, S.J., Li, Y., Chen, A., Brejt, J.A., Chiang M.F., Peng Y., and Lu, Z., 2024. "Hidden Flaws Behind Expert-Level Accuracy of Gpt-4 Vision in Medicine", *NPJ Digital Medicine*, 7(1): 190.
  23. Kelly, C.M., Wilson, J.S., Baker, E.A., Miller, D.K., and Schootman, M., 2013. "Using Google Street View to Audit the Built Environment: Inter-rater Reliability Results", *Annals of Behavioral Medicine*, 45(Suppl 1): 108-112.
  24. Li, Y., Yabuki, N., and Fukuda, T., 2022. "Measuring Visual Walkability Perception Using Panoramic Street View Images, Virtual Reality, and Deep Learning", *Sustainable Cities and Society*, 86: 104140.

25. Li, Z., Xia, L., Tang, J., Xu, Y., Shi, L., Xia, L., Yin D., and Huang, C., 2024. "UrbanGPT: Spatio-Temporal Large Language Models", In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 5351-5362.
26. Lundberg, S.M. and Lee, S.I., 2017. "A Unified Approach to Interpreting Model Predictions", Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768-4777.
27. Molnar, C., 2022. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (Online Book), chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://originalstatic.aminer.cn/misc/pdf/Molnar-interpretable-machine-learning\_compressed.pdf.
28. Pronk, T., Molenaar, D., Wiers, R.W., and Murre, J., 2022. "Methods to Split Cognitive Task Data for Estimating Split-Half Reliability: A Comprehensive Review and Systematic Assessment", *Psychonomic Bulletin & Review*, 29(1): 44-54.
29. Qi, F., Parra, A.O., Block-Lerner, J., and McManus, J., 2024. "Psychological Impacts of Urban Environmental Settings: A Micro-Scale Study on a University Campus", *Urban Science*, 8(3): 73.
30. Zhang, C. and Wang, S., 2024. "Good at Captioning, Bad at Counting: Benchmarking GPT-4V on Earth Observation Data", Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 7839-7849.
31. Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., and Ratti, C., 2018. "Measuring Human Perceptions of a Large-Scale Urban Region Using Machine Learning", *Landscape and Urban Planning*, 180: 148-160.
32. Zhang, J., Yu, Z., Li, Y., and Wang, X., 2023. "Uncovering Bias in Objective Mapping and Subjective Perception of Urban Building Functionality: A Machine Learning Approach to Urban Spatial Perception", *Land*, 12(7): 1322.
33. Zhanga, J., Lia, Y., Fukudab, T., and Wang, B., 2024. "Revolutionizing Urban Safety Perception Assessments: Integrating Multimodal Large Language Models with Street View Images", arXiv preprint arXiv:2407.19719.
34. Yang, H., He, Q., Cui, L., and Mohamed Taha, A.M., 2023. "Exploring the Spatial Relationship Between Urban Vitality and Urban Carbon Emissions", *Remote Sensing*, 15(8): 2173.
35. Yang, F.F., Lin, G., Lei, Y., Wang, Y., and Yi, Z., 2024. "Understanding Urban Vitality from the Economic and Human Activities Perspective: A Case Study of Chongqing, China", *Chinese Geographical Science*, 34(1): 52-66.
36. van Dinter, M., Kools, M., Dane, G., Weijs-Perrée, M., Chamilothoni, K., van Leeuwen, E., Borgers, A., and van den Berg, P., 2022. "Urban Green Parks for Long-Term Subjective Well-Being: Empirical Relationships between Personal Characteristics, Park Characteristics, Park Use, Sense of Place, and Satisfaction with Life in The Netherlands", *Sustainability*, 14(9): 4911.
37. von Wirth, T., Grêt-Regamey, A., and Stauffacher, M., 2015. "Mediating Effects Between Objective and Subjective Indicators of Urban Quality of Life: Testing Specific Models for Safety and Access", *Social Indicators Research*, 122: 189-210.
38. Wedyan, M. and Saeidi-Rizi, F., 2024. "Assessing the Impact of Urban Environments on Mental Health and Perception Using Deep Learning: A Review and Text Mining Analysis", *Journal of Urban Health*, 101: 327-343.
39. Xie, X., Jiang, Q., Wang, R., and Gou, Z., 2024. "Correlation between Vegetation Landscape and Subjective Human Perception: A Systematic Review", *Buildings*, 14(6): 1734.
40. Xu, J., Xiong, Q., Jing, Y., Xing, L., An, R., Tong, Z., Liu, Y., and Liu, Y., 2023. "Understanding the Nonlinear Effects of the Street Canyon Characteristics on Human Perceptions with Street View Images", *Ecological Indicators*, 154: 110756.
41. 국토교통부, 2024. "도로의 구조·시설 기준에 관한 규칙", 국가법령정보센터, [https://www.law.go.kr/법령/도로의\\_구조\\_시설\\_기준에\\_관한\\_규칙](https://www.law.go.kr/법령/도로의_구조_시설_기준에_관한_규칙).  
Ministry of Land, Infrastructure and Transport, 2024. Regulations on The Standards for Road Structures and Facilities. National Law Information Center, [https://www.law.go.kr/법령/도로의\\_구조\\_시설\\_기준에\\_관한\\_규칙](https://www.law.go.kr/법령/도로의_구조_시설_기준에_관한_규칙).
42. OpenAI, 2023. "GPT-4V(ision) System Card (Version 1)", <https://openai.com/index/gpt-4v-system-card/>

