

기계학습을 이용한 보행환경 정성적 평가에 영향을 미치는 거리영상 특성분석*

이지윤** · 강영옥*** · 김지연**** · 박지영*****

Analysis of Street View Image Object Affecting Perceived Walkability Using Machine Learning*

Jiyeon Lee** · Youngok Kang*** · Jiyeon Kim**** · Jiyoung Park*****

요약 : 보행은 녹색교통수단의 하나로 지속 가능한 도시로의 전환에 매우 중요하다. 또한 걷기 좋은 보행환경은 지역민의 건강증진에도 긍정적 효과를 나타냄이 확인되면서 세계 많은 도시들이 친환경적이고 사람중심의 보행환경 조성을 도시 계획에 있어 최우선 아젠더로 추진하고 있다. 보행환경의 중요성이 확대되면서, 보행환경을 구성하는 물리적 구성요소를 파악하기 위한 연구와 함께, 사람들이 걷기 좋다고 판단하는 보행환경이 어떤 것인지를 파악하고자 하는 연구도 이루어졌다. 정성적 보행환경을 분석하기 위한 방법으로 기존에는 주민이나 전문가를 통한 설문조사가 주를 이루면서 대표성의 한계를 보이고 있었지만, 최근 고해상도의 거리영상과 딥러닝 기술의 발전은 거리영상에 대한 쌍체비교 데이터를 훈련셋으로 하여 가로단위의 상세한 정성적 보행환경 평가점수 획득을 가능하게 하고 있다. 하지만 딥러닝 기술에 기반한 정성적 보행환경 평가 점수 예측은 왜 이러한 평가점수를 얻게 되었는지에 대한 해답을 제공하는데 한계가 있다. 본 연구의 목적은 거리영상을 기반으로 사람들이 걷기 좋다고 느끼는 보행환경에 영향을 미치는 도시경관의 특성을 분석하는 것이다. 이를 위해 거리영상의 정성적 보행환경 점수를 종속변수로, 거리영상의 시멘틱 세그멘테이션 비율 값을 독립변수로 하여 다양한 기계학습 모델을 실험하였다. 실험 결과 서포트 벡터 머신 회귀식이 가장 정확도가 높았고, 정성적 보행환경 점수를 예측함에 있어 도로와 보행로, 건물, 나무, 하늘 등의 순으로 객체 중요도가 도출되었다. 보행로(sidewalk), 가로등(streetlight), 길(road), 잔디(grass), 나무(tree) 등은 세그멘테이션 값이 높을수록 정성적 보행환경 평가를 높게 만들며, 쓰레기통(ashcan), 산(mountain), 트럭(truck), 담벼락(wall) 등은 세그멘테이션 값이 높을수록 평가점수를 낮게 만들 수 있었다. 본 연구는 기계학습모델을 통해 보행환경에 대한 정성적 평가에 영향을 주는 중요 객체와 객체의 방향성을 확인할 수 있었으며, 딥러닝 모델을 통해 예측한 보행환경 정성평가 점수를 일부 설명 가능하게 하였다는 점에 의의가 있다.

주요어 : 인지된 보행환경, 시멘틱 세그멘테이션, 기계학습, 설명가능한 AI

*본 연구는 2021년 한국국토정보공사 공간정보연구원의 산학협력R&D 자유공모과제 지원을 받아 수행된 연구임.

**이화여자대학교 빅데이터분석학협동과정 석박사통합과정(Integrative Program, Department of Big Data Analytics, Ewha Womans University, jiyoon821@ewhain.net)

***이화여자대학교 사회과교육과 교수(Professor, Department of Social Studies, Ewha Womans University, ykang@ewha.ac.kr)

****이화여자대학교 일반대학원 사회과교육과 지리학전공 박사과정(Ph.D. Student, Department of Social Studies, Ewha Womans University, pbmejy@gmail.com)

*****이화여자대학교 일반대학원 사회과교육과 지리학전공 박사과정정수료(Ph.D. Student, Department of Social Studies, Ewha Womans University, gisjy.park@gmail.com)

Abstract : Walking, one of the green modes of transportation, is very important for the transition to a sustainable city. In addition, as it has been confirmed that a pleasant walking environment has a positive effect on health of local residents, many cities around the world are promoting the creation of an eco-friendly and people-centered walking environment as the top agenda in urban planning. As the awareness of the importance of the walking environment has increased, many studies have been conducted to identify the physical components that constitute the walking environment and to find out what kind of walking environment people consider good for walking. Existing methods for analyzing the qualitative walking environment based on surveys targeting residents or experts showed limitations in their representativeness. On the other hand, the recent development of high-resolution street view images and deep learning technology makes it possible to obtain detailed perceived walkability scores by using paired comparison data for street view images as a training set. However, perceived walkability score prediction based on deep learning technology has limitations in providing an answer to why such evaluation score was obtained. The purpose of this study is to analyze the characteristics of the urban landscape that affect the perceived walkability based on street view images. In this study, we tested various machine learning models with the perceived walkability score of the street image as the dependent variable and the semantic segmentation ratio value of the street view image as the independent variable. According to our study, the regression equation of the support vector machine was the most accurate. And in predicting the perceived walkability score, the importance of the object was in the order of roads, sidewalks, buildings, trees, and the sky. Our study showed that the higher the segmentation value, the higher the perceived walkability score for sidewalk, streetlight, road, grass, and tree. On the contrary, the higher the segmentation value, the lower the score for ashcans, mountains, trucks, and walls. This study is meaningful in that it was possible to identify important objects and their direction that affect the evaluation of the walking environment through the machine learning model, and to partially explain the perceived walkability score predicted by the deep learning model.

Key Words : Perceived walkability, Semantic segmentation, Machine learning, Explainable AI

I. 서론

보행 활동을 유발하는 도시 환경은 개인의 건강 증진, 교통체증 완화, Co2 배출감소 등 환경적, 사회적 비용 감소에 도움이 될 뿐 아니라 인근 주택 가격 상승과도 유의한 상관관계가 있는 것으로 연구되고 있다. 보행 환경에 대한 연구는 보행안전 뿐 아니라(이수민·황기연, 2009) 지역 유대감을 높이고 커뮤니티 발전에 도움이 되며, 녹색 교통과 함께 탄소배출을 저감하는 등 지속 가능한 도시로의 전환에 매우 중요한 요소가 되는 것으로 평가되고 있다. 보행에 대한 중요성이 인식되면서 세계 많은 도시들이 친환경적이고 사람중심의 보행환경 조성을 도시 계획에 있어 최우선 아젠더(agenda)로 추진하고 있다. 우리나라에서도 1990년대 ‘걷고 싶은 도시 만들기’라는 구호 아래 보행권 운동이 시작되었고(김희철 등, 2014), 보행을 증진하기 위해 관련 법이 제정되는가 하면(행정안전부, 2012) 전국지자체는 5년마다 ‘보행안전 및 편의증진 기본계획’을 수립하고 보행안전 및 편의증

진 실태조사를 시행하도록 하는 등 보행을 증진시키기 위한 다양한 시책들이 추진되고 있다.

보행환경의 중요성이 확대되면서 이와 관련된 연구가 다양하게 진행되었다. 관련 연구로는 주로 보행환경을 구성하는 요소에 대한 연구(이경환·안건혁, 2008; 박소현 등, 2009; 박철영·이수기, 2016), 또는 보행환경 평가 지표를 개발하는 연구가 있는데(박지영 등, 2022), 이런 연구들이 도시의 물리적 요소에 초점을 두었다면 최근에는 사람들이 인지하는 보행환경을 기계학습이나 딥러닝 모델을 이용해 확인하고자 하는 연구들도 이뤄지고 있다(Zhou *et al.*, 2019). 이는 사람들이 인지적으로 느끼는 보행환경이 실제 보행과 연결되기 때문이다. 사람들이 정성적으로 느끼는 보행 환경과 관련하여 기존에는 주로 설문조사를 통해 어떤 요소가 중요한 지 표시하는 방법이 주를 이루었지만(성현곤 등, 2011; 김규리·이제선, 2016; 조혜민·이수기, 2016; Mateo-Babiano, 2016; 박영은·이우성, 2022), 최근에는 구글(Google)등에서 제공하는 거리뷰(street view) 자료를 활용하여 거리뷰 영

상을 보여주면서 해당 지점의 보행환경이 어느정도 되는지를 묻는 방식으로 조사를 하기도 한다. 이러한 설문 조사 결과는 해당 지점별 보행환경 점수를 매기는 데 사용되기도 하지만(Blečić *et al.*, 2018; Wang *et al.*, 2019), 이를 기반으로 학습데이터 셋을 구축한 후 딥러닝 모델을 활용하여 넓은 지역에 대한 정성적인 보행환경 점수를 산출할 수도 있다(Dubey *et al.*, 2016; 김지연·강영옥, 2022). 딥러닝 기술의 발전에 따라 인지적으로 느끼는 보행환경에 대한 평가가 가능해졌지만 보행환경에 대한 인지적 평가와 도시경관 요소의 상관성을 분석하는 연구는 시작단계이다.

