

빅데이터와 딥러닝 활용한 서울시 보행사고 분석과 시사점

한영준 이하식



빅데이터와 딥러닝 활용한 서울시 보행사고 분석과 시사점

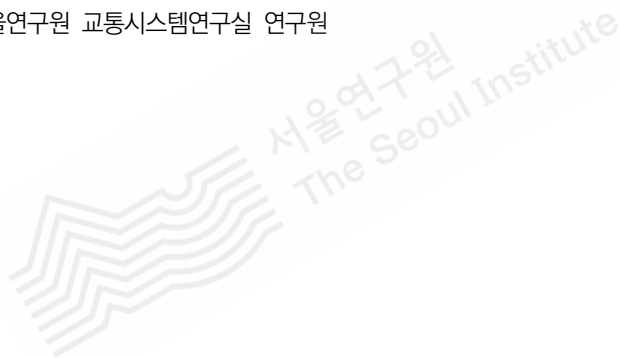


연구책임

한영준 서울연구원 교통시스템연구실 부연구위원

연구진

이하식 서울연구원 교통시스템연구실 연구원

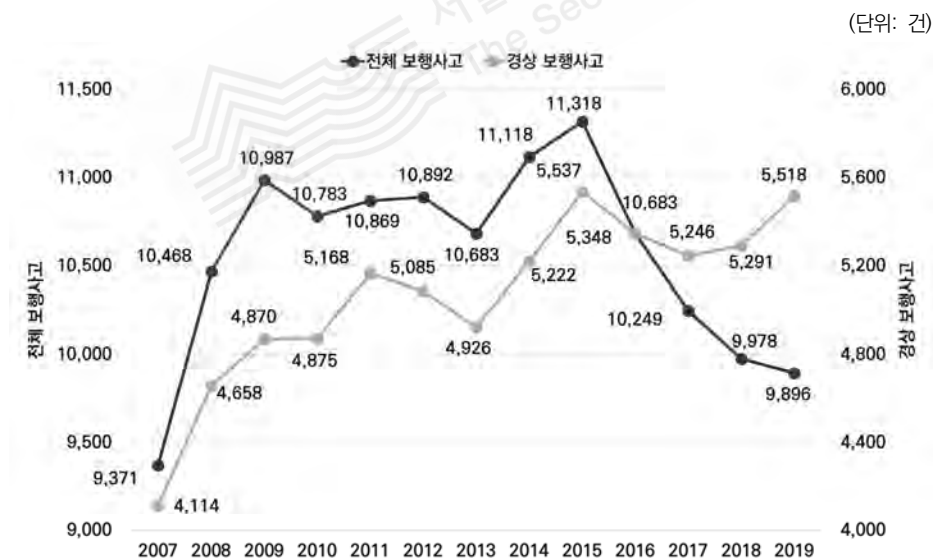


이 보고서의 내용은 연구진의 견해로서
서울특별시의 정책과는 다를 수도 있습니다.

보행환경 변화에 따른 빅데이터·딥러닝 활용으로 서울시 보행사고 예측력 향상 가능…과제도 산적

보행환경 변화에 따른 경상사고 급증으로 서울시 보행사고 감소추세 둔화

다양한 보행친화 정책을 추진한 서울시에서는 2015년 이후 보행사고가 감소하고 있으나 최근 경상사고가 급격히 증가하면서 그 감소 추세가 둔화되고 있다. 보행 중 스마트폰 이용이 활발해지면서 보행행태가 변화하고, 전동킥보드와 배달이륜차의 보도주행으로 보행환경이 악화되면서 경미한 보행사고가 증가하고 있는 것이다. 경상 보행사고는 상대적으로 심각도가 낮지만 언제든지 심각도가 높은 중상 및 사망사고로 발전할 수 있어 간과할 수 없는 현상이라 할 수 있다.



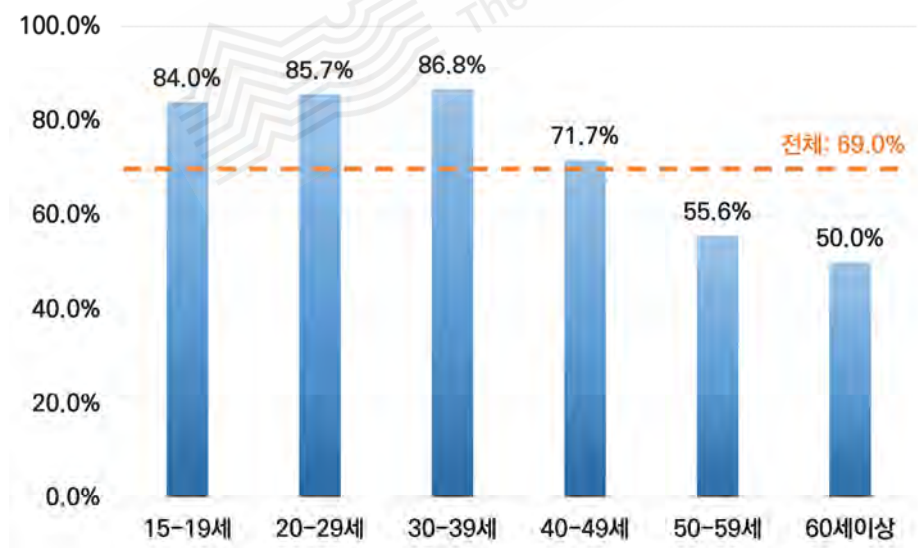
[그림 1] 서울시 보행사고 및 경상 보행사고 추이(2007~2019년)

보행사고 특성, 심각도·지역·연령대별 다양화 … 피해자 중 20대 비율 증가

서울시에서 지난 13년간 발생한 13.7만 건의 보행사고 데이터를 수집하고 보행사고의 변화를 분석한 결과, 서울시 보행사고의 특성이 심각도, 지역, 시기, 가해자 및 피해자 연령대별로 다양화되고 있는 것으로 나타났다. 특히, 최근 보행사고에서 20대 피해자 비율이 증가하고 있는데, 이는 보행 중 스마트폰이나 전동스쿠터 등 새로운 기기를 적극적으로 이용하는 연령대의 특징이 반영된 것으로 생각된다.

보행 중 스마트폰 이용률, 서울시 보행자의 69%, 30대 이하는 85% 넘어

서울시민의 69%가 보행 중 스마트폰을 이용하고 있었다. 특히 30대 이하의 사용률은 85% 이상을 차지하였고, 게임, 동영상, 만화 등 스마트폰 몰입으로 보행사고에 노출되기 쉬운 활동을 하는 응답자의 비율도 높게 나타났다. 조사 대상자의 대부분이 타인의 스마트폰 이용으로 인해 충돌 위험이나 보행 방해로 경험을 하였다고 응답하여, 보행 중 스마트폰 이용이 본인과 타인의 보행행태에 부정적인 영향을 주고 있음을 확인할 수 있었다.



[그림 2] 연령대별 보행 중 스마트폰 이용률

빅데이터 정확성 높이는 ‘딥러닝’ 모델 활용 결과, 보행사고 예측능력 향상

전통적인 통계분석 방법인 회귀분석 모형과 딥러닝 모형을 비교한 결과, 다양한 종속 변수를 추정하는 모든 경우에서 딥러닝 모형의 성능이 보다 우수하게 나타났다. 특히, 보행사고와 상관관계가 높은 생활인구의 영향을 배제한 보행사고율 추정 시 딥러닝 모형의 성능이 회귀분석 모형에 비해 현저히 높게 나타나 향후 보행사고 분석에 대한 가능성을 보여주었다. 또한, 회귀분석은 모형 구축 과정에서 빅데이터 상당 부분을 배제하게 되었지만, 딥러닝은 별도의 선별 작업 없이 모든 데이터를 활용할 수 있었다.

[표 1] 딥러닝과 회귀분석을 이용하여 예측한 보행사고 결과 비교

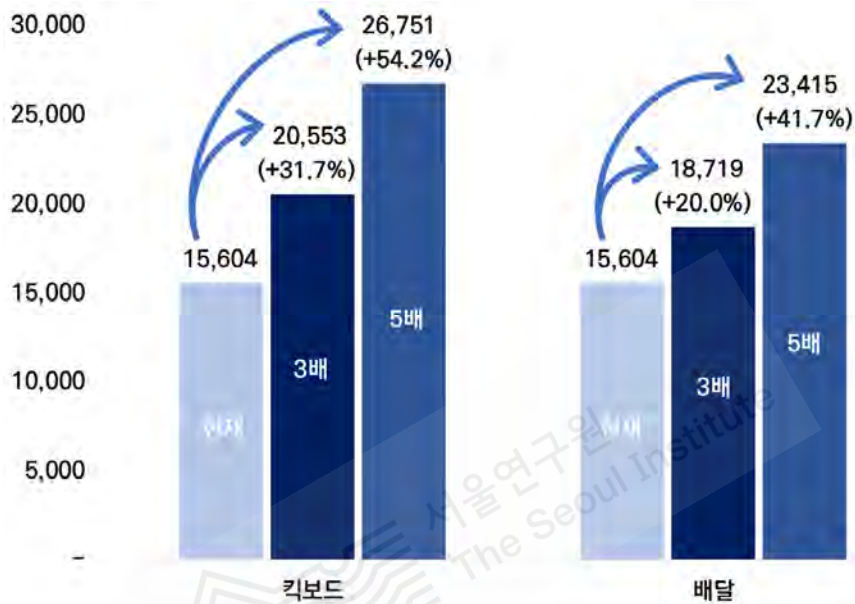
종속변수		모형	R ²		RMSE		MAE	
			Test	차이	Test	차이	Test	차이
보행사고 발생건수	전체	딥러닝	0.87	+0.19	12.02	-6.51	9.50	-3.95
		회귀분석	0.68		18.53		13.45	
	중상	딥러닝	0.72	+0.08	6.10	-1.45	4.78	-0.80
		회귀분석	0.64		7.55		5.58	
	경상	딥러닝	0.84	+0.22	7.81	-3.52	5.87	-2.39
		회귀분석	0.62		11.33		8.26	
보행사고율	전체	딥러닝	0.57	+0.33	2.25	-0.54	1.76	-0.39
		회귀분석	0.24		2.79		2.15	
	중상	딥러닝	0.42	+0.16	1.09	-0.12	0.85	-0.09
		회귀분석	0.26		1.21		0.94	
	경상	딥러닝	0.60	+0.40	1.33	-0.36	1.05	-0.22
		회귀분석	0.20		1.69		1.27	

서울시, 선제적 보행정책 수립 가능 … 자료 확보, 분석기법 등 과제도 많아

구축한 보행사고 추정 모형을 활용하여 장래 보행환경과 보행행태 변화에 따른 보행사고의 변화를 예측하였다. 보행환경 변화와 관련이 있는 키보드 앱 사용 횟수와 배달 앱 사용자 수가 증가하는 경우의 보행사고 변화를 추정한 결과, 해당 앱의 사용량이 증가할수록 보행사고도 증가하는 것으로 나타났다. 보행행태 변화를 대표하는 동영상과 음악 앱 사용자 수를 증가시키는 경우에도 보행사고가 증가하는 것을 확인할 수 있었다. 이러한 장래 보행사고 추정 결과를 바탕으로 향후 보행환경 변화에 따른 보행사고를 전망하여 선제적인 보행정책을 마련할 수 있을 것으로 기대된다. 개발된 모형의

성능을 향상시키기 위해 지속적이고 체계적인 보행 관련 빅데이터 수집과 시공간적 데이터를 통합 분석하는 최신 알고리즘의 적용이 필요할 것으로 생각된다. 또한, 이 연구에서 제시한 분석체계를 이용하여 도심 교통혼잡 예측, 대중교통 수요분석 등 다양한 분야에 적용이 가능할 것으로 기대된다.

(단위: 건)



[그림 3] 키보드, 배달 앱 사용 증가에 따른 보행사고 건수 추정

목차

01 연구개요	2
1_연구배경 및 목적	2
2_연구내용 및 방법	4
02 서울시 보행사고 현황 분석	6
1_이용된 보행사고 데이터	6
2_서울시 보행사고 현황 분석	9
3_서울시 보행사고 현황 분석의 시사점	20
03 설문조사를 통한 서울시민의 보행행태 분석	24
1_조사 개요	24
2_보행자의 보행 중 스마트폰 사용 실태	25
3_운전자의 운전 중 스마트폰 사용 실태	36
4_보행 중 스마트폰 사용억제 방법에 관한 시민인식	42
04 빅데이터와 딥러닝을 활용한 보행사고 추정	48
1_보행사고 설명을 위한 빅데이터 구축	48
2_보행사고 추정을 위한 딥러닝 모형 구축 및 적용	59
3_장래 보행환경·행태 변화에 따른 보행사고 변화 예측	75
4_딥러닝 분석의 시사점 및 향후 과제	77
05 맺음말	80
참고문헌	83
Abstract	84

표 목차

[표 2-1] TAAS 데이터 예시	8
[표 2-2] TAAS 보행사고의 사고유형 및 법규위반 종류	8
[표 3-1] 보행행태 분석을 위한 스마트폰 사용 실태 조사 개요	24
[표 3-2] 보행행태 분석을 위한 스마트폰 사용 실태 조사 표본 구성	25
[표 3-3] 보행 중 불편을 주는 요소	26
[표 3-4] 보행 중 타인의 스마트폰 사용으로 인한 불편 경험 유무	27
[표 3-5] 보행 중 스마트폰 사용이 타인에게 불편을 주는 이유	28
[표 3-6] 보행 중 스마트폰 사용 여부	29
[표 3-7] 보행 중 스마트폰 사용자의 사용 목적	31
[표 3-8] 보행 중 이용하는 스마트폰 활동의 중요성	34
[표 3-9] 보행 중 스마트폰 사용자의 이어폰 착용 여부	35
[표 3-10] 운전 경험 유무	36
[표 3-11] 운전 중 불편을 주는 요소	37
[표 3-12] 운전 중 스마트폰 사용 여부	38
[표 3-13] 운전 중 스마트폰 사용자의 사용 목적	39
[표 3-14] 운전 중 스마트폰 사용자의 부가장치 이용 여부	41
[표 3-15] 보행 중 스마트폰 사용억제 방법에 대한 선호도	42
[표 3-16] 스몸비-차량 사고 발생 시 스몸비를 가해자로 지정하는 것에 대한 의견	45
[표 4-1] 독립변수 목록	50
[표 4-2] 앱 데이터 카테고리 목록	51
[표 4-3] 최근 5년간 보행사고 발생 상위 10개 역세권	53
[표 4-4] 최근 5년간 보행사고 발생 하위 10개 역세권	53
[표 4-5] 변수 기초통계량	58

[표 4-6] 딥러닝 모델링을 위한 모형 종류	59
[표 4-7] 딥러닝 모형 평가	63
[표 4-8] 회귀분석을 위한 모든 독립변수의 분산팽창요인 지표	67
[표 4-9] 회귀분석을 위해 선별한 독립변수의 분산팽창요인 지표	68
[표 4-10] 전체사고 건수 추정 선형 회귀모형 결과	69
[표 4-11] 중상사고 건수 추정 선형 회귀모형 결과	69
[표 4-12] 경상사고 건수 추정 선형 회귀모형 결과	70
[표 4-13] 전체사고율 추정 선형 회귀모형 결과	70
[표 4-14] 중상사고율 추정 선형 회귀모형 결과	71
[표 4-15] 경상사고율 추정 선형 회귀모형 결과	71
[표 4-16] 회귀분석 모형 평가	71
[표 4-17] 딥러닝과 회귀분석을 이용하여 예측한 보행사고 결과 비교	74



그림 목차

[그림 1-1] OECD 주요 국가 인구 10만명당 보행사고 사망자 수(2017년)	3
[그림 1-2] 연구내용 및 방법	4
[그림 2-1] 교통사고 GIS 분석 시스템 예시	7
[그림 2-2] 연도별 서울시 보행사고 건수	9
[그림 2-3] 서울시의 보행정책(2012~2017년)	10
[그림 2-4] 사고내용별 서울시 보행사고 건수	11
[그림 2-5] 서울시 구별 2015년 대비 2019년 전체 보행사고 증감률	12
[그림 2-6] 서울시 구별 2015년 대비 2019년 중상사고 증감률	13
[그림 2-7] 서울시 구별 2015년 대비 2019년 경상사고 증감률	13
[그림 2-8] 연도별 서울시 보행사고 가해운전자 연령대 비율	14
[그림 2-9] 연도별 서울시 보행사고 피해운전자(보행자) 연령대 비율	15
[그림 2-10] 연도별 서울시 보행사고 가해자 차종 비율	16
[그림 2-11] 연도별 PM이 가해운전자 차종으로 분류된 서울시 보행사고 건수	17
[그림 2-12] 월별 서울시 보행사고 건수	17
[그림 2-13] 2018년 주차별 서울시 보행사고 건수	18
[그림 2-14] 요일별 서울시 보행사고 건수	19
[그림 2-15] 요일별 시간대별 서울시 보행사고 건수	20
[그림 3-1] 연령대별 보행 중 스마트폰 사용률	29
[그림 3-2] 보행 중 스마트폰 사용자의 사용 목적(전체)	32
[그림 3-3] 보행 중 스마트폰 사용자의 사용 목적(남성)	32
[그림 3-4] 보행 중 스마트폰 사용자의 사용 목적(여성)	33
[그림 3-5] 보행 중 이용하는 스마트폰 활동의 중요성	34

[그림 3-6] 보행 중 스마트폰 사용자의 운전 중 스마트폰 사용 비율	38
[그림 3-7] 보행 중 스마트폰 미사용자의 운전 중 스마트폰 사용 비율	38
[그림 3-8] 보행 및 운전 중 스마트폰 사용자의 사용 목적	40
[그림 3-9] 연령대별 보행 시 스마트폰 사용 금지 제도 마련 선택률	43
[그림 3-10] 연령대별 스마트폰 사용 보행자 사고 유발 시 과실 부과 선택률	43
[그림 3-11] 스몸비-차량 사고 발생 시 스몸비를 가해자로 지정하는 것에 대한 연령대별 찬성률	45
[그림 4-1] 역 반경별 포함되는 보행사고 비율	52
[그림 4-2] 역세권별 보행사고 발생 건수 히스토그램	52
[그림 4-3] 인스타그램 사용자 수 추정	57
[그림 4-4] 모델링 과정	61
[그림 4-5] 딥러닝 모형 구조	62
[그림 4-6] 스마트폰 이용 데이터 포함 여부에 따른 모형 성능 차이	63
[그림 4-7] 스마트폰 이용 데이터 미포함 전체사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-8] 스마트폰 이용 데이터 포함 전체사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-9] 스마트폰 이용 데이터 미포함 중상사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-10] 스마트폰 이용 데이터 포함 중상사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-11] 스마트폰 이용 데이터 미포함 경상사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-12] 스마트폰 이용 데이터 포함 경상사고 건수 추정 딥러닝 모형 테스트 결과	65
[그림 4-13] 스마트폰 이용 데이터 미포함 전체사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-14] 스마트폰 이용 데이터 포함 전체사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-15] 스마트폰 이용 데이터 미포함 중상사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-16] 스마트폰 이용 데이터 포함 중상사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-17] 스마트폰 이용 데이터 미포함 경상사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-18] 스마트폰 이용 데이터 포함 경상사고율 추정 딥러닝 모형 테스트 결과	66
[그림 4-19] 전체사고 건수 추정 회귀분석 모형 테스트 결과	72
[그림 4-20] 중상사고 건수 추정 회귀분석 모형 테스트 결과	72
[그림 4-21] 경상사고 건수 추정 회귀분석 모형 테스트 결과	72
[그림 4-22] 전체사고율 추정 회귀분석 모형 테스트 결과	72
[그림 4-23] 중상사고율 추정 회귀분석 모형 테스트 결과	72

[그림 4-24] 경상사고율 추정 회귀분석 모형 테스트 결과	72
[그림 4-25] 킥보드, 배달 앱 사용 증가에 따른 보행사고 건수 추정	76
[그림 4-26] 동영상, 음악 앱 사용 증가에 따른 보행사고 건수 추정	76

01

연구개요



1_연구배경 및 목적

2_연구내용 및 방법

01. 연구개요

1_연구배경 및 목적

1) 연구의 배경

서울시는 ‘보도블록 10계명’, ‘걷는 도시, 서울’ 등 다양한 보행친화 정책을 추진하였고, 그 결과 서울시의 보행사고는 지속적으로 감소하고 있다. 그러나, 2019년에도 서울에서 9,896건의 보행사고가 발생하여, 전체 교통사고의 25.2%를 차지하고 있다. 특히 서울시 인구 10만 명당 보행자 사망자 수는 2017년 현재 OECD 평균(1.0명)의 2배에 가까운 1.92명으로 서울시의 보행사고 심각도는 여전히 높은 편이다.

보행사고를 설명하기 위하여 기존의 연구들은 주로 교통량, 보행량, 도로 기하구조 등 교통 관련 변수나 물리적 환경을 이용하였다(백태현 외, 2014; 박철영·이수기, 2016; 이하식 외, 2019). 하지만, 최근 다양한 신기술의 등장으로 보행환경이 변화하면서 새로운 유형의 보행사고 발생이 증가하고 있다. 주요 보험사 접수 건수를 기준으로, 2017년 보행 중 스마트폰 사용으로 인해 발생한 사고는 177건으로 2011년과 비교하여 약 2배로 증가하였고,¹⁾ 전동킥보드와 관련된 교통사고도 2016년 49건에서 2018년 258건으로 큰 폭으로 증가하였다.²⁾ 따라서 최근의 보행사고를 이해하기 위해서는 변화하는 보행환경을 반영할 수 있는 데이터를 확보하여 분석할 필요가 있다. 또한, 교통사고에 영향을 미치는 변수는 아주 다양하고, 동일한 조건에서도 교통사고가 발생할 확률은 아주 낮기 때문에 통계 등 전통적인 방법으로는 교통사고의 특성을 명확히 설명하기 어려운 경우가 많았다. 최근에는 이를 극복하기 위해 다양한 기계학습(machine learning) 기법을 이용하는 노력이 증가하고 있다(곽호찬 외, 2014; 류

1) [https://blog.hi.co.kr/\(현대해상 공식블로그\)](https://blog.hi.co.kr/(현대해상 공식블로그))

2) 삼성화재, 2019.07.26., “삼성교통안전문화연구소 전동킥보드 교통사고 실태 및 예방대책”.

종득, 2018). 따라서 보행환경 변화를 반영한 다양한 데이터를 구축한 후에는 이를 효과적으로 분석하기 위해 최신 데이터 분석 기술을 적용하는 노력도 필요하다.



자료: 도로교통공단, 2019, OECD 회원국 교통사고 비교

[그림 1-1] OECD 주요 국가 인구 10만명당 보행사고 사망자 수(2017년)

2) 연구의 목적

본 연구의 첫 번째 목적은 서울시에서 발생한 보행사고를 분석하여 보행사고의 변화 추이를 도출하는 것이다. 단순한 보행사고 발생 건수뿐만 아니라, 심각도, 피해보행자 및 가해운전자에 따른 보행사고 현황을 상세히 분석하여 시사점을 도출하고자 하였다. 두 번째 목적은 스마트폰 이용, 전동스쿠터 활성화 등으로 변화하는 보행환경과 보행 행태를 파악하는 것이다. 이를 위해 서울시민 1,000명을 대상으로 설문조사를 시행하고 결과를 분석하였다.

세 번째 목적은 보행환경 및 보행행태와 관련된 다양한 데이터를 수집하여 보행사고 빅데이터를 구축하고, 최신 데이터 분석 기법인 딥러닝을 이용하여 보행사고 모형을 개발하는 것이다. 개발된 모형을 이용하여 장래 보행환경을 고려한 보행사고를 추정하고 시사점을 도출하고자 하였다.