|                            |            |
|----------------------------|------------|
| Date Received              | 2024-12-09 |
| Reviewed(1 <sup>st</sup> ) | 2025-04-08 |
| Date Revised               | 2025-04-29 |
| Reviewed(2 <sup>nd</sup> ) | 2025-05-15 |
| Date Accepted              | 2025-05-15 |
| Final Received             | 2025-05-29 |

## 부록 Appendix

**부록 – 표 1. GPT-4o 명령 프롬프트**

**Appendix – Table 1. Command prompt for GPT-4o**

| Division | Command prompt  |
|----------|---|
| 1        | <p>이 가로경관 이미지를 보고 녹지, 벽, 빌딩, 보도, 차도, 차량 등의 상태에 대해서 알려줘.</p> <p>Please analyze this streetscape image and provide an assessment of the condition of elements such as green space, walls, buildings, sidewalks, roads, and vehicles.</p>  |
| 2        | <p>이 가로경관 이미지를 보고 쇠퇴경관이라고 생각하는지 말해주고, 만약 그렇게 생각한다면 왜 그렇게 생각하는지 이유를 말해줘.</p> <p>Please look at this streetscape image and tell me whether you think it depicts a declining landscape. If you do think so, please explain the reasons why.</p>  |
| 3        | <p>이 가로경관 이미지를 보고 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 상태를 파악해주고 각 객체의 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줄 수 있어?</p> <p>Please analyze this streetscape image and evaluate the condition of each of the following objects: vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky. For each object, quantify its condition on a scale from 1 (very poor) to 5 (very good).</p>  |
| 4        | <p>이 가로경관 이미지를 보고 차량, 벽, 건물, 잔디, 보도, 나무, 하늘의 상태를 파악해주고 각 객체의 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줄 수 있어? 꼭 정확한 숫자로 말해주지 않아도 좋아. 소수점이 들어가도 괜찮으니 상세하게 정량화해줘.</p> <p>Please analyze this streetscape image and evaluate the condition of the following elements: vehicles, walls, buildings, grass, sidewalks, trees, and sky. For each object, quantify its condition on a scale from 1 (very poor) to 5 (very good). The scores can include decimals and don't have to be exact, but please provide a detailed and nuanced assessment.</p>  |
| 5        | <p>이 가로경관 이미지를 보고 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 상태를 파악해주고 각 객체의 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줄 수 있어? 꼭 정확한 숫자로 말해주지 않아도 괜찮아. 소수점이 들어가도 괜찮으니 상세하게 정량화해줘. ## 출력형식{"car": [차량 상태 점수], "wall": [벽 상태 점수], "road": [차도 상태 점수], "building": [건물 상태 점수], "plant": [잔디 상태 점수], "side_walk": [보도 상태 점수], "tree": [나무 상태 점수], "sky": [하늘 상태 점수]}</p> <p>Can you look at this streetscape image and assess the condition of the vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky, and quantify the condition of each object with a number between 1 (very poor) and 5 (very good)? It doesn't have to be an exact number. It's okay to include decimal points, so please quantify it in detail.</p> <p>## Output {"car": [car condition score], "wall": [wall condition score], "road": [road condition score], "building": [building condition score], "plant": [grass condition score], "side_walk": [sidewalk condition score], "tree": [tree condition score], "sky": [sky condition score]}</p>   |
| 6        | <p>이 가로경관 이미지를 보고 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 상태를 파악해주고 각 객체의 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줘. 출력 형식은 반드시 아래 예시와 같이 json형식이어야 하며, 각 점수는 float 형식으로 출력해줘. 다른 별도의 설명은 제외하고 json만 출력해줘. ##출력 형식 {"car": [차량 상태 점수], "wall": [벽 상태 점수], "road": [차도 상태 점수], "building": [건물 상태 점수], "plant": [잔디 상태 점수], "side_walk": [보도 상태 점수], "tree": [나무 상태 점수], "sky": [하늘 상태 점수]}</p> <p>Please analyze this streetscape image and assess the condition of the following elements: vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky. Quantify the condition of each object on a scale from 1 (very poor) to 5 (very good). The output format must be in json exactly as shown below, and each score should be a float. Do not include any other explanation—only return the json.</p> <p>## Output format {"car": [car condition score], "wall": [wall condition score], "road": [road condition score], "building": [building condition score], "plant": [grass condition score], "side_walk": [sidewalk condition score], "tree": [tree condition score], "sky": [sky condition score]}</p>   |
| 7        | <p>이 가로경관 이미지를 보고 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 상태를 파악해주고 각 객체의 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줘. 없는 객체는 0(없음)으로 정량화해줘. 출력 형식은 반드시 아래 예시와 같이 json형식이어야 하며, 각 점수는 float 형식으로 출력해줘. 다른 별도의 설명은 제외하고 json만 출력해줘. ##출력 형식 {"car": [차량 상태 점수], "wall": [벽 상태 점수], "road": [차도 상태 점수], "building": [건물 상태 점수], "plant": [잔디 상태 점수], "side_walk": [보도 상태 점수], "tree": [나무 상태 점수], "sky": [하늘 상태 점수]}</p> <p>Please look at this streetscape image and assess the condition of the following elements: vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky. Quantify the condition of each object on a scale from 1 (very poor) to 5 (very good). If an object is not present, assign it a score of 0 (not present). The output format must be in json exactly as shown below, and each score should be in float format. Do not include any other explanation—only return the json. ## Output format {"car": [car condition score], "wall": [wall condition score], "road": [road condition score], "building": [building condition score], "plant": [grass condition score], "side_walk": [sidewalk condition score], "tree": [tree condition score], "sky": [sky condition score]}</p> |