본 연구의 목적은 거리영상을 기반으로 사람들이 걷기 좋다고 느끼는 보행환경에 영향을 미치는 도시경관의 특성이 무엇인지를 분석하는 것이다. 이를 위해 거리영상의 인지적 보행환경 점수를 종속변수로, 거리영상의 시멘틱 세그멘테이션 값을 독립변수로 한 9개의 기계학습 모델을 이용하여 종속변수의 값을 가장 잘 예측하는 기계학습 모델을 선별하고, 이 모델에 의거하여 보행환경 평가에 중요한 영향을 미치는 객체 중요도와 그 방향성을 분석하고자 한다. 이러한 분석은 궁극적으로 딥러닝 모델의 결과를 보다 해석적으로 이해할 수 있도록 하며, 인지적 보행환경 개선 시 어떤 건조환경 요소의 개선이 필요한 지에 대한 대안제시에 도움을 줄 수 있을 것이라 판단된다.

II. 관련 연구 리뷰

보행환경의 중요성에 대한 인식과 함께 보행환경을 평가하기 위한 다양한 연구들이 진행되었다. Frank *et al.*(2005)은 근린의 물리적 환경인 주거 밀도, 토지 이용 혼합도, 가로망 연결성이 높을수록 보행활동이 높게 나타남을 밝힌 바 있으며, 이경환·안건혁(2008)은 토지이용 혼합도가 높고 근린 생활 편의시설 이용 및 쇼핑, 산책, 운동의 목적으로 보행이 이루어질 때 보행활동의 빈도가 높게 나타남을 밝혔다. 보행 환경의 물리적 특성을 분석하는 연구에서는 대부분 설문조사를 통해 보행 만족에 영향을 미치는 객관적, 물리적 요인을 추출하였는데 보행환경의 쾌적성, 안전성, 거리의 다양성, 가로의 연결성, 대중교통 접근성 등이 중요한 요인으로 밝혀지고 있다(Kim *et al.*, 2014; Quercia *et al.*, 2015; 이수기

등, 2016; Mateo-Babiano, 2016; 박근덕·이수기, 2018). 기존의 연구가 설문조사나 현장조사에 기반하여 이루어졌다면, 고해상도 거리영상에 시멘틱 세그멘테이션(semantic segmentation) 기법을 적용하여 보행환경 평가 요소를 개발하고 이를 종합하여 물리적 보행환경을 평가하는 연구도 이루어졌는데, 가로 쾌적성, 보행 혼잡도, 시각적 도로포장, 개방감 등이 주요 요소로 평가되고 있다(Blečić *et al.*, 2018; Zhou *et al.*, 2019; Li *et al.*, 2020; 박지영 등, 2022).

보행환경을 구성하는 물리적 구성요소를 파악하기 위한 연구와 함께, 사람들이 걷기 좋다고 판단하는 보행환경이 어떤 것인지를 파악하고자 하는 연구도 이루어졌다(박소현 등, 2009; 김규리·이제선, 2016; 이수기 등, 2016; Mateo-Babiano, 2016). 정성적 보행환경을 분석하기 위한 연구로 박소현 등(2009)의 연구에서는 가로의 쾌적성이 보행만족도에 큰 영향을 주는 요소라고 밝혔으며, Mateo-Babiano(2016)는 보행자를 만족시키는 환경요인으로 이동성, 안전성, 용이성, 접근성, 정체성, 즐거움 등의 요소가 주요함을 밝힌 바 있다. 그러나 주민들이 느끼는 정성적 보행환경에 대한 조사는 소규모 지역을 대상으로 현장조사를 수행하거나, 해당지역 주민 일부를 대상으로 설문조사를 수행하고 있어(김규리·이제선, 2016; 이지원·유다은, 2021), 대표성에 있어 한계를 보이고 있다.

한편 구글에서 거리영상(street view) 서비스를 시작하고, 딥러닝(deep learning) 기술이 빠르게 발전하면서 도시에 대한 감성분석(Salleses *et al.*, 2013; Dubey *et al.*, 2016), 보행환경 평가(Wang *et al.*, 2019; Zhou *et al.*, 2019; Li *et al.*, 2020), 녹지평가(Li and Ratti, 2018) 등 도시 건조환경에 대한 사람들의 인지 및 감성을 분석하려는 연구가 시작되고 있다. 거리영상은 위성이나 항공 사진과 같이 하늘에서 촬영한 수직적 영상과는 다르게 인간 관점에서의 거리의 보행 환경을 나타내기 때문에 사람들이 느끼는 도시 환경의 주관적 인식을 평가하기에 적합하다고 연구되고 있다(Biljecki and Ito, 2021; Kim *et al.*, 2021).

미국 MIT대학교 미디어랩(media lab)에서는 웹사이트를 구축해 웹사이트를 자발적으로 방문한 평가자들에게 거리영상을 보고 도시환경을 평가하는 설문을 진행하였다. 이 설문에서 참여자는 제시된 두 개의 거리영상 중 안전성, 생동감, 아름다움과 같은 정성적 관점에서 더 우

수한 것을 선택하도록 하였다. 이 응답결과를 데이터로 구축하여 도시건조환경에 대한 정성적 평가에 사용할 수 있도록 한 것이 플레이스 펄스(Place Pulse) 데이터셋이다. Place Pulse 1.0(Salesses *et al.*, 2013)은 4,109개의 이미지를 미국의 뉴욕(New York)과 보스턴(Boston), 오스트리아(Austria)의 린츠(Linz)와 잘츠부르크(Salzburg)에서 수집하고 안전성, 독창성, 부유함에 대해 20만여 번의 쌍별 비교를 한 데이터 셋이다. Place Pulse 2.0은 버전 1.0의 확장버전으로, 110,988개 이미지를 28개국 56개 도시에서 수집하여 “어느 장소가 더 안전하게, 활기차게, 아름답게, 부유하게, 우울하게, 지루하게 보이십니까?”하는 6개의 감성에 대해 117만 번의 쌍별 비교를 한 데이터셋이다(Dubey *et al.*, 2016).

쌍별 비교 데이터셋의 제공과 함께 이를 활용한 딥러닝 모델 연구도 활성화되었다. Dubey *et al.*(2016)은 Place Pulse 2.0 데이터를 활용해 이미지 쌍 입력 시 안전함, 활기참, 아름다움, 부유함, 우울함, 지루함에 대한 쌍별 비교를 예측하는 street score-CNN(SS-CNN), ranking SS-CNN(RSS-CNN)이라고 명명한 합성곱신경망(convolutional neural network: CNN)을 제안하였다. 이후 거리 영상으로 연구대상 지역에 대한 사람들의 감성을 예측하는 딥러닝 모델 구조에 대한 여러 연구가 진행되었다(Blečić *et al.*, 2018; Santani *et al.*, 2018; Min *et al.*, 2019; Xu *et al.*, 2019; Guan *et al.*, 2021; 김지연·강영옥, 2022).

거리영상에 기반한 도시건조환경에 대한 정성적 평가는 도시의 어떠한 특징이 특정 감성과 연계되는지 분석하는 연구로 이어졌다. Joglekar *et al.*(2020)은 구글 거리 영상을 사용해 각 영상이 아름다운지 아닌지를 평가하고, 아름다운 영상의 구성 요소가 무엇인지 분석하는 연구를 진행하였다. 해당연구에서는 두 개의 딥러닝 모델을 활용하였는데, 플레이스넷(PlaceNet)을 사용하여 거리 영상을 205개의 객체로 분석하고, 이를 다시 ‘건축(architectural)’, ‘걷기 좋음(walkable)’, ‘랜드마크(landmark)’, ‘자연경관(natural)’의 4개 카테고리로 분류하였다. 아름다움은 ‘걷기 좋음’ 카테고리에 가장 많이 속해 있었고, ‘걷기 좋음’ 카테고리를 구성하는 객체 중에는 마당(yard), 식물원(botanical garden), 숲길(forest path), 횡단보도(crosswalk) 등의 객체가 중요한 비중을 차지하고 있음을 밝혔다. 한 편 모든 거리영상을 세그넷(Segnet)을 활용하여 12종류의 객체(road, sky, trees, buildings, poles, signage, pedestrians, vehicles, bicycles, pavement, fences,

road markings)별 비율 값으로 분할한 후, 로지스틱 회귀(logistic regression)로 각 객체의 회귀 계수(beta coefficient)를 구하였는데, 그 결과 ‘나무’가 아름다운 경관과 가장 큰 양의 관계를 갖고 있었고 ‘길’과 ‘건물’이 아름다움과 가장 큰 음의 관계를 갖는 것을 확인하였다. Zhang *et al.*(2018)은 Place Pulse 2.0 데이터 셋을 이용하여 도시건조환경에 대한 사람들의 정성적 평가를 지도화 하였다. 이후 거리영상을 정성적으로 평가한 결과와 거리영상의 시멘틱 세그멘테이션 결과값을 활용하여 6개 카테고리별로 영향력이 높은 변수 10개를 선정하고 다중회귀분석을 실시하였다. 그 결과 여섯 개 카테고리 모두 ‘하늘’과 가장 큰 관계가 있었고 그 외에 ‘벽’, ‘건물’, ‘나무’, ‘자동차’, ‘보행로’가 의미 있는 객체로 분석되었다. Nagata *et al.*(2020)은 보행환경의 물리적 특성이 보행활동과 관련이 있는지 분석하고자 했다. 65세이상 노인을 대상으로 GSV영상을 보고 보행환경 점수를 부여하도록 한 후, 보행환경 점수를 종속변수로 GSV영상의 시멘틱 세그멘테이션 객체 비율값을 독립변수로 한 보행 점수 예측 회귀모델을 도출하였다. 도출된 모델의 정확도는 MSE가 0.66, R² 값이 0.51이었으며, 여성노인에 있어 모델의 적합도가 더 높게 나왔다. 보행환경과 관련하여 건물에 둘러싸인 도로는 안정감을 부여하며, 도로에 접한 식물, 잔디 등이 보행 활동을 증가시키는 반면, 하늘은 보행에 부정적 영향을 미치는 것으로 나타났다. 한편 국내 관련연구로 유승재 등(2021)은 서울특별시의 거리뷰 영상에 대해 여섯 가지 인지 지표(아름다움, 안전함, 활기참, 우울함, 지루함, 전체적인 만족)별로 5점 척도로(매우 불만족, 불만족, 보통, 만족, 아주 만족) 설문조사를 하였다. 그리고 설문조사를 통해 얻은 정성 평가 자료와 거리뷰에 대해 의미론적 분할기법을 적용해 얻은 객체 별 비율 값을 비교하여 정성적 평가 자료에 어떤 객체가 영향을 주었는지 분석하였다. 그 결과 하늘과 녹지가 전체적인 만족도에 긍정적인 영향을 준 것으로 확인하였다.

