



자전거 교통사고 다발지역 예측을 위한 딥러닝 모형의 적용*

An Application of Deep Learning for Detecting Bicycle Collision Hotspots

전희정** · 강서윤*** · 강승엽**** · 조철호***** · 허재필*****

Jun, Hee-Jung · Kang, Seoyoon · Kang, Seungyeoup · Cho, Cheol-Ho · Heo, Jae-Pil

Abstract

As the paradigm shift in transportation has led to an increased use of bicycles, the need to ensure the bicycle users' safety by preventing bicycle collisions has also gained importance. This study aims to predict bicycle collision hotspots in Korea by using collected Google Street View (GSV) images of bicycle collision hotspots and non-hotspots. We conducted experiments with five deep learning models (VGG16, 19, ResNet50, 101, and Inception), and employed Class Activation Map (CAM) analysis to visualize the factors contributing to bicycle collisions. Results suggest that the VGG19 model is the best model for predicting bicycle collision hotspots. Additionally, the CAM analysis indicates that larger roads with more vehicles and more physical facilities (such as crosswalks, traffic lights, and pillars) are related to bicycle collision hotspots. This study indicates the effectiveness of using GSV images together with a deep learning model and suggests more concrete and specific policy implications for building a safer environment for bicycle users.

주제어 자전거, 교통사고, 딥러닝
Keywords Bicycle, Collision, Deep Learning

1. 서론

기후변화, 에너지 고갈 등 전 세계적인 환경문제에 대응하여 친환경 녹색성장의 개념이 대두되었다. 교통 분야에서는 지속가능한 교통체계를 구현하는 것이 제1 당면과제로서 주목받고 있다. 이를 위해 화석연료를 기반한 자동차와 같은 동력교통수단에서 녹색교통수단으로의 패러다임 전환이 필요한 실정이다. 특히 자전거는 대표적인 무동력·비탄소 교통수단으로, 지속가능한 교통체계 구현에 매우 중요한 역할을 할 수 있다. 뿐만 아니라 자전거 이용에 따른 신체활동의 증가로 비만을 및 사망률 감소에 긍정

적인 영향을 준다는 연구결과가 보고되고 있다(De Hartog et al., 2010; Oja et al., 2011; Rabl and De Nazelle, 2012; Rojas-Rueda et al., 2011; Wen and Rissel, 2008). 따라서 우리나라 역시 국민들의 자전거 이용을 독려하기 위해 공공자전거의 도입, 자전거도로 등 인프라의 확장과 같이 다양한 제도를 시행하고 있으며, 1990년대 중반부터 자전거 시범도시를 선정하고 자전거 이용 활성화를 위한 법률을 제정하는 등의 노력을 기울여 왔다(김세용 외, 2008). 이러한 자전거 이용 활성화 정책에 힘입어 2017년에는 국내 자전거 이용인구가 1340만명으로, 전체 인구의 33.5%를 차지하는 것으로 나타났으며(한국교통연구원,

* 이 논문은 성균관대학교의 2021학년도 AI융합연구비에 의하여 연구되었음.

** Associate Professor, Department of Public Administration/Graduate School of Governance, Sungkyunkwan University (First/Corresponding Author: hjun@skku.edu)

*** Master's Student, Graduate School of Governance, Sungkyunkwan University (syoony9811@g.skku.edu)

**** Ph.D. Candidate, Graduate School of Governance, Sungkyunkwan University (kkangyoup@g.skku.edu)

***** Ph.D. Student, Department of Artificial Intelligence, Sungkyunkwan University (gersys@skku.edu)

***** Assistant Professor, Department of Computer Science and Engineering/Department of Artificial Intelligence, Sungkyunkwan University (jaepilheo@skku.edu)

2017), 서울시의 공유자전거인 '따릉이'의 경우 2022년에 들어 도입한 지 약 6년 만에 누적이용량이 1억 건에 돌파하였다(서울경제, 2022.04.27; 이투데이, 2022.04.27). 그러나 이러한 성과 역시 해외의 사례와 비교하였을 때는 미미한 수준에 그친다고 할 수 있다.

자전거 이용 활성화를 위해서는 무엇보다 자전거 이용자의 안전확보가 중요하다. 실제로 세계적으로 많은 자전거 이용자를 보유하고 있는 네덜란드, 독일, 일본 등은 자전거 이용의 안전확보를 위해 우리나라보다 강력하고 체계적인 법제도를 시행하고 있다(정경옥 외, 2019). 도로교통공단의 보도자료에 따르면 우리나라의 자전거 교통사고는 2018년 4,771건에서 2019년 5,633건, 2020년 5,667건까지 꾸준히 증가하는 양상을 보이며, 2021년에는 5,509건으로 소폭 감소하였으나 여전히 5천건을 넘으며 자전거 이용자의 안전이 크게 위협받고 있음을 확인할 수 있다. 반면, 교통체계의 패러다임 전환으로 자전거와 더불어 주목받고 있는 보행의 경우, 보행교통사고가 2019년 46,682건에서 2020년과 2021년 각각 36,601건, 35,665건으로 집계되며, 코로나19로 인한 외부활동 감소를 감안하더라도 자전거 교통사고와 달리 교통사고 건수가 크게 감소했음을 확인할 수 있다(TAAS, 2022.08.09). 특히, 보행교통사고 사망자는 2016년부터 2021년 사이에 약 41.1% 대폭 감소하였다. 이는 「보행안전 및 편의 증진에 관한 법률」의 제정이나 「보행안전 종합대책」의 시행 등 보행자의 안전확보를 위한 다양한 노력의 결과라고 할 수 있다. 또한, 안전속도 5030과 민식이법, 어린이 보호구역 안전속도 및 노인 보호구역 확대 등 교통안전 종합대책의 성과로 볼 수 있다(국무조정실 국무총리비서실, 2022). 이처럼 안전확보를 위한 법과 제도의 중요성 및 효과성을 확인할 수 있으나, 아직까지 자전거 이용자의 안전을 위한 구체적인 법이나 제도가 없다는 점에서 앞으로 자전거 교통사고를 예방하기 위한 대책을 강구할 필요가 있다.

이에 그간 연구들은 도로교통공단이 운영하는 교통사고분석시스템(Traffic Accident Analysis System, 이하 TAAS)에서 제공하는 자전거 교통사고 자료를 통해 그 영향요인을 규명하고자 노력해왔다. 선행연구들은 일반적으로 자전거 교통사고 영향요인으로서 자전거 운전자의 인적사항과 법규위반 등의 인적 요인과 가로환경, 인근시설, 토지이용 등의 환경적 요인을 제시하고 있다. 이 중 인적 요인의 경우 각 사고의 특징으로서 교통환경 개선을 위한 일반화에 한계가 있다는 점에서(류종득, 2018; 전희정 외, 2021), 본 연구는 자전거 교통사고 예방을 위한 시사점 도출을 위해 환경적 요인에 초점을 맞추었다. 특히 그중에서도 가로환경에 집중하고 수도권과 비수도권 간 지역적 차이를 탐구하고자 하였다.

한편 그간 교통사고에 영향을 미치는 환경적 요인을 다룬 연구들은 행정동 수준의 데이터를 사용해왔다는 점에서 사고를 유발하는 보다 세부적인 영향요인의 파악에는 한계가 존재한다. 이에

본 연구는 자전거 교통사고 다발지역의 위치정보를 활용하여 Google Street View(이하 GSV) 이미지를 수집하여 분석을 진행함으로써 선행연구들이 갖는 분석단위의 한계를 극복하고자 하였다. 보다 세부적으로 본 연구는 GSV 이미지의 분석에 있어 컴퓨터 비전과 딥러닝 기법을 활용하여 자전거 교통사고 다발지역을 예측하는 모형을 구축하고자 한다. 또한, 자전거 교통사고 다발지역과 비다발지역을 구분하는 이미지상의 특징(feature)을 시각화하는 Class Activation Map(이하 CAM)을 통해 다발적인 자전거 교통사고를 유발하는 환경요인을 탐색하고자 한다.

이러한 최신 기법의 적용으로 자전거 교통사고의 유발요인을 도출하여 스마트 도시관리적 측면에서 자전거 이용자의 안전확보를 위한 보다 구체적인 정책적 시사점을 도출하고자 한다. 이는 나아가 자전거와 더불어 녹색교통수단으로 주목 받고 있는 전동휠, 전동킥보드 등의 개인형 이동수단(Personal Mobility, 이하 PM) 이용의 안전확보에도 기여할 것이라 기대된다. 따라서 본 연구의 분석을 통해 도출한 자전거 교통사고 예방책이 앞으로 더욱 다양해질 교통수단의 안전확보에 함의를 제공할 것이라 기대할 수 있다.

II. 이론적 논의

1. 자전거 교통사고 영향요인

자전거 교통사고에 관한 선행연구들은 사고의 발생건수 및 심각도에 영향을 미치는 요인을 크게 인적 요인과 환경적 요인으로 구분하여 제시하고 있다. 먼저 자전거 교통사고의 인적 요인에는 자전거 운전자의 성별과 연령 등 인적사항과 법규위반 등이 포함된다. 성별의 경우 선행연구들은 공통적으로 남성일 때 자전거 교통사고 발생건수가 높으며(신혜정 외, 2012; 조현우·김동진, 2020; 최승희·이구연, 2016), 심각도에 있어서 최승희와 이구연(2016)은 성별에 따라 자전거 교통사고 사상자 수의 차이가 있을 수 있음을 밝히기도 했다. 연령의 경우 청소년일 때 사고 발생건수 및 심각도가 더 높다고 밝힌 연구들과(Brown et al., 2002; 신희철·박재영, 2012), 고령층에서 더 많은 자전거 교통사고가 발생하며 그 심각도 역시 높다고 밝힌 연구들이(Rodgers, 1997; 신혜정 외, 2012) 혼재하며, 사고에 취약한 자전거 운전자의 연령대는 청소년과 고령자임을 확인할 수 있다(김태양·박병호, 2017). 법규위반의 경우 자전거 이용자가 법규를 위반하였을 때 더 심각한 사고가 발생할 수 있으며(신혜정 외, 2012), 그간 자전거 교통사고들의 법규위반 상황을 살펴보면 안전의무 불이행이 가장 많았고, 그 뒤로 중앙선 침범, 신호위반, 교차로 통행방법 위반, 안전거리 미확보, 보행자 보호의무 위반 등이 따랐다.