2_연구내용 및 방법

본 연구에서는 서울시에서 발생한 13.7만 건의 보행사고 데이터를 분석하여 보행사고 현황과 특성을 도출하였다. 특히, 보행사고의 시·공간적 분포와 보행사고 심각도에 따른 특성을 분석하였고, 보행사고 가해운전자와 피해보행자를 상세하게 분석하여 최근 보행사고의 발생 원인을 유추하였다.

또한, 변화하는 보행환경과 보행행태를 조사하기 위해 서울시민 1,000명을 대상으로 설문조사를 시행하였다. 설문조사 결과를 분석하여 서울시민이 인식하는 보행환경과 선호하는 보행환경 개선 방향을 도출하였다. 변화하는 보행행태를 추정하기 위하여 보행 중 스마트폰 이용 여부, 보행 중 스마트폰 기반 활동 등을 성별, 연령대별로 상세히 분석하였고, 운전 중 스마트폰 이용행태와 비교하여 보행환경 개선을 위한 시사점을 도출하였다.

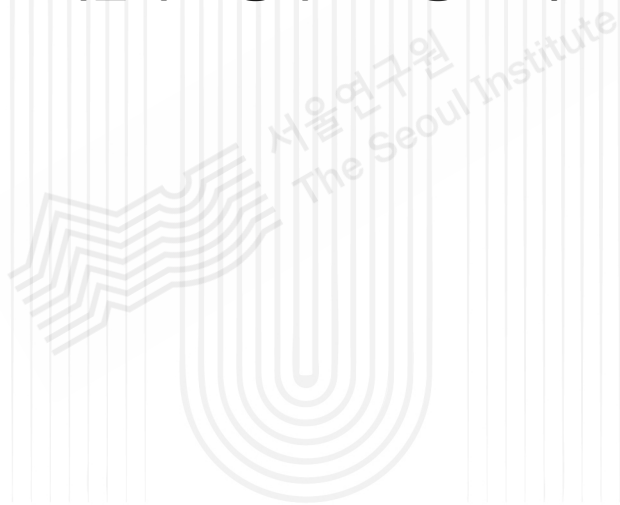
보행사고 현황 조사와 보행행태 설문조사를 통해 보행환경이 변화함을 확인하고, 이를 반영한 모형을 구축하여 보행사고를 설명하였다. 우선, 교통데이터, 사회경제지표 등 기존 연구에 활용되었던 공공데이터와 스마트폰 이용정보를 포함하는 민간데이터를 통합하여 보행사고 빅데이터를 구축하고, 최신 데이터 분석 기법인 딥러닝을 적용하여 모형을 구축하였다. 딥러닝을 이용한 분석 결과를 기존의 회귀분석 결과와 비교하여 모형의 우수성을 입증하고, 구축된 모형을 이용하여 장래 보행환경 변화 시 보행사고의 변화를 추정하여 보행사고 감소를 위한 시사점을 제시하였다.



[그림 1-2] 연구내용 및 방법

02

서울시 보행사고 현황 분석



1_이용된 보행사고 데이터

2_서울시 보행사고 현황 분석

3_서울시 보행사고 현황 분석의 시사점

02. 서울시 보행사고 현황 분석

1_이용된 보행사고 데이터

서울시 보행사고 분석을 위해 도로교통공단 교통사고분석시스템(Traffic Accident Analysis System, TAAS)의 데이터를 이용하였다. 도로교통공단에서는 경찰에 접수된 교통사고 자료를 TAAS 홈페이지를 통해 제공하고 있다. 특히, ‘교통사고 GIS 분석 시스템’³⁾을 이용하면, [그림 2-1]과 같이 교통사고의 공간적 분포와 특성을 확인할 수 있다. TAAS 교통사고 데이터는 사고 식별 번호인 사고번호, 사고 발생 일시와 요일 등 시간적 정보, 발생 시군구 및 좌표 등 공간적 정보⁴⁾, 사고유형, 사고내용, 가해운전자 및 피해운전자 정보, 노면 및 기상 상태 등 환경 정보를 포함하고 있다. TAAS교통사고 데이터의 예시는 [표 2-1]과 같다.

TAAS 교통사고 데이터의 사고유형은 ‘차대사람’, ‘차대차’, ‘차량단독’, ‘철길건널목’으로 구분된다. 이 중에서 사고유형이 ‘차대사람’인 데이터를 추출하여 보행사고 데이터를 구축하였다.

TAAS 교통사고 데이터의 사고내용은 ‘사망자 수’, ‘중상자 수’, ‘경상자 수’, ‘부상신고자 수’를 포함하고 있다. 본 연구에서는 각각의 사고내용 발생 여부에 따라 ‘사망사고’, ‘중상사고’, ‘경상사고’, ‘부상신고사고’로 분류하였으며, 심각도가 높은 사고내용을 우선하였다. 보행사고의 사고유형, 법규위반 종류는 [표 2-2]와 같다.

보행사고 데이터의 시간적 범위는 2007년부터 2019년까지이고, 공간적 범위는 서울

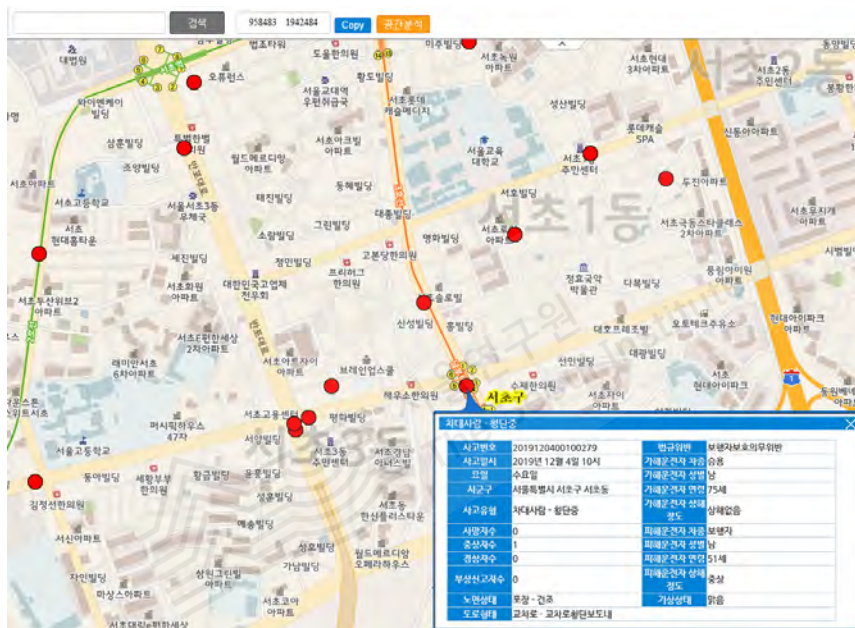
3) 교통사고 GIS 분석 시스템,

http://taas.koroad.or.kr/gis/mcm/mcl/initMap.do?menuId=GIS_GMP_STS_RSN

4) TAAS의 교통사고 GIS 분석 시스템에서 표시된 사고지점을 선택하면 사고지점의 좌표를 알 수 있다. 하지만, GIS 분석 시스템에 표시된 사고지점은 실제 사고지점과 다를 수 있어, 사고지점 좌표를 공식적인 자료로는 제공하지 않고 있다. 본 연구에서는 권역별, 역세권별 등 일정 지역에서 발생하는 사고 현황 분석을 위해 TAAS의 교통사고 GIS 분석 시스템에서 표시되는 교통사고 좌표를 추출하여 활용하였다.

특별시 전체로 하였다. 해당 기간 서울시에서 발생한 전체 교통사고 데이터 526,954건 중 약 26.1%인 137,295건의 보행사고 데이터를 수집하였다.

구축된 보행사고 데이터를 이용하여 2장 2절에서는 서울시의 보행사고 발생 현황을 분석하고 시사점을 도출하였으며, 3장에서는 보행행태 관련 시민 설문조사 결과를 통해 변화된 보행환경과 2장 2절에서 도출된 시사점을 검증하였다. 4장에서는 수집된 보행사고 데이터를 이용하여 보행 빅데이터를 구축하고 딥러닝을 이용하여 보행사고 예측 모형을 개발하였다.



[그림 2-1] 교통사고 GIS 분석 시스템 예시

[표 2-1] TAAS 데이터 예시

구 분	내 용	구 분	내 용
사고번호	2019120400100279	법규위반	보행자보호의무위반
사고일시	2019년 12월 4일 10시	내용	중상사고
요일	수요일	가해운전자 차종	승용
시군구	서울특별시 서초구 서초동	가해운전자 성별	남
사고유형	차대사람 - 횡단중	가해운전자 연령	75세
사망자수	0	가해운전자 상해정도	상해없음
중상자수	1	피해운전자 차종	보행자
경상자수	0	피해운전자 성별	남
부상신고자수	0	피해운전자 연령	51세
노면상태	포장 - 건조	피해운전자 상해정도	중상
도로형태	교차로 - 교차로 횡단보도내	기상상태	맑음

[표 2-2] TAAS 보행사고의 사고유형 및 법규위반 종류

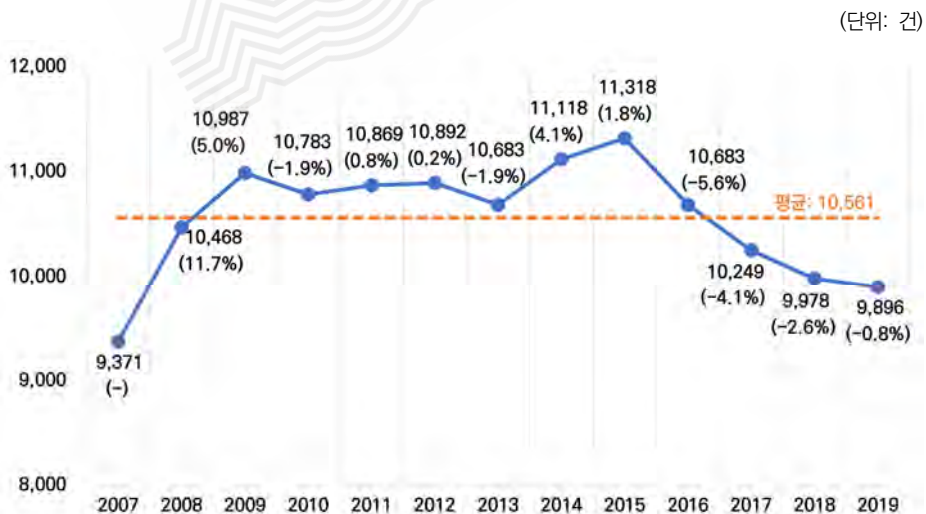
구 분	내 용	구 분	내 용
사고유형	길가장자리구역 통행중 보도 통행중 차도 통행중 횡단중 기타	법규위반	과속 교차로운행방범위반 보행자보호의무위반 불법유턴 신호위반 안전거리 미확보 안전운전불이행 중앙선침범 직진우회전 진행방해 차로위반 미분류 기타

2_서울시 보행사고 현황 분석

1) 연도별 변화 분석

(1) 서울시 전체 보행사고 건수 변화

2007년부터 2019년까지 서울시에서 발생한 전체 보행사고 건수를 연도별로 분석하였다. 분석 기간의 서울시 보행사고의 연평균 발생 건수는 10,561건으로, 시기별로 증가(2007~2009년), 소폭 증가(2009~2015년), 감소(2015년 이후)하는 특성을 보여주었다. 우선, 2007년부터 2009년까지는 사고 건수가 급격히 증가하였다. 특히 2008년에는 보행사고가 전년과 비교하여 1,097건이 증가하여 11.7%의 증가율을 나타내었다. 2010년부터 2015년까지는 증가율 -1.9 ~ 4.1%로 보행사고 건수가 소폭 상승하였고, 2015년에는 11,318건으로 분석 기간 중 가장 높은 보행사고 건수를 나타내었다. 반면에, 2015년 이후로는 보행사고가 지속적으로 감소하였다. 이는 2012년 이후 추진된 서울시의 다양한 보행친화 정책의 성과가 나타난 결과로 판단된다([그림 2-3 참조]). 다만, 사고감소율이 2016년(-5.6%)과 2017년(-4.1%)에는 상대적으로 높았지만, 2018년에는 -2.6%, 2019년에는 -0.8%로 감소하고 있어 최근 보행사고 감소 추세가 둔화된 것으로 나타났다.



[그림 2-2] 연도별 서울시 보행사고 건수



자료: 서울연구원, 2019, '걷는 도시, 서울' 정책효과와 향후 정책방향.

[그림 2-3] 서울시의 보행정책(2012~2017년)

(2) 보행사고 심각도의 변화

전체 보행사고를 심각도에 따라 구분하여 각각의 추세를 분석하였다(그림 2-4) 참조). 우선 사망사고와 중상사고는 전반적으로 감소하는 추세를 보였다.

사망사고는 2008년과 2011년을 제외하고는 지속 감소하였고, 중상사고는 2009년 5,438건을 기록한 이후 2019년까지 계속하여 감소하였다. 특히, 최근 중상사고 감소율은 더욱 커졌는데, 2019년의 중상사고 건수는 2015년과 비교하여 24.0%가 감소하였다. 부상신고사고도 2015년 이후 감소하는 추세를 보여주고 있다. 반면에 경상사고는 분석 기간 동안 전반적으로 증가하였다. 특히 다른 심각도의 사고가 모두 감소하는 2015년 이후에도 경상사고는 소폭 감소하거나 증가하고 있어 최근 전체 보행사고 중 경상사고의 비중이 크게 증가하고 있음을 알 수 있다.

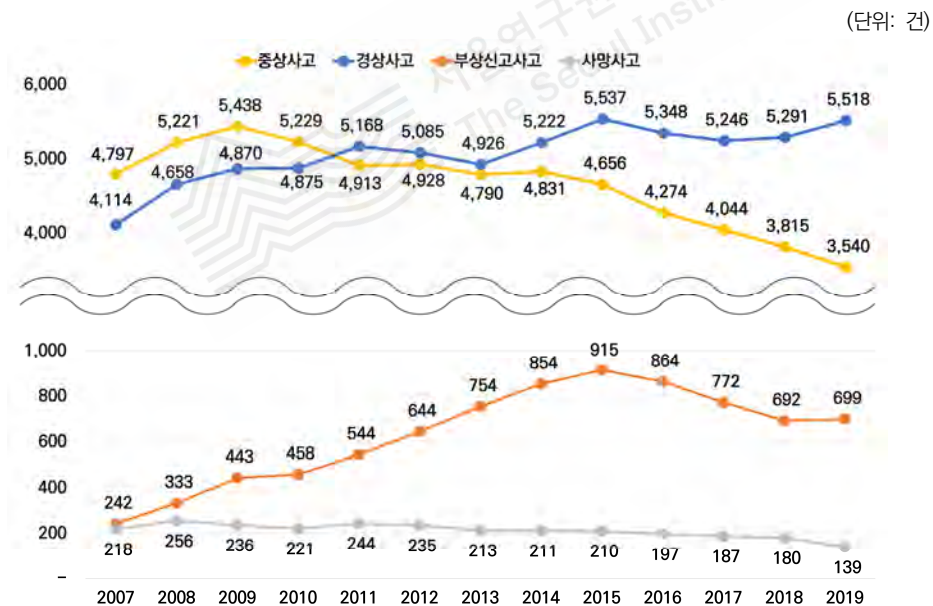
사망사고와 부상신고사고가 전체 보행사고에서 차지하는 비중이 상대적으로 낮기 때문에, 서울시 보행사고 심각도에 따른 추세는 '중상사고 감소, 경상사고 증가'로 정리할 수 있다.

중상사고와 경상사고의 추세가 다른 원인은 다음의 두 가지 경우로 구분하여 예상할 수 있다. 첫 번째 경우는 보행사고의 심각도가 낮아져 기존에는 중상사고로 발생할 사고가 경상사고로 전환되는 경우이다. 이는 차량의 보행자 보호 기술 발전과 다양한 보행친화 교통정책의 영향으로 설명할 수 있다. 예를 들어, 최근 차량에 액티브 후드 시스템⁵⁾ 등 보행자 보호 기능이 적용되면서 사고 발생 시 보행자 사고의 심각도를

낮추고 있다. 또한, ‘차대사람’ 사고에서 보행자의 충격량을 결정하는 차량 속도가 낮아지면 사고 심각도가 낮아질 수 있는데, 5030안전속도, 어린이보호구역 강화 등 다양한 보행친화 정책이 시행되면서 차량 속도가 낮아지고, 그 결과 사고의 심각도가 낮아졌다고 할 수 있을 것이다.

두 번째는 보행자 행태와 보행환경이 변화하면서 보행자 부주의로 발생하는 경미한 사고가 증가하는 경우이다. 최근 보행 중 스마트폰 이용이나 공유형 전동킥보드 활성화로 발생하는 보행사고가 크게 증가하고 있는데, 새롭게 발생하는 사고들이 대부분 낮은 심각도로 발생하여 경상사고 건수가 증가한다고 할 수 있다.

보행사고의 원인은 복합적이기 때문에 앞에서 제시한 두 가지 경우 모두 경상사고 증가에 영향을 미쳤을 수 있다. 전자의 경우를 위해서는 현재와 같이 적극적인 보행친화 정책을 지속 추진하고 다양한 보행자 안전기술을 개발해야 할 것이다. 하지만, 후자의 경우는 보행사고의 발생 원인이 추가적으로 생성된 것이기 때문에, 변화하는 보행행태를 반영한 보행사고 분석을 시행하고 변화된 보행환경에 맞는 정책을 마련해야 할 것이다.



[그림 2-4] 사고내용별 서울시 보행사고 건수

5) 액티브 후드 시스템: 보행자 충돌 감지 시 차량의 후드를 들어 올려 사람이 받는 충격량을 줄여주는 기술

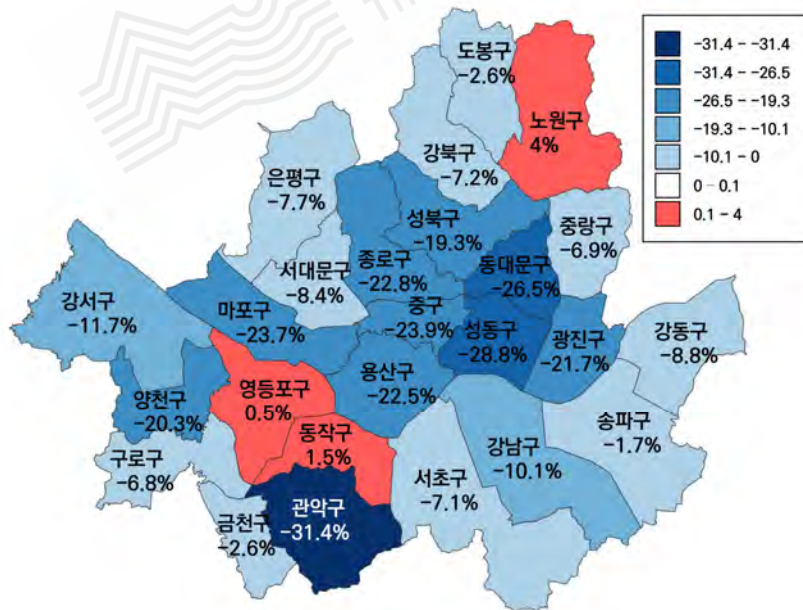
(3) 지역별 보행사고 건수와 심각도

보행사고는 지역별로도 다른 변화 양상을 보여주었다. [그림 2-5]부터 [그림 2-7]은 2015년과 2019년의 전체, 중상, 경상 보행사고의 증감률을 자치구별로 표시한 것이다. [그림 2-5]에 나타난 것처럼, 전체 보행사고는 대부분의 자치구에서 감소하였고, 감소하지 않은 3개 자치구에서도 증가율이 아주 낮게 나타났다. 이는 앞서 언급한 서울시의 보행친화 정책의 효과로 전반적인 보행사고 발생 건수가 줄어든 것과 같은 추세를 보여준다.

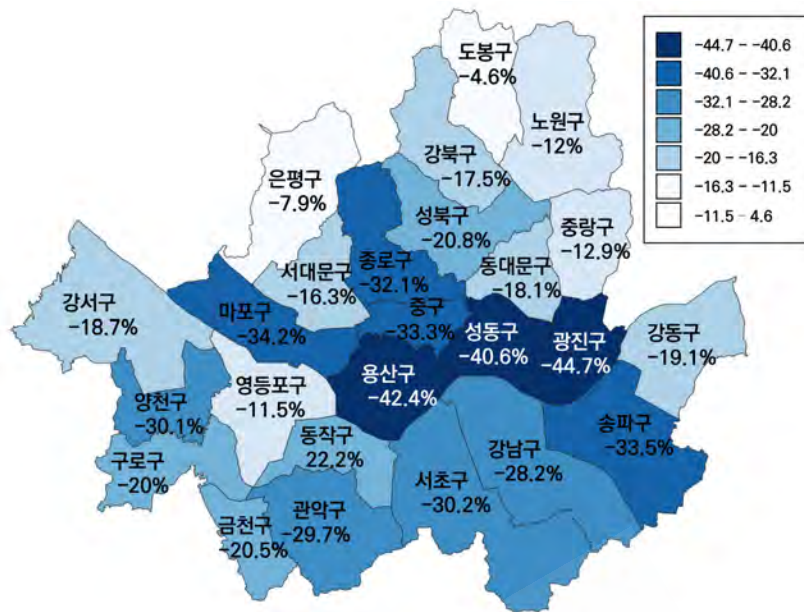
[그림 2-6]에 나타난 중상 보행사고의 경우 모든 자치구에서 감소한 것으로 나타났다. 하지만 자치구별로 감소 정도는 차이가 있었는데 광진구에서는 44.7%가 감소한 반면 도봉구에서는 4.6%만 감소하였다.

반면에 경상 보행사고는 지역별 차이가 두드러지게 나타났다. [그림 2-7]에 나타난 바와 같이, 25개 자치구 중 11개 구에서는 -34.3 ~ -3.1% 감소한 반면, 14개 구에서는 0.6 ~ 32.7% 증가하였다.

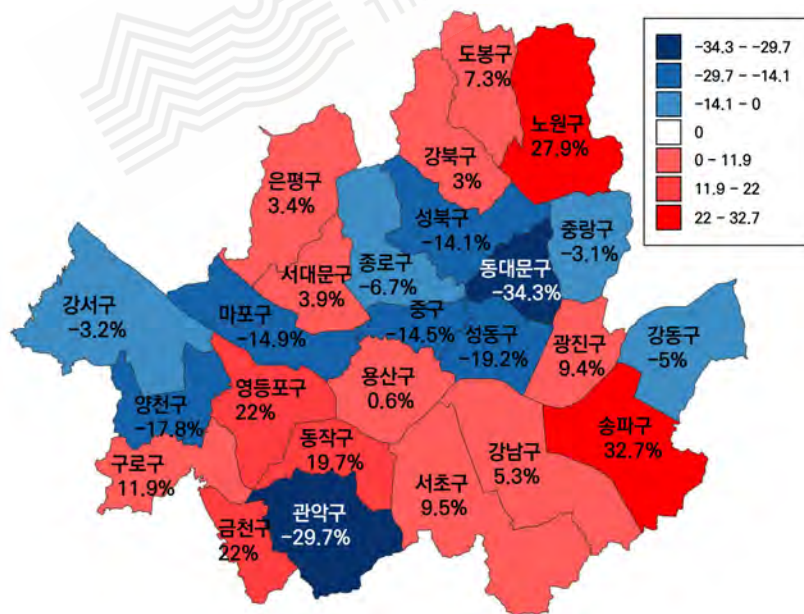
이처럼 보행사고의 추세는 지역별로 큰 차이를 나타내고 있는데, 이는 최근 보행환경과 보행행태가 지역별로 다르게 변화하고 있기 때문인 것으로 판단된다. 따라서, 지역별 보행사고를 분석하고 이에 대응하는 정책을 시행하기 위해서는 지역별 보행환경 및 보행행태를 설명할 수 있는 데이터 확보가 우선 필요할 것이다.