보행환경 연구에 있어 거리영상 자료와 딥러닝 기술의 발전은 기존의 보행환경 연구와는 다른 방식의 연구를 가능하게 하고 있고, 사람들이 느끼는 정성적 평가에 대해서도 설문지에 근거한 것이 아닌 거리영상을 보여주며 평가하도록 하여 시각적 특성에 근거한 평가를 가능하게 하고 있다. 이러한 시각적 평가 자료는 왜 이러한 평가결과를 갖게 되는지를 분석하는 연구와 연결되

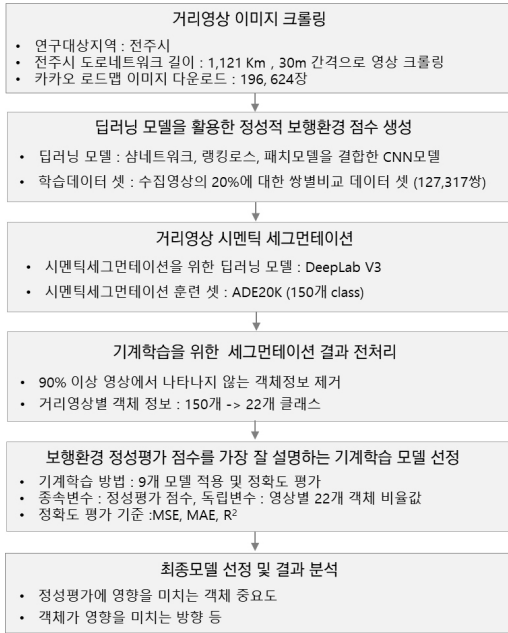


그림 1. 연구흐름도

는 추세이다. 선행연구 리뷰에서 보는 바와 같이 정성적 평가에 영향을 미치는 객체특성을 분석하기 위한 연구가 일부 시도되었지만, 보행환경에 영향을 미치는 객체에 대한 연구도 소수이며, 종속변수에 해당하는 정성적 평가결과를 설명함에 있어 다양한 기계학습 알고리즘을 테스트할 수 있지만 이에 대한 연구도 진행된 바 없다.

III. 연구방법

본 연구의 흐름은 그림 1과 같다. 보행환경에 대한 정성적 평가 점수를 구축한 연구 지역은 대한민국 전주시이다. 거리영상은 전주시 가로를 따라 30미터 간격으로 카카오 거리영상 자료를 수집한 후 어느 쪽의 보행환경이 더 우수한지 쌍별 비교를 하도록 하여 이를 훈련데이터 셋으로 구축하였다(Yoo *et al.*, 2021). 구축된 훈련데이터 셋을 활용하여 딥러닝 모델을 개발하고, 개발된 모델을 통해 수집된 모든 거리영상에 대한 정성적 보행환경 점수를 산출하였고, 이 점수를 종속변수로 활용하였다. 수집된 모든 영상에 대해서는 시멘틱 세그멘테이션 기법을 적용하여 영상에서 객체가 차지하는 비율 값을 생성하였고, 이를 독립변수로 활용하였다. 기계학습 적

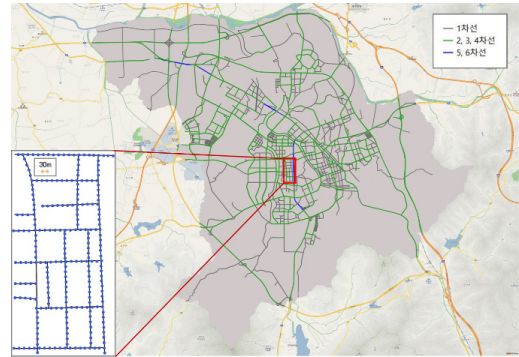


그림 2. 연구대상지역의 거리영상 수집간격

용에 앞서 독립변수에 대한 객체를 선별하는 전처리작업을 수행하였으며, 이후 기계학습 모델 정확도 평가를 통해 종속변수의 특징을 가장 잘 설명하는 모델을 선별하고 해당모델에 근거하여 정성적 보행환경에 영향을 미치는 객체 특징을 분석하였다.

1. 데이터 수집

전주시 거리 영상은 카카오 맵에서 제공하는 카카오 거리영상 자료를 활용하였다. 전주시 도로길이는 총 1,121km인데, 도로망을 따라 30m 간격으로 49,156개 포인트에서 영상을 추출하였다(그림 2). 거리뷰 영상은 한 지점에서 360도의 파노라마 사진으로 촬영되는데, 거리영상의 특징이 방향별로 상이하기 때문에 한 지점당 4방향(0, 90, 180, 270도)의 이미지를 수집하여, 이미지 수로는 총 196,624장의 거리뷰 영상을 크롤링하였다(그림 3).

2. 보행환경 정성적 평가 점수 데이터 구축

보행환경에 대한 정성적 평가 점수를 예측하는 모델을 훈련시키기 위해서는 연구지역의 거리영상과 그에 대한 사람들의 선호를 라벨로 하는 훈련 데이터셋이 필요하다. 훈련데이터 셋의 구축은 거리영상 취득지점 49,156개의 약 20%에 해당하는 10,590지점으로 정하였으며, 10,590지점의 선택은 전주시 거리특징을 반영할 수 있도록 도로유형(대로, 로, 길)과 토지이용유형(공업시설지, 상업 및 업무시설지, 고층주택지, 일반주택지, 기타)을 고려하여 층화 추출하였다. 10,590지점에서 4방향의 영상을 대상으로 하여 총 42,360장이 훈련데이터

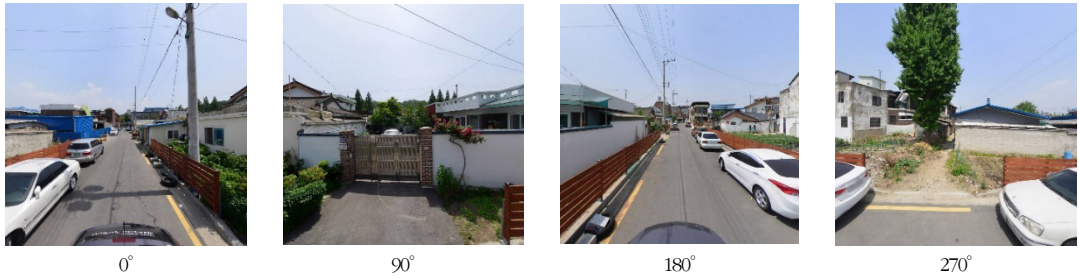


그림 3. 한 지점에 대한 4방향 영상 예시

* 숫자는 동일지점에 대한 영상촬영각도임.

셋 구축 대상이 되었다(김지연·강영욱, 2022). 거리영상에 대한 선호를 측정하는 방법은 참여자들이 각 영상에 직접 점수를 부여하는 방식보다 상대비교된 거리영상에 대한 선호도를 수집하는 것이 보다 정확하고 효율적인 방법으로 알려져 있다(Bijmolt and Wedel, 1995; Stewart *et al.*, 2005). 거리영상에 대한 평가는 웹사이트를 구축해 임의의 거리영상을 두 장 보여준 후 상대적으로 보행 환경이 더 우수하다고 느껴지는 이미지를 선택하는 상대비교 방식을 사용하였다. 쌍별 비교 데이터 셋 구축 시 한 개의 이미지는 적어도 6회 이상 평가될 수 있도록 하여 점수가 안정적으로 생성될 수 있도록 하였다. 2021년 12월부터 2022년 3월까지 총 52명의 응답자가 설문에 답하였으며, 이를 통해 총 127,317개의 응답자료가 수집되었고, 이 중 '같음(=)'으로 응답한 6,422개를 제외하고, 모델 훈련에는 120,895개의 응답자료를 사용하였으며 이 중 80%를 훈련데이터 셋으로 사용하고, 20%는 테스트셋으로 활용하였다.