자전거 교통사고의 환경적 요인에는 가로환경, 인근시설, 토지이용 등이 포함된다. 먼저 가로환경 요인과 관련하여 많은 연구

들이 교차로에서의 자전거 교통사고를 대상으로 연구를 진행해 왔다. 교차로 주변에서 자전거 교통사고가 증가한다는 점이 밝혀졌으며, 이는 교차로에서 교통량이 증가하여 자전거와 차량 간 상충이 일어나기 때문이라고 해석된다(김도훈 외, 2008; 김숙희 외, 2017; 김응철 외, 2007; 오주택, 2018; 오주택 외, 2007; Lee et al., 2017; Wang and Nihan, 2004; Wei and Lovegrove, 2013). 이와 유사한 맥락에서 Harris et al.(2013)은 교차로 지점에서 도로 수가 많거나 자전거 및 자동차의 이동량이 많은 경우 자전거 사고의 심각도가 높음을 밝혔다. 또 다른 가로환경 요인으로 횡단보도가 있는데, 대부분의 연구들이 횡단보도의 경우 자전거, 자동차, 보행자가 교차하며 모두 빈번하게 접촉할 수 있는 도로 시설이기 때문에 사고 발생률이 높다는 점을 밝혔다(김응철 외, 2007; 조항훈, 2019; 황선근·이수기, 2021). 자전거 이용의 편의성을 향상하고 이용을 장려하는 자전거도로나 자전거 대여소 등의 자전거 인프라 역시 자전거 교통사고에 영향을 미치는 가로환경 요인으로서 보다 안전한 자전거 이용을 가능하게 하기도 하지만(Kaplan and Prato, 2015; Lusk et al., 2011; Marqués and Hernández-Herrador, 2017), 자전거 이용량을 유발해 자전거 사고의 발생 및 심각도에 영향을 미침이 밝혀지기도 하였다(오주택, 2018; 조항훈, 2019).

가로환경과 더불어 자전거 교통사고에 영향을 미치는 환경적 요인으로 인근시설이 있다. 인근시설 요인 중 버스정류장은 자전거 사고의 발생 가능성을 감소시키는 역할을 하는 것을 확인할 수 있는데, 그 이유를 충돌을 방지하기 위해 자전거 및 자동차, 그리고 버스 운전자가 사전에 감속하기 때문이라 보고 있다(김도훈 외, 2008; 김응철 외, 2007; Wei and Lovegrove, 2013). 또 다른 인근시설 요인 공원은 자전거 이용 및 진입의 증가를 야기해 자전거 교통사고의 발생 및 그 심각도에 영향을 미치는 것으로 나타났다(오주택, 2018; 황선근·이수기, 2018). 나아가 자전거 교통사고에 영향을 미치는 환경적 요인으로 토지이용이 있다. 토지이용 요인에는 주거용·상업용·공업용 토지이용과 토지이용혼합도가 고려되는데, 이들 요인이 자전거 교통사고에 미치는 영향에 대해서는 선행연구들마다 서로 다른 결과를 도출하고 있다(김응철 외, 2007; 오주택, 2018; 황선근·이수기, 2018; 2021; Chen, 2015; Ding et al., 2020; Kamel et al., 2019).

이처럼 선행연구들은 자전거 교통사고의 발생 및 심각도에 다양한 인적 요인과 환경적 요인이 영향을 미친다는 것을 밝히고 있다. 그러나 인적 요인의 경우 성별과 연령, 법규위반 등 각 사고의 특징에 그치므로 일반화에 한계가 존재하며(류종득, 2018; 전희정 외, 2021), 자전거 교통사고를 예방하기 위한 정책적 시사점을 도출하는 데 한계가 있다. 자전거 교통사고가 다발적으로 발생하는 인접지역에 대한 환경의 분석은 자전거 교통사고를 예방할 수 있는 보다 실용적인 시사점 도출을 가능하게 한다.

2. 인공지능을 활용한 자전거 교통사고 연구

최근 교통사고를 예측하는 데 있어 머신러닝(Machine Learning)이나 딥러닝(Deep Learning) 등 인공지능(Artificial Intelligence, AI) 기법을 적용한 연구들이 늘어나고 있다. 전통적 통계기법을 활용한 교통사고 연구들은 상대적으로 넓은 지역을 포괄한 지역 수준의 평균 데이터를 활용하여 교통사고의 요인을 분석하였기 때문에 교통사고 지점과 직접적으로 관련된 요인을 파악하기 어려운 경향이 있다. 인공지능 기법은 대규모의 데이터 혹은 복잡한 자료를 효율적으로 처리할 수 있어 사고 지점에 초점을 맞추어 요인을 도출할 수 있어, 기존 연구들이 가진 연구분석 단위의 한계를 극복할 수 있다는 점에서 주목받고 있다(전희정 외, 2021; Ibrahim et al., 2021).

이에 머신러닝을 활용한 자전거 교통사고 연구는 몇몇 존재하나, 이 경우 대부분 환경적 요인보다 인적 요인을 다루고 있다(최승희·이구연, 2016; Lu et al., 2022). 개별 교통사고 데이터는 주로 운전자의 성별과 연령 등 인적사항과 법규위반 등의 인적 요인이나 사고 당시 날씨, 도로의 표면상태 등의 내용만 제공하고 있어 자전거 교통사고가 다발적으로 일어나는 환경개선을 위한 정책적 접근이 어렵다.

이를 극복하기 위해 직접 현장에 나가 데이터를 수집하는 방법이 있으나, 이는 많은 시간과 비용의 투입을 요구한다(Brownson et al., 2009; Ewing and Dumbaugh, 2009). 이에 최근의 연구동향은 도시환경을 나타내는 GSV 이미지의 활용이 늘어나고 있다. GSV 이미지를 통해 특정 지역의 환경에 대한 데이터를 수집하는 것은 직접 현장에 나가는 것보다 효율적이며, 연구자의 주관이 덜 개입된다는 점에서 유용성이 높다(Chen et al., 2020; Griew et al., 2013; Li et al., 2015; Lu, 2019; Wang et al., 2019). 실제 이러한 GSV 이미지를 사용한 연구들이 있어왔는데(Hanson et al., 2013; Kita-Wojciechowska and Kidziński, 2019a; 2019b), Shirani-Bidabadi et al.(2020)의 연구는 GSV 이미지를 통해 도로의 형태 및 유형(차로 수, 교차로 수, 직선/곡선 여부, 신호등 존재 여부, 자전거 도로 여부 등)이나 주변 시설(주차장, 버스정류장 등)과 같은 환경적 요인을 수집하여 분석을 진행해 자전거 교통사고에 영향을 미치는 요인들을 규명하기도 하였다.

GSV 이미지는 딥러닝 기반 컴퓨터 비전기법을 통해 분석이 되어 왔다. 같은 머신러닝의 한 분야인 딥러닝은 인공신경망(Artificial Neural Network) 이론을 기반으로 복잡한 비선형 문제를 기계가 스스로 학습 해결하는 기법이다(김윤진, 2017). 이에 영상 객체 인식, 음성 인식, 자율 주행, 로봇공학, 자연어처리 등 다양한 분야에서 활용되고 있다. 특히 딥러닝 기법은 최근 컴퓨터비전 분야에서 광범위하게 적용되고 있다. 컴퓨터비전 분야는 영상과 관련된 많은 문제(Task)를 다루는데, 그중 대표적인 것

이 이미지 분류(Image Classification) 문제이다. 현재 컴퓨터비전 분야에서 분류 문제를 해결하기 위해 사용하는 벤치마크(benchmark) 데이터셋인 Imagenet-1K 분류 문제의 경우, Top-5 error rate 기준으로 딥러닝 적용 전 26%에서 적용 후 1.3%까지 성능이 개선된 상황이며(Pham et al., 2021), 이는 사람의 Top-5 error rate인 5%보다도 낮은 수치이다. 따라서 교통사고 관련 연구에 있어 다발지역 분류 문제 및 예측 모델 구축에 있어 딥러닝을 활용하는 이유는 영상 분류(Image Classification) 문제에서 딥러닝 기법의 성능이 높기 때문이라고 할 수 있다. 딥러닝 기법을 활용하여 교통사고를 분석한 연구로 전희정 외(2021)는 수도권의 노인 및 어린이 보행교통사고 다발지역의 GSV 이미지를 수집하고 합성곱 신경망 네트워크(Convolutional Neural Network, 이하 CNN) 구조의 여러 딥러닝 모형을 통해 다발지역을 예측하기도 하였다. 그러나 자전거 교통사고 다발지역 예측을 위해 GSV 이미지에 대하여 딥러닝 기법을 적용을 분석한 연구는 이루어지지 않아왔다.

III. 연구설계

1. 연구자료 및 범위

본 연구는 자전거 교통사고가 다발적으로 일어난 지역의 GSV 이미지 자료를 기반으로 컴퓨터비전과 딥러닝 기법을 활용하여 자전거 교통사고 다발지역 예측모형을 구축하고자 한다. 이를 위한 기초 연구자료는 도로교통공단에서 제공하는 2013년에서 2021년까지 자전거 교통사고 다발지역 정보를 활용하였다. 해당 자료는 전국 3,956곳의 자전거 교통사고 다발지역 정보를 포함하고 있으며, 이때 자전거 교통사고는 차 대 차, 차 대 사람, 차량단독사고로 사고유형을 구분할 수 있다. 구체적으로 자전거 교통사고 지역에 대한 위치정보와 자전거 교통사고 발생건수, 사상자수, 사망자수, 중상자수, 경상자수, 부상신고자 수 정보를 포함하고 있다.

자전거 교통사고 다발지역 예측을 위한 딥러닝 모형의 구축과 검증에 위해서 자전거 교통사고 지역에 대한 위치정보를 활용하였다. 도로교통공단에서 정의하는 자전거 교통사고 다발지역의 기준은 반경 200m에서 자전거사고가 4건 이상 또는 사망사고 3건 이상 지점이다(TAAS, 2022.06.15). 해당 자료에 포함된 구체적 공간적 범위는 전국 3,956개 지역으로 서울, 경기, 인천지역을 포함하는 수도권 지역은 2,285곳이며 비수도권 지역은 1,671곳이다.

자전거 교통사고 다발지역 예측을 위한 딥러닝 모형 구축을 위하여 앞서 논의한 자전거 교통사고 다발지역의 GSV 이미지를 수집하였다. 또한, 자전거 교통사고 다발지역의 대조군으로서 자전거 교통사고가 다발적으로 일어나지 않은 인접지역(이하 자전거 교통사고 비다발지역)의 GSV 이미지 또한 수집하였다. 이때 자전거 교통사고 비다발지역은 각각의 자전거 교통사고 다발지역에

서 제공되는 위치정보를 기준으로 자전거 교통사고 다발지역 200m 반경의 두 배에 해당하는 반경 400m의 원을 생성한 뒤, 원의 둘레에서 유효한 좌표 한 개를 무작위 추출하여 그에 해당하는 GSV 이미지를 수집하였다. 자전거 교통사고 다발지역과 인접한 지역을 자전거 교통사고 비다발지역으로 선정하여 광역적인 지역 특성으로 인한 자전거 교통사고율의 차이가 아닌 보다 세부적인 교통환경 요인에 집중할 수 있을 것으로 판단하였다.