[그림 2-5] 서울시 구별 2015년 대비 2019년 전체 보행사고 증감률



[그림 2-6] 서울시 구별 2015년 대비 2019년 중상사고 증감률

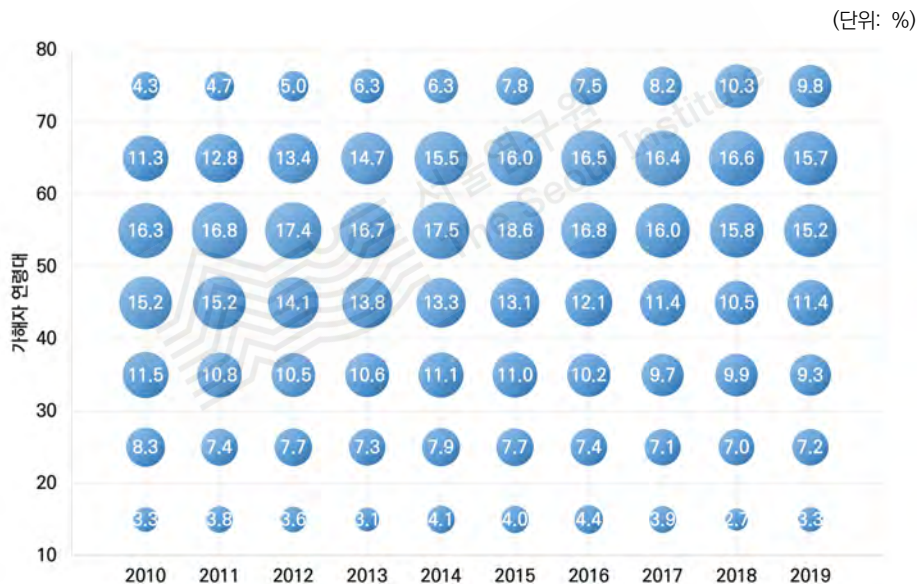


[그림 2-7] 서울시 구별 2015년 대비 2019년 경상사고 증감률

(4) 가해운전자 연령대 변화

2010년부터 2019년까지 서울시에서 발생한 보행사고 가해운전자의 연령대 비율을 분석하였다. 연령대 비율은 해당연도에 발생한 보행사고의 모든 가해운전자 수 중 해당 연령대의 가해운전자 수가 차지하는 비율로 정의하였다. 가해운전자 연령대 비율에서 9세 이하와 80세 이상은 현저히 드물어 제외하였다.

연령대 비율 변화를 연도별로 살펴보면 2015년 이후 60대와 70대의 비율이 상대적으로 증가하고 30~50대의 비율은 감소하는 것을 알 수 있다. 특히 가해운전자 중 70대가 차지하는 비율이 2010년 4.3%에서 2019년 9.8%로 2배 이상 높아졌다. 같은 기간 서울의 70대 인구수와 비율도 444,099명(4.3%)에서 684,087명(7.0%)으로 증가하였기 때문에 고령 운전자 비율이 상대적으로 높아지면서 가해운전자 중에 차지하는 비율이 증가한 것으로 판단된다.



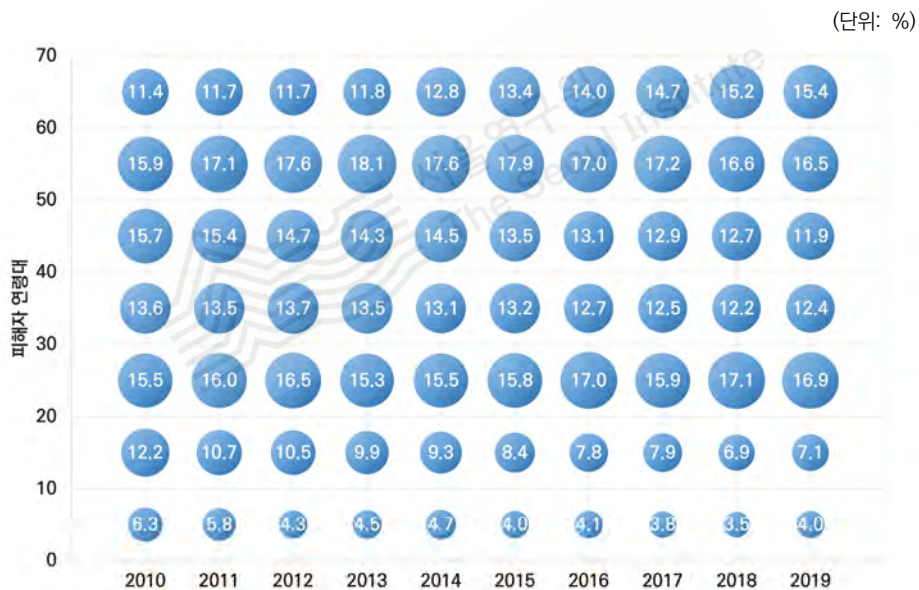
[그림 2-8] 연도별 서울시 보행사고 가해운전자 연령대 비율

(5) 피해보행자 연령대 변화

차대사람 사고의 경우 피해운전자는 피해보행자를 의미한다. 가해운전자와 같은 방법으로 피해보행자의 연령대 비율을 연도별로 살펴보았다.

가해운전자의 경우와 마찬가지로 60대 이상의 비율이 높아지고 있으며, 30대와 40대의 비율은 다소 낮아지고 있다. 10대와 9세 이하의 비율도 낮아지고 있는데, 이는 어린 이보호구역 강화 등의 정책적 효과와 출산율 저하로 인한 유년층 인구감소⁶⁾가 모두 영향을 미쳤을 것으로 판단된다.

반면에 20대가 차지하는 비율은 전반적으로 감소하지 않고 있으며 특히, 2017년 이후 15.9~17.0%로 40대 이하의 다른 연령대와 비교하여 상대적으로 높은 비율을 유지하고 있다. 최근 스마트폰 및 전동킥보드 관련 사고가 크게 증가하고 있는데, 새로운 기기를 적극적으로 이용하는 20대의 보행행태가 상대적으로 높은 피해운전자 비율에 영향을 주었을 것으로 판단된다.⁷⁾



[그림 2-9] 연도별 서울시 보행사고 피해운전자(보행자) 연령대 비율

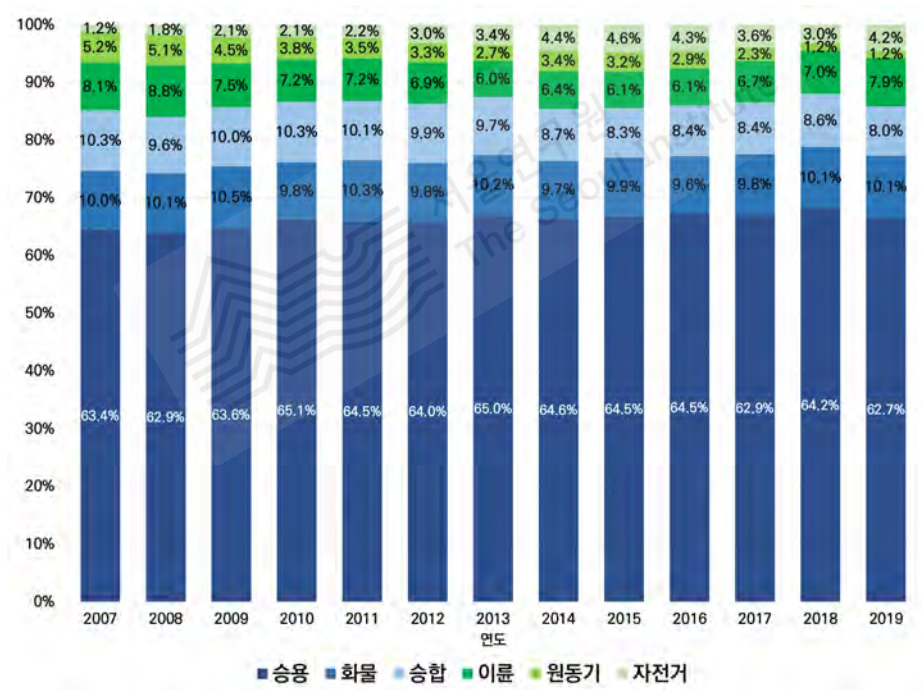
6) 10대와 9세 이하의 서울시 인구는 2010년 2,107,510명(20.4%)에서 2019년 1,482,264명(15.2%)으로 감소하였다.

7) 3장에서 분석한 설문조사 결과, 보행 중 20대의 스마트폰 사용, 이어폰 착용 등이 다른 연령대에 비해 높은 것으로 나타났다. 자세한 내용은 3장에서 제시하였다.

(6) 가해운전자 차종 변화

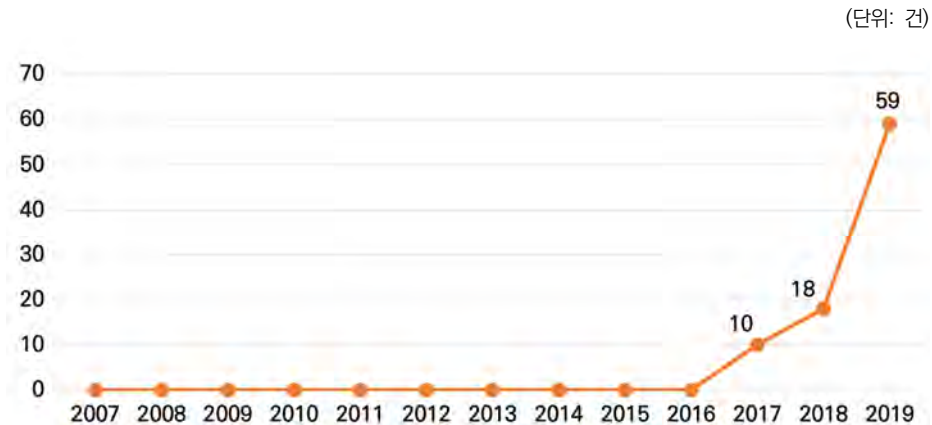
연도별 가해운전자의 차종 비율 변화를 분석한 결과, [그림 2-10]과 같이 승용차, 화물차의 비율은 큰 변화가 없었고, 승합차와 원동기 비율은 감소하였다. 반면에 이륜차와 자전거의 비율은 2017년 이후 증가하고 있다. 최근 배달문화가 확산되면서 배달이륜차의 통행이 증가하고⁸⁾, 2017년 이후 서울 전역으로 시행이 확대된 공유 자전거 따릉이의 영향인 것으로 판단된다.

그 외에 절대적인 발생 건수가 작아 [그림 2-10]에 표시되지 못한 전동킥보드 등 PM(Personal Mobility)이 유발한 보행사고의 경우, 2017년에 10건이 최초 발생하였으며 2019년에 59건으로 사고 건수가 급격히 증가하였다. 본 연구에서 구축한 보행사고 데이터는 경찰청에 접수된 사고만 포함하고 있기 때문에, PM과 관련된 실제 보행사고 건수는 더 많을 것으로 예상된다.



[그림 2-10] 연도별 서울시 보행사고 가해자 차종 비율

⁸⁾ 4장에서 스마트폰 이용 데이터를 분석한 결과 대한민국 전체의 배달 관련 애플리케이션 실행 횟수는 2016년 30.5억 건에서 2019년 163.2억 건으로 434% 증가하였다.

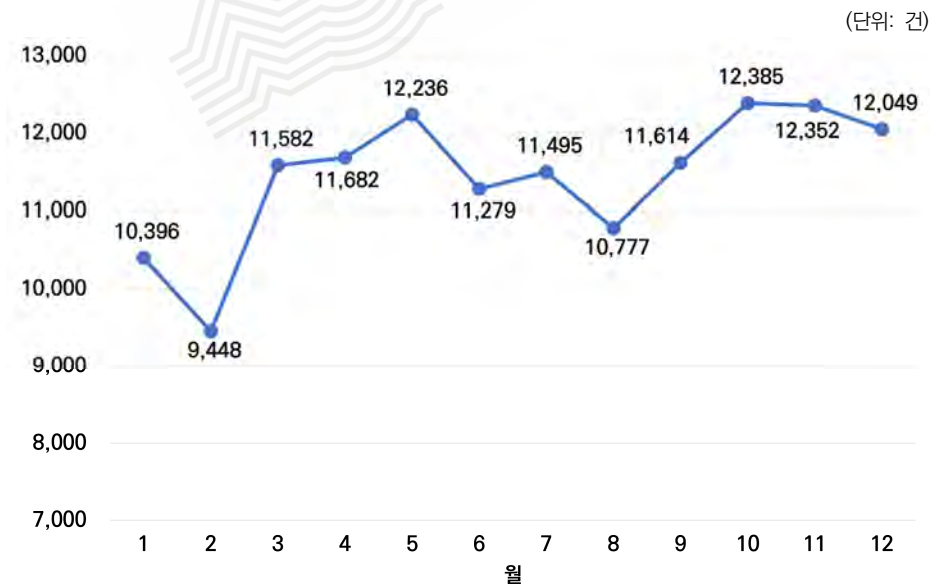


[그림 2-11] 연도별 PM이 가해운전자 차종으로 분류된 서울시 보행사고 건수

2) 연중 변화 분석

(1) 월 단위 변화 분석

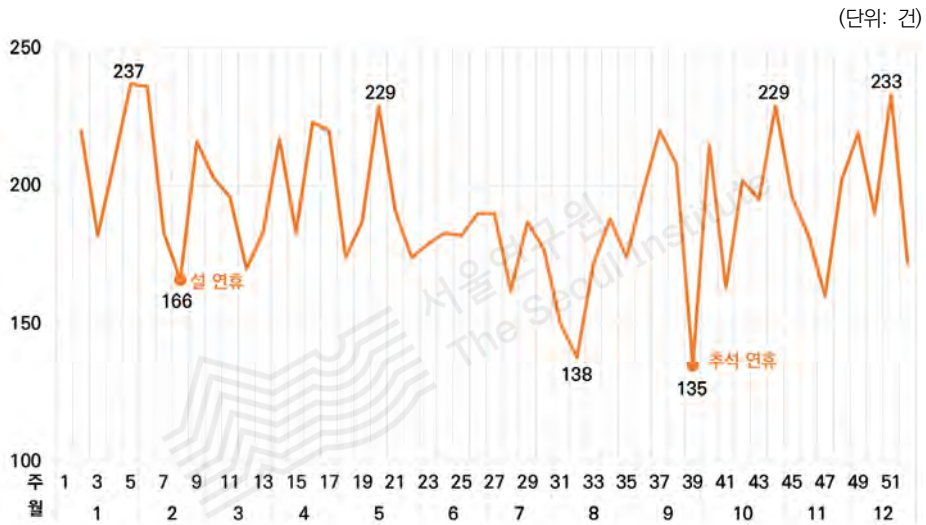
보행사고의 월별 특성을 분석하였다. 2007년부터 2019년까지 서울시에서 발생한 보행사고를 월별로 집계하면, 야외활동이 활발한 봄(3-5월)과 가을(9-11월)에 상대적으로 보행사고가 많이 발생하고, 2월에 가장 적게 발생하는 것을 알 수 있다. 더위로 야외활동이 줄어드는 8월에도 보행사고가 감소하는 것으로 나타났다.



[그림 2-12] 월별 서울시 보행사고 건수

(2) 주 단위 변화 분석

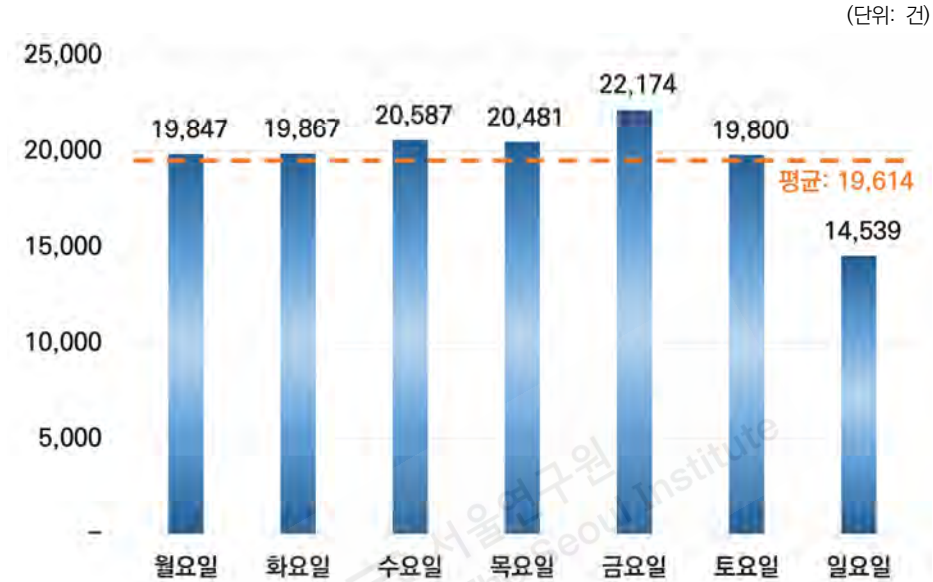
2018년 서울시에서 발생한 보행사고를 주 단위로 간격을 좁혀 상세히 분석하였다. 설 연휴가 포함된 8주차와 여름휴가가 집중되고 무더위로 야외활동이 제한되는 31-33주차, 추석 연휴가 포함된 39주차에 보행사고 발생 건수가 낮았다. 반면에, 폭설, 빙판 등으로 차량의 제동 능력이 저하되는 겨울(5, 6주차)과 야외활동이 활발한 봄(20주차)과 가을(44주차)에는 보행사고 발생 건수가 높게 나타났다. 월별 특성과 마찬가지로 연휴와 휴가 등으로 출퇴근 통행량이 감소하는 경우 보행사고가 감소하고 날씨의 영향으로 인한 야외활동량의 변화도 보행사고 발생에 영향을 미치는 것을 확인할 수 있었다.



[그림 2-13] 2018년 주차별 서울시 보행사고 건수

(3) 요일 단위 변화 분석

2007년부터 2019년까지 발생한 보행사고를 요일별로 집계하여 특성을 분석하였다. 보행사고가 가장 많이 발생하는 요일은 금요일이었고, 일요일의 보행사고 발생 건수가 가장 낮았다. 금요일과 일요일을 제외한 요일의 보행사고 건수는 유사하게 나타났다.

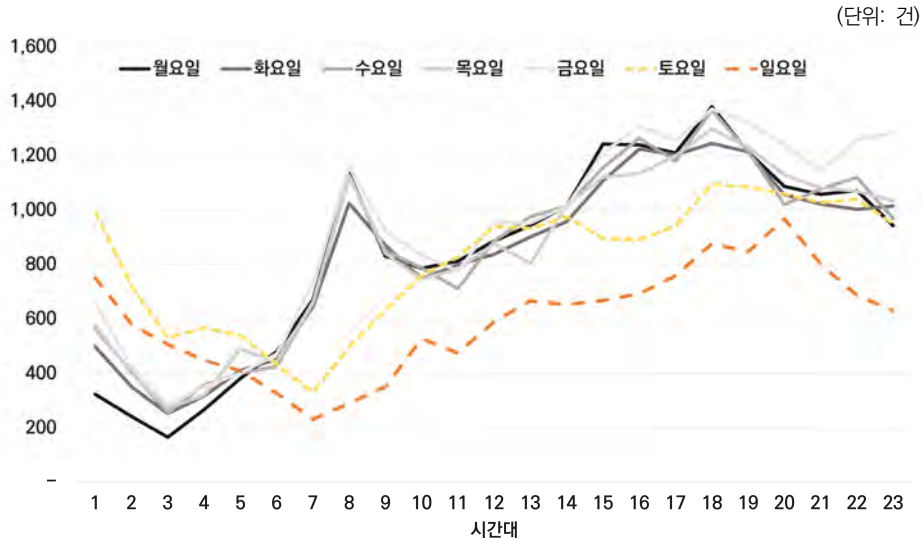


[그림 2-14] 요일별 서울시 보행사고 건수

(4) 시간 단위 변화 분석

주중과 주말의 시간대별 보행사고 발생 분포는 서로 다르게 나타났다. 주중에는 오전 8시와 오후 3-7시에 보행사고가 많이 발생하였으며, 보행사고가 가장 많이 발생하는 금요일은 오후 6시 이후 심야시간까지 보행사고 발생이 이어졌다. 주말에는 오전에 발생하는 보행사고 건수는 상대적으로 적었으나, 오후 6시 이후 보행사고 발생이 증가하기 시작하여 자정 이후 심야시간의 보행사고 발생 건수는 주중과 비교하여 높게 나타났다.

요일 및 시간대별 사고 발생 추세가 다양하게 나타나기 때문에 보행사고 감소를 위해서는 요일과 시간에 따른 맞춤형 대책이 필요할 것으로 판단된다.



[그림 2-15] 요일별 시간대별 서울시 보행사고 건수

3_서울시 보행사고 현황 분석의 시사점

보행사고 데이터를 활용하여 서울시 보행사고의 추세와 특성을 분석하였다. 전반적으로 보행사고는 감소하는 경향을 보여주었지만 최근 감소 추세가 둔화되고 있었다. 특히 심각도가 낮은 경상사고는 사망이나 중상사고 등 심각도가 높은 보행사고와 달리 발생 건수가 증가하고 있는 것을 확인할 수 있었다. 최근 스마트폰 이용 등으로 보행자의 행태가 보행사고에 취약하게 변화하고 있는데, 이러한 변화가 보행사고 증가의 새로운 원인으로 작용하고 있는 것은 아닌지 면밀한 검토가 필요한 것으로 판단된다. 2015년 대비 2019년 구별 보행사고 증감률을 분석한 결과, 지역별로 큰 차이를 보였는데, 이는 지역별로 상이한 보행환경 및 보행행태의 영향이라고 판단되며, 분석에 있어서 해당 내용을 대변하는 변수의 중요성이 높음을 나타낸다.

피해보행자의 연령대별 특성을 분석한 결과, 20대 피해보행자 비율이 상대적으로 높게 나타나는 등 일반적인 인구구조 변화로는 설명할 수 없는 경향을 보여주었다. 보행 중 돌발 상황 발생 시 신체적으로 대응이 용이한 젊은 층의 피해보행자 비율이 증가하고 있는 결과는 연령대별로 다양화되고 있는 보행행태를 반영한 연구의 필요성을 보여주고 있다.

연중 보행사고 건수 변화를 분석한 결과, 야외활동이 활발한 봄, 가을, 평일, 낮 시간대에 보행사고가 많이 발생하였다. 이렇게 월, 주, 요일, 시간대별 발생 패턴이 다르기 때문에 계절, 요일, 시간대에 따른 맞춤형 대책을 마련해야 할 것으로 판단된다. 서울시 보행사고의 발생 추세와 특성이 심각도, 지역, 시기별로 다양하게 나타나고 있기 때문에 이를 반영한 데이터 수집과 분석이 필요하고, 보행정책의 목표와 추진방안도 보다 다양화될 필요가 있는 것으로 판단된다.



03

설문조사를 통한 서울시민의 보행행태 분석



1_조사 개요

2_보행자의 보행 중 스마트폰 사용 실태

3_운전자의 운전 중 스마트폰 사용 실태

4_보행 중 스마트폰 사용억제 방법에 관한 시민인식

03. 설문조사를 통한 서울시민의 보행행태 분석

1_조사 개요

변화하는 보행행태를 분석하기 위해 온라인 설문조사를 수행하였다([표 3-1] 참조). 조사대상은 서울시에 거주하는 만 15세 이상 남녀 1,000명으로, 서울시 다섯 개 권역의 성별, 연령별 기준에 맞춰 조사대상자를 모집하였다. 최종적으로 선정된 대상자 분포는 [표 3-2]와 같다.