거리영상의 정성적 보행환경 점수 예측을 위한 딥러닝 모델은 삼 네트워크와 랭크넷을 베이스라인으로 구성하되, 연구 대상지인 전주시와 같이 중소도시에서 도시 경관의 차이가 뚜렷하지 않은 지역에서도 거리영상 이미지의 차이를 효율적으로 학습할 수 있는 구조를 추가적으로 개발하였다(김지연·강영욱, 2022). 삼 네트워크는 딥러닝 모델에 있어 입력 이미지가 쌍으로 들어가며 입력된 데이터의 공통점을 학습할 수 있는 구조를 갖으며, 랭크넷은 쌍별 비교 결과를 통해 최종 순위를 점수로 출력하는 구조를 갖는다. 이러한 베이스라인 모델에 영상에서 나타나는 전역적 특징과 지역적 특징을 함께 학습할 수 있는 구조를 추가하였다. 제안된 모델은 크게 3가지 부분으로 구성되며, 지역적 특징을 학습하는 패치 브랜치(patch branch), 전역적 특징을 학습하는 글로벌

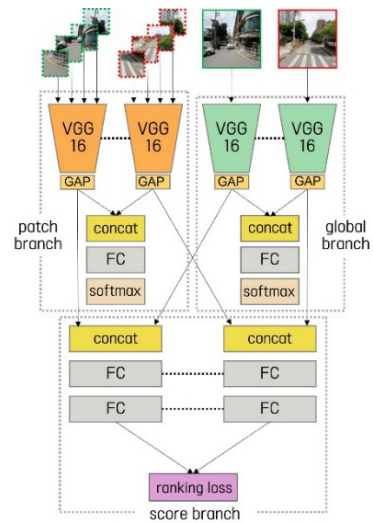


그림 4. 딥러닝 모델 아키텍처

출처 : 김지연·강영욱, 2022.

브랜치(global branch), 앞서 추출된 전역적·지역적 특징을 바탕으로 점수를 평가하는 스코어 브랜치(score branch)로 구성된다(그림 4). 입력 데이터는 쌍별 비교 대상인 두 장의 거리 영상이며, 모델의 최종 결과값은 영상별로 평가된 보행환경 점수이다. 이러한 멀티 패치(multi-patch)는 이미지의 특징을 정밀하게 추출할 때 유용한 방법이다(Lu *et al.*, 2014; Lu *et al.*, 2015; Zou *et al.*, 2021). 모델의 정확도는 총 24,179개의 쌍으로 이루어진 테스트셋을 통해 각 쌍에 대한 실제 우열 값과 예측된 점수의 우열 값이 일치하는 비율로 계산하였다. 구축한 글로벌-패치 모델의 정확도는 75.01%이고, 글로벌-패치 모델로 전체 196,626장의 거리영상에 대해 보행점수를 예측한 후 그 분포를 나타낸 결과는 그림 5와 같다. 평균인 -0.21 주위에 가장 많은 데이터가 몰려 있고, 모

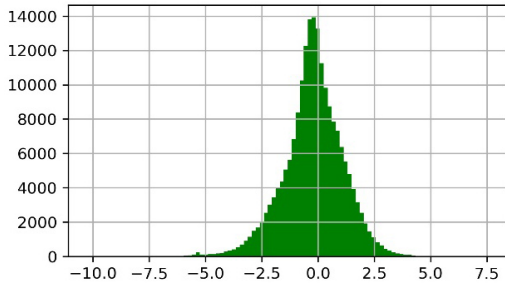


그림 5. 정성적 보행환경 평가 점수 분포
출처 : 김지연·강영옥, 2022.

양은 대칭과 가까우며 최솟값은 -10.21, 최대값은 7.52이다. 점수는 양수일수록 보행환경이 좋다고 평가한 것을 나타내며, 음수일수록 보행환경이 좋지 않다고 인식함을 나타낸다. 그림 6은 거리영상 이미지에 대한 보행환경 예측 점수를 예시로 나타낸 것이다.

3. 거리영상의 시멘틱 세그먼테이션

거리영상의 특징을 분석하기 위해 시멘틱 세그먼테이션 기법을 활용하였다. 시멘틱 세그먼테이션은 이미지의 모든 픽셀마다 그에 해당하는 카테고리 레이블(categorical label)을 부여하여 이미지를 픽셀 단위로 분

류, 인식하는 딥러닝 기법이다(Wang *et al.*, 2018). 시멘틱 세그먼테이션을 위한 딥러닝 모델은 여러가지가 있지만 본 연구에서는 DeepLab v3모델을 활용하였다. 거리영상의 이미지 분할을 위해서는 거리영상을 의미 있게 분할할 수 있는 훈련데이터 셋이 필요한데 가장 대표적인 훈련데이터 셋으로는 cityscape와 ADE20K가 있다. cityscape는 객체를 30개의 클래스로, ADE20K는 실내외 이미지를 150개의 클래스로 구분한 것이다. 본 연구에서 객체 분류는 ADE20K를 기준으로 하였다(그림 7). 수집된 196,624장의 거리영상은 150개 객체별 비율 값이 생성되며, 전처리를 거쳐 22개의 객체가 거리영상의 독립 변수 값으로 사용될 예정이다(표 1). 세그먼테이션의 비율 값은 최소가 0.0, 최대가 1이다.

4. 데이터 전처리 및 기술통계

ADE20K는 거리영상을 150개 객체로 분할하지만, 실제 ADE20K에는 실외뿐 아니라 실내 영상에 대한 클래스를 포함하고 있어서 전주시를 대상으로 수집된 영상에서 나타나지 않는 클래스도 다수 존재한다. 그렇기 때문에 대부분의 영상에서 나타나지 않는 객체는 제거하는 작업을 수행하였다. 196,624개의 관측치 중 90% 이상의 관측치에서 0으로 나타난 클래스는 제거하였으며, 이



그림 6. 거리영상에 대한 정성적 평가 점수 예시



그림 7. 거리영상의 시멘틱 세그멘테이션 결과 예시

표 1. 그림 7의 거리 영상 별 객체 비율 값

객체	(a)	(b)	(c)	(d)
road	0.3999	0.117	0.2829	0.0
sky	0.2655	0.1958	0.3074	0.0799
tree	0.2095	0.0206	0.0854	0.0099
building	0.0354	0.3151	0.1214	0.3219
sidewalk	0.0233	0.0185	0.0988	0.0528
car	0.0188	0.0811	0.0	0.0
wall	0.0159	0.0951	0.0393	0.1261
signboard	0.009	0.011	0.0056	0.035
fence	0.0072	0.0011	0.0017	0.0017
earth	0.0044	0.0092	0.0029	0.0007
pole	0.0013	0.0	0.0	0.0119
railing	0.0008	0.0	0.0	0.0039
streetlight	0.0008	0.0	0.0004	0.0
grass	0.0	0.0	0.001	0.0
plant	0.0	0.016	0.0089	0.0362
field	0.0	0.0	0.0	0.0
ashcan	0.0	0.0	0.0	0.0
truck	0.0	0.0	0.0	0.0
mountain	0.0	0.0002	0.0023	0.0
awning	0.0	0.0067	0.0013	0.0
trade name	0.0	0.0053	0.0	0.0
person	0.0	0.0	0.0005	0.0243

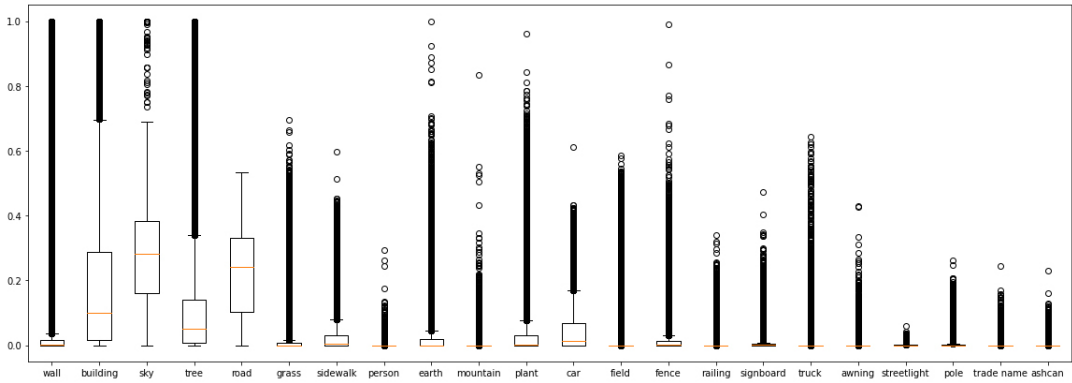


그림 8. 독립 변수의 박스 플롯

* y축은 0~1의 세그먼테이션 비율 값임.

과정을 통해 22개 클래스(wall(벽, 판유리), building(건물), sky(하늘), tree(나무), road(길), grass(풀, 잔디), sidewalk(보행로), person(사람), earth(흙, 땅), mountain(산), plant(식물), car(자동차), field(들판), fence(펜스, 울타리), railing(철조망, 난간), signboard(간판), truck(트럭), awning(차양, 가림막))만 독립변수로 포함하였다.

거리 영상은 한 장 당 가로 1,200 픽셀, 세로 1,200 픽셀이고 총 1,440,000개의 픽셀로 이루어져 있다. 세그먼테이션 값은 1,440,000개의 픽셀에서 22개의 객체가 각각 차지하는 비율 값이다. 독립 변수로 사용된 22개 객체의 박스 플롯을 통해(그림 8) 객체마다 이상치(outlier)가 많음을 알 수 있다. 각 객체의 이상치를 제거할 경우 모델의 정확도가 높아지는 장점이 있지만, 박스에서 벗어난 값도 정성적 평가 점수에 영향을 주었을 가능성이 높다고 판단하고 이상치 제거 작업을 하지 않았다. 예를 들어 한 영상의 90퍼센트가 나무로 이루어져 있을 경우 해당 영상의 'tree' 세그먼테이션 값은 평균에서 많이 벗어나지만 나무가 많은 지역을 걷기 불편하다고 판단하여 정성적 평가 점수가 낮다면 이상치도 종속 변수에 영향을 주는 요소로 의미 있다고 판단하였다.