자전거 교통사고 다발지역과 대조군 지역에 대한 위치정보를 바탕으로 각각의 지역이 다발지역으로 선정된 가장 근접한 시점의 GSV 이미지를 GSV API key를 활용하여 수집하였다. GSV 이미지의 경우 비정기적으로 해당 지역에 대한 이미지가 업데이트 된다. 본 연구에서 주요자료의 시간적 범위는 2013년에서 2021년으로 시간적 범위가 분산되어 있는데, 이때 기본설정으로 제공되는 최근 시점의 사진을 수집하는 것은 각각의 지점이 자전거 교통사고 다발지역으로 지정된 시점의 도로환경을 제대로 반영하지 못할 수 있다. 이에, 시점 간 차이를 줄이고 정확한 분석을 위하여 자전거 교통사고가 발생한 가장 근접한 시점의 이미지를 활용하였다. 예를 들어, 자전거 교통사고 다발지역으로 2014년에 선정된 A 지점에 대해서 GSV 이미지가 2010년과 2015년의 이미지가 존재할 경우 시간상 더 가까운 2015년에 해당하는 이미지를 수집하여 분석의 자료로 활용하였다.

이상의 자전거 교통사고 다발지역과 이에 대한 대조군인 자전거 교통사고 비다발지역을 포함하여 최종적으로 이미지를 수집한 지역은 6,574개의 지점이다(다발 및 비다발지역 각각 3,287지점). 기초 자료의 3,956개 지점의 장소와 차이가 발생하는 이유는 수집된 이미지 가운데 GSV 이미지가 존재하지 않거나, 실내에서 촬영된 사진, 어두운 사진 등과 같이 자전거 교통사고 및 도로 상황을 반영하지 않는 지역에 대해서는 분석에서 제외하였기 때문이다. 전체 6,574개 지점 가운데 수도권은 4,352지점(다발 및 비다발지역 각각 2,176)이며, 비수도권은 2,222지점(다발 및 비다발지역 각각 1,111)이다. 우리나라는 수도권과 비수도권 지역의 교통환경과 차량, 인구 규모의 차이가 존재하기 때문에 이러한 광역적인 지역 특성에 따른 차이를 통제하기 위하여 수도권과 비수도권 지역을 구분하여 분석을 실시하였다.

해당 자료를 바탕으로 수도권과 비수도권 자전거 교통사고 다발지역과 비다발지역을 구분하는 분류 딥러닝 모형의 적용을 위해서 수집된 자료를 <표 1>과 같이 구분하여 실험 데이터 세트를 구성하였다. 일차적으로 자전거 교통사고 다발지역과 비다발지역을 구분하고, 다발/비다발지역 각각을 수도권과 비수도권으로 나누었다. 마지막으로, 구분된 지역에 대해서 딥러닝 모형을 위해 학습데이터(train data)와 검증데이터(validation data)로 각각 할당하였다. 검증데이터의 경우 데이터의 정도에 따라 배분되는 정도가 다르지만, 머신러닝과 같은 많은 양의 데이터를 활용할 경우 검증데이터는 전체 데이터의 0.25%으로도 검증이 충분하

표 1. 자전거 교통사고 다발/비다발지역 테스트 데이터와 검증 데이터의 구조

Table 1. Data structure of training and validation for bicycle collision hotspot/non-hotspot

Unit: spot

Division	Capital region		Non capital region		Total
	Train data	Validation data	Train data	Validation data	
Hotspot	2,076	100	1,011	100	3,287
Non hotspot	2,076	100	1,011	100	3,287
Sub total	4,152	200	2,022	200	6,574

다(Andrew Ng, 2017). 본 연구에서는 전체 데이터의 약 1.53%를 딥러닝 모델 성능 검증에 할당하였다.

2. 이미지 분류를 위한 딥러닝 모형의 활용과 CAM 분석을 통한 시각화

자전거 교통사고 다발지역과 비다발지역에 대한 분류를 위해 본 연구에서 CNN 딥러닝 구조를 활용하였다. CNN은 이미지를 딥러닝 모델에 적용하기 위해서 고안된 방식으로 인간이 시각적으로 물체를 인식하는 시신경 구조에 착안하고 있다. LeCun et al.(1989)은 손글씨를 숫자로 인식하는 방법에 대한 논문에서 LeNet 모델을 제시하였으며, 그 이후 CNN구조의 다양한 이미지 정보를 통한 딥러닝 모델들이 등장하였다.

심층신경망(Deep Neural Network, DNN)은 구조의 특성상 1차원의 데이터를 사용한다. 이미지 데이터와 같은 다차원의 데이터를 활용할 때, 작은 이미지의 경우와 달리 큰 이미지의 경우 처리해야 하는 정보의 양과 파라미터가 증가한다. 그 결과 딥러닝 모델에서 더욱 많은 데이터와 학습시간이 소모된다(Géron, 2017). CNN은 DNN이 가지고 있는 단점을 해결하기 위해서 고안된 방식이다. CNN 모형은 크게 특징을 추출하고, 데이터의 크기를 축소시키는 Feature learning 단계와 해당 특징을 바탕으로 데이터를 구분하는 classification 단계로 구분 지을 수 있다. Feature learning 단계에서는 합성곱 형태의 필터를 통해서 데이터의 특징이 손실되지 않으면서 데이터 분석이 용이하도록 하는 단계이다. 즉, Convolution 층에서는 이미지에 행렬(matrix) 필터를 적용하여 이미지의 특징(feature)을 추출하고, Pooling 레이어를 통해서 데이터의 크기를 축소 또는 강조하는 형태를 거치게 된다. 이 과정을 통해서 정제된 데이터는 완전히 연결된(fully connected) 데이터의 특징을 바탕으로 분류(classifier)한다. 다음의 <그림 1>은 CNN 모형의 구조를 도식화한 자료이다.

본 연구에서 자전거 교통사고 다발/비다발 지역을 분류하기 위해 활용한 CNN 모형은 VGG16/19, ResNet50/101, Inception

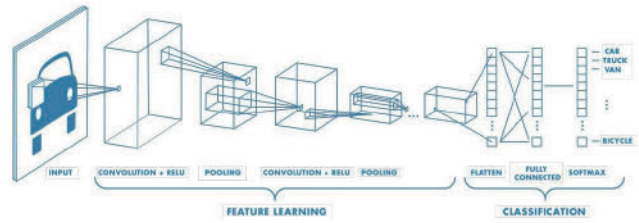


그림 1. CNN 모형의 구조
Figure 1. CCN Structure

출처(Source): MATLAB website

이다. 해당 모델들은 이미지 인식 및 분류 성능을 평가하는 Imagenet Large Scale Visual Recognition Competition(ILSVRC)에서 성능에 대한 입증과 수상을 획득한 모델들이다. VGG 모형의 경우 Oxford Visual Geometry Group에서 개발한 모형으로서 간단한 구조와 단일네트워크에서 뛰어난 성능을 보이는 모델로 평가되고 있다(Kaiming et al., 2016). ResNet 모형은 Microsoft에서 개발한 딥러닝 모형으로 기존의 딥러닝 모형은 모델의 깊이가 깊을수록 학습이 어렵다는 단점을 극복하였다는 특징을 가지고 있다. Inception 모형은 GoogLeNet을 기반으로 하는 모델로서 서로다른 크기의 필터를 활용하여 이미지의 특징을 추출하는 방법을 통해서 성능을 향상시킨 모형이다(Karen and Zisserman, 2014).

본 연구는 이상의 5개 모델을 통해 자전거 교통사고 다발/비다발 지역의 GSV 이미지 분류를 검증하였다. 이때 각 딥러닝 모델의 성능과 정확도는 수치로 표현된다. 그러나 모델별 성능은 이미지에서 어떠한 부분을 통해서 분류와 예측을 하는지 파악하기 어렵다는 한계를 가진다. 따라서, 한계점을 극복하기 위해서 Class Activation Map(이하 CAM)의 알고리즘을 딥러닝 모형에 적용하였다. CAM은 딥러닝 모델이 이미지에서 어떠한 부분의 특징이 자전거 교통사고 다발/비다발 지역을 구분하는 과정에서 가장 주요하게 작용하는 특징 부분을 색상으로 표현하는 방법이다. CNN 모형의 구조에서 Pooling layer 단계서 강조되는 이미지의 주요 특징이 CAM 분석결과에서 표현되며, 본 연구에서는 이미지에서 노란색이 진하게 표현된 부분이 이미지를 분류하는 데 주요하게 작용하는 부분이다. 부분이다. 다음의 <그림 2>는 CAM

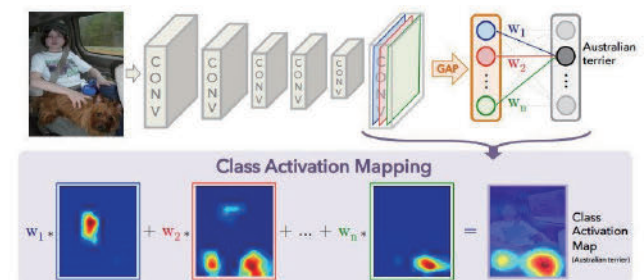


그림 2. Class Activation Map (CAM) 구조
Figure 2. Class Activation Map (CAM) structure

출처(Source): Learning deep features for discriminative localization

방식에 대한 도식화 자료이다.