본 설문조사는 스마트폰 사용으로 인한 보행행태의 변화를 확인하는 것을 주요 목적으로 하였다. 우선, 보행자가 보행 중 불편을 느끼는 보행환경 요소에 대해 질문하고 그 중 스마트폰을 사용하는 보행자로 인해 겪은 불편에 대해 조사하였다. 다음으로 보행자의 보행 중 스마트폰 사용 실태를 파악하여 보행행태 변화를 추정하고, 보행 중 스마트폰 사용억제 방법에 관한 인식 조사를 수행하였다. 추가로 과거부터 이슈가 되었던 운전자의 운전 중 스마트폰 사용 실태를 조사하여 보행 중 스마트폰 이용행태와 비교하고 시사점을 도출하였다.

[표 3-1] 보행행태 분석을 위한 스마트폰 사용 실태 조사 개요

구분	내용
모집단	서울시에 거주하는 만 15세 이상 남녀
표집방법	권역별, 성별, 연령별 기준 비례할당추출
표본크기	1,000명
표본오차	무작위추출을 전제할 경우, 95% 신뢰수준에서 최대허용 표집오차는 ±3.1%p
응답(협조)률	조사요청 12,565명, 조사참여 1,349명, 조사완료 1,000명 (요청대비 8.0%, 참여대비 74.1%)
조사방법	웹조사(휴대전화 문자와 이메일을 통해 url 발송)
조사내용	보행 및 운전 중 스마트폰 사용 여부, 활동, 중요도 등
조사기간	2020년 6월 1일 ~ 6월 8일

[표 3-2] 보행행태 분석을 위한 스마트폰 사용 실태 조사 표본 구성

(단위: 명, %)

구분		인원	비율
계		1,000	100.0
성별	남자	483	48.3
	여자	517	51.7
연령	15-19세	50	5.0
	20-29세	168	16.8
	30-39세	174	17.4
	40-49세	180	18.0
	50-59세	178	17.8
	60세 이상	250	25.0
거주지역(서울)	도심권	54	5.4
	동북권	310	31.0
	동남권	210	21.0
	서북권	121	12.1
	서남권	305	30.5

2_보행자의 보행 중 스마트폰 사용 실태

1) 보행자가 보행 중 불편을 느끼는 보행환경 요소

보행자가 보행 시 불편을 느끼는 원인에 대해 조사하였다. [표 3-3]과 같이 총 5개의 선택항목을 제시하고 복수 선택이 가능하도록 하였다. 조사 결과 전체 응답자의 53.4%가 ‘좁은 보도 등 물리적 보행환경’을 선택하여, 지속적인 보행환경 개선사업이 필요함을 알 수 있었다. 특히, 최근 보행환경 개선사업을 중점적으로 시행한 도심권 거주자의 응답률은 46.3%로 상대적으로 낮은 반면, 서북권(59.5%)과 서남권(55.7%) 거주자는 높은 응답률을 보여주어 보행환경 개선사업의 시행지역을 확대해야 할 필요성을 알 수 있었다.

뒤를 이어, 전체 응답자 비율에서는 ‘배달원 등 이륜차’(50.2%)와 ‘킥보드 등 PM’(43.1%)이 높은 선택률을 나타내었다. 최근 배달문화가 발달하면서 이륜차의 보도 주행이 자주 발생하고 있는데, 이에 대한 대책이 필요한 것으로 판단된다. 또한, ‘킥보드

등 PM'은 최근 공유형 전동킥보드가 활성화된 동남권에서 54.8%의 높은 응답률을 나타내어, 보도를 이용하는 다양한 교통수단에 대한 대책을 시급히 마련할 필요가 있는 것으로 나타났다.

'다른 보행자'가 보행에 불편을 준다는 응답률은 전체 응답자의 41.5%로 나타났으며, 특히 도심권(48.1%)과 서남권(43.9%)에서 높은 선택률을 보여주었다.

연령대별로 살펴보면 10대는 다른 보행자, 킥보드, 이륜차에 대한 불편을 다른 연령대에 비해 상대적으로 적게 느끼는 것으로 나타났고, 30대를 제외한 전 연령층은 물리적 보행환경이 보행에 가장 불편을 준다고 응답했다. 30대는 킥보드와 이륜차가 가장 많이 불편을 주는 요소라고 선택하여, 다른 교통수단의 보도 이용에 대해 가장 민감하게 반응하고 있음을 알 수 있었다.

[표 3-3] 보행 중 불편을 주는 요소

(단위: %)

구분		물리적 보행환경	보도 침범 차량	타 보행자	킥보드 등 PM	배달원 등 이륜차
계		53.4	38.1	41.5	43.1	50.2
성별	남자	51.6	36.4	43.3	44.5	51.1
	여자	55.1	39.7	39.8	41.8	49.3
연령	15-19세	56.0	40.0	28.0	24.0	34.0
	20-29세	51.2	39.3	36.9	41.7	41.7
	30-39세	51.1	40.8	41.4	55.2	54.0
	40-49세	49.4	36.1	44.4	47.2	49.4
	50-59세	57.9	36.0	45.5	43.3	51.7
	60세 이상	55.6	38.0	42.4	36.4	56.0
거주지역(서울)	도심권	46.3	33.3	48.1	46.3	59.3
	동북권	49.0	38.7	40.6	40.0	45.2
	동남권	54.8	42.9	38.1	54.8	51.4
	서북권	59.5	30.6	40.5	40.5	51.2
	서남권	55.7	38.0	43.9	38.7	52.5

2) 타인의 보행 중 스마트폰 사용으로 인해 경험한 불편과 원인

스마트폰 이용과 보행환경의 영향을 구체적으로 확인하기 위하여, 보행 중 타인의 스마트폰 사용으로 인해 불편을 겪은 경험을 조사하였다. [표 3-4]와 같이 전체 응답자의 78.3%가 불편을 겪은 적이 있다고 응답하여, 보행자 대부분이 타인의 스마트폰 사용으로 인해 보행 시 불편을 경험한 것으로 나타났다.

[표 3-4] 보행 중 타인의 스마트폰 사용으로 인한 불편 경험 유무

(단위: %)

구분		예	아니오
계		78.3	21.7
성별	남자	79.7	20.3
	여자	77.0	23.0
연령	15-19세	66.0	34.0
	20-29세	70.8	29.2
	30-39세	80.5	19.5
	40-49세	82.8	17.2
	50-59세	79.2	20.8
	60세 이상	80.4	19.6
거주지역(서울)	도심권	70.4	29.6
	동북권	74.8	25.2
	동남권	78.6	21.4
	서북권	82.6	17.4
	서남권	81.3	18.7

보행 중 타인의 스마트폰 사용으로 인해 불편을 겪은 응답자를 대상으로, 타인의 어떤 행동이 가장 보행에 불편을 주었는지 조사하였다. [표 3-5]에 나타난 바와 같이, 조사 대상자의 73.9%가 스마트폰 이용으로 전방을 확인하지 않아 충돌의 위험이 있었다고 응답하였다. 18.3%의 응답자는 ‘느린 보행속도’를 선택하여 스마트폰에 몰입하며 걷는 행위가 동일방향 보행 시에도 영향을 주는 것으로 나타났다. 특히, 20대와 30대에서 ‘느린 보행속도’를 선택한 비율이 22.7%와 24.3%로 상대적으로 높게 나타났다. 소음 유발을 선택한 응답자는 2.4%로 낮아 대부분의 보행자가 타인의 스마트폰 사용으로 인한 물리적 충돌과 방해를 더 불편하게 느끼는 것을 알 수 있었다.

[표 3-5] 보행 중 스마트폰 사용이 타인에게 불편을 주는 이유

(단위: %)

구분		느린 보행속도	사선으로 걸는 행위	전방확인하지 않아 충돌 위험	소음 유발	기타
계		18.3	5.1	73.9	2.4	0.3
성별	남자	18.7	6.5	72.7	1.8	0.3
	여자	17.8	3.8	75.1	3.0	0.3
연령	15~19세	18.2	3.0	78.8	0.0	0.0
	20~29세	22.7	4.2	71.4	0.8	0.8
	30~39세	24.3	11.4	62.9	0.7	0.7
	40~49세	12.8	5.4	79.9	2.0	0.0
	50~59세	14.9	2.8	78.7	3.5	0.0
	60세 이상	17.9	3.0	74.6	4.5	0.0
거주지역(서울)	도심권	13.2	5.3	81.6	0.0	0.0
	동북권	16.8	4.3	76.3	2.6	0.0
	동남권	15.2	7.9	74.5	1.8	0.6
	서북권	26.0	3.0	69.0	1.0	1.0
	서남권	19.4	4.8	72.2	3.6	0.0

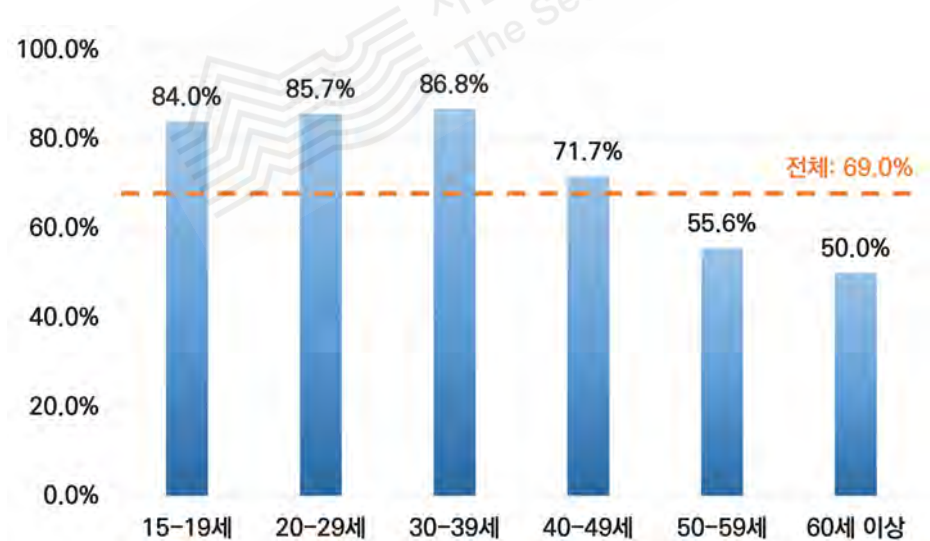
3) 보행 중 스마트폰 사용 여부

보행자의 스마트폰 사용 실태를 파악하기 위해 보행 중 스마트폰 사용 여부를 물었다. 조사대상자의 69.0%가 보행 중 스마트폰을 사용한다고 응답하였다. 연령대별로는 30대 이하의 사용률이 84.0~86.8%로 높아 30대 이하 대부분은 보행 중 스마트폰을 사용하고 있는 것으로 나타났다([그림 3-1] 참조). 반면에 50대 이상의 보행 중 스마트폰 사용률은 50.0~55.6%로 상대적으로 낮았다. 거주 지역별로는 동남권의 사용률이 71.9%로 가장 높았고, 동북권의 사용률이 66.5%로 가장 낮게 나타났다.

[표 3-6] 보행 중 스마트폰 사용 여부

(단위: %)

구분		예	아니오
계		69.0	31.0
성별	남자	69.2	30.8
	여자	68.9	31.1
연령	15-19세	84.0	16.0
	20-29세	85.7	14.3
	30-39세	86.8	13.2
	40-49세	71.7	28.3
	50-59세	55.6	44.4
	60세 이상	50.0	50.0
거주지역(서울)	도심권	68.5	31.5
	동북권	66.5	33.5
	동남권	71.9	28.1
	서북권	69.4	30.6
	서남권	69.5	30.5



[그림 3-1] 연령대별 보행 중 스마트폰 사용률

4) 보행 중 스마트폰을 사용하는 이유

보행 중 스마트폰을 사용한다고 대답한 응답자를 대상으로 보행 중 스마트폰으로 어떤 활동을 하는지 조사하였다. [표 3-7]에 나타난 것과 같이 총 12개의 활동을 제시하고 중복응답이 가능하도록 하였다.

제시된 12개의 활동 중 ‘통화’(70.1%), ‘메신저’(68.8%), ‘지도’(63.2%)가 60% 이상의 높은 응답률을 나타내었고, ‘음악’(48.7%)과 ‘동영상’(44.9%)도 응답률이 높았다. ‘뉴스’(36.7%), ‘소셜’(36.1%), ‘쇼핑’(34.8%), ‘금융’(33.8%)이 30% 이상의 응답률을 나타내었고, ‘게임’(23.6%)과 ‘만화’(18.8%), ‘배달주문’(17.7%)이 뒤를 이었다.

성별로는 남성과 여성이 서로 다른 결과를 보여준다. 남성의 경우 ‘동영상’(51.8%), ‘게임’(30.8%), ‘뉴스’(41.3%), ‘금융’(38.9%)의 응답률이 여성보다 9.0~14.0% 높았다. 여성의 경우 가장 높은 응답률을 보인 활동은 ‘메신저’(68.8%)로 ‘통화’(67.1%)보다 높게 나타났다. ‘음악’(50.8%), ‘소셜’(37.1%), ‘쇼핑’(36.2%)도 상대적으로 높은 응답률이 나타났고, 남성의 응답률이 높았던 ‘동영상’(38.5%)과 ‘게임’(16.9%)은 상대적으로 낮은 응답률을 보여주었다.

연령대별로도 보행 중 이용하는 스마트폰 활동이 서로 다르게 나타났다. 40대 이상에서는 ‘통화’가 가장 응답률이 높은 활동이었지만, 20대와 30대는 ‘메신저’를 가장 많이 이용하였고, 10대는 ‘소셜’이 69.0%로 가장 높은 응답률을 보였다. 30대 이하에서는 ‘게임’, ‘동영상’, ‘음악’, ‘배달주문’의 응답률이 상대적으로 높은 반면, 40대 이상에서는 ‘뉴스’, ‘금융’이 다른 연령대에 비해 높게 나타났다. ‘쇼핑’은 30대와 40대가 다른 연령대보다 높은 응답률을 나타내었다.

성별 및 연령대별로 보행 중 이용하는 스마트폰 활동이 상이하여 보행자의 보행행태도 과거와 비교하여 다양화되었을 것으로 추정된다. 특히, 지속적으로 스마트폰에 몰입하게 하는 ‘동영상’, ‘게임’, ‘메신저’ 등의 활동을 하는 보행자 비율이 높을수록 보행사고 발생 가능성이 증가할 것으로 예상되어 이에 대한 대책을 마련해야 할 것으로 판단된다.

[표 3-7] 보행 중 스마트폰 사용자의 사용 목적

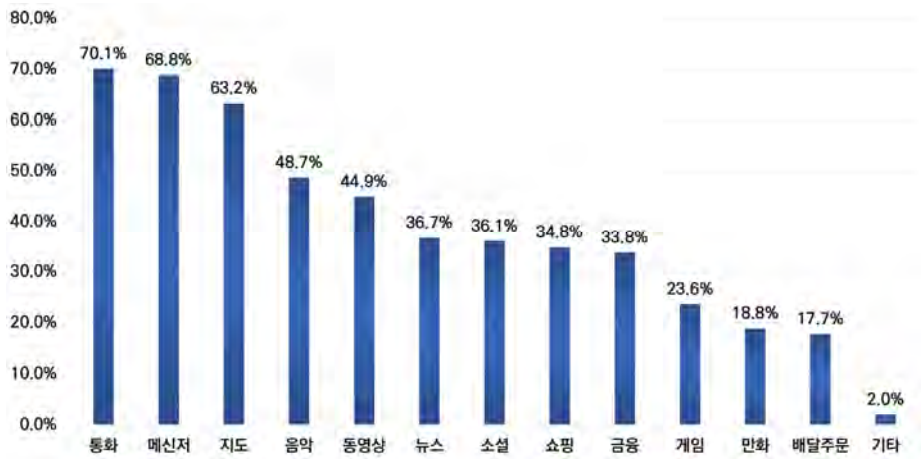
(단위: %)

구분		통화	지도	쇼핑	게임	동영상	음악
계		70.1	63.2	34.8	23.6	44.9	48.7
성별	남자	73.4	62.9	33.2	30.8	51.8	46.4
	여자	67.1	63.5	36.2	16.9	38.5	50.8
연령	15-19세	50.0	26.2	21.4	38.1	52.4	59.5
	20-29세	56.3	58.3	34.0	29.9	53.5	66.7
	30-39세	69.5	68.2	43.0	31.8	50.3	55.6
	40-49세	70.5	61.2	42.6	22.5	41.9	48.8
	50-59세	82.8	71.7	29.3	13.1	37.4	37.4
	60세 이상	83.2	70.4	26.4	11.2	35.2	24.8
거주지역(서울)	도심권	73.0	67.6	35.1	21.6	51.4	40.5
	동북권	71.4	60.7	33.5	23.3	43.7	50.0
	동남권	70.2	70.9	36.4	22.5	47.7	51.0
	서북권	65.5	56.0	31.0	25.0	42.9	44.0
	서남권	70.3	62.3	36.3	24.5	43.9	49.1

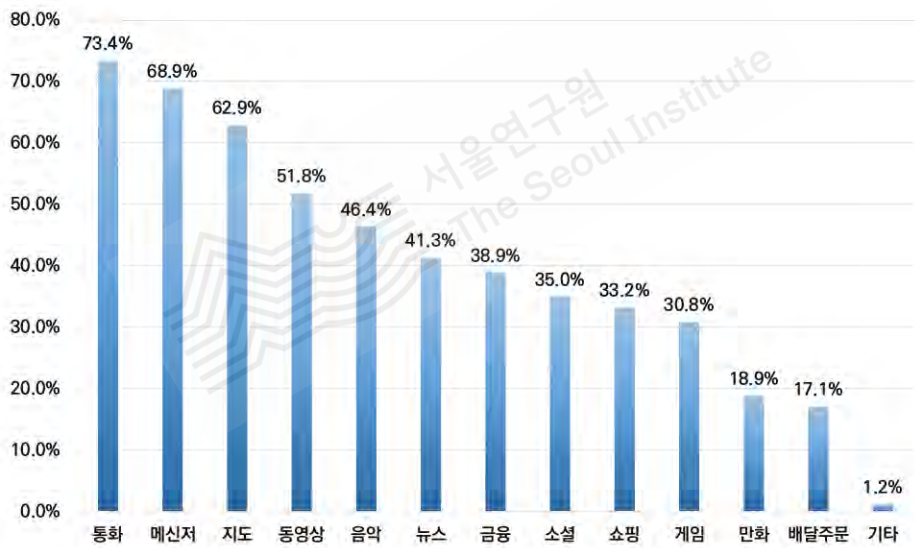
[표 3-7 계속] 보행 중 스마트폰 사용자의 사용 목적

(단위: %)

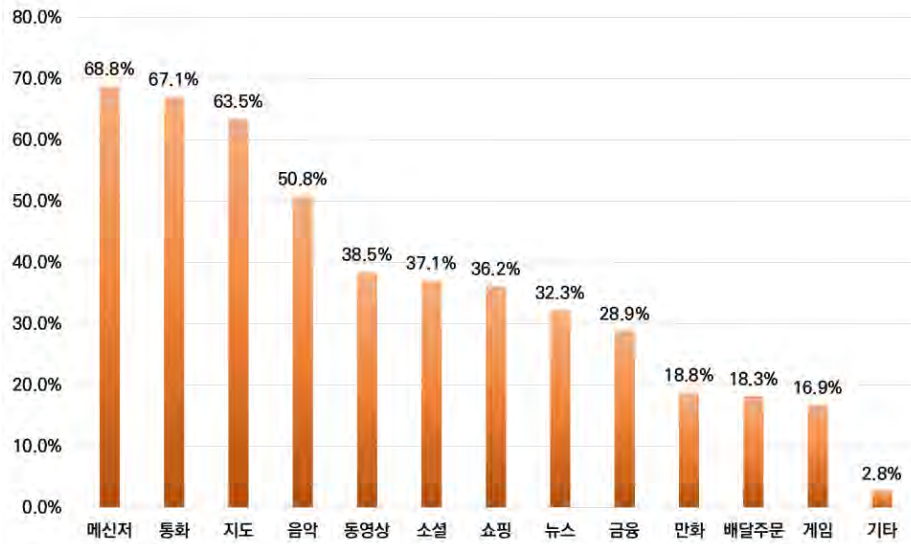
구분		메신저	소셜	뉴스	만화	배달	금융
계		68.8	36.1	36.7	18.8	17.7	33.8
성별	남자	68.9	35.0	41.3	18.9	17.1	38.9
	여자	68.8	37.1	32.3	18.8	18.3	28.9
연령	15-19세	61.9	69.0	9.5	35.7	7.1	2.4
	20-29세	75.0	50.0	20.8	28.5	20.1	26.4
	30-39세	76.8	45.0	40.4	27.8	29.1	38.4
	40-49세	68.2	24.0	51.9	17.8	18.6	40.3
	50-59세	64.6	27.3	40.4	7.1	11.1	41.4
	60세 이상	58.4	17.6	40.8	1.6	8.8	34.4
거주지역(서울)	도심권	64.9	35.1	51.4	27.0	13.5	27.0
	동북권	65.5	36.9	31.6	15.0	16.5	35.0
	동남권	74.2	41.1	43.0	18.5	22.5	41.1
	서북권	69.0	29.8	35.7	19.0	10.7	26.2
	서남권	68.9	34.4	34.9	21.2	18.9	31.6



[그림 3-2] 보행 중 스마트폰 사용자의 사용 목적(전체)



[그림 3-3] 보행 중 스마트폰 사용자의 사용 목적(남성)



[그림 3-4] 보행 중 스마트폰 사용자의 사용 목적(여성)

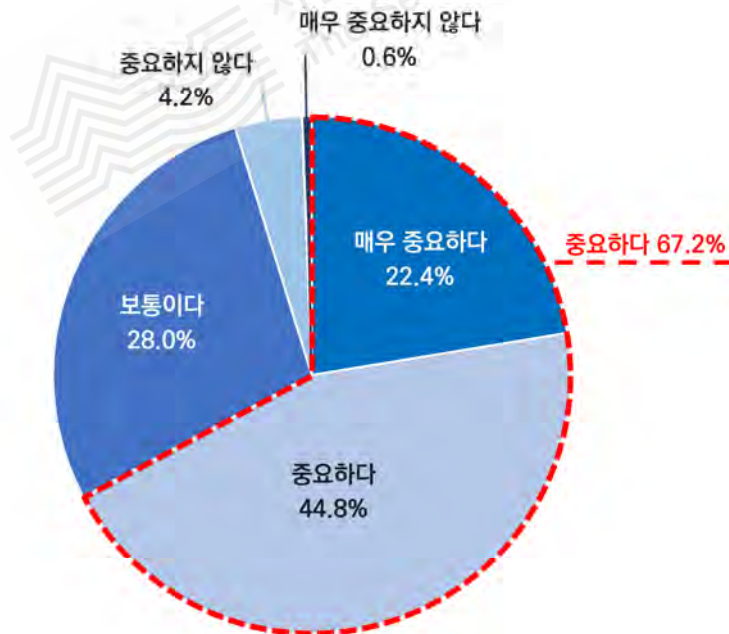
5) 보행 중 스마트폰을 이용한 활동의 중요성

보행 중 스마트폰을 사용하는 조사대상자에게 보행 중 스마트폰 활동의 중요성을 물었다. [표 3-8]에 나타난 것처럼 응답자의 67.2%는 중요도가 높다고 하였고, 28.0%는 보통으로 대답하였다. 중요도가 낮다고 대답한 응답자는 4.8%로 대부분의 보행자들이 보행 중 스마트폰 활동을 필요한 활동으로 인식하고 있었다. 이러한 결과를 통해 향후에도 보행자의 스마트폰 사용은 크게 감소하지 않을 것으로 예상할 수 있다. 남성과 여성의 응답률은 유사하게 나타나 성별에 따른 유의미한 특성은 나타나지 않았다. 반면에 연령별 특성에서는 50대의 스마트폰 활동이 중요하다는 응답률이 75.8%로 가장 높았고, 10대는 54.8%로 가장 낮았다. 특히, 10대의 11.9%가 스마트폰 활동이 중요하지 않다고 응답하였는데, 이는 10대가 보행 중 주로 이용하는 스마트폰 활동이 '소셜', '동영상', '음악', '게임' 등 엔터테인먼트적 요소가 높기 때문인 것으로 보인다.