5. 기계학습 모델과 정확도 평가 기준

거리 영상별 22개 객체의 면적 비율로 나타난 독립변수 자료를 활용하여 종속변수인 보행환경 정성평가 점수를 가장 잘 설명할 수 있는 기계학습 모델을 찾기 위해 9개의 회귀모델을 테스트하였고 마지막으로 의사결정나무 모델을 통해 객체가 점수 도출에 어떻게 작용하였

는지 세부적으로 살펴보았다. 사용한 기계학습 모델은 다중회귀분석(multilinear regression: MLR), 라쏘 회귀(LASSO regression), 릿지 회귀(Ridge regression), 엘라스틱 넷 회귀(elastic net regression), XGBoost 회귀(XGB), 랜덤 포레스트(Random Forest: RF), 서포트 벡터 회귀(Support Vector Regression: SVR), 인공신경망(Artificial Neural Network: ANN), 의사결정나무(decision tree regression: DT)이다.

기계학습 모델의 정확도 평가를 위해 거리영상 196,624개의 관측치 중 70%는 모델 훈련에 사용하고 30%는 검증용으로 활용하였으며, 세그먼테이션 결과 값이 0에서 1 사이의 비율 값이기 때문에 정규화 작업은 진행하지 않았다. 최종모델의 선택은 평균제곱오차(mean squared error; MSE)(Eq. (1)), 평균절대오차(mean absolute error; MAE)(Eq. (2)), 결정계수(R-squared)(Eq. (3)) 값을 모두 고려하였다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \tag{1}$$

평균제곱오차는 실제 값과 예측 값 간 오차의 평균을 제공한 것이다. 값이 작을수록 정확도가 높다고 판단한다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \tag{2}$$

평균절대오차는 실제 값과 예측 값 간 절대오차의 평균이다.

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST} \quad (3)$$

결정계수는 0에서 1사이의 값으로 0에 가까울수록 회귀식의 정확도는 낮고 1에 가까울수록 회귀식이 유의하다고 판단한다.

$$SST = \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (4)$$

Sum of Squares Total(SST)는 각 y값의 편차제곱의 합이다.

$$SSE = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad (5)$$

Explained Sum of Squares(SSE)는 회귀식의 추정 y값과 관측값 평균인 편차제곱의 합이다.

$$SSR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (6)$$

Sum of Squares Residual(SSR)은 Residual Sum of Squares(RSS)라고 부르기도 하며 잔차제곱의 합이다.

1) 다중선형회귀

다중선형회귀(Eq. (7))는 하나의 종속변수에 대한 다중독립변수의 선형모형이다. 학술 연구에 가장 널리 사용되는데, 다양한 유형의 데이터에 대한 적용 가능성과 해석의 용이성 등으로 인해 많이 사용되는 모델이라 할 수 있다(Mason *et al.*, 1991).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon \quad (7)$$

2) 릿지, 라쏘 회귀모델

릿지와 라쏘 회귀모델은 선형회귀모형에 추가적인 제약 조건을 걸어 MSE를 최소화하는 L1-norm과 L2-norm을 사용한다(Melkumova *et al.*, 2017). 두 모델의 가장 큰 특징은 λ 값을 회귀계수에 곱하여 회귀계수의 값이 너무 커지거나, 유의하지 않은 계수의 값을 낮추는 역할을 하는 것이다. 즉, 라쏘는 λ 값을 조정함으로써 유의

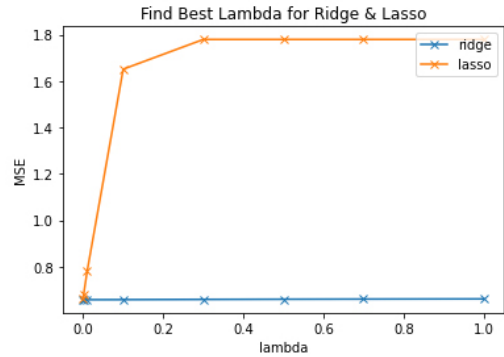


그림 9. 릿지와 라쏘의 λ 값 테스트 결과

하지 않은 가중치를 0이 되게 하여 모델을 일반화하고 특성을 제거한다. 릿지는 라쏘와 비슷하지만 가중치를 0이 아니라 0에 가깝게 만드는 모델이다.

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (8)$$

릿지 회귀식은 β(가중치)를 제공하여 사용하고 λ 값으로 가중치의 크기를 조정한다. 이를 통해 큰 가중치를 λ 값을 통해 큰 폭으로 조정하여 식을 일반화할 수 있다.

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (9)$$

라쏘 회귀식은 β를 절대 값으로 취해 회귀식에서 유의하지 않은 변수의 계수를 0에 가깝도록 추정한다.

릿지와 라쏘 모델의 파라미터인 λ 값을 구하기 위해 0.0001, 0.001, 0.01, 0.1, 0.3, 0.5, 0.7, 1을 λ 값으로 적용해 MSE 값을 측정하였다(그림 9). 그 결과 릿지와 라쏘 모두 λ 값이 0.0001일 때의 MSE값이 각각 0.6580과 0.6637으로 가장 낮았다.

3) 엘라스틱 넷 회귀모델

엘라스틱 넷 회귀모델은 릿지와 라쏘의 절충안으로, L1-norm과 L2-norm을 모두 제약조건으로 갖는 모형이다. 특히 엘라스틱 넷은 그룹화 효과를 주도하고 예측 정확도가 높은 희소 모델(sparse model)을 생성하며 라쏘보다 우수한 성능을 보여준다(Zou *et al.*, 2005). 이 모델에서의 파라미터는 α 값과 L1-ratio이며 α 값은 릿지,

라쏘에서의 λ 와 같이 가중치에 대한 규제 값이고 L1-ratio는 0일 경우 L1-norm만 사용, 1일 경우 L2-norm만 사용한다는 의미이다. 0에서 1사이의 값을 설정함으로써 L1-norm과 L2-norm을 적절히 혼합하여 사용하게 된다. 엘라스틱 넷의 튜닝 역시 릿지와 라쏘처럼 α 값과 L1-ratio값을 0.0001, 0.001, 0.01, 0.1, 0.3, 0.5, 0.7, 1과 0.2, 0.5, 0.8로 두고 최적의 파라미터를 찾는 방법을 사용했다. 그 결과 α 값 0.0001, L1-ratio값 0.8일 때 MSE가 가장 높았다.

4) 의사결정나무

의사결정나무 회귀모델은 의사결정나무 분류기의 변형으로 데이터를 파티션으로 분할하는데 이진 재귀 분할을 기반으로 한다(Xu *et al.*, 2005). Xu *et al.*(2005)에 의하면 데이터를 두 개의 파티션으로 분리할 때 평균에서 제곱 편차의 합을 최소화하는 방법으로 분할하고 이 방법은 다음 분기의 파티션에 각각 적용된다. 본 연구에서는 트리의 깊이를 5로 설정해 분할이 5번 이루어지게 하고 각 파티션의 결과에 해당하는 거리 영상을 살펴보았다.

5) XGBoost

XGBoost는 확장 가능한 엔드 투 엔드 트리 부스팅(end to end tree boosting) 시스템으로 앙상블 기법인 그래디언트 트리 부스팅(Gradient Tree Boosting)을 이용한 모델이다(Chen and Guestrin, 2016). XGBoost 모델은 회귀(regression)와 분류(classification)를 모두 지원한다. 본 연구에서는 부스터 파라미터(booster parameter)로 'gblinear'를 사용하였고 트리의 깊이는 6으로 설정하였다.

6) 랜덤 포레스트

랜덤 포레스트 모델은 분류, 회귀 분석 등에 사용되는 앙상블 학습 방법으로 m개의 무작위 회귀 트리의 모음으로 구성된다(Biau and Scomet, 2016). 하나의 결정 트리만 사용하면 규칙 중간에 오류가 생길 경우 그 오류가 다음 노드로 계승된다는 단점이 있다. 이를 보완하기 위해 랜덤 포레스트는 동일한 데이터로 여러 개의 결정 트리를 만든다. 서로 독립된 여러 결정 트리를 생성하고 각 결정 트리에 사용되는 데이터의 중복 사용을 허용하는 배깅(bagging)과, 이전 결정 트리의 오차를 보완해 다

음 결정 트리에 적용하는 연속적 방법인 부스팅(boosting)으로 랜덤 포레스트를 구성하는 다수의 결정 트리로부터 평균 예측치를 출력하기 때문에 높은 예측 성능을 기대할 수 있는 모델이다. 본 연구에서는 랜덤 포레스트의 최적의 파라미터를 찾기 위해 GridSearchCV 기법을 사용하였다. 교차 검증(cross validation)을 5 fold로 설정하였고 추정기(estimator), 최대 깊이(max depth), 최소 샘플 리프(min sample leaf), 최소 샘플 분할(min sample split)에 대한 파라미터 테스트를 진행했다. 그 결과 추정기, 최대깊이, 최소 샘플 리프, 최소 샘플 분할에 대한 최적 값은 200, 8, 8, 8로 도출되었다. 이를 랜덤 포레스트 회귀 모델의 파라미터로 적용하였다.