IV. 분석결과

1. 자전거 교통사고 다발/비다발 지역 예측을 위한 딥러닝 모델 성능 검증

〈표 2〉는 본 연구에서 실시한 딥러닝 모형들의 자전거 교통사고 다발/비다발 지역 분류 모델의 성능에 대한 결과표이다. 자전거 교통사고 다발지역과 비다발지역 이미지를 수도권과 비수도권 지역으로 구분하고 각각 구분에 따라 딥러닝 모형을 적용하여 이미지 분류 모형 성능을 검증한 결과, 전체 모델 평균 약 70.3% 수준으로 자전거 교통사고 다발지역과 비다발지역을 예측하는 것으로 나타났다. 5개의 모델 가운데 VGG19 모델의 분류 성능이 가장 뛰어난 것으로 나타났다. VGG19 모델은 전체 지역에서 자전거 교통사고 다발지역과 비다발지역을 각각 분류하는 데 평균 약 73.75%의 성능을 나타냈다. 보다 구체적으로 VGG19 모델의 성능을 살펴보면 수도권 자전거 교통사고 다발지역에 대한 예측률은 약 60.0%이며, 비다발 지역은 약 87.0%로 나타난다. 이는 수도권에서 약 10장의 자전거 교통사고 다발지역 이미지 가운데 6장을 다발지역으로 맞게 예

측하며, 비다발지역 이미지 10장 가운데 8장 이상을 비다발지역으로 맞게 분류하는 것을 의미한다. 비수도권의 경우 다발지역을 예측하는 성능은 약 69%이며, 비다발지역을 예측하는 성능은 79%로 나타난다. 개별 모델의 성능을 볼 때, Inception 모델에서 비수도권 다발지역을 예측하는 성능이 79.0%로 가장 높게 나타났다.

2. 자전거 교통사고 다발/비다발 지역에 대한 CAM분석 결과

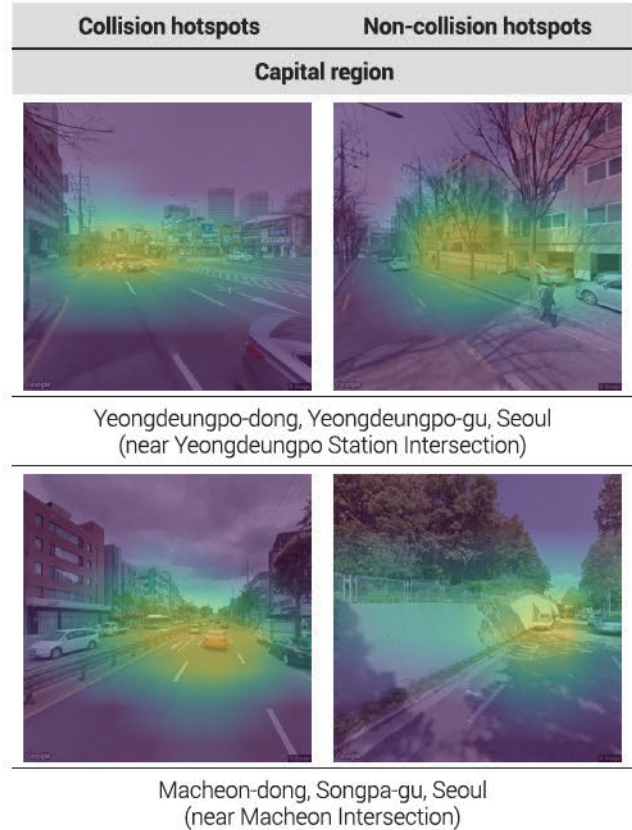
본 연구는 평균 성능이 가장 높게 나타난 VGG19 모델이 이미지의 어떤 특징을 중심으로 자전거 교통사고 다발지역과 비다발지역 여부를 판단하는지 확인하기 위해 CAM 분석을 추가로 실시하였다. 〈그림 4〉는 수도권 지역의 자전거 교통사고 다발지역과 비다발지역 CAM 분석결과이며, 〈그림 5〉는 비수도권 지역의 자전거 교통사고 다발지역과 비다발지역 CAM 분석결과이다. 이때 이미지에서 형광색으로 진하게 표현된 부분은 딥러닝 모델이 이미지를 분류하는데 주요하게 활용한 특징에 대해서 주의(attention)가 많이 가해진 부분이다.

본 연구는 앞서 모든 모델에서 수도권 비다발지역과 비수도권 다발지역의 성능의 차이가 존재한다는 점을 바탕으로 수도권 다발/비다발, 비수도권 다발/비다발 지역에 대한 CAM 결과를 대조하여 특징을 파악하였다. CAM 분석결과를 살펴보

표 2. 자전거 교통사고 다발/비다발 지역 분류 모델 성능

Table 2. Performance of CNN model for bicycle collision hotspot/non-hotspot classification

Model	Location	Hotspot (%)	Non-hotspot (%)	Average (%)
ResNet 50	Capital region	58.00	81.00	69.50
	Non-capital region	74.00	66.00	70.00
	Pooled	66.00	73.50	69.75
ResNet 101	Capital region	54.00	83.00	68.50
	Non-capital region	64.00	77.00	70.50
	Pooled	59.00	80.00	69.50
VGG16	Capital region	71.00	71.00	71.00
	Non-capital region	77.00	60.00	68.50
	Pooled	74.00	65.50	69.75
VGG19	Capital region	60.00	87.00	73.50
	Non-capital region	69.00	79.00	74.00
	Pooled	64.50	83.00	73.75
Inception	Capital region	61.00	71.00	66.00
	Non-capital region	79.00	64.00	71.50
	Pooled	70.00	67.50	68.75



다음 페이지에 계속(Continue on next page)

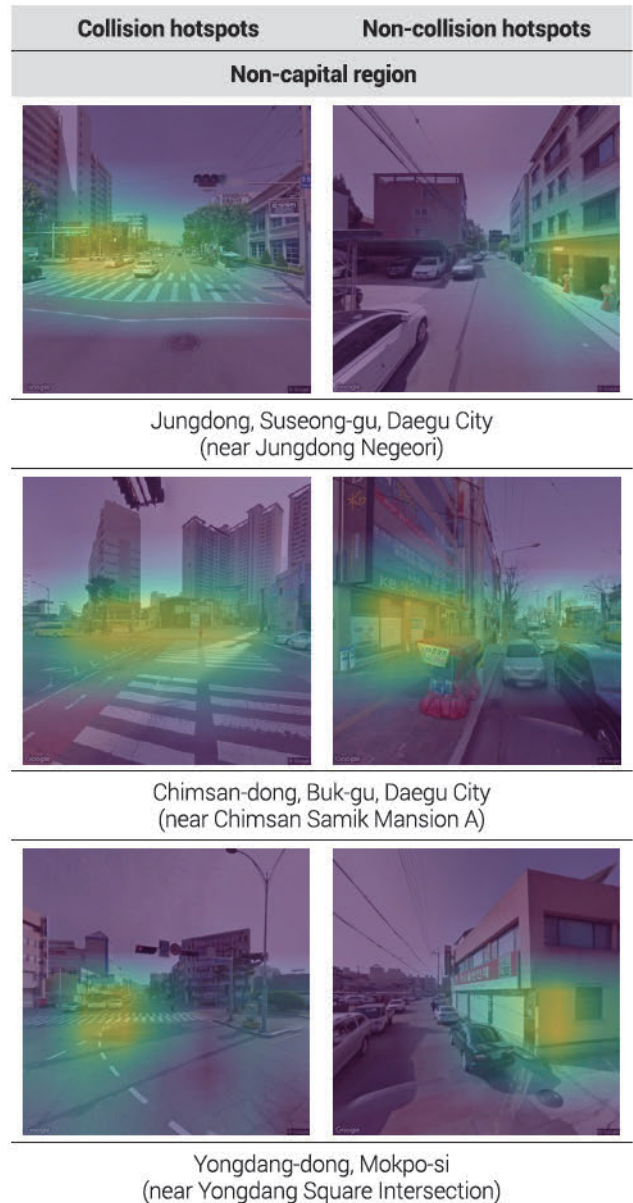


그림 4. 수도권 CAM 분석결과
Figure 4. CAM results in the captial region

면 자전거 교통사고 다발지역과 비다발지역에 대한 다음과 같은 특징과 차이점을 추론할 수 있다. 먼저, 수도권 다발지역에 대한 CAM 분석결과 차량, 도로 구조물, 가로수에 대한 주의가 집중되었으며, 수도권 비다발지역에 비해 상대적으로 넓은

폭의 도로 형태를 확인할 수 있었다. 반면, 수도권 비다발지역의 경우 건물의 벽면, 펜스, 화단, 나무, 하늘, 지면에 주의가 집중되었으며, 다발지역과 달리 도로의 폭이 좁다는 특징이 나타났다. 비수도권 다발지역과 비다발지역의 특징을 비교 검토한 결과, 비수도권 다발지역에서는 수도권 다발지역과 유사하게 차량, 도로 구조물에 주의가 집중되었다. 반면, 비수도권 비다발지역은 건물 벽면, 펜스에 주의가 집중되고 도로의 폭이 좁은 특징을 파악할 수 있다.

이를 바탕으로 다발지역과 비다발지역의 특징을 다음과 같이 정리할 수 있다. 첫째, 도로의 규모에서 차이를 확인할 수 있다. <그림 5>의 1행의 비수도권 다발지역과 비다발지역(대구시 수성구 중동네거리 부근)의 두 사진을 비교해보면 다발지역의 경우 차량의 폭이 비다발지역에 비해 상대적으로 넓다. 이는 CAM 분석결과 전반에서 가장 두드러지는 특징 가운데 하



다음 페이지에 계속(Continue on next page)