[표 3-8] 보행 중 이용하는 스마트폰 활동의 중요성

(단위: %)

구분		중요하다		보통이다	중요하지 않다	
		매우 중요하다	중요하다	보통이다	중요하지 않다	매우 중요하지 않다
계		22.4	44.8	28.0	4.2	0.6
성별	남자	21.3	47.3	26.6	4.2	0.6
	여자	23.6	42.4	29.2	4.2	0.6
연령	15-19세	9.5	45.2	33.3	9.5	2.4
	20-29세	25.7	43.1	27.1	3.5	0.7
	30-39세	25.2	43.7	23.8	6.0	1.3
	40-49세	23.3	41.9	33.3	1.6	0.0
	50-59세	16.2	59.6	20.2	4.0	0.0
	60세 이상	24.0	39.2	32.8	4.0	0.0
거주지역(서울)	도심권	10.8	51.4	37.8	0.0	0.0
	동북권	24.8	45.6	24.3	5.3	0.0
	동남권	23.8	46.4	25.2	4.0	0.7
	서북권	14.3	54.8	28.6	2.4	0.0
	서남권	24.5	37.7	31.6	4.7	1.4



[그림 3-5] 보행 중 이용하는 스마트폰 활동의 중요성

6) 보행 중 스마트폰 사용자의 이어폰 착용 여부

보행 중 스마트폰 활동 시 이용하는 부가장치의 대표적 사례로 이어폰 착용 여부를 조사하였다. 응답자의 48.0%가 양쪽 이어폰을, 21.6%는 한쪽 이어폰을 사용한다고 응답하였고, 30.4%는 사용하지 않는다고 하였다. 보행자의 약 70%가 이어폰을 착용하여 주변 소리에 의한 상황 판단이 저하되고 보행사고 가능성이 높아질 수 있을 것으로 예상된다.

연령대별 특성을 살펴보면 60세 이상은 절반 이상인 54.4%가 이어폰을 착용하지 않지만, 20대는 16.7%만이 이어폰을 착용하지 않는다고 응답하였다. 양쪽 이어폰을 모두 착용하는 20대의 비율도 65.3%로 전체 연령대 중 가장 높았는데, 2장에서 분석한 서울시 보행사고 피해보행자의 20대 비율이 상대적으로 높은 것도 이어폰 사용 등으로 청각 인지능력이 저하되는 20대의 보행행태와 관련이 있어 보인다.

[표 3-9] 보행 중 스마트폰 사용자의 이어폰 착용 여부

(단위: %)

구분		양쪽 착용	한쪽 착용	미착용
계		48.0	21.6	30.4
성별	남자	50.0	18.9	31.1
	여자	46.1	24.2	29.8
연령	15-19세	38.1	35.7	26.2
	20-29세	65.3	18.1	16.7
	30-39세	60.3	15.2	24.5
	40-49세	47.3	24.0	28.7
	50-59세	36.4	30.3	33.3
	60세 이상	26.4	19.2	54.4
거주지역(서울)	도심권	40.5	29.7	29.7
	동북권	51.9	19.9	28.2
	동남권	50.3	22.5	27.2
	서북권	46.4	16.7	36.9
	서남권	44.3	23.1	32.5

3_운전자의 운전 중 스마트폰 사용 실태

운전 중 스마트폰 사용 문제는 휴대폰이 일반화되기 시작한 20여 년 전부터 이슈가 되었고, 해당 문제로 인해 발생하는 사고의 심각도가 높아 개선 노력이 활발히 이루어 졌다. 따라서 보행 중 스마트폰 사용 실태와 비교하고 시사점을 도출하기 위하여 운전 중 스마트폰 사용 실태를 조사하였다. 조사대상은 [표 3-12]와 같이 보행 중 스마트폰 사용 실태 조사에서 운전 경험이 있다고 응답한 76%(760명)를 대상으로 하였다.

[표 3-10] 운전 경험 유무

(단위: %)

구분		예	아니오
계		76.0	24.0
성별	남자	87.2	12.8
	여자	65.6	34.4
연령	15-19세	6.0	94.0
	20-29세	57.7	42.3
	30-39세	82.2	17.8
	40-49세	88.9	11.1
	50-59세	83.7	16.3
	60세 이상	83.2	16.8
거주지역(서울)	도심권	90.7	9.3
	동북권	71.6	28.4
	동남권	83.3	16.7
	서북권	65.3	34.7
	서남권	77.0	23.0

1) 운전 중 불편을 주는 요소

운전자에게 운전 중 불편을 주는 요소에 대해 복수로 응답하도록 하였다. 그 결과 운전 중 불편을 주는 요소로 ‘배달원 등 이륜차’(65.1%)를 가장 많이 선택하였고, ‘다른 자동차’(43.9%), ‘좁은 차도 등 물리적 환경’(40.0%)이 뒤를 이었다. ‘키패드 등 PM’(38.2%)과 ‘보행자’(32.0%)도 30% 이상의 응답자가 선택했다. 배달문화 활성화로 이륜차 이용이 증가하고 공유 전동키패드 도입으로 다양한 교통수단이 도로를 함께 사용하면서 운전자들이 불편을 겪고 있음을 알 수 있었다.

[표 3-11] 운전 중 불편을 주는 요소

(단위: %)

구분		물리적 환경	다른 차량	보행자	키패드 등 PM	배달원 등 이륜차
계		40.0	43.9	32.0	38.2	65.1
성별	남자	37.8	46.6	29.0	36.1	64.1
	여자	42.8	40.7	35.7	40.7	66.4
연령	20-29세	45.4	50.5	38.1	36.1	56.7
	30-39세	36.4	55.2	28.7	36.4	64.3
	40-49세	32.5	44.4	33.8	38.8	66.9
	50-59세	38.3	44.3	28.9	43.6	66.4
	60세 이상	46.6	33.2	32.7	36.5	67.8
거주지역(서울)	도심권	40.8	51.0	38.8	38.8	59.2
	동북권	38.7	42.3	30.6	36.0	67.1
	동남권	42.3	44.6	26.9	42.9	61.1
	서북권	34.2	46.8	35.4	41.8	69.6
	서남권	41.3	42.6	34.5	35.3	66.0

2) 운전자의 운전 중 스마트폰 이용 여부

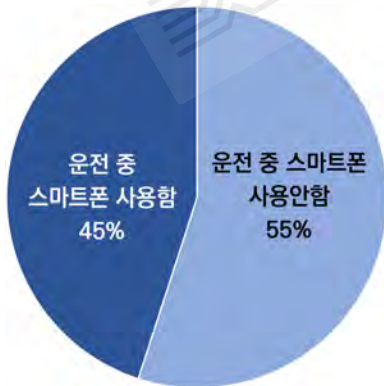
운전 중 스마트폰 사용 여부를 물었다. 응답자의 36.7%가 운전 중 스마트폰을 사용한다고 대답하여 보행 중 스마트폰 사용 수치(69.0%)보다 현저히 낮은 것으로 나타났다. 남성(40.4%)이 여성(32.2%)보다 운전 중 스마트폰 사용률이 높았고, 30대, 40대, 50대의 사용률이 상대적으로 높게 나타났다.

보행 중 스마트폰 사용 여부와 운전 중 스마트폰의 사용 여부도 연관성을 나타냈다. 보행 중 스마트폰을 사용하는 경우에는 45.0%의 비율로 운전 중에도 스마트폰을 사용하였지만, 보행 중 스마트폰을 사용하지 않는 경우는 18.2%만이 운전 중 스마트폰을 사용하였다.

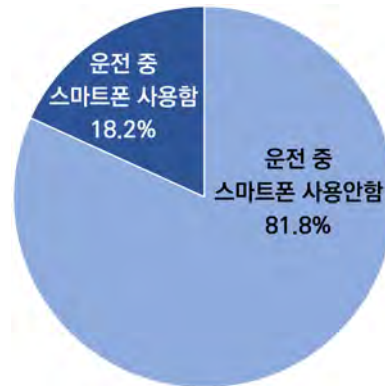
[표 3-12] 운전 중 스마트폰 사용 여부

(단위: %)

구분		예	아니오
계		36.7	63.3
성별	남자	40.4	59.6
	여자	32.2	67.8
연령	15-19세	0.0	100.0
	20-29세	32.0	68.0
	30-39세	43.4	56.6
	40-49세	41.3	58.8
	50-59세	36.9	63.1
	60세 이상	31.3	68.8
거주지역(서울)	도심권	34.7	65.3
	동북권	40.1	59.9
	동남권	32.6	67.4
	서북권	32.9	67.1
	서남권	38.3	61.7
보행 중 스마트폰 사용 여부	사용함	45.0	55.0
	사용하지 않음	18.2	81.8



[그림 3-6] 보행 중 스마트폰 사용자의 운전 중 스마트폰 사용 비율



[그림 3-7] 보행 중 스마트폰 미사용자의 운전 중 스마트폰 사용 비율

3) 운전 중 스마트폰 사용자의 사용 목적

보행 중 스마트폰 활동 조사와 동일한 12개의 활동을 제시하고 운전 중 스마트폰을 이용한 활동을 조사하였다. 중복선택이 가능하도록 설계하여 조사한 결과, ‘지도’의 응답률이 86%로 가장 높았고, ‘통화’(72.8%), ‘음악’(35.8%), ‘메신저’(26.5%)의 순으로 나타났다. 보행 중 스마트폰 활동 조사 결과와 비교하면, [그림 3-11]에 나타난 것처럼 ‘지도’, ‘통화’, ‘음악’을 제외한 모든 활동의 응답률이 크게 감소하였다.

연령대별 특성에서는, 운전 중 주의 산만이 발생하기 쉬운 ‘동영상’, ‘쇼핑’, ‘게임’, ‘뉴스’의 20대 응답률이 다른 연령대에 비해 현저히 높게 나타났다. 2장 [그림 2-8]에 나타난 것처럼 최근 서울시 보행사고 가해운전자의 20대 비율이 소폭 증가하고 있는데, 주의 산만을 유발하는 운전행태와 관련이 있을 것으로 예상된다. 60세 이상에서는 ‘뉴스’의 응답률이 20대 다음으로 높게 나타나 서울시 보행사고 가해운전자 비율이 증가하는 60세 이상의 운전행태도 개선이 필요한 것으로 판단된다.

[표 3-13] 운전 중 스마트폰 사용자의 사용 목적

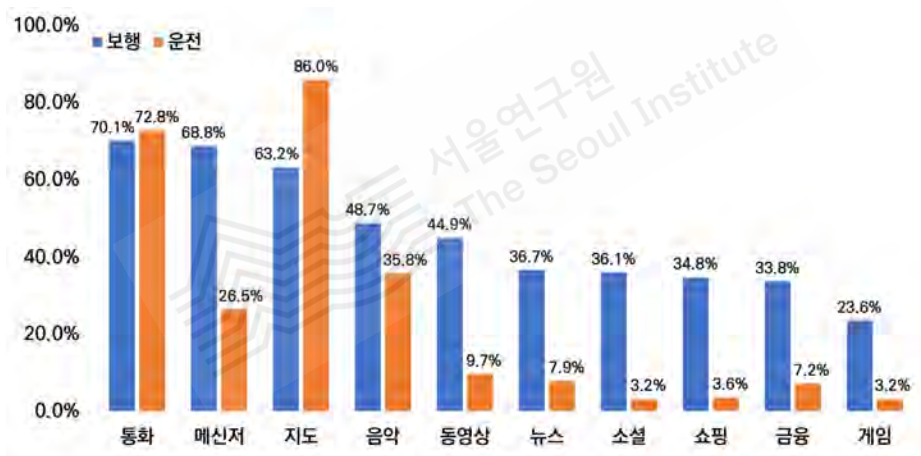
(단위: %)

구분		통화	지도	쇼핑	게임	동영상	음악
계		72.8	86.0	3.6	3.2	9.7	35.8
성별	남자	74.1	87.6	4.7	3.5	11.2	35.9
	여자	70.6	83.5	1.8	2.8	7.3	35.8
연령	20-29세	67.7	80.6	16.1	9.7	19.4	45.2
	30-39세	58.1	85.5	1.6	6.5	12.9	50.0
	40-49세	72.7	89.4	0.0	3.0	12.1	28.8
	50-59세	76.4	85.5	1.8	0.0	5.5	32.7
	60세 이상	86.2	86.2	4.6	0.0	3.1	27.7
거주지역(서울)	도심권	76.5	88.2	0.0	11.8	11.8	35.3
	동북권	70.8	89.9	2.2	0.0	10.1	42.7
	동남권	70.2	77.2	7.0	10.5	12.3	36.8
	서북권	69.2	84.6	3.8	0.0	7.7	26.9
	서남권	76.7	87.8	3.3	1.1	7.8	31.1

[표 3-13 계속] 운전 중 스마트폰 사용자의 사용 목적

(단위: %)

구분		메신저	소셜	뉴스	만화	배달	금융
계		26.5	3.2	7.9	1.4	2.5	7.2
성별	남자	30.0	3.5	8.8	2.4	3.5	8.8
	여자	21.1	2.8	6.4	0.0	0.9	4.6
연령	20~29세	25.8	6.5	12.9	3.2	3.2	6.5
	30~39세	24.2	4.8	4.8	1.6	1.6	8.1
	40~49세	33.3	3.0	6.1	3.0	3.0	6.1
	50~59세	23.6	0.0	7.3	0.0	1.8	7.3
	60세 이상	24.6	3.1	10.8	0.0	3.1	7.7
거주지역(서울)	도심권	23.5	0.0	5.9	0.0	0.0	0.0
	동북권	25.8	0.0	4.5	0.0	0.0	7.9
	동남권	28.1	8.8	10.5	7.0	8.8	10.5
	서북권	34.6	3.8	11.5	0.0	0.0	11.5
	서남권	24.4	3.3	8.9	0.0	2.2	4.4



[그림 3-8] 보행 및 운전 중 스마트폰 사용자의 사용 목적

4) 운전 중 스마트폰 사용자의 부가장치 이용 여부

2000년대 초반부터 운전 중 휴대폰 이용으로 인한 사고가 증가하면서 이를 보완하기 위한 다양한 기술이 개발되었다. 운전 시 블루투스 연결, 음성인식 등 부가기능 사용 행태에 대해 복수 선택이 가능하도록 설계하여 조사한 결과, 조사대상자의 79.9%가 핸드프리 등 블루투스 연결을 사용한다고 대답하였다. 최근 활성화되기 시작한 스마트폰 또는 차량의 음성인식 기능을 사용하는 응답자도 25.1%였고, 스마트폰을 운전

자모드로 전환하는 운전자도 20.4%였다.

성별 특성을 살펴보면, 블루투스 사용률은 남성이 더 높았고, 음성인식과 운전자모드 전환 기능 사용률은 여성의 비율이 더 높았다. 부가기술을 사용하지 않는 응답자는 전체의 10.0%였는데, 여성의 비율이 남성에 비해 조금 더 높은 것으로 나타났다. 핸즈프리 기술은 모든 연령대에서 많이 사용하고 있으나 음성인식, 애플 카플레이나 안드로이드 오토와 같은 신기술은 50대 이상의 사용률이 낮게 나타났다.

[표 3-14] 운전 중 스마트폰 사용자의 부가장치 이용 여부

(단위: %)

구분		블루투스	음성인식	운전자모드 전환	애플 카플레이, 안드로이드 오토	없음
계		79.9	25.1	20.4	6.8	10.0
성별	남자	84.1	22.4	17.6	7.6	8.2
	여자	73.4	29.4	24.8	5.5	12.8
연령	20-29세	83.9	35.5	16.1	12.9	3.2
	30-39세	79.0	27.4	14.5	9.7	11.3
	40-49세	86.4	34.8	24.2	9.1	6.1
	50-59세	80.0	21.8	21.8	0.0	9.1
	60세 이상	72.3	10.8	23.1	4.6	16.9
거주지역(서울)	도심권	76.5	11.8	11.8	11.8	17.6
	동북권	80.9	22.5	16.9	6.7	11.2
	동남권	84.2	28.1	21.1	8.8	8.8
	서북권	84.6	26.9	23.1	3.8	7.7
	서남권	75.6	27.8	24.4	5.6	8.9

5) 운전 중 스마트폰 사용 실태 조사의 시사점

스마트폰을 사용하는 운전자는 여전히 존재하지만, 그 비율은 보행자에 비해 현저히 낮았으며, 스마트폰을 사용하는 경우에도 운전에는 필요한 ‘지도’, 상대적으로 스마트폰 몰입도가 낮은 ‘통화’, ‘음악’ 활동으로 한정되었다. 또한, 다양한 부가장치와 첨단기능을 이용하여 운전 중 스마트폰 이용으로 인한 주의 산만이 발생하지 않도록 하고 있어 보행 중 스마트폰 이용과 차별화된 결과를 보여주었다.

이러한 차이점은 운전 중 휴대폰 사용으로 증가한 교통사고에 대응하여 운전자 의식 개선 캠페인, 단속 등 제도적 조치, 신기술을 이용한 적극적인 운전 보조 장치 개발 등의 효과로 판단된다.

4_보행 중 스마트폰 사용억제 방법에 관한 시민인식

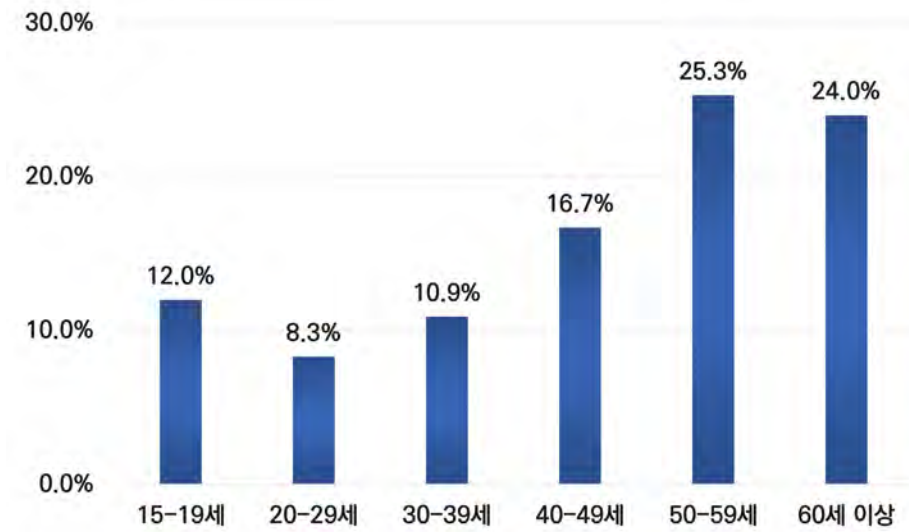
‘보행 시 스마트폰 사용 금지 제도 마련’ 같은 스마트폰 사용 억제 방법에 관해 시민 인식을 조사하였다. 조사 결과, ‘캠페인 등 시민인식 개선’이 36.5%로 가장 높은 선호도를 보였고, ‘보행 안전을 위한 스마트폰 경고 기술 개발’(25.3%)과 ‘사고 유발 시 보행자의 과실 비율 부과’(20.5%)가 뒤를 이었다. 보행 중 스마트폰 이용을 제도적으로 제한하는 ‘보행 시 스마트폰 이용 금지 제도 마련’은 17.4%의 선호도를 나타내었다. 연령대별로 살펴보면 50대 이상에서는 ‘보행 시 스마트폰 이용 금지 제도 마련’의 선호도가 다른 연령대에 비해 높았고, 30대 이하에서는 ‘사고 유발 시 보행자의 과실 비율 부과’ 선호도가 40대 이상보다 최대 3배 이상 높게 나타났다.

전반적으로는 캠페인 등 소극적인 개선방안을 선호하지만, 고령층은 제도를 통한 적극적인 개선방안을 상대적으로 더 선호하고, 젊은 층에서는 사고를 유발하는 보행자 개인에게 책임을 부과하는 방식을 더 선호하는 것으로 나타났다.

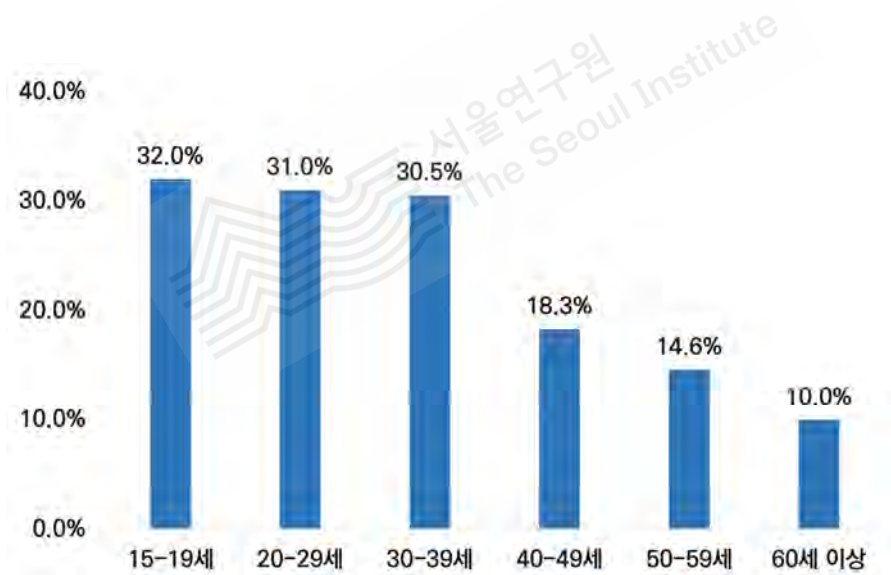
[표 3-15] 보행 중 스마트폰 사용억제 방법에 대한 선호도

(단위: %)

구분		보행 시 스마트폰 사용 금지 제도 마련	사고 유발 시 과실 비율 부과	캠페인 등 시민인식 개선	보행 안전을 위한 스마트폰 경고 기술개발	기타
계		17.4	20.5	36.5	25.3	0.3
성별	남자	17.0	19.7	40.4	22.8	0.2
	여자	17.8	21.3	32.9	27.7	0.4
연령	15-19세	12.0	32.0	30.0	24.0	2.0
	20-29세	8.3	31.0	40.5	19.0	1.2
	30-39세	10.9	30.5	35.6	23.0	0.0
	40-49세	16.7	18.3	33.9	31.1	0.0
	50-59세	25.3	14.6	37.6	22.5	0.0
	60세 이상	24.0	10.0	36.8	29.2	0.0
거주지역(서울)	도심권	7.4	27.8	27.8	37.0	0.0
	동북권	19.0	20.6	35.5	24.5	0.3
	동남권	20.0	19.0	38.6	21.9	0.5
	서북권	10.7	20.7	42.1	26.4	0.0
	서남권	18.4	20.0	35.4	25.9	0.3



[그림 3-9] 연령대별 보행 시 스마트폰 사용 금지 제도 마련 선택률



[그림 3-10] 연령대별 스마트폰 사용 보행자 사고 유발 시 과실 부과 선택률

‘사고 유발 시 보행자의 과실 비율 부과’와 관련하여 최근 이슈가 되고 있는 ‘차대사람’ 사고에서 스몐비(Smombie)⁹⁾의 가해자 지정 여부에 대한 의견을 조사하였다. 지나친 스마트폰 활동은 보행사고를 유발할 수 있지만, 현행 ‘자동차사고 과실비율 인정 기준’ 등에서는 ‘차대사람’ 사고 발생 시 보행자가 스마트폰 사용을 한 경우에도 가해자로 지정하거나 과실 비율을 부과하는 것은 제한되고 있다. 스몐비의 가해자 지정 여부를 조사한 결과, 응답자의 74.9%가 가해자 지정에 찬성하였고 13.5%가 반대하였다. 특히, 응답자의 연령대가 높거나 운전 경험이 있는 경우, 또는 보행 중 스마트폰을 사용하지 않는 응답자의 찬성 비율이 높게 나타났다.