7) 서포트 벡터 회귀모델

서포트 벡터 회귀모델은 서포트 벡터 머신의 회귀모델로, 예측 값과 실제 값의 차이를 최소화하는 선형식을 찾고 최적화하는 모델이다(Awad and Khanna, 2015). 서포트 벡터 회귀모델은 실제 값과 예측 값의 차이를 마진(margin) 안의 오차($\pm\epsilon$) 이내로 유지하는 것이 목적이다(김준봉 등, 2016). 그렇기 때문에 마진의 넓이를 어느 정도로 설정하는 지가 중요하다. 서포트 벡터 회귀모델을 튜닝(tuning)할 때 주요 파라미터(parameter)는 커널(kernel)이다. 커널의 종류는 선형 커널(linear kernel), 가우시안 커널(gaussian kernel, RBF(radial basis function)), 다항 커널(polynomial kernel)이 있고 본 연구에서는 RBF를 커널로 채택하여 사용하였다. 또한 오차를 얼마나 허용하는지에 따라 소프트 마진(soft margin)과 하드 마진(hard margin)으로 나눌 수 있다(그림 10). 서포트 벡터 모델은 이상치에 민감하나 본 연구에 사용된 데이터는 이상치도 정성적 평가 점수 예측에 영향을 주었을 것으로 보고 이상치를 제거하지 않았기 때문에 마진을 조정하는 파라미터인 C(cost)는 기본 설정 값인 1.0으로 하였다.

8) 인공신경망

인공신경망은 인간의 뉴런을 본떠 패턴화 한 컴퓨팅 시스템이다(Saritas and Yasar, 2019). 뉴런과 유사한 기능을 하는 노드가 다층의 layer를 형성하고 있다. 이러한 구조는 복잡한 선형적 회귀 또는 분류 작업을 하는 데에 용이한 장점을 가지고 있다. 본 연구에서는 앞선 기계학습 기법과의 비교를 위해 인공신경망을 사용해 상관관

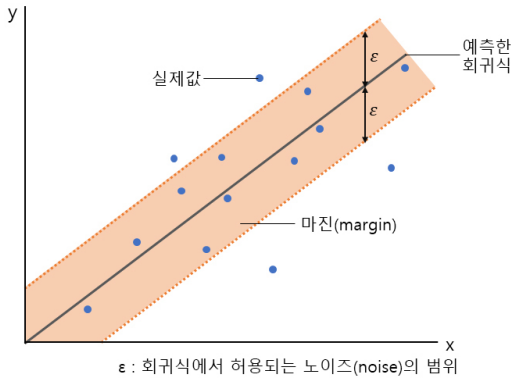


그림 10. 서포트 벡터 모델의 기본 아이디어

* 마진의 넓이를 좁게 할수록 하드 마진, 넓게 할수록 소프트 마진임.

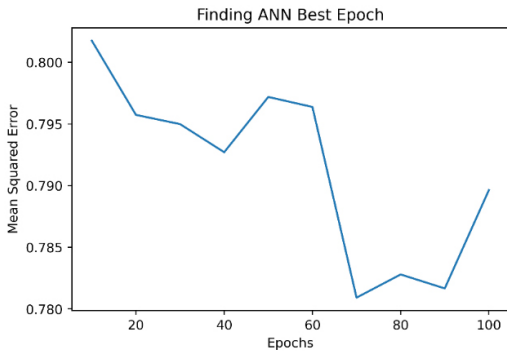


그림 11. 인공신경망 모델 에포크 최적화 실험 결과

계를 살펴보았다. 모델은 총 2개의 은닉층(hidden layer)으로 구성하였고 각 층은 15, 20개의 노드(node)로 형성하였다. 최적의 성능을 위해 에포크(epoch)를 10부터 10단위로 올려 100까지 적용하는 실험을 하였고, 그 결과 MSE가 가장 낮은 에포크는 70이었고 이를 반영해 결과를 도출하였다.

IV. 정확도 비교 및 객체 중요도 평가

기계학습 모델의 정확도를 비교한 결과는 표 2와 같다. 서포트 벡터 회귀모델의 MSE와 MAE 값이 0.4697과 0.4982로 MSE와 MAE가 두 번째로 낮은 랜덤 포레스트 모델보다 약 0.1정도 낮은 수치이고 인공신경망 모델의 MSE와는 0.3172 차이로, 모델 구축과 해석이 복잡한 인공신경망을 사용하는 것 보다 서포트 벡터 회귀를 사용하는 것이 더 효율적이고 정확도도 높은 것으로 나타났다. 서포트 벡터 회귀의 R² 값은 0.6106으로 다중선형회귀, 릿지, 라쏘에 비해 낮지만 0.01~0.02 차이이므로 세 가지 지표 모두를 고려했을 때 값이 모두 고르게 잘 도출된 서포트 벡터 회귀모델이 종속변수를 가장 잘 예측하는 모델로 판단하였다.

정성적 평가 점수를 가장 잘 예측한다고 판단된 서포트 벡터 회귀모델의 결과를 바탕으로 변수 중요도를 살펴본 것은 그림 12와 같다. 변수 중요도는 순열 중요도(permutation importance) 기법을 통해 도출하였는데 이는 훈련된 모델의 독립변수 중 모델에 가장 영향력이 큰 변수를 파악하는 방법이다. 모델이 특정 변수를 사용하지 않았을 때 성능이 얼마나 저하되는지에 따라 변수의 중요도를 결정한다. 순열 중요도로 본 정성적 보행환경 평가 지수는 road, sidewalk, building, tree, sky의 순서로 영향을 받은 것으로 보였다. 특히 road와 sidewalk의 변수 중요도가 각각 0.6437과 0.4074로 다른 객체보다 월등히 높게 나옴을 확인할 수 있다.

서포트 벡터 회귀식에서 객체에 대한 회귀계수의 방향을 통해서도 객체가 정성적 평가에 미치는 영향력을 확인할 수 있다(그림 13). 양의 회귀계수를 보이는 객체로는 보행로(sidewalk), 가로등(streetlight), 길(road), 잔디(grass), 나무(tree) 등이며 이는 세그멘테이션 값이 높을수록 보행환경의 정성적 평가에 높은 점수를 만들도록 한다. 반면 쓰레기통(ashcan), 산(mountain), 트럭(truck), 담벼락(wall) 등은 음의 방향으로 역할을 함을 알 수 있

표 2. 9개 모델의 MSE, MAE, R² 결과

	MLR	Ridge	LASSO	Elastic	XGB	RF	SVR	ANN	DT
MSE	0.6580	0.6580	0.6637	0.6657	0.6768	0.5285	0.4697	0.7869	0.6889
MAE	0.5916	0.5916	0.5939	0.5947	0.5958	0.5320	0.4982		0.6120
R ²	0.6303	0.6303	0.6271	0.3891	0.3798	0.5492	0.6106		0.6129

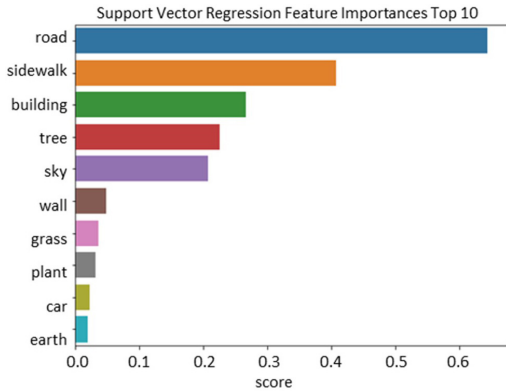


그림 12. 서포트 벡터 회귀모델 변수 중요도 그래프

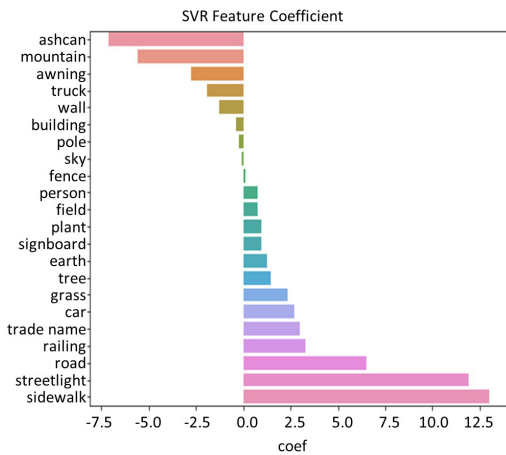


그림 13. 서포트 벡터 회귀 계수 그래프

다. 이는 거리영상에서 이들 객체가 차지하는 비율이 높을수록 정성적 보행환경 평가점수는 낮음을 의미한다. 각 객체의 중요도와 함께 방향을 고려하면 도로(road)와 보행로(sidewalk)가 정성평가에 중요한 요소이며 양의 방향으로 작용함을 알 수 있으며, 건물(building)이나 하

늘(sky)은 중요도는 높게 나오는데 이들이 거리영상에서 미치는 영향은 음의 방향임을 알 수 있다. 건물(building)이 보행환경의 정성적 평가에 미치는 영향은 잘 정비된 지역의 건물, 혹은 랜드마크적 요소가 있는 건물을 생각할 수도 있지만 낙후되고 노후화된 기존시가지의 건물은 음의 방향으로 작용하는 것으로 보인다. 이와 함께 보행환경 영상이 도로의 앞뒤 진행방향과 함께 양 옆을 촬영한 영상도(그림 3) 포함하고 있음을 고려했을 때 건물과 벽에 해당하는 세그멘테이션 비율 값이 높음은 상대적으로 보행을 할 수 있는 공간인 길과 보행로가 적음을 의미하고 이는 보행환경의 정성적 평가에 부정적 영향을 미치는 것으로 해석할 수 있다. 하늘(sky)과 관련해서도 해외 연구에서 보행환경의 개방감은 정성적 평가에 긍정적인 것으로 평가되는 것이 일반적이다. 그러나 전주시와 같이 도시 외곽지역에 도로가 발달되지 않고, 나지 형태로 있는 곳은 하늘 객체의 비율이 높아 하늘 개방감 지수는 좋을 수 있지만, 보행여건이 좋지 않아 정성적 보행환경 평가 점수는 낮게 평가되는 결과와 연계되는 것으로 판단된다.