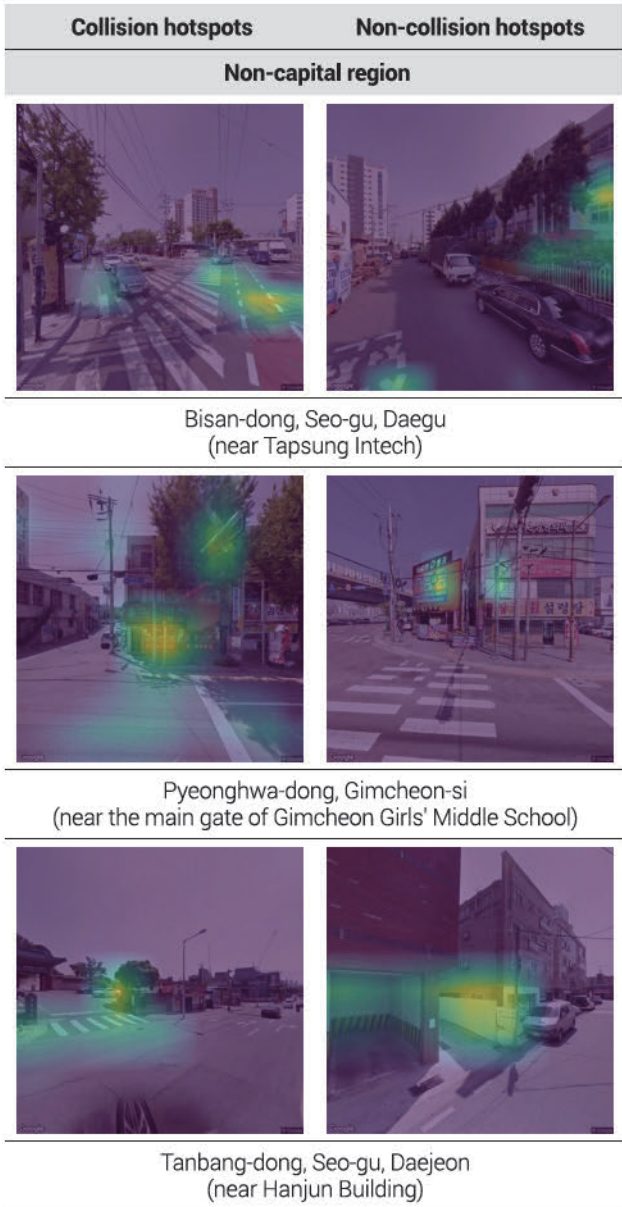


그림 5. 비수도권 CAM 분석결과
Figure 5. CAM results in the non-capital region

나이다. 다발지역으로 바르게 분류된 지역들의 이미지는 도로의 폭이 넓은 광로(光路)들이 다수를 이루는 것과는 달리, 비다발지역의 경우 협소한 골목길의 형태나 좁은 이면도로가 다수를 이루고 있다. 이는 전희정 외(2021)의 연구에서 개방감이 높은 지역이 보행교통사고 다발지역으로 분류되는 경향이 있다는 분석결과와 결을 같이 한다. 또한, 앞서 Harris et al.(2013)의 연구에서 교차로와 같은 도로의 수가 많은 지역일 경우 차량 또는 자전거의 통행량이 많아 자전거 교통사고의 심각도가 높게 나타난다는 결과와 맥락을 같이하고 있다.

둘째, 자전거 교통사고 다발지역의 경우 도로상의 차량과 횡단보도, 인도 경계석, 블라드(기둥) 등의 물체는 자전거 교통사고 다발지역을 예측하는 주요 특징으로 추정된다. 예시사진을 통해 보면 다발지역의 경우, 횡단보도 근처의 차량 또는 도로

주변의 신호등과 일부 인도의 경계석 부근에 주의가 집중되고 있다. 특히, <그림 4>의 1과 3행의 다발지역인 영등포구 영등포동과 안양시 호계동, <그림 5>의 대구 북구 침산동, 목포 용당동에 대한 CAM 분석결과를 살펴보면, 횡단보도를 따라 주의가 집중되는 결과를 확인할 수 있다. 또한, 김천시 평화동의 사진에서 도로 주변의 블라드(기둥) 및 인도와 도로의 경계석 부근에 주의가 집중되는 것을 볼 수 있다. 도로 주변의 횡단보도는 자전거와 보행자 또는 차량과의 접촉 확률이 빈번하게 발생할 수 있는 도로의 시설이다. 선행연구들(김용철 외, 2007; 조항훈, 2019; 황선근·이수기, 2021) 역시 횡단보도를 자전거 교통사고를 발생시키는 주요 원인으로 파악하고 있다. 또한, 도로 주변의 신호등과 기둥과 같은 객체들은 자전거 주행 시 정상적인 주행을 방해하는 장애물로 작용할 수 있다는 점에서 자전거 교통사고 다발지역으로 판단하는 데 주요 요인으로 볼 수 있다.

한편, 자전거 전용도로가 자전거 교통사고가 발생한 지역의 주요 특징으로서 CAM 분석결과에 나타나고 있다. 대구시 비산동과 대구시 북구 침산동의 사진을 볼 때, 도로 옆으로 존재하는 자전거 전용도로의 일부는 해당 지점이 자전거 교통사고 다발지역으로 구분되는 주요 특징으로 나타난다. 이는 일반적으로 자전거도로가 자전거 교통사고를 감소시킨다는 통념 및 선행연구의 결과(Kaplan and Prato, 2015; Lusk et al., 2011; Marqués and Hernández-Herrador, 2017)에 반한다. 그러나 자전거 전용도로와 같은 자전거 인프라의 경우, 자전거 통행 유발을 증가시키는 요인으로서 보행자와 자전거 주행자가 접촉할 가능성이 높일 수 있기 때문이라 그 이유를 추론할 수 있다(조항훈, 2019). 또한 해당 자전거 도로의 경우 자전거 주행 전용도로이지만, 실제 자전거가 주행할 때, 자전거 이용자의 안전을 보호받을 수 있는 시설이 전무하다는 점에서 자전거 교통사고에 부정적인 영향을 미치는 것으로 볼 수 있다.

도로 주변으로 위치한 건물과 도로 주변의 녹지 공간, 버스정류장과 같은 근린 환경은 자전거 교통사고 비다발지역으로 판단하는 주요 특징으로 작용하고 있다. 자전거 교통사고 비다발지역으로 분류되는 지역들의 전반적인 특징은 해당 지역들이 주택가 주변 혹은 아파트와 같은 주거용 토지이용 지역으로 분포하고 있다. 황선근과 이수기(2018; 2021)의 연구는 주거용 토지이용의 경우 자전거 교통사고에 영향을 미치지 않는다는 결과를 제시하고 있는데, 이는 토지이용이 상대적으로 혼잡도가 낮다는 점이 보다 안전한 자전거 이용에 기여할 수 있음을 의미한다. 또한, 도로 주변의 가로수나 녹지, 버스정류장 등의 경우 자전거 교통사고 비다발지역으로 구분하는 주요 요소로 나타나고 있다. 앞서 다발지역의 경우 개방감이 높다는 특징으로 인해서 교통사고 다발지역으로 분류되는 것과 달리, 비다발지역의 녹지와 건물의 벽면 등은 보행자나 도로 상황에서 발생

할 수 있는 상황에 더욱 주의를 집중시키기에 완충작용을 하는 결과로 볼 수 있다. 또한 버스정류장의 경우, 자전거 운전자와 버스와의 충돌 이전에 안전확보를 위해서 감속을 하는 행태가 존재하기 때문에 자전거 사고 발생의 가능성을 낮추는 것으로 볼 수 있다(김도훈 외, 2008; 김응철 외, 2007; Wei and Lovegrove, 2013)

딥러닝 모형을 통해 다발/비다발지역을 결정하는 주요 특징에 대해서 검토한 결과, 도로의 규모가 상대적으로 크고, 차량을 포함하여 횡단보도, 신호등, 기둥 등과 같은 구조물, 자전거 전용도로 등이 존재하는 환경은 자전거 교통사고 다발지역으로 구분되고 있음을 확인할 수 있다. 반면, 도로의 폭이 상대적으로 좁고, 토지이용 혼합도가 낮은 주거지 주변과 녹지, 벽면, 버스정류장은 자전거 교통사고 비다발지역으로 지점을 구분하는 주요 요인으로 작용하고 있는 것으로 결과를 정리할 수 있다. 가로 공간 위에서의 밀도와 교통안전의 관계를 연구한 Ewing과 Dumbaugh(2009)는 도로 공간의 높은 밀도는 교통사고의 확률을 낮춘다고 주장하고 있다(Ewing and Dumbaugh, 2009; 전희정 외, 2021). 본 연구에서도 자전거 교통사고 다발지역과 비다발지역은 개방감을 포함하여 도로 주변의 객체들의 차이에 따라 도로의 구분이 나누어지는 것을 확인하였다. 이처럼 각 지역 분류에 작용하는 주요 특징에 대한 면밀한 검토를 바탕으로, 추후 도시계획의 측면에서 자전거 안전을 비롯하여 도로안전의 문제를 해결하기 위한 방향과 의사결정에 딥러닝 모형이 주요하게 활용될 수 있는 가능성을 확인할 수 있었다.

V. 논의

1. 자전거 교통사고 심각도 예측을 위한 딥러닝의 적용

자전거 교통사고와 더불어 교통사고를 다룬 그간의 연구들은 단순히 사고가 발생했는지를 넘어서 심각도를 고려하는 양상을 보였다. TAAS는 교통사고의 심각도를 부상신고, 경상, 중상, 사망, 이상의 네 단계로 구분하여 제시하고 있으며, 선행연구들은 이를 경상이하(부상신고와 경상)와 중상이상(중상과 사망)으로 이진화하여 심각도에 대한 연구를 진행해왔다. 자전거 교통사고의 경우 심각도에 대한 연구가 거의 진행되지는 않았지만 대부분 인적 요인에 초점을 맞추었으며(신혜정 외, 2012; 최승희·이구연, 2016), 황선근과 이수기의 연구(2018)가 서울시의 집계구를 분석단위로 삼아 자전거 교통사고 심각도에 영향을 미치는 환경적 요인을 다루었지만, 여전히 분석단위의 한계에서 자유롭지 못하였다.

본 연구는 앞서 자전거 교통사고 다발/비다발지역을 예측하는 딥러닝 모형을 구축하는 데 GSV 이미지를 활용함으로써 이러한 한계의 극복 가능성을 확인하였다. 이에 선행연구들에서 다루어

진 사고 심각도 역시 딥러닝 모형을 통해 예측가능한지 확인하고자, 사고 심각도를 기준으로 동일한 방법으로 모형을 구축하고 정확도를 확인하였다. 그 결과 자전거 교통사고의 심각도를 예측하는 모델의 정확도가 약 50% 수준으로 매우 낮은 것을 확인할 수 있었다. 본 연구가 구축한 자전거 교통사고 다발/비다발지역을 분류하는 모델 약 70%의 정확도로, 현장에서 해당 모델을 자전거 교통사고 다발지역 예측 모델로 바로 사용하기는 어려운 수준이었는데 심각도 예측 모델은 더 낮게 나타난 것이다.

본 연구에서 활용한 딥러닝 모형들은 지도학습(Supervised learning)의 형태이다. 지도학습은 학습데이터가 가지고 있는 정답(label)에 따라서 학습 모델이 결정되는데 본 연구에서 자전거 교통사고의 심각도 예측을 위한 분석의 한계가 발생하는 이유는 실질적으로 GSV 이미지를 통해서 교통사고 심각도에 따라 명시적으로 구분되는 특징(feature)을 찾아내기 어렵다는 추측을 할 수 있다. 우선적으로 실시한 다발/비다발지역을 구분하는 실험에서 수도권과 비수도권의 차이를 구분하는 난이도가 낮은 과제의 수행결과에서도 특이점을 파악하는 것은 난이도가 높은 것임을 고려할 때 GSV 이미지를 통해서 세부적인 지역의 특성과 심각도의 수준을 파악하는 것은 인간의 인지능력에 준하는 분류 능력과 수준을 요구한다는 점에서 분석의 결과에 한계가 존재한다고 볼 수 있다.