운전자는 도로교통법 제27조 등에 의해 보행자를 보호할 의무가 있으므로 해당 조사 결과를 근거로 보행자에게 책임을 부과하는 것은 한계가 있다. 하지만, 본 연구에서 살펴본 바와 같이 보행행태와 보행환경이 변화하고 있고 보행자의 부주의한 행동으로 인한 불편사항과 보행사고도 증가하고 있다. 따라서 보행사고 감소를 위해 변화된 보행행태와 보행환경을 반영한 제도적 개선은 지속적으로 추진해야 할 것이다.

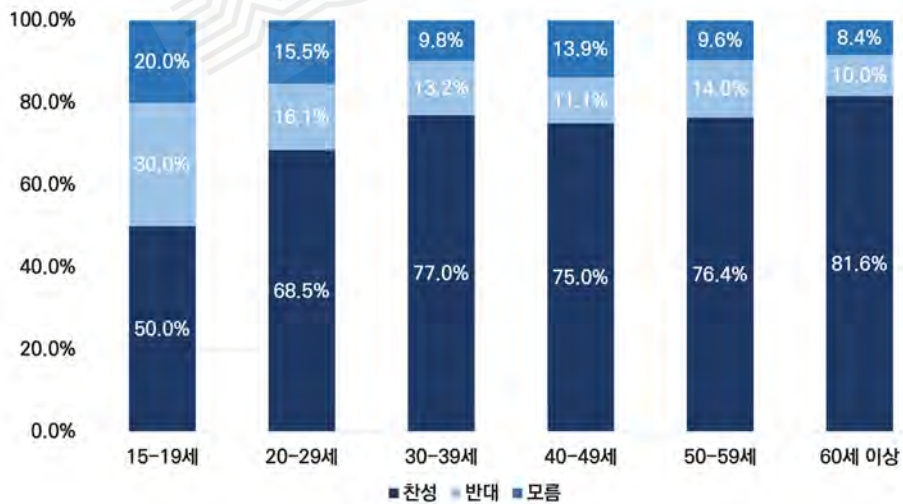
최근 스마트폰 사용으로 인한 보행사고가 증가하고 있지만 사고 심각도가 상대적으로 낮아 적극적인 조치는 이루어지지 않고 있다. 하지만, 본 연구에서 분석한 바와 같이 사고 위험성을 높이는 보행행태가 지속적으로 증가하고 있고 이로 인한 사고 발생도 증가하고 있기 때문에, 더욱 적극적인 시민인식 개선 노력과 제도적인 조치, 다양한 부가장치의 개발이 필요한 것으로 판단된다.

⁹⁾ 스몐비란 스마트폰(Smartphone)과 좀비(Zombie)의 합성어로 보행 중 스마트폰 화면만 주시하며 걸어가는 사람을 일컫는 신조어이다.

[표 3-16] 스몸비-차량 사고 발생 시 스몸비를 가해자로 지정하는 것에 대한 의견

(단위: %)

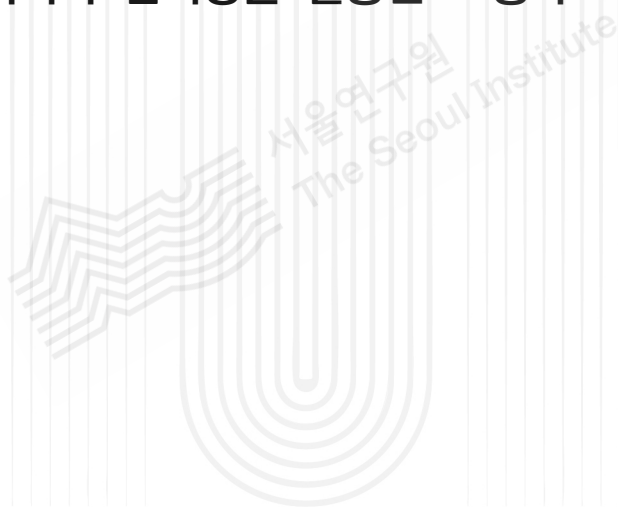
구분		찬성	반대	모름
계		74.9	13.5	11.6
성별	남자	76.6	13.9	9.5
	여자	73.3	13.2	13.5
연령	15-19세	50.0	30.0	20.0
	20-29세	68.5	16.1	15.5
	30-39세	77.0	13.2	9.8
	40-49세	75.0	11.1	13.9
	50-59세	76.4	14.0	9.6
	60세 이상	81.6	10.0	8.4
거주지역(서울)	도심권	79.6	13.0	7.4
	동북권	73.2	15.8	11.0
	동남권	73.8	14.8	11.4
	서북권	71.9	11.6	16.5
	서남권	77.7	11.1	11.1
보행 중 스마트폰 사용 여부	사용함	72.5	15.9	11.6
	사용하지 않음	80.3	8.1	11.6
운전 경험	있음	78.7	12.2	9.1
	없음	62.9	17.5	19.6
운전 중 스마트폰 사용 여부	사용함	79.2	12.9	7.9
	사용하지 않음	78.4	11.9	9.8



[그림 3-11] 스몸비-차량 사고 발생 시 스몸비를 가해자로 지정하는 것에 대한 연령대별 찬성률

04

빅데이터와 딥러닝을 활용한 보행사고 추정



- 1_보행사고 설명을 위한 빅데이터 구축
- 2_보행사고 추정을 위한 딥러닝 모형 구축 및 적용
- 3_장래 보행환경·행태 변화에 따른 보행사고 변화 예측
- 4_딥러닝 분석의 시사점 및 향후 과제

04. 빅데이터와 딥러닝을 활용한 보행사고 추정

1_보행사고 설명을 위한 빅데이터 구축

보행사고는 지역, 보행환경, 보행자 성별 및 연령 등 다양한 변수의 영향을 받는다. 특히, 3장에서 분석한 바와 같이 최근 보행행태는 아주 다양하게 변화하고 있어 이를 반영하기 위한 새로운 데이터 확보가 필요하다. 본 연구에서는 TAAS의 보행사고 데이터, 서울 열린데이터 광장과 통계청의 교통 및 사회경제 데이터, 민간의 스마트폰 이용 데이터 등 다양한 데이터를 통합하여 보행사고를 설명하기 위한 빅데이터를 구축하였다.

1) 독립변수로 활용된 데이터

(1) 공공데이터: 교통 관련 데이터 및 사회경제지표

서울 열린데이터 광장, 통계청 등을 통해 보행사고와 관련된 공공데이터를 수집하였다. 우선, 보행환경과 관련된 교통 데이터는 자전거도로길이, 자동차주행거리, 보도면적, 생활인구 등 11개의 데이터를 [표 4-1]과 같이 선정하였다. 생활인구 데이터는 서울시가 보유한 공공데이터와 통신데이터로 추정한 특정 시점에 서울의 특정 지역에 존재하는 인구¹⁰⁾로 정의되어 있고 ‘서울 열린데이터 광장’을 통해 일평균 연령대별, 성별 생활인구 수치로 제공되고 있다. 본 연구에서는 0시부터 23시까지의 시간대별 생활인구 24개 데이터를 구분하여 사용하였다. 공간적 범위는 자치구, 행정동, 집계구로 구분되어 있는데, 본 연구에서는 집계구 단위의 생활인구를 이용하였다.

10) 서울 생활인구 내국인, <http://data.seoul.go.kr/dataList/OA-14979/F/1/datasetView.do>

사회경제지표 역시 서울 열린데이터 광장, 통계청 등에서 보행환경과 관련이 있는 행정동별 평균연령, 사업체 수, 종사자 수, 주민등록인구를 수집하여 사용하였다.

(2) 민간데이터: 스마트폰 앱 사용 데이터

본 연구에서는 보행 중 스마트폰 이용이 보행행태의 변화와 관련이 있다고 판단하고 스마트폰 이용 데이터를 반영하였다. 분석에 사용된 데이터는 ㈜아이디어웨어의 스마트폰 앱 사용자 행태 분석 데이터 제공 서비스인 ‘와이즈앱’을 통해 구축하였다. 스마트폰 이용 시의 보행 여부를 구분하는 것은 불가능하여, 본 연구에서는 전체 스마트폰 이용 현황이 보행 중 스마트폰 이용을 대표한다고 가정하였다.

‘와이즈앱’은 대한민국의 10세 이상 개인 중 안드로이드 OS 스마트폰을 사용하는 약 4만 명(2019년 기준)의 패널을 구성하여 패널의 스마트폰 이용행태를 조사하고, 통계적 방법을 통해 전국 단위 스마트폰 앱 사용 현황을 추정하여 제공하고 있다. 구체적으로 제공하는 정보는 스마트폰 앱별 설치자 수, 사용자 수, 사용 시간, 실행 횟수, 데이터 사용량, 이용자 정보(성별, 연령 비율) 등이며, 데이터는 월별, 주간별, 시간별로 집계된다.

본 연구에서는 카카오톡, 라인을 포함하는 커뮤니케이션, 넷플릭스, 웨이브를 포함하는 엔터테인먼트 등 34개 카테고리별 상위 10개(총 346개) 앱에 대한 데이터를 수집하였고, 그중에서 설문조사 결과를 기반으로 12개 카테고리(메신저, 지도, 음악, 동영상, 뉴스, 소셜, 쇼핑, 금융, 게임, 만화, 배달주문, 키보드)를 재구성하여 85개 앱의 데이터를 분석에 반영하였다(표 4-2 참조). 애플리케이션의 특성을 고려하여 키보드 카테고리 앱은 실행횟수를, 키보드를 제외한 11개의 카테고리 앱은 사용자 수를 독립변수로 사용하였다.

[표 4-1] 독립변수 목록

구분	예시	집계 단위	수집시기
교통 자료	자전거 도로 길이	자치구	2015년 ~ 2018년
	자동차 하루 평균 주행 거리		
	보도 면적		
	보도 길이		
	자동차 등록 대수		2015년 ~ 2019년
	보행환경 만족도		
	차량 평균 통행 속도		
	역 승하차 인원	역	
	도로 링크 길이	링크	
	보행환경 개선 사업비	블록	2015년 ~ 2017년
	시간대별 생활인구 (24시간대)	통계청 집계구	2017년 ~ 2019년
사회경제지표	평균연령	행정동	2015년 ~ 2019년
	사업체 수		2015년 ~ 2018년
	종사자 수		
	주민등록인구		2015년 ~ 2019년
스마트폰 앱 사용 민간 데이터	지도 앱 사용자 수 집계	전국	2016년 ~ 2019년 ¹¹⁾
	메신저 앱 사용자 수 집계		
	음악 앱 사용자 수 집계		
	동영상 앱 사용자 수 집계		
	뉴스 앱 사용자 수 집계		
	소셜 앱 사용자 수 집계		
	쇼핑 앱 사용자 수 집계		
	금융 앱 사용자 수 집계		
	게임 앱 사용자 수 집계		
	만화 앱 사용자 수 집계		
	배달 앱 사용자 수 집계		
	키보드 앱 실행 횟수 집계		

11) 앱에 따라 수집시기는 상이함

[표 4-2] 앱 데이터 카테고리 목록

카테고리	앱
메신저	Messenger, LINE, 밴드, 카카오톡
지도	T map, 네이버 지도, 원내비, 지도, 카카오내비, 카카오맵
음악	FLO, Google Play Music, 삼성 뮤직, YouTube Music, 네이버 뮤직, 멜론, 알송 모바일, 지니 뮤직, 카카오톡뮤직, 팟캐스트
동영상	Google Play 무비, MX 플레이어, My oksusu, Netflix, SBS - 온에어, Seezn, SPOTV NOW, U+모바일tv, wavve, YouTube, 삼성 비디오, 아프리카TV
뉴스	Flipboard, Google 뉴스, 뉴스픽
소셜	Facebook, Instagram, 카카오톡스토리, 트위터
쇼핑	11번가, CJmall, GS SHOP, 옥션, G마켓, 쿠팡, 티몬, 한샘몰, 홈앤쇼핑
금융	우리은행, KB국민은행, NH스마트뱅킹, 신한은행, 카카오뱅크, 토스
게임	Google Play 게임, Minecraft, Pokemon GO, Rider, ROBLOX, 무한의 계단, 배틀그라운드, 브롤스타즈, 클래시 로얄
만화	FlipaClip, SERIES, 네이버 웹툰, 다음웹툰, 라프텔, 레진코믹스, 마루, 미스터블루, 투믹스
배달	맥딜리버리, 배달요기요, 배달의민족, 배달통
킥보드	Lime, 스윙, 쌍쌍, 일레클, 지빌리티, 킥고잉

2) 종속변수로 활용된 역세권별 보행사고 건수와 생활인구당 발생률

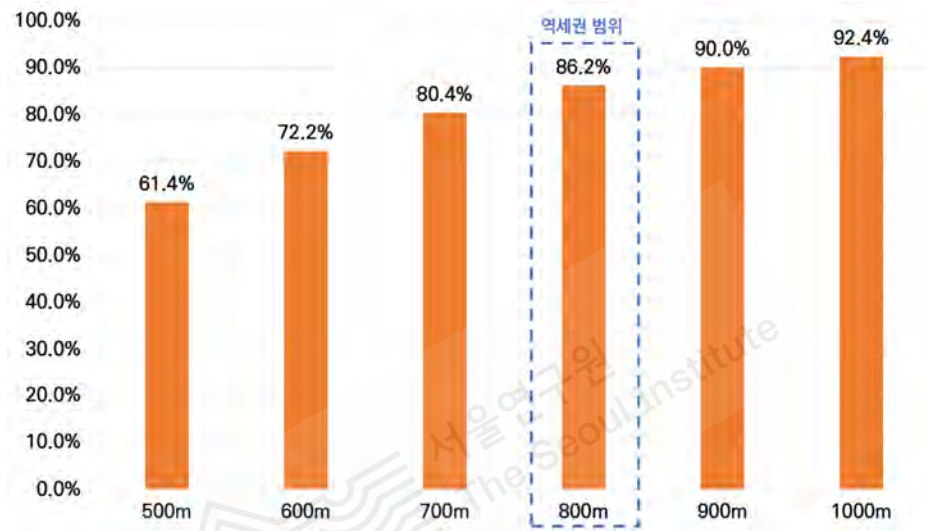
분석에 사용된 종속변수는 TAAS를 통해 수집된 보행사고 데이터이다. 시간적 범위는 독립변수의 수집시기와 최근 스마트폰 상용화 시기, 전동 스쿠터 도입 시기 등을 고려하여 2015년부터 2019년으로 설정하였다.

종속변수의 공간적 단위는 서울 도시철도 역세권으로 설정하고 해당 역세권의 종속변수를 추정하였다. 서울시 노선별 도시철도역의 합은 385개이나, 동일한 이름의 환승역을 하나의 역으로 통합하여 총 300개 역을 기준으로 하였다.

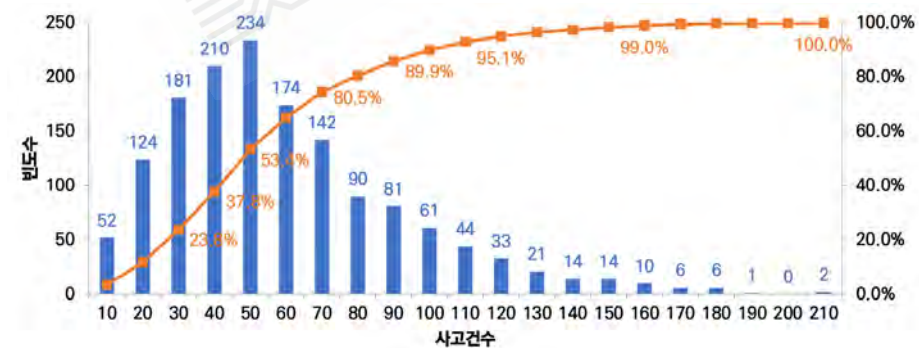
일반적으로 역세권은 도시철도역을 중심으로 반경 500~1,000m 내 지역으로 정의한다. 본 연구에서는 역세권 범위 설정에 따라 포함되는 전체 보행사고 비율과 역세권의 중복성을 고려하여 85% 이상의 보행사고가 포함되는 최소 거리인 반경 800m를 역세권으로 정의하였다 ([그림 4-1] 참조).

분석 기간 동안 서울시 300개 역세권에서 발생한 연도별 보행사고 건수를 살펴보면, 역세권별로 평균 55건의 보행사고가 발생하였고 2019년 신논현역과 강남역 역세권에서는 200건 이상의 보행사고가 발생하였다¹²⁾. 보행사고 발생 건수가 높은 역세권은

[표 4-3]에 나타난 것처럼 유동인구가 많은 지역임을 알 수 있다. 즉, 역세권 내 보행량과 보행사고가 밀접한 관련이 있을 것으로 예상되어 보행량과 연관된 생활인구가 보행사고의 주요 설명 변수인 것으로 예상할 수 있었다. 본 연구에서는 역세권별 보행사고 발생 건수 이외에도 역세권 보행환경의 영향을 받는 보행사고율도 종속변수로 설정하였다. 보행사고율은 주간(오전 8시~오후 6시) 생활인구 10만 명당 보행사고발생 건수로 정의하였다. 보행사고 심각도를 고려하여 전체사고, 중상사고, 경상사고 각각의 보행사고 건수와 보행사고율을 포함하는 총 6개의 종속변수를 설정하였다.



[그림 4-1] 역 반경별 포함되는 보행사고 비율



[그림 4-2] 역세권별 보행사고 발생 건수 히스토그램

- 12) 신논현역과 강남역의 직선거리는 약 800m로, 본 연구에서 설정한 역세권 범위(800m)를 고려하면 중복 지역이 발생하게 된다. 본 연구에서는 역세권이 중복된 경우 중복 지역에서 발생하는 보행사고도 각각 해당 역세권 내 보행사고에 포함되는 것으로 분석하였다.

[표 4-3] 최근 5년간 보행사고 발생 상위 10개 역세권

(단위: 역, 건)

순위	2015년	2016년	2017년	2018년	2019년
1	강남역, 179	신논현역, 179	강남역, 174	강남역, 190	신논현역, 209
2	신논현역, 176	을지로입구역, 169	신논현역, 163	신논현역, 173	강남역, 205
3	종각역, 172	종각역, 169	을지로입구역, 161	홍대입구역, 149	강동역, 155
4	을지로입구역, 165	강남역, 162	홍대입구역, 157	종각역, 142	을지로입구역, 147
5	홍대입구역, 159	수유역, 151	수유역, 154	수유역, 132	종각역, 142
6	강동역, 158	홍대입구역, 150	종각역, 152	강동역, 128	수유역, 140
7	역삼역, 157	을지로3가역, 149	영등포역, 151	송파나루역, 122	을지로3가역, 127
8	을지로3가역, 151	청량리역, 144	강동역, 145	을지로입구역, 121	영등포역, 126
9	시청역, 146	강동역, 142	영등포사당역, 145	영등포사당역, 120	논현역, 123
10	청량리역, 145	제기동역, 139	청량리역, 143	논현역, 119	홍대입구역, 121

[표 4-4] 최근 5년간 보행사고 발생 하위 10개 역세권

(단위: 역, 건)

순위	2015년	2016년	2017년	2018년	2019년
1	청계산입구역, 1	청계산입구역, 1	남태령역, 4	청계산입구역, 0	남태령역, 0
2	개화역, 2	개화역, 2	개화역, 5	남태령역, 2	개화역, 1
3	마곡나루역, 3	남태령역, 2	녹번역, 5	개화역, 3	청계산입구역, 2
4	남태령역, 5	마곡나루역, 4	청계산입구역, 5	서빙고역, 5	일원역, 4
5	서빙고역, 6	서빙고역, 5	마곡나루역, 6	동작역, 7	북한산우이역, 5
6	북정역, 8	녹천역, 6	동작역, 8	둔촌오륜역, 9	서빙고역, 8
7	북한산우이역, 8	북한산우이역, 8	북한산우이역, 8	마곡나루역, 9	마곡나루역, 8
8	일원역, 8	온수역, 8	상일동역, 8	북정역, 9	북정역, 9
9	녹천역, 9	동작역, 9	김포공항역, 9	북한산우이역, 9	신목동역, 9
10	동작역, 11	마곡역, 10	녹천역, 9	구파발역, 11	이촌역, 10

3) 데이터 전처리

본 연구에서 수집한 데이터는 집계되는 공간적 범위와 수집 시기가 서로 상이하다. 따라서, 본 연구에서 설정한 시간적 범위(1년)와 공간적 단위(역세권)별 데이터셋을 구축하기 위하여 데이터 전처리를 수행하였다.

(1) 역세권별 교통 및 사회경제지표 구축

[표 4-1]에 나타난 것처럼 교통 및 사회경제 지표들의 집계 단위는 자치구, 행정동, 집계구 등으로 다양하다. 본 연구에서 설정한 공간적 분석 단위는 역세권으로, 하나의

역세권은 대부분 2개 이상의 자치구, 행정동, 집계구를 포함하고 있다. 따라서 공간적으로 집계 범위가 다양한 교통 및 사회경제지표를 역세권 단위로 변환하기 위해 데이터 전처리를 수행하였다.

데이터 전처리를 위해 ‘교통 및 사회경제지표는 집계 단위 내에서 동일하다’고 가정하였다. 즉, 데이터 속성값을 집계 단위의 면적으로 나눈 단위면적당 속성값이 해당 집계구를 대표할 수 있다고 가정한 것이다. 위의 가정을 근거로 생활인구를 제외한 교통 및 사회경제지표의 역세권별 속성을 아래 식과 같이 추정하였다.