거리영상 객체가 보행환경 정성 평가 예측에 어떻게 영향을 미치는지 더 세부적으로 살펴보기 위해 의사결정나무 모델을 이용해 확인해 보았다(그림 14). 나무의 깊이를 최대 5로 설정하고 모델을 실험한 결과 첫 번째 가지에서는 '도로(road)'의 세그멘테이션 값에 따라 노드가 나뉘었고 두 번째 가지에서는 두 노드 모두 '보행로(sidewalk)'의 값이 노드를 나누는 기준으로 사용되었다. 나머지 세 번의 분할에서도 road와 sidewalk의 등장 빈도가 가장 높았다. 네 번째와 다섯 번째 가지에서는 '나무(tree)'가 분할 조건으로 빈번하게 등장했고 그 외에 '건물(building)'과 '잔디/초원(grass)', '자동차(car)'가 분할 조건으로 사용되었다. 이는 서포트 벡터 회귀모델 변수 중요도의 상위권에 위치한 객체가 의사결정나무 노

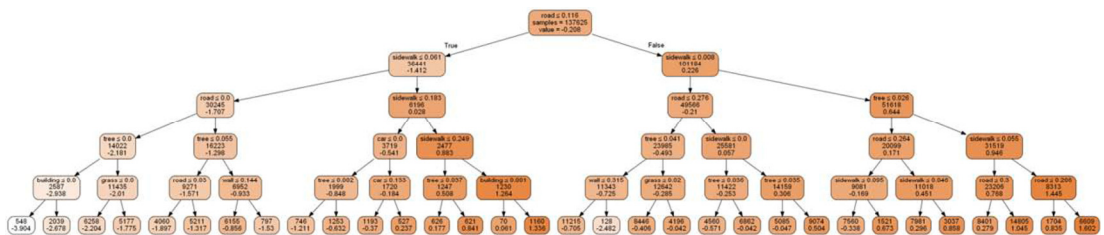


그림 14. 의사결정나무 시각화 그래프

* 노드 중앙은 관측 값 개수, 하단은 정성적 평가 예측 점수임.

드 분할 규칙에도 중요하게 적용되었음을 보여준다.

몇 가지 예를 살펴보면, 규칙을 처음 시작하는 노드에서 road의 비율 값이 0.116을 기준으로 설정되어 있다. 이는 영상에서 도로가 10% 이상 존재하는지 여부가 가장 먼저 고려됨을 의미한다. 영상에서 sidewalk가 차지하는 비율이 0.183보다 많을 경우 정성적 평가 점수가 모두 양수로 도출되었고 0.183보다 적을 경우 정성적 평가 점수는 음수로 도출되었다.

V. 결론

최근 고해상도의 거리영상과 딥러닝 기술의 발전은 거리영상에 대한 쌍별 비교 데이터를 훈련셋으로 하여 가로단위의 상세한 정성적 보행환경 평가점수 획득을 가능하게 하고 있다. 하지만 딥러닝 기술에 기반한 정성적 보행환경 평가 점수 예측은 왜 이러한 평가점수를 얻게 되었는지에 대한 해답을 제공하는데 한계가 있다. 본 연구의 목적은 거리영상을 기반으로 사람들이 걷기 좋다고 느끼는 보행환경에 영향을 미치는 도시경관의 특성을 분석하는 것이다. 이를 위해 거리영상의 정성적 보행환경 점수를 종속변수로, 거리영상의 시멘틱 세그멘테이션 비율 값을 독립변수로 하여 다양한 기계학습 모델을 실험하였다.

연구대상지는 전주시이며, 전주시를 대상으로 수집된 196,624장의 영상에 딥러닝 모델 적용을 통해 예측된 정성적 보행환경 평가 점수를 종속변수로, 196,624장에 시멘틱 세그멘테이션 기법을 적용하여 도출된 22개 객체별 비율 값을 독립변수로 9개의 기계학습 모델을 적용하고 정확도를 평가하였다. 모델의 성능지표는 MAE, MSE, R^2 등 3개의 값을 사용하였으며, 9개의 기계학습 모델 가운데 서포트 벡터 회귀모델의 정확도가 가장 높게 나왔다. 서포트 벡터 회귀모델을 통해 각 객체의 중요도를 확인할 수 있었는데, 도로와 보행로, 건물, 나무, 하늘 등의 순으로 객체 중요도가 도출되었다. 변수의 방향성을 고려하였을 때 보행로(sidewalk), 가로등(streetlight), 길(road), 잔디(grass), 나무(tree) 등은 양의 방향으로 작용하였으며, 이는 세그멘테이션 값이 높을수록 정성적 보행환경 평가를 높게 만들며, 쓰레기통(ashcan), 산(mountain), 트럭(truck), 담벼락(wall) 등은 음의 방향으로, 세그멘테이션 값이 높을수록 평가점수를 낮게 만들

을 뜻한다.

각 객체의 중요도와 함께 방향을 고려하면 도로(road)와 보행로(sidewalk)는 정성평가에 중요한 요소이며 양의 방향으로 작용함을 알 수 있고, 건물(building)이나 하늘(sky)은 중요도는 높게 나오는데 반해 이들이 거리영상에서 미치는 영향은 음의 방향임을 알 수 있었다. 건물(building)이 보행환경의 정성적 평가에 부정적으로 평가되는 이유는 전주시의 낙후되고 노후화된 기존시가지 건물의 영향, 그리고 건물과 벽에 해당하는 세그멘테이션 비율 값이 높음은 상대적으로 보행을 할 수 있는 공간인 길과 보행로가 작기 때문에 정성적 보행환경 점수에는 부정적 영향을 미친 것으로 해석된다. 하늘(sky)과 관련해서도 전주시와 같이 도시 외곽지역에 도로가 발달되지 않고, 나지 형태로 있는 곳은 하늘 객체의 비율 값은 높게 나타나지만 정성적 보행환경 점수는 높게 평가되지 않아 하늘 객체 비율이 높은 것이 정성적 보행환경 평가에 음의 방향으로 작용한 것으로 판단된다.

본 연구는 기계학습 모델을 통해 보행 환경의 정성적 평가에 영향을 주는 중요 객체와 객체의 방향성을 확인할 수 있었으며, 딥러닝 모델을 통해 예측한 보행환경 정성평가 점수를 일부 설명 가능하게 하였다라는 점에 의의가 있다. 향후 세그멘테이션 값과 함께 다른 연구를 통한 수치 데이터를 추가해 모델을 보완하거나, CAM(class activation map)을 통해 정성평가에 영향을 미치는 객체를 분석하는 방법을 실험할 수 있을 것으로 보인다. 아울러 다른 딥러닝 모델을 사용하거나 전처리에서 제거하지 않은 이상치 제거 등 전처리 과정을 변형함으로써 추가적인 연구가 이루어질 수 있을 것으로 예상된다.

참고문헌

- 김규리·이제선, 2016, “보행공간 요소에 대한 보행자의 인지 및 보행만족도에 관한 연구” 한국도시설계학회, 17(3), 89-103.
- 김지연·강영옥, 2022, “거리영상 기반 보행환경의 정성적 평가 예측을 위한 딥러닝 모델 개발” 대한공간정보학회지, 30(2), 45-56.
- 김준봉·오승철·서기성, 2016, “MLR 및 SVR 기반 선형과 비선형회귀분석의 비교: 풍속 예측 보정” 전기학회논문지, 65(5), 851-856.