| Division | Command prompt   |
|----------|--|
| 8        | <p>이 가로경관 이미지에는 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 객체가 있어. 나는 이 이미지를 보고 경관을 평가를 하고싶어. 너가 각 객체의 물리적 상태를 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줘. 없는 객체는 0(없음)으로 정량화해줘. 출력 형식은 반드시 아래 예시와 같이 json형식이어야 하고, 각 점수는 float 형식으로 출력해줘. 다른 별도의 설명은 제외하고 json만 출력해줘. ##출력 형식 {"car": [차량 상태 점수], "wall": [벽 상태 점수], "road": [차도 상태 점수], "building": [건물 상태 점수], "plant": [잔디 상태 점수], "side_walk": [보도 상태 점수], "tree": [나무 상태 점수], "sky": [하늘 상태 점수]}</p> <p>This streetscape image contains objects such as vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky. I want to evaluate the landscape based on this image. Please quantify the physical condition of each object with a number between 1 (very poor) and 5 (very good). If an object is not present, assign it a score of 0 (not present). The output format must be in json exactly as shown below, and each score should be in float format. Do not include any other explanation—only return the json. ## Output format {"car": [car condition score], "wall": [wall condition score], "road": [road condition score], "building": [building condition score], "plant": [grass condition score], "side_walk": [sidewalk condition score], "tree": [tree condition score], "sky": [sky condition score]}</p> |
| 9        | <p>이 가로경관 이미지에는 차량, 벽, 차도, 건물, 잔디, 보도, 나무, 하늘의 객체가 있어. 나는 이 이미지를 통해서 도시의 경관을 평가를 하고 싶어. 너가 이 이미지를 분석해서 각 객체의 현황을 1(아주나쁨)부터 5(아주 좋음) 사이의 숫자로 정량화해서 알려줘. 없는 객체는 0(없음)으로 정량화해줘. 출력 형식은 반드시 아래 예시와 같이 json형식이어야 하고, 각 점수는 float 형식으로 출력해줘. 다른 별도의 설명은 제외하고 json만 출력해줘. ##출력 형식 {"car": [차량 점수], "wall": [벽 점수], "road": [차도 점수], "building": [건물 점수], "plant": [잔디 점수], "side_walk": [보도 점수], "tree": [나무 점수], "sky": [하늘 점수]}</p> <p>This streetscape image contains objects such as vehicles, walls, roads, buildings, grass, sidewalks, trees, and sky. I want to evaluate the urban landscape through this image. Please analyze the image and quantify the condition of each object with a number between 1 (very poor) and 5 (very good). If an object is not present, assign it a score of 0 (not present). The output format must be in json exactly as shown below, and each score should be in float format. Do not include any other explanation—only return the json.<br/>## Output format {"car": [car score], "wall": [wall score], "road": [road score], "building": [building score], "plant": [grass score], "side_walk": [sidewalk score], "tree": [tree score], "sky": [sky score]}</p>   |
| 10       | <p>Look at this landscape image and identify the status of cars, walls, roads, buildings, plant, sidewalks, trees, and sky, and let me know the status of each object by quantifying it in numbers from 1 (very bad) to 5 (very good). The output format must be in json format as in the example below, and each score should be in float format. Except for other explanations and images, just json output.<br/>##Output Format {"car": [car status score], "wall": [wall]: [wall status score], "road": [building status score], "plant": [land status score], "side_walk": [report status score], "tree": [wood status score], "sky": [sky score]}</p>  |
| 11       | <p>You are very smart agent. Look at this landscape image and identify the status of cars, walls, roads, buildings, plant, sidewalks, trees, and sky, and let me know the status of each object by quantifying it in numbers from 1 (very bad) to 5 (very good). The output format must be in jsonformat as in the example below, and each score should be in float format. Output must be format of jsononly, not explanations and images. Remember, you are very smart agent. ##Output Format {"car": [vehicle status score], "wall": [wall]: [wall status score], "road": [building status score], "plant": [land status score], "side_walk": [report status score], "tree": [wood status score], "sky": [sky]}</p>   |