아래 식에서는 속성의 특성에 따라 두 가지 경우로 속성값을 추정한다. 첫 번째 경우는 생활인구, 자동차 등록대수와 같이 역세권 내에 포함되는 다양한 집계 단위의 합이 속성의 값을 나타내는 경우이다. 반면에, 보행환경 만족도, 평균연령과 같이 역세권 내 집계 단위의 평균값이 필요한 경우에는 두 번째 경우로 산정하였다.

$$X_i^f = \begin{cases} \sum_{j=1}^{n_i} \frac{x_j^f}{a_j} a_{ij} & (\text{속성 값이 집계 단위 구성 요소의 합인 경우}) \\ \sum_{j=1}^{n_i} \frac{x_j^f}{A_i} a_{ij} & (\text{속성 값이 집계 단위 구성 요소의 평균인 경우}) \end{cases}$$

$$\sum_{j=1}^{n_i} a_{ij} = A_i$$

X_i^f : 역세권 i 의 속성 f (f = 교통 및 사회경제 지표, $i = 1, \dots, 300$)

x_j^f : 집계단위(자치구, 행정동) j 의 속성 f

n_i : 역세권 i 에 포함되는 집계단위의 수

a_j : 집계단위 j 의 면적 ($j = 1, \dots, J$)

J : 집계단위의 전체 수, $J = 25$ (자치구), $J = 424$ (행정동)

a_{ij} : 역세권 i 내 집계단위 j 의 면적

A_i : 역세권 i 의 면적

생활인구는 집계구를 공간적 단위로 하여 해당 연도의 시간대별, 연령대별 일평균 수치로 제공되고 있다. 역세권별로 해당 연도의 시간대별 생활인구를 산출하기 위하여 아래 식과 같은 전처리 과정을 수행하였다. 우선, 연령대별로 구분된 생활인구 데이터를 해당연도의 시간대별 생활인구로 변환하고, 위에서 제시한 교통 및 사회경제 지표 데이터의 경우와 동일하게 역세권 내 포함되는 집계구 면적을 이용한 가중평균을 통해 해당 역세권의 시간대별 생활인구 데이터를 추정하였다.

$$FP[y, h]_i = \sum_{j=1}^{n_i} \frac{fp[y, h]_j}{a_j} a_{ij}$$

$$fp[y, h]_j = \frac{1}{365} \sum_{d=1}^{365} \sum_{AA=1}^{14} fp[y, d, h]_j^{AA}$$

$$\sum_{j=1}^{n_i} a_{ij} = A_i$$

$FP[y, h]_i$: 역세권 i 의 y 년 h 시간대 하루 평균 생활인구
$fp[y, h]_j$: 집계구 j 의 y 년 h 시간대 하루 평균 생활인구
$fp[y, d, h]_j^{AA}$: 집계구 j 의 y 년 d 일 h 시간대 AA 연령대의 하루 평균 생활인구
d	: 365일 체계로 환산한 일자(1~365)
h	: 24시간 체계로 환산한 시간대(0, ..., 23)
AA	: 연령대(0세, 10세, 15세, 20세, 25세, ..., 70세)
n_i	: 역세권 i 에 포함되는 집계구 수
a_j	: 집계구 j 의 면적($j = 1, \dots, 19153$)
a_{ij}	: 역세권 i 에 포함되는 집계구 j 의 면적
A_i	: 역세권 i 의 면적($i = 1, \dots, 300$)

(2) 역세권별 스마트폰 이용 데이터

‘와이즈앱’에서 제공하는 스마트폰 이용 데이터는 전국을 공간적 범위로 하고 있어 이를 역세권별 정보로 변환하기 위한 전처리 과정을 수행하였다. 데이터 가공을 위해 ‘연령대별 스마트폰 이용 특성은 전국이 동일하다’고 가정하였다. 예를 들어 특정 앱을 사용하는 20대 비율이 전국 인구 대비 10%인 경우 특정 집계구 내 20대 생활인구 중 10%는 해당 앱을 사용한다고 가정하여 집계구의 해당 앱 사용자 수를 계산하였다. 우선, 생활인구 데이터를 월단위로 변환하고 이를 이용하여 월단위 전국 앱 사용자 수를 연단위 집계구 앱 사용자 수로 변환하였다.

$$wa_{APP}^{user}[y]_j = \sum_{m=1}^{12} \sum_{AA=1}^5 \frac{wa_{APP}^{user}[y,m]^{AA}}{Pop[y,m]^{AA}} fp[y,m]_j^{AA}$$

$$fp[y,m]_j^{AA} = \sum_{h=0}^{23} fp[y,m,h]_j^{AA}$$

$wa_{APP}^{user}[y]_j$: 집계구 j 의 y 년 APP 앱 사용자 수
$wa_{APP}^{user}[y,m]^{AA}$: 전국의 y 년 m 월 AA 연령대의 하루 평균 APP 앱 사용자 수
$wa_{APP}^{user}[y,m]$: 전국 y 년 m 월 하루 평균 APP 앱 사용자 수
$Pop[y,m]^{AA}$: 전국 y 년 m 월 AA 연령대 주민등록인구 수
$fp[y,m]_j^{AA}$: 집계구 j 의 y 년 m 월 평균 AA 연령대 하루 평균 생활인구
AA	: 연령대(10대, 20대, 30대, 40대, 50대)

(3) 누락 데이터 추정

본 연구의 분석기간은 독립변수의 수집 시기를 고려하여 2015년부터 2019년까지 5년으로 설정하였다. 하지만, 분석기간 동안 일부 독립변수의 수집 시기가 누락된 부분이 있어 해당기간에 대해서는 데이터 추정이 필요하였다. 예를 들어, 생활인구 데이터는 2017년 이후 자료가 제공되기 시작하였고, ‘와이즈앱’에서 제공하는 스마트폰 이용 데이터도 앱에 따라 가용한 데이터 시기가 상이하여 누락 기간이 발생하였다. 본 연구에서는 시계열 데이터 추정에 널리 활용되는 지수평활법을 사용하여 누락 데이터를 추정하였다. 지수평활법은 보유하고 있는 데이터를 기반으로 누락 기간의 데이터 값을 추정하게 되는데, 추정하려는 시간대와 근접한 관측값에 높은 가중치를 부여하여 누락기간의 데이터를 추정한다. 데이터 추정의 예는 [그림 4-3]과 같다.



[그림 4-3] 인스타그램 사용자 수 추정

위의 과정을 거쳐 2015년부터 2019년까지 연도별 역세권별 데이터셋을 구축하였다 ([표 4-5] 참조).

[표 4-5] 변수 기초통계량

구분	예시	평균	표준편차	최댓값	최솟값
교통	자전거 도로 길이(km)	2.20	1.18	5.48	0.32
	자동차 하루 평균 주행 거리 (km/대/일)	36.56	1.70	42.50	31.95
	보도 면적(m ²)	23,788	7,997	54,730	8,952
	보도 길이(m)	5,842	1,870	12,782	2,586
	자동차 등록 대수(대)	10,536	2,320	17,358	4,132
	보행환경 만족도	6.04	0.23	6.65	5.55
	차량 평균 통행 속도(km/h)	24.30	1.67	29.30	19.50
	역 승하차 인원(명/년)	25, 281,367	27, 023,707	172, 445,063	228,621
	도로 링크 길이(m)	24,006	6,721	41,284	3,274
	보행환경 개선 사업비 (백만 원)	1,982	10,212	86425	0
	시간대별 생활인구(명)	54,466	23,357	242,684	2,105
사회 경제 지표	평균연령(세)	41.95	2.20	50.90	37.48
	사업체 수(개수)	4,202	3,496	24,942	233
	종사자 수(명)	27,202	26,081	174,582	1,167
	주민등록인구(명)	40,042	17,488	87,120	4,463
민간 데이 터	지도 앱 사용자 수 집계(명)	26,838	11,361	83,904	1,574
	메신저 앱 사용자 수 집계(명)	57,887	22,639	151,322	3,642
	음악 앱 사용자 수 집계(명)	26,806	10,515	73,719	1,917
	동영상 앱 사용자 수 집계(명)	43,013	17,761	129,884	2,503
	뉴스 앱 사용자 수 집계(명)	8,133	3,604	23,773	411
	소셜 앱 사용자 수 집계(명)	32,711	13,031	90,340	2,139
	쇼핑 앱 사용자 수 집계(명)	33,892	14,258	100,257	2,068
	금융 앱 사용자 수 집계(명)	17,236	11,117	85,304	575
	게임 앱 사용자 수 집계(명)	23,617	9,336	62,376	1,552
	웹툰 앱 사용자 수 집계(명)	8,052	3,429	26,437	497
	배달 앱 사용자 수 집계(명)	7,877	4,854	40,465	293
	키보드 앱 실행 횟수 집계(횟수)	1,089, 905,507	658, 479,718	3,252, 152,415	440,542
보행 사고 (종속 변수)	전체사고 건수(건수)	55.02	33.35	209.00	0.00
	중상사고 건수(건수)	21.12	12.30	67.00	0.00
	경상사고 건수(건수)	28.57	19.48	154.00	0.00
	전체사고율(건수/10만 명)	8.53	3.30	22.93	0.00
	중상사고율(건수/10만 명)	3.35	1.41	9.53	0.00
	경상사고율(건수/10만 명)	4.37	2.03	15.71	0.00

2_보행사고 추정을 위한 딥러닝 모형 구축 및 적용

1) 딥러닝 모형 구축

본 연구에서는 역세권 보행사고 추정을 위하여 기계학습(Machine Learning) 기법 중 하나인 딥러닝(Deep Learning)을 사용하였다. 기계학습은 데이터 간의 규칙을 컴퓨터가 스스로 찾고 필요한 값을 예측하거나 분류하는 알고리즘을 총칭하며, 본 연구에서는 딥러닝의 다양한 알고리즘 중 복잡한 빅데이터 분석에 널리 이용되는 인공 신경망(Artificial Neural Network)을 사용하였다.

모델링은 파이썬의 numpy, pandas, scikit-learn 그리고 keras 라이브러리를 사용하여 수행하였다. 독립변수의 포함 범위와 앞에서 제시한 6종의 종속변수에 따라 총 12개의 모형을 구축하고 성능을 비교하였다([표 4-6] 참조).

[표 4-6] 딥러닝 모델링을 위한 모형 종류

모형	독립변수	종속변수	
1	교통, 사회경제지표	보행사고 건수	전체
2			중상
3			경상
4		보행사고율	전체
5			중상
6			경상
7	교통, 사회경제지표, 앱 사용 데이터	보행사고 건수	전체
8			중상
9			경상
10		보행사고율	전체
11			중상
12			경상

딥러닝 모형은 전체 독립변수를 입력층(input layer)을 통해 입력하면 은닉층(hidden layer)을 거쳐 출력층(output layer)에서 보행사고 건수 또는 보행사고율의 값을 출력하는 형태로 구성하였다. 은닉층과 은닉층 별 노드(node) 개수를 결정하기 위하여 다양한 은닉층과 노드 조합을 구성하여 결과를 비교하였고, 연산 속도의 효율성과 모형의 정확도를 고려하여 사고건수 예측 모형은 3개의 은닉층과 은닉층별 64,

32, 16개의 노드로 모형을 구성하였다. 사고율 예측 모형은 3개의 은닉층과 은닉층별 32, 16, 8개의 노드로 구성하였다¹³⁾. 입력층(input layer)의 노드 개수는 독립변수의 개수와 동일하며 출력층의 노드는 종속변수의 개수와 같은 1개이다. 은닉층 간의 활성화함수(activation function)는 ReLU(Rectified Linear Unit)를 사용하였다. 모형의 목적 함수(objective function)는 모형을 통해 추정된 결과(보행사고 건수 또는 보행사고율)와 실제값의 차이를 비교하는 평균제곱오차(Mean Squared Error, MSE)를 최소화하도록 하였다.

구축된 딥러닝 모형을 적용하기 위해 앞에서 구축한 역세권별 데이터 셋의 독립변수에 대해 표준화를 시행한 후¹⁴⁾, 데이터 셋을 학습(training)과 시험(test)데이터 셋으로 구분하고 학습데이터 셋을 이용하여 모형을 학습하였다. 본 연구에서는 5년간의 300개 역세권을 대상으로 하여 총 1,500개의 데이터 셋을 구축하였고, 학습과 시험 데이터 셋은 7:3의 비율로 분리하였다. 모형 학습시에는 과대 적합(overfitting)을 방지하기 위해 8:2의 비율로 학습과 검증(validation)데이터 셋을 별도로 구분하여 적용하였다.

학습 데이터 셋을 통해 모형이 학습되면, 학습된 모형과 시험 데이터 셋을 이용하여 시험 데이터 셋에 대한 종속변수를 추정하였다. 모형의 성능을 비교하기 위한 평가지표는 RMSE(Root Mean Square Error)와 MAE(Mean Absolute Error)를 사용하였다. 전반적인 모형의 구축 과정은 [그림 4-4]와 같다.

13) 은닉층, 노드의 개수가 너무 적으면 과소적합, 너무 많으면 과대적합이 발생한다. 본 연구에서는 은닉층의 개수를 2~5개, 은닉층별 노드의 개수를 5~4,096개로 다양한 조합을 시험하여 결정하였다.

14) 독립변수들의 크기와 단위가 상이하고 이상치(outlier)의 영향을 줄이기 위하여 표준화 과정이 필요하다. 표준화를 통해 각각의 독립변수는 평균이 0이고 분산이 1인 값으로 변환된다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

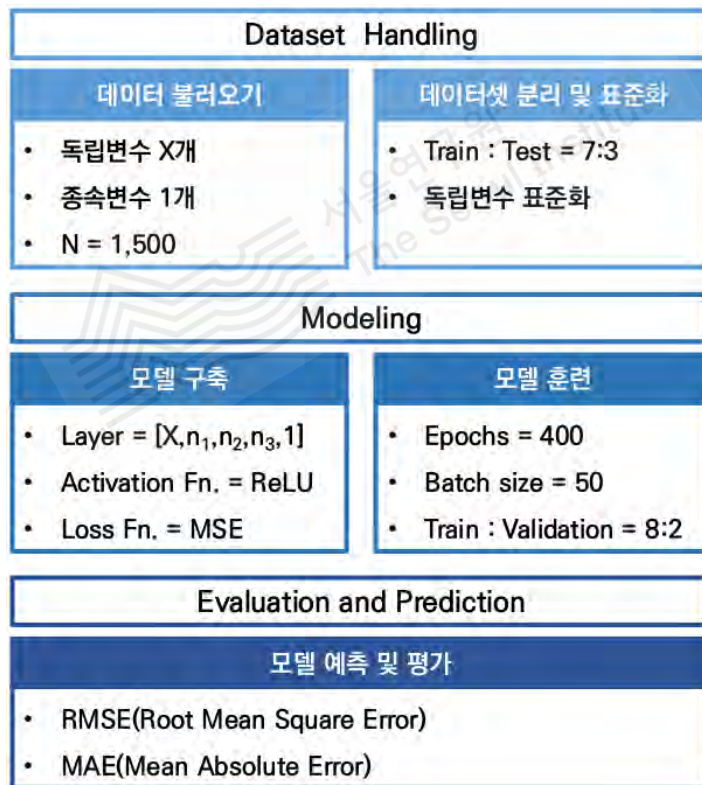
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

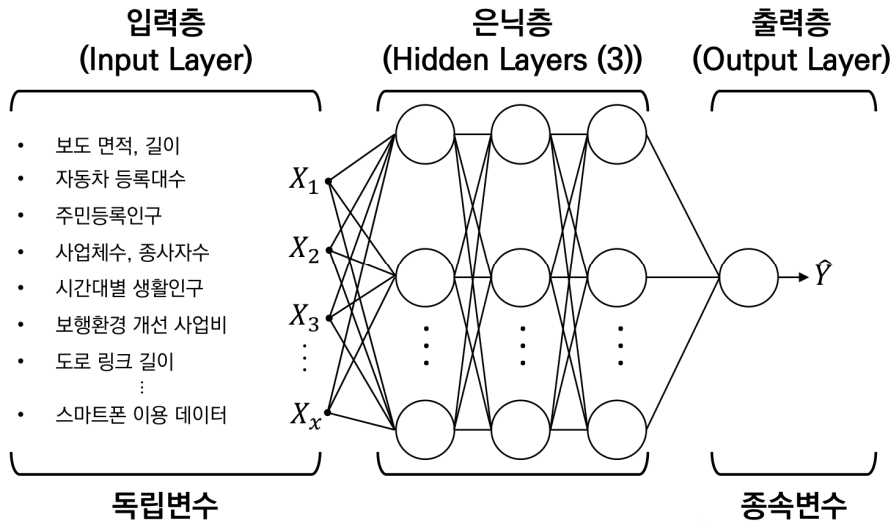
n : 샘플수

y_i : i 번째 관측값

\hat{y}_i : i 번째 예측값



[그림 4-4] 모델링 과정



[그림 4-5] 딥러닝 모형 구조

2) 보행사고 추정 결과

본 연구에서는 스마트폰 이용 데이터 포함 여부에 따른 독립변수 2종과 종속변수 6종의 조합으로 총 12개의 모형을 구축하였다. 각 모형의 성능은 위에서 제시한 RMSE, MAE와 일반적으로 모형의 설명력을 표현하는 결정계수(R^2)를 이용하여 [표4-7]과 같이 비교하였다.

모형의 성능 비교를 통해 다음과 같은 시사점을 도출하였다. 첫째, 보행사고 건수를 추정하는 모형이 보행사고율을 추정하는 모형보다 전반적으로 정확도가 높았다. 2장의 보행사고 발생 현황에서 분석한 것처럼 보행사고 발생 건수는 생활인구와 밀접한 관련이 있어 상대적으로 추정이 쉬운 반면, 보행사고율 추정은 보행량 이외의 보행환경 영향을 이용하여 반영하기 때문에 정확도가 다소 낮아진 것으로 판단된다¹⁵⁾.

둘째, 중상사고를 추정하는 모형보다 전체 또는 경상사고를 추정하는 모형의 정확도가 상대적으로 높게 나타났다. 특히, 경상사고율을 추정하는 모형은 중상사고율을 추정하는 모형보다 현저히 높은 정확도를 보여주었다. 본 연구에서 활용한 사회경제 지표, 스마트폰 이용 데이터 등이 최근 증가하는 경상사고를 유발하는 보행환경을 심각

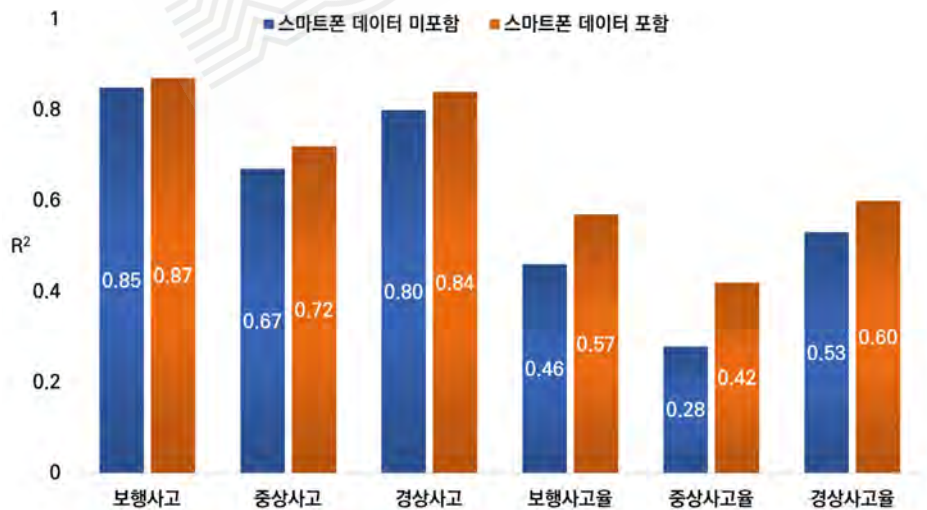
¹⁵⁾ 본 연구에서 딥러닝을 적용하여 보행사고율을 추정한 모형의 성능은 전통적인 방법인 회귀분석을 이용한 경우보다 현저히 높은 것으로 나타났다. 자세한 사항은 [3] 회귀분석을 이용한 보행사고 추정 및 비교에서 제시하였다.

도가 높은 중상사고의 경우보다 더 잘 반영한 결과로 판단된다.

마지막으로, [그림 4-6]에 나타난 것처럼 스마트폰 이용 데이터를 포함하여 보행사고를 추정하는 모형이 더 좋은 성능을 보여주었다. 본 연구에서는 보행 중 스마트폰 이용이 보행행태와 보행환경을 변화시키고 보행사고에 영향을 미칠 것으로 가정하였는데, 모형 간 성능 비교를 통해 보행 중 스마트폰 이용 데이터가 보행사고 분석에 필요하다는 것을 확인할 수 있었다.

[표 4-기] 딥러닝 모형 평가

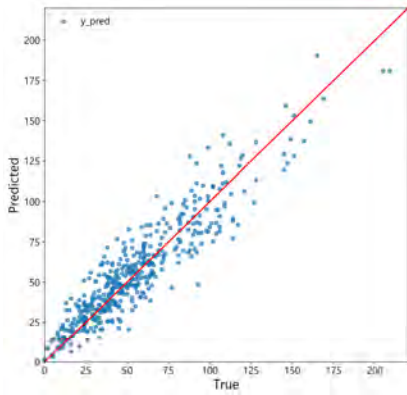
모형	독립변수	종속변수		R ²		RMSE		MAE	
				Train	Test	Train	Test	Train	Test
1	교통, 사회경제지표	보행 사고 건수	전체	0.95	0.85	7.70	12.88	5.44	9.96
2			중상	0.87	0.67	4.50	6.72	3.26	5.12
3			경상	0.91	0.80	5.73	8.72	4.11	6.67
4		보행 사고 율	전체	0.70	0.46	1.78	2.52	1.37	1.95
5			중상	0.69	0.28	0.78	1.21	0.57	0.93
6			경상	0.77	0.53	0.96	1.44	0.72	1.10
7	교통, 사회경제지표, 앱 사용 데이터	보행 사고 건수	전체	0.94	0.87	8.30	12.02	5.85	9.50
8			중상	0.91	0.72	3.84	6.10	2.60	4.78
9			경상	0.95	0.84	4.42	7.81	2.94	5.87
10		보행 사고 율	전체	0.78	0.57	1.53	2.25	1.13	1.76
11			중상	0.72	0.42	0.74	1.09	0.53	0.85
12			경상	0.79	0.60	0.90	1.33	0.67	1.05



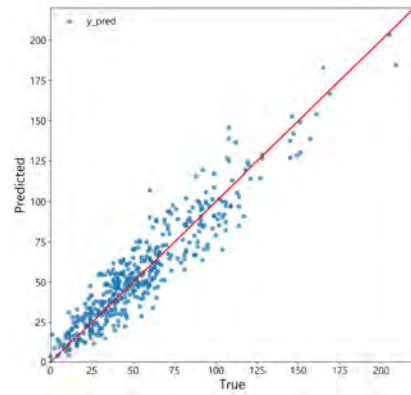
[그림 4-6] 스마트폰 이용 데이터 포함 여부에 따른 모형 성능 차이

모형의 추정값을 실제 관측값과 시각적으로 비교하기 위하여 산점도(Scatter Plot)를 작성하였다. [그림 4-7]~[그림 4-18]의 x 축은 관측값(역세권별 실제 보행사고 발생 건수 또는 발생률), y 축은 모형을 통해 추정한 값을 나타낸다. 산점도상의 지점들이 중앙 직선 ($y = x$ 그래프)에 가까울수록 관측값과 추정값의 오차가 작은 결과를 보여 준다. 산점도 비교에서도 스마트폰 이용 데이터를 포함하는 우측열 모형의 결과가 좌측열보다 오차가 적게 발생하며, 특히 보행사고율을 추정하는 모형에서 스마트폰 이용 데이터 반영 시 중앙직선에 더 잘 수렴함을 확인할 수 있다.