2. 자전거 교통사고 다발지역과 보행 교통사고 다발지역의 비교

지속가능한 교통체계와 패러다임 전환으로 자전거와 함께 보행이 주목받고 있다. 이에 Rahman의 일련의 연구들은 다양한 머신러닝 기법을 활용하여 자전거와 더불어 도로안전에 취약하다고 여겨지는 보행자를 함께 다루어, 보행교통사고와 자전거 교통사고에 영향을 미치는 환경적 요인을 연구하기도 하였다(Rahman, 2018; Rahman et al., 2019). 이에 자전거 교통사고 다발지역을 분석한 본 연구 역시 유사한 방법을 활용하여 수도권 지역의 어린이와 노인 보행 교통사고를 다룬 전희정 외(2021)의 연구결과와 비교하였다.

전희정 외(2021)의 연구는 어린이와 노인 보행 교통사고 다발지역은 도로의 규모가 크고, 차량, 교통 관련 시설물에 주의(attention)가 집중되는 특징을 확인하였다(전희정 외, 2021). 다음의 <그림 6>은 보행교통사고 다발지역에 대한 CAM 분석결과이다. 앞서 자전거 교통사고 다발지역은 차량, 도로 구조물, 가로수와 같은 객체에 주의가 집중되는 특징과 함께 도로의 규모가 상대적으로 큰 것을 확인하였다. 이를 바탕으로 자전거 교통사고 다발지역과 보행교통사고 다발지역을 비교하면 도로의 규모, 차량, 도로 주변의 구조물, 가로수와 같은 측면의 특징에 주의가 집중된다는 공통점을 발견할 수 있다. 또한, 전희정 외(2021)의 보

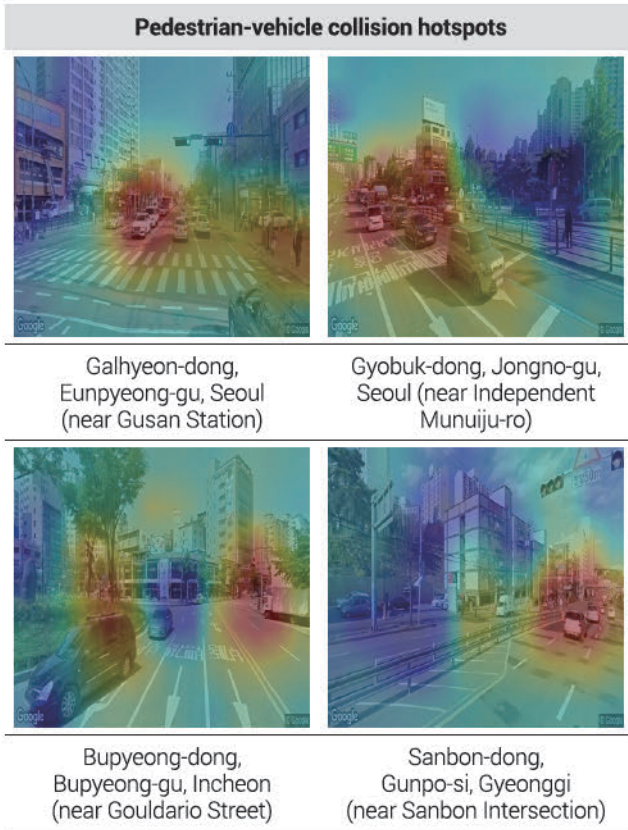


그림 6. 보행자 교통사고 다발지역 CAM 분석결과 예시

Figure 6. CAM results for pedestrian-vehicle collision hotspots

출처: 전희정 외, 2021

Source: Jun et al., 2021

행 교통사고 다발지역 예측 연구에서 CNN 모형의 분석결과와 비교를 해보면 VGG16모형에서 약 73%로 비슷한 수준의 모델 정확도가 도출되었다.

도로교통법상 자전거는 차(車)에 속한다. 그럼에도 불구하고 보행교통사고 다발지역과 자전거 교통사고 다발지역 각각이 딥러닝 모형에서 판단되는 특징이 유사하다는 점은 자전거 주행 안전과 보행 안전을 위한 도로환경 또한 유사할 수 있음을 의미한다. 이는 차량은 보행자와 자전거 운전자 모두에게 잠재적인 위협의 요소로 작용할 뿐 아니라 우리나라에서 자전거 이용자가 보행자가 다니는 인도를 많이 이용하기 때문일 수도 있다. 하지만 실정법상 자전거가 인도로 보행하는 것은 불법이라는 전제하에 교통시설물의 관점에서 두 주체의 안전성을 어떻게 확보하는가에 대한 문제는 보다 심도 깊은 연구를 통해서 파악할 필요가 있다.

VI. 결론

본 연구의 목적은 실제 도로환경의 이미지를 딥러닝 모형에 적용하여 자전거 교통사고 다발지역과 비다발지역을 구분하는 모델을 도출하고, 다발/비다발지역으로 판단되는 지역의 특징의 파악을 통해서 도로 환경의 차이를 파악하는 것이다. 분석을 위해

서 도로교통공단의 자전거 교통사고 다발지역을 바탕으로 총 6,574곳의 GSV 이미지를 활용하여 분석을 하였다.

분석에 사용한 CNN 딥러닝 모형은 VGG16, VGG19, ResNet50, ResNet100, Inception이며, 해당 모델들에 대해서 지도 학습(supervised learning)을 실시하였으며, VGG19 모델의 이미지 분류 성능이 약 73.75%로 가장 높게 나타났다. 또한, 자전거 교통사고 다발지역 분류의 결과를 시각화하기 위해서 CAM 분석을 실시하였다. 분석의 결과 자전거 교통사고 다발지역과 비다발지역에 대한 도로환경의 차이점을 일부 확인할 수 있었다. 자전거 교통사고 다발지역의 경우 보행자를 포함하여 차량, 도로 주변의 구조물, 가로수의 특징에 주의를 집중되는 것을 확인할 수 있으며, 도로의 규모가 상대적으로 큰 것으로 나타났다. 반면, 비다발지역은 상대적으로 도로의 규모가 협소하고 도로 주변의 건물 벽면, 펜스, 화단, 나무 등에 주의를 집중되고 있었다. 즉, 지역적으로 다발지역과 비다발지역을 구분하는 특성에는 차이를 발견할 수 있었다.

이상의 결과와 논의들을 통해 본 연구는 다음과 같은 이론적 함의를 갖는다. 첫째, 본 연구는 자전거 교통사고가 다발적으로 일어난 지역의 GSV 이미지를 활용하여 자전거 교통사고 다발지역을 분석하였다는 점에서 그 의미가 있다. 그간 교통사고와 관련된 연구들은 행정동과 같이 보다 광범위한 지역의 환경적 특성을 고려하여 교통사고가 발생한 지점의 환경을 정확히 반영하지 못했다는 한계를 갖는다. 본 연구는 객관적인 환경을 반영하고 교통사고가 발생한 지역의 직접 방문할 필요가 없도록 하는 GSV 이미지를 활용하여 자전거 교통사고의 예방과 관리에 있어 보다 효과적으로 기여할 수 있을 것이라 기대한다.

둘째, 딥러닝과 컴퓨터비전 등의 최신 기법들을 활용하여 자전거 교통사고 다발지역을 예측하는 모델을 구축하였다는 이론적 함의를 갖는다. 자전거 교통사고에 대한 연구들은 지역의 발생 건수 및 심각도에 영향을 미치는 요인들을 탐구하는 데 그쳤다. 본 연구는 실제 도로환경을 기반으로 교통사고의 다발지역과 비다발지역을 구분하고, 보다 객관적으로 사고 발생에 영향을 미치는 요인들을 분석했다는 점에서 의미 있다. 특히 최근 교통사고 연구에서 딥러닝과 컴퓨터비전 등을 적용하고자 하는 시도가 나타나고 있지만, 자전거 교통사고와 관련한 연구는 거의 진행된 바 없다. 본 연구는 자전거 교통사고의 예방 및 관리에 있어서도 이들 기법을 적용할 수 있음을 확인했다는 점에서 의미가 크다고 할 수 있다.

이에 더하여 본 연구는 다음과 같은 정책적 시사점을 제시한다. 첫째, 자전거 교통과 안전을 위한 의사결정에 빅데이터와 딥러닝과 같은 도구 활용을 보다 적극적으로 고려하여야 한다는 것이다. 인간이 실증적으로 모든 도로를 점검하고 예측하는 것은 시간과 비용이 많이 소모된다. 또한, 합리적인 자전거 또는 도로 교통을 위한 의사결정에 있어서 제약이 따른다. 이때, 기술적으

로 딥러닝과 같은 도구의 활용을 통해서 도시설계나 유지를 위한 의사결정의 도구로 활용하여 보다 효율적인 스마트 도시관리가 가능해질 것으로 보인다.

둘째, 자전거 통행의 안전을 위한 환경 조성이 필요하다. 본 연구의 분석결과에서 알 수 있듯이 가로 상황에서 보행자, 도로 구조물, 차량은 자전거 교통사고 다발지역으로 판단되는 특징으로 나타났다. 따라서, 자전거 주행이 있어서 주행에 방해가 되지 않는 자전거 주행 인프라의 구축이 필요하다. 실제 자전거 전용도로가 존재하지만, 우리나라에 있는 대부분의 자전거도로는 차도의 가장자리에 위치하고 있으며, 이러한 특성상 보행자, 도로위 장애물, 보행섬 등과 같은 장애물에 직접적으로 영향을 받을 수밖에 없다. 또한, 자전거도로 지면의 색상을 달리 표시하는 정도에 그치며, 그나마 존재하는 자전거도로 역시 단절되어 연속성이 매우 떨어진다. 따라서, 자전거 주행이 안정적으로 이루어질 수 있고 다른 요소에 의한 방해가 발생하지 않도록 구조적 개선의 필요성을 제시한다. 이를 위해서 자전거 전용도로가 차량과 보행자가 분리되어 운행하는 형태로 변화시키는 방안 등을 고려해 볼 수 있다.