- 김희철·안건혁·권영상, 2014, “개인의 보행 확률에 영향을 미치는 거주지 환경요인” 한국도시설계학회지, 15(3), 5-18.
- 박근덕·이수기, 2018, “근린환경특성과 일상보행활동 그리고 주관적 건강수준의 구조적 관계 분석: 경로모형의 적용,” 국토계획, 53(1), 255-272.
- 박소현·최이명·서한림·김준형, 2009, “주거지 보행환경 인지가 생활권 보행만족도에 미치는 영향에 관한 연구” 대한건축학회 논문집-계획계, 25(8), 253-261.
- 박영은·이우성, 2022, “보행환경 평가지표를 활용한 통한 노인의 보행환경 인식 분석” 한국공간디자인학회 논문집, 17(3), 200-212.
- 박지영·강영옥·김지연, 2022, “거리 영상과 시멘틱 세그먼테이션을 활용한 보행환경 평가 지표 개발” 한국지도학회지, 22(1), 53-68.
- 박철영·이수기, 2016, “가로환경 특성이 보행자 교통사고에 미치는 영향 분석” 한국도시설계학회지 도시설계, 17(3), 105-121.
- 성현곤·김태호·강지원, 2011, “구조방정식을 활용한 보행환경 계획요소의 이용만족도 평가에 관한 연구: 종로 및 강남일대를 대상으로,” 국토계획, 46(5), 275-288.
- 유승재·하정원·김혜준·기동환·이수기, 2021, “서울시 가로경관 이미지에 대한 주관적 인지에 영향을 미치는 가로환경 요인분석” 국토계획, 56(2), 79-93.
- 이경환·안건혁, 2008, “지역 주민의 보행 활동에 영향을 미치는 근린 환경 특성에 관한 실증 분석: 서울시 12개 행정동을 대상으로” 대한건축학회 논문집 - 계획계, 24(6), 293-302.
- 이수기·고준호·이기훈, 2016, “근린환경특성이 보행만족도에 미치는 영향 분석: 서울서베이 2013 년 자료를 중심으로” 국토계획, 51(1), 169-187.
- 이수민·황기연, 2009, “보행환경요인이 보행안전에 미치는 영향분석” 대한교통학회지, 27(1), 107-114.
- 이지원·유다은, 2021, “보행자의 주관적 인식과 성별이 보행환경 만족도에 미치는 영향분석” 대한건축학회논문집, 37(10), 51-62.
- 조혜민·이수기, 2016, “보행목적별 보행활동시간에 영향을 미치는 근린환경 특성분석-주관적 인지환경과 객관적 측정환경의 차이를 중심으로” 대한국토·도시계획학회, 51(4), 105-122.
- 행정안전부, 2012, 「보행안전 및 편의증진에 관한 법률」.
- Awad, M. and Khanna, R., 2015, Support vector regression, in Awad, M. and Khanna, R., eds., *Efficient Learning Machines*, Berkeley, CA: Apress, 67-80.
- Biau, G. and Scornet, E., 2016, A random forest guided tour, *Test*, 25(2), 197-227.
- Biljecki, F. and Ito, K., 2021, Street view imagery in urban analytics and GIS: A review, *Landscape and Urban Planning*, 215, 104217.
- Bijmolt, T.H. and Wedel, M., 1995, The effects of alternative methods of collecting similarity data for multidimensional scaling, *International Journal of Research in Marketing*, 12(4), 363-371.
- Blečić, I., Cecchini, A., and Trunfio, G.A., 2018, Towards automatic assessment of perceived walkability, *International Conference on Computational Science and Its Applications*, Cham: Springer, 351-365.
- Chen, T. and Guestrin, C., 2016, Xgboost: A scalable tree boosting system, *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, San Francisco, CA, USA, 785-794.
- Dubey, A., Naik, N., Parikh, D., Raskar, R., and Hidalgo, C.A., 2016, Deep learning the city: Quantifying urban perception at a global scale, *European conference on computer vision*, Cham: Springer, 196-212.
- Frank, L.D., Schmid, T.L., Salis J.F., Chapman, J., and Saelems, B.E., 2005, Linking Objectively Measured Physical Activity with Objectively Measured Urban Form: Findings from SMARTRAQ, *American Journal of Preventive Medicine*, 28(2), 117-125.
- Guan, W., Chen, Z., Feng, F., Liu, W., and Nie, L., 2021, Urban perception: sensing cities via a deep interactive multi-task learning framework, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(1s), 1-20.
- Joglekar, S., Quercia, D., Redi, M., Aiello, L.M., Kauer, T., and Sastry, N., 2020, FaceLift: a transparent deep learning framework to beautify urban scenes, *Royal Society Open Science*, 7(1), 190987.
- Kim, J.H., Lee, S., Hipp, J.R., and Ki, D., 2021, Decoding

- urban landscapes: Google street view and measurement sensitivity, *Computer, Environment and Urban Systems*, 88, 101626.
- Kim, S., Park, S., and Lee, J.S., 2014, Meso-or micro-scale? Environmental factors influencing pedestrian satisfaction, *Transportation Research Part D: Transport and Environment*, 30, 10-20.
- Li, X. and Ratti, C., 2018, Mapping the spatial distribution of shade provision of street trees in Boston using Google Street View panoramas, *Urban Forestry & Urban Greening*, 31, 109-119.
- Li, Y., Yabuki, N., Fukuda, T., and Zhang, J., 2020, A big data evaluation of urban street walkability using deep learning and environmental sensors-a case study around Osaka University Suita campus, *The Cognitive City(AI)*, 2, 319-328.
- Lu, X., Lin, Z., Jin, H., Yang, J., and Wang, J.Z., 2014, Rapid: Rating pictorial aesthetics using deep learning, *Proceedings of the 22nd ACM International Conference on Multimedia*, Orlando, Florida, USA, 457-466.
- Lu, X., Lin, Z., Shen, X., Mech, R., and Wang, J.Z., 2015, Deep multi-patch aggregation network for image style, aesthetics, and quality estimation, *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 990-998.
- Mason, C.H. and Perreault Jr, W.D., 1991, Collinearity, power, and interpretation of multiple regression analysis, *Journal of Marketing Research*, 28(3), 268-280.
- Mateo-Babiano, I., 2016, Pedestrian's needs matter: Examining Manila's walking environment, *Transport Policy*, 45, 107-115.
- Melkumova, L.E. and Shatskikh, S.Y., 2017, Comparing Ridge and LASSO estimators for data analysis, *Procedia Engineering*, 201, 746-755.
- Min, W., Mei, S., Liu, L., Wang, Y., and Jiang, S., 2019, Multi-task deep relative attribute learning for visual urban perception, *IEEE Transactions on Image Processing*, 29, 657-669.
- Nagata, S., Nakaya, T., Hanibuchi, T., Amagasa, S., Kikuchi, H., and Inoue, S., 2020, Objective scoring of streetscape walkability related to leisure walking: Statistical modeling approach with semantic segmentation of Google Street View images, *Health & Place*, 66, 102428.
- Quercia, D., Aiello, L.M., Schifanella, R., and Davies, A., 2015, The digital life of walkable streets, *Proceedings of the 24th international conference on World Wide Web*, Florence, Italy, 875-884.
- Salesses, P., Schechtner, K., and Hidalgo, C.A., 2013, The collaborative image of the city: mapping the inequality of urban perception, *PLOS ONE*, 8(7), e68400.
- Santani, D., Ruiz-Correa, S., and Gatica-Perez, D., 2018, Looking south: Learning urban perception in developing cities, *ACM Transactions on Social Computing*, 1(3), 1-23.
- Saritas, M.M. and Yasar, A., 2019, Performance analysis of ANN and Naive Bayes classification algorithm for data classification, *International Journal of Intelligent Systems and Applications in Engineering*, 7(2), 88-91.
- Stewart, N., Brown, G.D., and Chater, N., 2005, Absolute identification by relative judgment, *Psychological Review*, 112(4), 881-911.
- Xu, M., Watanachaturaporn, P., Varshney, P.K., and Arora, M.K., 2005, Decision tree regression for soft classification of remote sensing data, *Remote Sensing of Environment*, 97(3), 322-336.
- Xu, Y., Yang, Q., Cui, C., Shi, C., Song, G., Han, X., and Yin, Y., 2019, Visual Urban Perception with Deep Semantic-Aware Network, *International Conference on Multimedia Modeling*, Cham: Springer, 28-40.
- Yoo, K., Ryu, D., Lee, C., Nam, K., and Kang, Y., 2021, Creation of a paired comparison set for preference evaluation based on crowdsourcing, *Proceedings of Korea Industrial Information Systems Research*, Incheon, Korea, 2-6.
- Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., and Cottrell, G., 2018, Understanding convolution for semantic segmentation, *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Piscataway, NJ: IEEE, 1451-1460.
- Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., and

- Grekousis, G., 2019, Perceptions of built environment and health outcomes for older Chinese in Beijing: A big data approach with street view images and deep learning technique, *Computers, Environment and Urban Systems*, 78, 101386.
- 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.
- Zhou, H., He, S., Cai, Y., Wang, M., and Su, S., 2019, Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning, *Sustainable Cities and Society*, 50, 101605.
- Zou, H. and Hastie, T., 2005, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society: Series B (statistical methodology)*, 67(2), 301-320.
- Zou, W., Zhang, D., and Lee, D.J., 2021, A new multi-feature fusion based convolutional neural network for facial expression recognition, *Applied Intelligence*, 52(3), 2918-2929.
- Cityscapes Dataset, <https://www.cityscapes-dataset.com/>
ADE20K, <https://groups.csail.mit.edu/vision/datasets/ADE20K/>
- 교신 : 강영옥, 03760, 서울시 서대문구 이화여대길 52, 이화여자대학교 사회과교육과 지리전공(이메일: ykang@ewha.ac.kr)
- Correspondence : Youngok Kang, 03760, 52 Ewhayodae-gil, Seodaemun-Gu, Seoul, Korea, Department of Social Studies, Geography major, Ewha Womans University (Email: ykang@ewha.ac.kr)
- 투고접수일: 2022년 8월 19일
심사완료일: 2022년 9월 11일
게재확정일: 2022년 9월 19일