한편, 본 연구는 자전거 교통사고 다발지역 예측을 위해 딥러닝 모델을 사용한 연구로서 의의가 있지만, 현재 모델의 정확도는 70% 정도로 현장에서 해당 모델을 자전거 교통사고 다발지역 예측 모델로 바로 사용하기는 힘들다는 한계점이 있다. 이 이유는 다양할 수 있으나, 본 연구진이 제시하는 가능성은 자전거 교통사고 다발지역 예측 문제가 본 연구에서 사용한 딥러닝 분류 모델로는 완벽히 해결하기는 힘든 문제일 수 있다는 점이다. 이는 단순히 크로스-엔트로피(Cross-entropy) 손실 함수로 학습한 딥러닝 모델은 입력 영상에서 특정한 특징 유무로만 클래스(Class)를 판별하는 경향이 있기 때문이다. 예를 들어 CAM 결과 분석에서 언급한 것과 같이 횡단보도, 인도 경계석, 기둥 같은 물체가 있는 경우, 딥러닝 모델은 해당 지역이 자전거 교통사고 다발지역이라고 예측하는 경향을 보였지만, 이와 같은 특징은 비다발지역에서도 나타날 수 있다. 이러한 경우에는 단순히 영상에서 해당 물체나 특징의 유무가 아닌, 해당 물체의 수나 영상 내에서 차지하는 비율과 같이 복합적인 형태의 요인을 고려한 분석이 적용되어야 해당 지역을 비다발지역이라고 분류할 수 있다. 그러나 본 연구에서 사용한 딥러닝 모델은 이와 같은 점은 고려하기 어려운 모델이기에 70%대의 성능이 나왔다고 생각해볼 수 있다. 이외에도 본 연구에서 활용한 연구자료에 있어 GSV 이미지 자료는 특정시점에 촬영이 되었으며, 촬영각도에 따라 물리적 환경이 다르게 보일 수 있어 자전거 교통사고가 다발적으로 발생한 지역의 정확한 환경을 반영하지 않을 수 있다. 이외에도 GSV 이미지 자료만으로 자전거 교통사고량에 영향을 줄 수 있는 통행량을 정확히 반영하지 못한다는 한계점이 있다.

이러한 연구의 한계점을 극복하고 자전거 교통사고 다발지역 환경에서 안전의 영향관계를 보다 명확히 파악하기 위해서 객체

탐지(Object detection)나 의미론적 분할 기법(Scene segmentation)을 통해서 인과관계를 파악하는 것을 후속연구의 방향으로 제시하고자 한다. 이와 더불어 본 연구는 특정 시점에 찍힌 도로환경 이미지 자료를 활용하여 자전거 교통사고 다발지역을 분석하였지만, CCTV 등으로 일정 기간 동안 촬영된 동영상을 기반으로 차량, 보행자, 자전거 등의 통행량 및 속도, 조도 등 시간에 따라 변화하는 변수들을 딥러닝을 활용하여 도출해내고, 분석에 활용한다면 더 정확한 예측 모델을 구현할 수 있을 것이다. 나아가 이런 분석법을 CCTV 영상에 실시간으로 적용한다면 현재 시점의 자전거 교통사고 위험도를 즉각적으로 예측하고 경고할 수 있어 사고를 예방하는 데에도 기여할 수 있을 것이라 예상된다.

인용문헌

References

1. 국무조정실 국무총리비서실, 2022. 「지난 5년간(’16~’21) 교통사고 사망자 32.4% 감소」, 서울.
Office for Government Policy Coordination Prime Minister's Secretariat, 2022. *Reduction of '16~'21 Traffic Accidents Deaths 32.4% for the Last 5 Years*, Seoul.
2. 김도훈·조한선·김응철, 2008. "도시부 교차로에서의 자전거 사고유형 분석에 관한 연구", 「한국도로학회논문집」, 10(4): 117-125.
Kim, D.H., Jo, H.S., and Kim, E.C., 2008. "A Study on Bicycle Accident Patterns at Urban Intersections", *International Journal of Highway Engineering*, 10(4): 117-125.
3. 김세용·배기목·이수일, 2008. "자전거 이용 활성화를 위한 법제도 개선 방향", 「국토지리학회지」, 42(1): 35-48.
Kim, S.Y., Bae, K.M., and Lee, S.I., 2008. "A Study on the Institutional Improvement Suggestion for Bicycle Use", *The Geographical Journal of Korea*, 42(1): 35-48.
4. 김숙희·심태일·정가형, 2017. "수원시 자전거 교통사고 특성분석: TAAS 및 시민보험자료 비교 중심으로", 「교통기술과정책」, 14(5): 6-13.
Kim, S.H., Shim, T.I., and Jung, G.H., 2017. "Analysis of Bicycle Traffic Accident Characteristics in Suwon City", *Transportation Technology and Policy*, 14(5): 6-13.
5. 김윤진, 2017. "딥러닝(Deep Learning)을 활용한 이미지 빅데이터(Big Data) 분석 연구", 중앙대학교 대학원 박사학위논문.
Kim, Y.J., 2017. "Analysis of Image Big Data Using Deep Learning", Ph.D. Dissertation, Chung-ang University.
6. 김응철·오주택·김명섭, 2007. "비선형 회귀모형을 이용한 자전거 거사고 요인분석에 관한 연구", 「국토계획」, 42(3): 261-272.
Kim, E.C., Oh, J.T., and Kim, M.S., 2007. "A Study on Crash Causations for Bicycles in Signalized Intersections Using Non-linear Regression Models", *Journal of Korea Planning Association*, 42(3): 261-272.

7. 김태양·박병호, 2017. “지역 특성을 고려한 국내 자전거 사고 분석”, 『대한교통학회지』, 35(5): 447-457.
Kim, T.Y. and Park, B.H., 2017. “Analysis of Bicycle Accidents in Korea based on Regional Characteristics”, *Journal of Korean Society of Transportation*, 35(5): 447-457.
8. 류종득, 2018. “딥 러닝을 이용한 고속도로 교통사고 예측모델 개발”, 아주대학교 대학원 박사학위논문.
Ryu, J.D., 2018. “Development of Expressway Traffic Accident Prediction Model Using Deep Learning”, Ph.D. Dissertation, AJou University.
9. 신혜정·홍지연·이수범, 2012. “자전거 사고 심각도에 영향을 미치는 요인 연구”, 『교통연구』, 19(1): 67-78.
Shin, H.J., Hong, J.Y., and Lee, S.B., 2012. “A Study on the Factors Affecting the Severity of Bicycle-Related Accidents”, *Journal of Transport Research*, 19(1): 67-78.
10. 신희철·박재영, 2012. “연령과 도시규모에 따른 자전거 교통사고 분석”, 『월간교통』, 68-72.
Shin, H.C. and Park, J.Y., 2012. “Analysis of Bicycle Traffic Accidents according to Age and City Size”, *Monthly Traffic*, 68-72.
11. 오주택·김응철·지민경, 2007. “도시부 신호교차로에서의 자전거 사고 분석”, 『한국도로학회논문집』, 9(2): 1-11.
Oh, J.T., Kim, E.C., and Ji, M.K., 2007. “A Study of Bicycle Crash Analysis at Urban Signalized Intersections”, *International Journal of Highway Engineering*, 9(2): 1-11.
12. 오주택, 2018. “PLS 구조방정식을 이용한 자전거사고 요인분석”, 『한국 ITS 학회 논문지』, 17(4): 26-40.
Oh, J.T., 2018. “A Factor Analysis for Bicycle Accidents Using the PLS Structural Equation”, *The Journal of The Korea Institute of Intelligent Transportation Systems*, 17(4): 26-40.
13. 전희정·강승엽·정수영·김태완·조철호·주원영·김지영·허재필, 2021. “보행교통사고 다발지역 예측을 위한 딥러닝의 적용: 수도권 노인과 어린이 보행교통사고를 중심으로”, 『국토계획』, 56(7): 115-127.
Jun, H.J., Kang, S.Y., Jung, S.Y., Kim, T.Y., Cho, C.H., Jhoo, W.Y., Kim, J.Y., and Heo, J.P., 2021. “A Deep Learning Approach to Predict Elderly and Child Pedestrian Crash Hot Spots in the Seoul Metropolitan Area”, *Journal of Korea Planning Association*, 56(7): 115-127.
14. 정경옥·박성용·신희철, 2019. 「국내외 자전거 법제도 및 정책 추진 체계 조사 분석」, 한국교통연구원 이슈페이퍼.
Jung, K.O., Park, S.Y., and Shin, H.C., 2019. *Research and Analysis of Domestic and Foreign Bicycle Laws and Policy Systems*, The Korea Transport Institute Issue Paper.
15. 조항훈, 2019. “자전거 대 사람 충돌사고 발생지점의 공간적 특성 분석: 서울 한강공원 자전거도로를 대상으로”, 한양대학교 석사학위 논문.
Jo, H.H., 2019. “A Study on the Spatial Characteristics of Bicycle to Person Collision Accidents: Focused on Bicycle Road of Han River Park in Seoul”, Master’s Dissertation, Hanyang University.
16. 조현우·김동진, 2020. “국내 자전거 사고 위험도 분석-자전거 운전자의 과실로 발생한 사고를 중심으로-”, 『Crisisonomy』, 16(11): 61-75.
Cho, H.U. and Kim, D.J., 2020. “A Risk Analysis of Bicycle Accidents in Korea –Focused on Driver’s Human Error–”, *Crisisonomy*, 16(11): 61-75.
17. 최승희·이구연, 2016. “자전거 교통 사고 현황 및 예측 분석”, 『전자공학회논문지』, 53(9): 89-96.
Choi, S.H. and Lee, G.Y., 2016. “Analysis and Prediction of Bicycle Traffic Accidents in Korea”, *Journal of the Institute of Electronics and Information Engineers*, 53(9): 89-96.
18. 황선근·이수기, 2018. “서울시 자전거 교통사고와 사고 심각도에 영향을 미치는 근린환경 요인 분석”, 『국토계획』, 53(7): 49-66.
Hwang, S.G. and Lee, S., 2018. “Analysis of Neighborhood Environmental Factors Affecting Bicycle Accidents and Accidental Severity in Seoul, Korea”, *Journal of Korea Planning Association*, 53(7): 49-66.
19. 황선근·이수기, 2021. “서울시 자전거 교통사고의 물리적 환경요인 분석”, 『국토계획』, 56(5): 83-96.
Hwang, S.H. and Lee, S., 2021. “Analysis of Physical Environment Factors of Bicycle Accidents in Seoul, Korea”, *Journal of Korea Planning Association*, 56(5): 83-96.
20. Brown, R.L., Koepplinger, M.E., Mehlman, C.T., Gittelman, M., and Garcia, V.F., 2002. “All-terrain Vehicle and Bicycle Crashes in Children: Epidemiology and Comparison of Injury Severity”, *Journal of Pediatric Surgery*, 37(3): 375-380.
21. Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., and Sallis, J.F., 2009. “Measuring the Built Environment for Physical Activity: State of the Science”, *American Journal of Preventive Medicine*, 36(4): S99-123.
22. Chen, L., Lu, Y., Sheng, Q., Ye, Y., Wang, R., and Liu, Y., 2020. “Estimating Pedestrian Volume using Street View Images: A Large-scale Validation Test”, *Computers, Environment and Urban Systems*, 81: 101481.
23. Chen, P., 2015. “Built Environment Factors in Explaining the Automobile-involved Bicycle Crash Frequencies: A Spatial Statistic Approach”, *Safety Science*, 79: 336-343.
24. De Hartog, J.J., Boogaard, H., Nijland, H., and Hoek, G., 2010. “Do the Health Benefits of Cycling Outweigh the Risks?”, *Environmental Health Perspectives*, 118(8): 1109-1116.
25. Ding, H., Sze, N.N., Li, H., and Guo, Y., 2020. “Roles of Infrastructure and Land Use in Bicycle Crash Exposure and Frequency: A Case Study using Greater London Bike Sharing Data”, *Accident Analysis and Prevention*, 144: 105652.
26. Ewing, R. and Dumbaugh, E., 2009. “The Built Environment and Traffic Safety: A Review of Empirical Evidence”, *Journal of Planning Literature*, 23(4): 347-367.
27. Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., and Wilkinson, P., 2013. “Developing and Testing a Street Audit Tool using Google Street View to Measure Environmental Supportiveness for Physical Activity”, *International Journal of Behavioral Nutrition and Physical Activity*, 10(1): 1-7.
28. Géron, A., 2017. *Hands-on Machine Learning with Scikit-learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems*, Newton: O’Reilly Media.
29. Hanson, C.S., Noland, R.B., and Brown, C., 2013. “The

- Severity of Pedestrian Crashes: An Analysis using Google Street View Imagery”, *Journal of Transport Geography*, 33: 42-53.
30. Harris, M.A., Reynolds, C.C., Winters, M., Crompton, P.A., Shen, H., Chipman, M.L., Michael, D., Cusimano, B.S., Jeffrey, B.J.R., Friedman, S.M., Hunte, G., Monro, M., Vernich, L., and Teschke, K., 2013. “Comparing the Effects of Infrastructure on Bicycling Injury at Intersections and Non-intersections using a Case-crossover Design”, *Injury Prevention*, 19(5): 303-310.
 31. Ibrahim, M.R., Haworth, J., Christie, N., Cheng, T., and Hailes, S., 2021. “Cycling Near Misses: A Review of the Current Methods, Challenges and the Potential of an AI-embedded System”, *Transport Reviews*, 41(3): 304-328.
 32. Kamel, M.B., Sayed, T., and Osama, A., 2019. “Accounting for Mediation in Cyclist-vehicle Crash Models: A Bayesian Mediation Analysis Approach”, *Accident Analysis and Prevention*, 131: 122-130.
 33. Kaplan, S. and Prato, C.G., 2015. “A Spatial Analysis of Land Use and Network Effects on Frequency and Severity of Cyclist-motorist Crashes in the Copenhagen Region”, *Traffic Injury Prevention*, 16(7): 724-731.
 34. Kita-Wojciechowska, K. and Kidziński, Ł., 2019a. “Google Street View Image of a House Predicts Car Accident Risk of Its Resident”, *arXiv Preprint arXiv*, 1904.05270.
 35. Kita-Wojciechowska, K. and Kidziński, Ł., 2019b. “Google Street View Image Predicts Car Accident Risk”, *Central European Economic Journal*, 6(53): 151-163.
 36. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., and Jackel, L.D., 1989. “Backpropagation Applied to Handwritten Zip Code Recognition”, *Neural Computation*, 1(4): 541-551.
 37. Lee, J., Abdel-Aty, M., and Cai, Q., 2017. “Intersection Crash Prediction Modeling with Macro-level Data from Various Geographic Units”, *Accident Analysis and Prevention*, 102: 213-226.
 38. Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., and Zhang, W., 2015. “Assessing Street-level Urban Greenery using Google Street View and a Modified Green View Index”, *Urban Forestry and Urban Greening*, 14(3): 675-685.
 39. Lu, W., Liu, J., Fu, X., Yang, J., and Jones, S., 2022. “Integrating Machine Learning into Path Analysis for Quantifying Behavioral Pathways in Bicycle-motor Vehicle Crashes”, *Accident Analysis and Prevention*, 168: 106622.
 40. Lu, Y., 2019. “Using Google Street View to Investigate the Association between Street Greenery and Physical Activity”, *Landscape and Urban Planning*, 191: 103435.
 41. Lusk, A.C., Furth, P.G., Morency, P., Miranda-Moreno, L.F., Willett, W.C., and Dennerlein, J.T., 2011. “Risk of Injury for Bicycling on Cycle Tracks versus in the Street”, *Injury Prevention*, 17(2): 131-135.
 42. Marqués, R. and Hernández-Herrador, V., 2017. “On the Effect of Networks of Cycle-tracks on the Risk of Cycling. The Case of Seville”, *Accident Analysis and Prevention*, 102: 181-190.
 43. Oja, P., Titze, S., Bauman, A., De Geus, B., Krenn, P., Reger-Nash, B., and Kohlberger, T., 2011. “Health Benefits of Cycling: A Systematic Review”, *Scandinavian Journal of Medicine & Science in Sports*, 21(4): 496-509.
 44. Pham, H., Dai, Z., Xie, Q., and Le, Q.V., 2021. “Meta Pseudo Labels”, paper presented at the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 11557-11568.
 45. Rabl, A. and De Nazelle, A., 2012. “Benefits of Shift from Car to Active Transport”, *Transport Policy*, 19(1): 121-131.
 46. Rahman, M.S., 2018. “Applying Machine Learning Techniques to Analyze the Pedestrian and Bicycle Crashes at the Macroscopic Level”, Master’s Dissertation, University of Central Florida.
 47. Rahman, M.S., Abdel-Aty, M., Hasan, S., and Cai, Q., 2019. “Applying Machine Learning Approaches to Analyze the Vulnerable Road-users’ Crashes at Statewide Traffic Analysis Zones”, *Journal of Safety Research*, 70: 275-288.
 48. Rodgers, G.B., 1997. “Factors Associated with the Crash Risk of Adult Bicyclists”, *Journal of Safety Research*, 28(4): 233-241.
 49. Rojas-Rueda, D., De Nazelle, A., Tainio, M., and Nieuwenhuijsen, M.J., 2011. “The Health Risks and Benefits of Cycling in Urban Environments Compared with Car Use: Health Impact Assessment Study”, *BMJ*, 343.
 50. Shirani-Bidabadi, N., Mallipaddi, N., Haleem, K., and Anderson, M., 2020. “Developing Bicycle-vehicle Crash-specific Safety Performance Functions in Alabama using Different Techniques”, *Accident Analysis and Prevention*, 146: 105735.
 51. Karen, S. and Zisserman, A., 2014. “Very Deep Convolutional Networks for Large-scale Image Recognition”, *arXiv*, 1409.1556 (in press).
 52. 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.
 53. Wang, Y. and Nihan, N.L., 2004. “Estimating the Risk of Collisions between Bicycles and Motor Vehicles at Signalized Intersections”, *Accident Analysis and Prevention*, 36(3): 313-321.
 54. Wei, F. and Lovegrove, G., 2013. “An Empirical Tool to Evaluate the Safety of Cyclists: Community Based, Macro-level Collision Prediction Models using Negative Binomial Regression”, *Accident Analysis and Prevention*, 61: 129-137.
 55. Wen, L.M. and Rissel, C., 2008. “Inverse Associations between Cycling to Work, Public Transport, and Overweight and Obesity: Findings from A Population based Study in Australia”, *Preventive Medicine*, 46(1): 29-32.
 56. 서울경제, 2022.04.27. “서울 공공자전거 ‘따릉이’, 6년 5개월 만에 누적 이용 1억건 돌파”, <https://www.sedaily.com/NewsView/264UUETKRC>
Seoul Economic Daily, 2022, Apr 27. “Seoul’s Public Bicycle ‘Ttareungyi’, after 6 Years and 5 Months Exceeded 100 Million Cumulative Usage on”, <https://www.sedaily.com/NewsView/264UUETKRC>

57. 이투데이, 2022.04.27. “따릉이’ 누적이용량 1억 건 돌파…코로나 이후 대폭 증가”, <https://www.etoday.co.kr/news/view/2127974>
 Etoday, 2022, Apr 27. “The Cumulative Usage of ‘Ttareungyi’ Exceeded 100 Million. . . Significant Increase after Covid19”, <https://www.etoday.co.kr/news/view/2127974>

58. 한국교통연구원, 2017. “자전거 이용 인구 1340만 명 - 그중 330만 명은 매일 자전거 이용,” 2022.12.14. 읽음. https://www.koti.re.kr/user/bbs/BD_selectBbs.do?q_bbsCode=1005&q_bbscttSn=20170309140233834
 Korea Transport Institute, 2017. “13.4 Million Cyclists – 3.3 Million of Whom Ride a Bike Every Day,” Accessed Dec 14, 2022. https://www.koti.re.kr/user/bbs/BD_selectBbs.do?q_bbsCode=1005&q_bbscttSn=20170309140233834

59. TAAS, “교통사고 부문별 사고현황 최근 5년,” 2022.08.09. 읽음. http://taas.koroad.or.kr/sta/acs/gus/selectTfcacdOccrrncSttusAnnual.do?menuId=WEB_KMP_OVT_MVT_TAC_TOA

TAAS, “Traffic Accident Status by Category Last 5 Years”, Accessed Sep 8, 2022. http://taas.koroad.or.kr/sta/acs/gus/selectTfcacdOccrrncSttusAnnual.do?menuId=WEB_KMP_OVT_MVT_TAC_TOA

60. TAAS, “교통사고 분석시스템”, 2022.06.15. 읽음. http://taas.koroad.or.kr/gis/mcm/mcl/initMap.do?menuId=GIS_GMP_ABS
 TAAS, “Raffic Accident Analysis System”, Accessed Jun 15, 2022. http://taas.koroad.or.kr/gis/mcm/mcl/initMap.do?menuId=GIS_GMP_ABS

61. Andrew Ng, 2017, Aug 26. “Train/Dev/Test Sets(C2W1L01)”, <https://www.youtube.com/watch?v=1waHlpKiNyY>

Date Received 2022-10-14
 Date Reviewed 2022-11-22
 Date Accepted 2022-11-22
 Date Revised 2022-12-16
 Final Received 2022-12-16