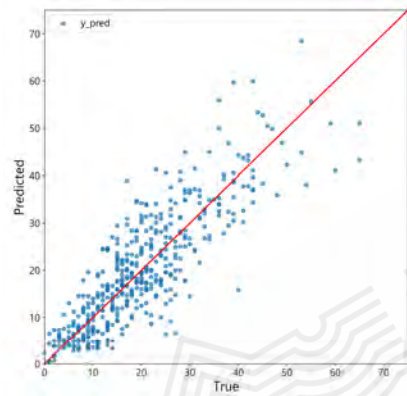




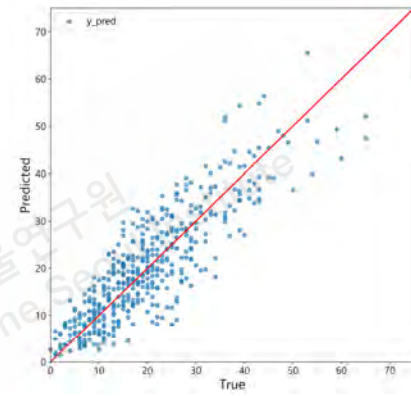
[그림 4-7] 스마트폰 이용 데이터 미포함
전체사고 건수 추정 딥러닝 모형 테스트 결과



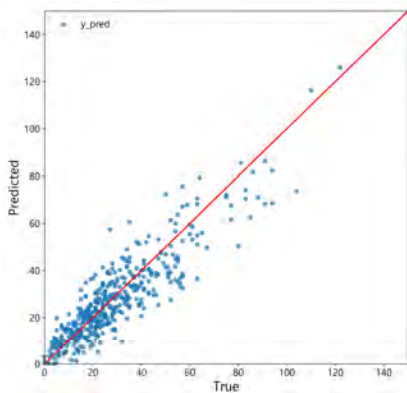
[그림 4-8] 스마트폰 이용 데이터 포함
전체사고 건수 추정 딥러닝 모형 테스트 결과



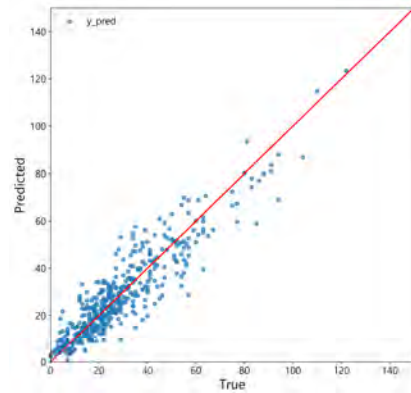
[그림 4-9] 스마트폰 이용 데이터 미포함
중상사고 건수 추정 딥러닝 모형 테스트 결과



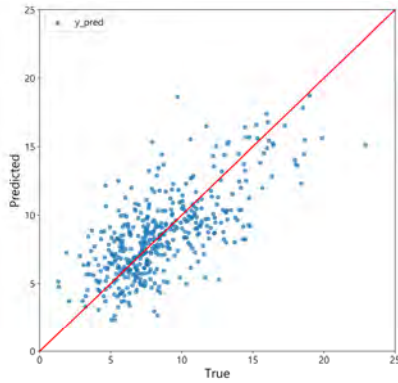
[그림 4-10] 스마트폰 이용 데이터 포함
중상사고 건수 추정 딥러닝 모형 테스트 결과



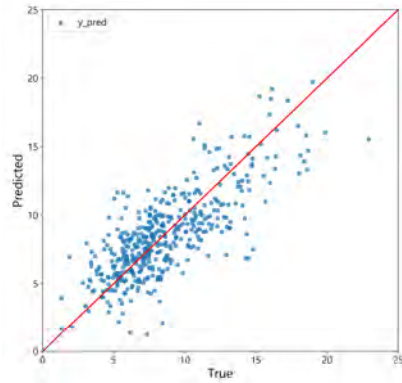
[그림 4-11] 스마트폰 이용 데이터 미포함
경상사고 건수 추정 딥러닝 모형 테스트 결과



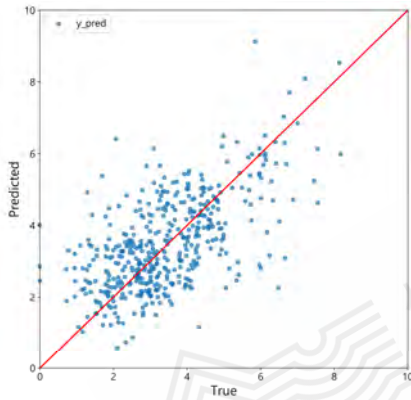
[그림 4-12] 스마트폰 이용 데이터 포함
경상사고 건수 추정 딥러닝 모형 테스트 결과



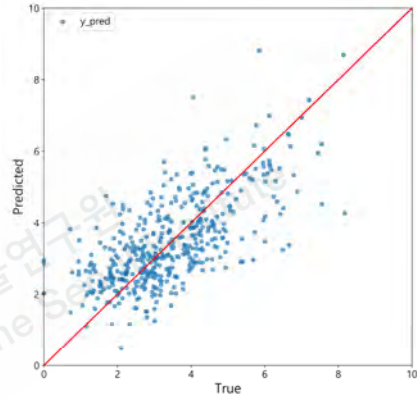
[그림 4-13] 스마트폰 이용 데이터 미포함
전체사고를 추정 딥러닝 모형 테스트 결과



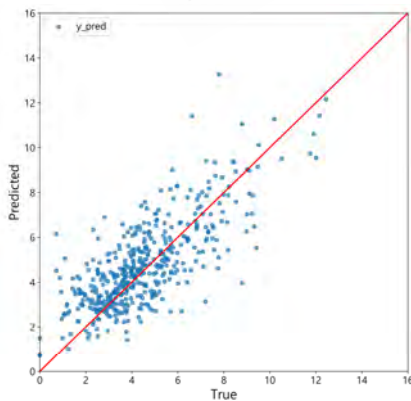
[그림 4-14] 스마트폰 이용 데이터 포함
전체사고를 추정 딥러닝 모형 테스트 결과



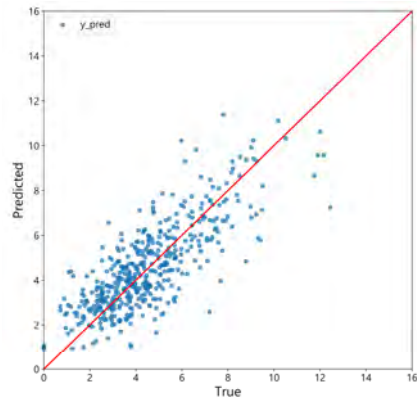
[그림 4-15] 스마트폰 이용 데이터 미포함
중상사고를 추정 딥러닝 모형 테스트 결과



[그림 4-16] 스마트폰 이용 데이터 포함
중상사고를 추정 딥러닝 모형 테스트 결과



[그림 4-17] 스마트폰 이용 데이터 미포함
경상사고를 추정 딥러닝 모형 테스트 결과



[그림 4-18] 스마트폰 이용 데이터 포함
경상사고를 추정 딥러닝 모형 테스트 결과

3) 회귀분석을 이용한 보행사고 추정 및 비교

(1) 회귀분석을 이용한 보행사고 추정

본 연구에서는 다양한 변수를 포함하는 빅데이터를 분석하기 위하여 딥러닝을 이용하였다. 이 장에서는 기존의 연구에서 주로 이용하였던 다중 회귀분석을 통해 보행사고를 추정하고 딥러닝 적용 결과와 비교하여 차이점을 도출하였다.

회귀분석에 적용할 독립변수 결정을 위해 딥러닝에 적용된 독립변수에 대해 다중공선성을 확인하였다. 회귀분석의 다중공선성 판단 지표인 분산팽창요인(Variance Inflation Factor, VIF)을 이용하여 일반적인 다중공선성 판단 기준인 10~15보다 낮은 값을 갖는 독립변수를 도출하였다([표 4-8] 참조). 독립변수 간 상관관계를 분석하여 상관관계가 높은 독립변수 집단 중 하나의 독립변수만 분석에 사용하였다. 예를 들어 보도 면적과 보도 길이 중에서는 보도면적을 사용하였다. 생활인구는 24개의 시간대별 생활인구 중 보행량과 가장 관련이 높은 것으로 추정되는 최댓값을 사용하였다. 다중공선성 판단에 의해 선별된 독립변수는 [표 4-9]와 같다.

[표 4-8] 회귀분석을 위한 모든 독립변수의 분산팽창요인 지표

변수	VIF	변수	VIF
자전거 도로 길이	3.0	11시 생활인구	21430
자동차 하루 평균 주행 거리	2.4	12시 생활인구	32114
보도 면적	13.7	13시 생활인구	59111
보도 길이	12.0	14시 생활인구	61332
자동차 등록 대수	3.8	15시 생활인구	54272
보행환경 만족도	2.4	16시 생활인구	30572
차량 평균 통행 속도	2.5	17시 생활인구	11977
평균 연령	5.3	18시 생활인구	5788
사업체 수	17.0	19시 생활인구	12823
주민등록인구	6.6	20시 생활인구	14464
종사자 수	23.7	21시 생활인구	8415
도로 링크 길이	1.8	22시 생활인구	1726
보행환경 개선 사업비	2.1	23시 생활인구	6771
역 승하차 인원	7.4	메신저 앱 사용자 수 집계	13471
0시 생활인구	37942	지도 앱 사용자 수 집계	10466
1시 생활인구	119288	음악 앱 사용자 수 집계	3496
2시 생활인구	178970	동영상 앱 사용자 수 집계	26004
3시 생활인구	85965	뉴스 앱 사용자 수 집계	222
4시 생활인구	22664	소셜 앱 사용자 수 집계	40651
5시 생활인구	14802	쇼핑 앱 사용자 수 집계	3205
6시 생활인구	5860	금융 앱 사용자 수 집계	7508
7시 생활인구	1587	게임 앱 사용자 수 집계	599
8시 생활인구	3100	웹툰 앱 사용자 수 집계	1780
9시 생활인구	15766	배달 앱 사용자 수 집계	2227
10시 생활인구	22228	키보드 앱 실행 횟수 집계	3.0

[표 4-9] 회귀분석을 위해 선별한 독립변수의 분산팽창요인 지표

변수	VIF	변수	VIF
자전거 도로 길이	1.9	종사자 수	6.9
자동차 하루 평균 주행 거리	1.8	도로 링크 길이	1.4
보도 면적	2.8	보행환경 개선 사업비	1.5
자동차 등록 대수	2.8	역 승하차 인원	5.0
보행환경 만족도	2.0	생활인구 최댓값	6.8
차량 평균 통행 속도	1.9	뉴스 앱 사용자 수 집계	5.1
평균 연령	2.2	금융 앱 사용자 수 집계	5.5
사업체 수	5.0	키패드 앱 실행 횟수 집계	2.5
주민등록인구	2.7		

선정된 독립변수를 적용하여 보행사고 건수(전체, 중상, 경상)와 보행사고율(전체, 중상, 경상) 6개 종속변수에 대한 회귀모형을 구축하였다. 각각의 회귀모형 구축 시에는 딥러닝의 경우와 마찬가지로 전체 데이터 셋 70%를 이용하였고, 독립변수 계수의 유의확률을 확인하여 통계적으로 유의한 독립변수를 기반으로 모형을 구축하였다. 구축된 모형의 결과는 [표 4-10]~[표 4-15]와 같다. 모형 구축에 이용되지 않은 나머지 30% 데이터 셋으로 구축된 모형을 테스트한 결과는 [표 4-16], [그림 4-19]~[그림 4-24]에서 제시하였다.

구축한 선형 회귀모형의 결과를 살펴보면, 6가지 모형 전부에서 자전거 도로 길이가 음의 계수를 갖는 것으로 나타났고, 보도 면적, 자동차 하루 평균 주행 거리, 보행환경 개선 사업비 등이 종속변수에 음의 영향을 주는 변수들로 나타났다. 결과를 해석해보면 자전거 도로 길이가 길어지면 보행사고가 줄어든다는 뜻인데, 이는 자전거 도로 환경 개선에 따른 수단 전환(자동차, 도보 → 자전거)이 발생하여 보행사고 노출률이 낮아진 효과로 판단된다. 또한 보도 면적과 보행환경 개선 사업비의 경우, 보도 면적과 보행환경 개선 사업비가 증가하면 보행환경이 개선되어 상대적으로 보행자와 차가 상충할 가능성이 낮아지기 때문에 보행사고가 줄어드는 것으로 추정된다. 종속변수에 양의 영향을 주는 변수로는 생활인구 최댓값, 자동차 등록 대수, 사업체 수, 주민등록 인구 수, 역 승하차 인원 등 주로 보행량, 통행량과 관련된 변수들이었고, 그 중 생활인구 최댓값의 계수가 가장 크게 나타난 것으로 보아, 생활인구가 보행사고에 가장 많은 영향을 미치는 것으로 판단된다.

[표 4-10] 전체사고 건수 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	54.889	0.600	91.471	0.000	***
자전거 도로 길이	-2.286	0.812	-2.815	0.005	***
자동차 하루 평균 주행 거리	-1.613	0.720	-2.239	0.025	**
보도 면적	-5.097	0.980	-5.203	0.000	***
자동차 등록 대수	5.455	0.970	5.621	0.000	***
평균 연령	2.823	0.795	3.550	0.000	***
사업체 수	8.154	1.344	6.067	0.000	***
주민등록인구	2.626	0.966	2.718	0.007	***
종사자 수	-3.520	1.563	-2.252	0.025	**
보행환경 개선 사업비	-4.024	0.729	-5.523	0.000	***
역 승하차 인원	2.537	1.309	1.939	0.053	*
생활인구 최댓값	22.882	1.453	15.753	0.000	***
금융 앱 사용자 수 집계	2.967	1.066	2.783	0.005	***
키보드 앱 실행 횟수 집계	-3.413	0.822	-4.154	0.000	***

*p<0.1; **p<0.05; ***p<0.01

[표 4-11] 중상사고 건수 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	20.969	0.228	92.058	0.000	***
자전거 도로 길이	-0.746	0.303	-2.464	0.014	**
보도 면적	-1.034	0.367	-2.817	0.005	***
자동차 등록 대수	1.740	0.345	5.041	0.000	***
평균 연령	2.252	0.300	7.493	0.000	***
사업체 수	3.400	0.486	6.999	0.000	***
주민등록인구	2.081	0.339	6.145	0.000	***
종사자 수	-2.186	0.594	-3.681	0.000	***
도로 링크 길이	0.670	0.267	2.515	0.012	**
보행환경 개선 사업비	-0.727	0.275	-2.641	0.008	***
생활인구 최댓값	9.006	0.489	18.422	0.000	***
금융 앱 사용자 수 집계	-1.867	0.384	-4.860	0.000	***
키보드 앱 실행 횟수 집계	-0.718	0.311	-2.311	0.021	**

*p<0.1; **p<0.05; ***p<0.01

[표 4-12] 경상사고 건수 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	28.594	0.393	72.707	0.000	***
자전거 도로 길이	-1.759	0.528	-3.327	0.001	***
자동차 하루 평균 주행 거리	-1.388	0.479	-2.899	0.004	***
보도 면적	-3.925	0.633	-6.200	0.000	***
자동차 등록 대수	3.601	0.594	6.061	0.000	***
사업체 수	3.771	0.753	5.009	0.000	***
종사자 수	-1.495	0.863	-1.733	0.083	*
도로 링크 길이	-8.07	0.468	-1.723	0.085	*
보행환경 개선 사업비	-3.026	0.461	-6.559	0.000	***
역 승하차 인원	1.893	0.795	2.380	0.017	**
생활인구 최댓값	12.094	0.7665	15.800	0.000	***
금융 앱 사용자 수 집계	5.022	0.694	7.231	0.000	***
키보드 앱 실행 횟수 집계	-2.445	0.539	-4.539	0.000	***

*p<0.1; **p<0.05; ***p<0.01

[표 4-13] 전체사고율 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	8.499	0.091	93.748	0.000	***
자전거 도로 길이	-0.342	0.122	-2.808	0.005	***
자동차 하루 평균 주행 거리	-0.294	0.105	-2.8184	0.005	***
보도 면적	-0.755	0.145	-5.206	0.000	***
자동차 등록 대수	1.003	0.145	6.941	0.000	***
평균 연령	1.065	0.120	8.910	0.000	***
사업체 수	0.778	0.197	3.958	0.000	***
주민등록인구	0.533	0.135	3.943	0.000	***
종사자 수	-0.692	0.236	-2.935	0.003	***
보행환경 개선 사업비	-0.404	0.109	-3.706	0.000	***
생활인구 최댓값	0.835	0.179	4.654	0.000	***
키보드 앱 실행 횟수 집계	-0.374	0.095	-3.930	0.000	***

*p<0.1; **p<0.05; ***p<0.01

[표 4-14] 중상사고율 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	3.339	0.037	89.123	0.000	***
자전거 도로 길이	-0.146	0.050	-2.899	0.004	***
보도 면적	-0.194	0.059	-3.261	0.001	***
자동차 등록 대수	0.345	0.056	6.100	0.000	***
평균 연령	0.486	0.048	10.171	0.000	***
사업체 수	0.337	0.077	4.351	0.000	***
주민등록인구	0.393	0.058	6.791	0.000	***
종사자 수	-0.393	0.096	-4.093	0.000	***
도로 링크 길이	0.091	0.044	2.074	0.038	**
생활인구 최댓값	0.306	0.087	3.494	0.000	***
뉴스 앱 사용자 수 집계	-0.218	0.087	-2.520	0.012	**
금융 앱 사용자 수 집계	-0.178	0.084	-2.219	0.027	**
키보드 앱 실행 횟수 집계	-0.179	0.059	-3.046	0.002	***

*p<0.1; **p<0.05; ***p<0.01

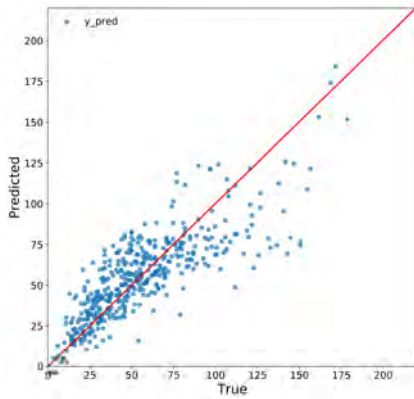
[표 4-15] 경상사고율 추정 선형 회귀모형 결과

변수	Coef.	std. err	t-value	p-value	유의수준
상수	4.359	0.057	75.875	0.000	***
자전거 도로 길이	-0.206	0.077	-2.668	0.008	***
자동차 하루 평균 주행 거리	-0.297	0.069	-4.307	0.000	***
보도 면적	-0.536	0.093	-5.738	0.000	***
자동차 등록 대수	0.596	0.093	6.434	0.000	***
평균 연령	0.460	0.076	6.062	0.000	***
사업체 수	0.347	0.125	2.763	0.006	***
주민등록인구	0.168	0.086	1.964	0.050	*
종사자 수	-0.284	0.149	-1.901	0.058	*
보행환경 개선 사업비	-0.319	0.070	-4.570	0.000	***
생활인구 최댓값	0.507	0.123	4.134	0.000	***
금융 앱 사용자 수 집계	0.345	0.101	3.427	0.001	***
키보드 앱 실행 횟수 집계	-0.240	0.078	-3.059	0.002	***

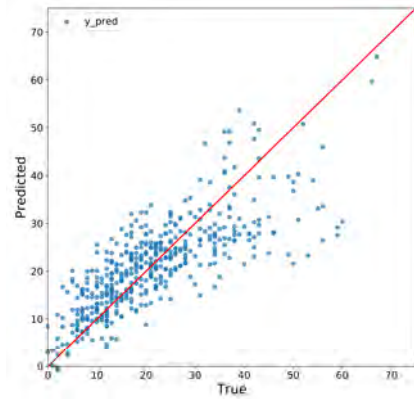
*p<0.1; **p<0.05; ***p<0.01

[표 4-16] 회귀분석 모형 평가

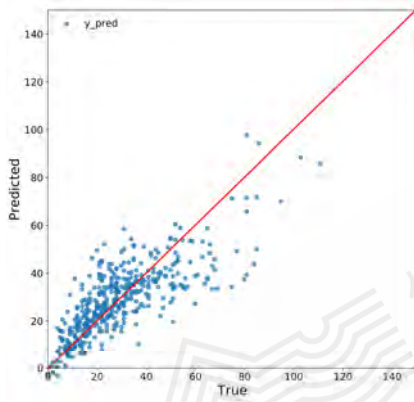
모형	종속변수		R ²		RMSE		MAE	
			Train	Test	Train	Test	Train	Test
선형 회귀분석 모형	보행 사고 발생 건수	전체	0.67	0.68	19.30	18.53	14.19	13.45
		중상	0.64	0.64	7.33	7.55	5.45	5.58
		경상	0.60	0.62	12.64	11.33	9.10	8.26
	보행 사고 발생률	전체	0.24	0.24	2.91	2.79	2.22	2.15
		중상	0.28	0.26	1.20	1.21	0.92	0.94
		경상	0.21	0.20	1.85	1.69	1.39	1.27



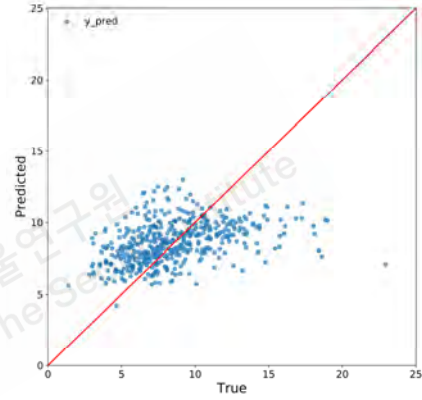
[그림 4-19] 전체사고 건수 추정 회귀분석
모형 테스트 결과



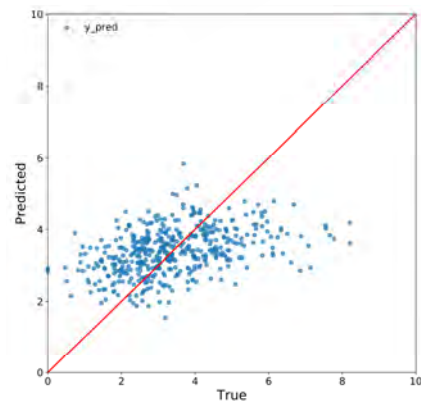
[그림 4-20] 중상사고 건수 추정 회귀분석
모형 테스트 결과



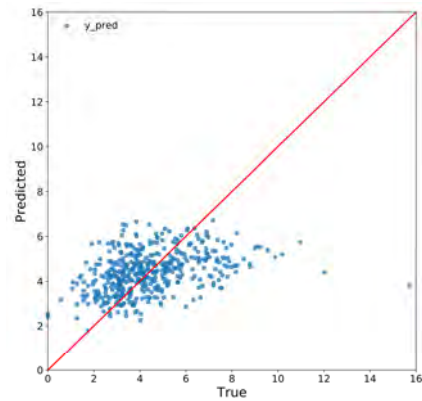
[그림 4-21] 경상사고 건수 추정 회귀분석
모형 테스트 결과



[그림 4-22] 전체사고율 추정 회귀분석
모형 테스트 결과



[그림 4-23] 중상사고율 추정 회귀분석
모형 테스트 결과



[그림 4-24] 경상사고율 추정 회귀분석
모형 테스트 결과

(2) 회귀분석과 딥러닝 분석 비교 및 시사점

회귀분석과 딥러닝 모형의 구축 과정과 결과를 비교하여 다음과 같은 시사점을 도출하였다. 우선, 모형의 구조를 결정하는 과정은 회귀분석이 딥러닝과 비교하여 상대적으로 명확하였다. 딥러닝은 은닉층과 노드 개수 등을 다양하게 조합하며 적합한 모형 구조를 결정하여야 하지만, 회귀분석은 선형, 비선형 등 독립변수와 종속변수의 기본적인 관계 분석만을 통해 모형 구조를 결정할 수 있었다.

또한, 회귀분석은 독립변수와 종속변수의 관계를 명확히 설명할 수 있는 장점을 가지고 있었다. 딥러닝은 일반적으로 블랙박스(Black Box)와 같이 인식되어 독립변수와 종속변수의 관계를 설명하는 것에 제한이 있었다. 하지만, 최근 해석가능한 기계학습(Interpretable Machine Learning)에 대한 연구가 증가하고 있어 향후에는 딥러닝 적용 시에도 변수 간의 관계확인이 가능할 것으로 기대된다.

딥러닝을 이용하면 수집된 데이터를 충분히 활용할 수 있는 장점이 있었다. 회귀분석 모형은 구축하는 과정에서 독립변수 간 다중공선성, 독립변수 계수의 유의확률 등을 확인하는 사전 작업이 필요하였고, 이 과정에서 많은 독립변수 데이터가 최종 모형에 활용되지 못했다. 일반적으로 다중 회귀분석에서 반영할 수 있는 독립변수 개수에 한계가 있음을 고려할 때, 보행사고와 관련된 다양한 데이터를 확보하여도 회귀분석 등 기존의 방법으로는 구축된 빅데이터를 모두 활용하지 못하는 문제가 발생할 수 있음을 알 수 있었다. 반면에, 딥러닝을 이용하면 가용한 데이터를 모두 독립변수(input data)로 활용할 수 있었고 별도의 선정과정도 필요치 않았다.¹⁶⁾

또한, 모형의 성능을 비교하면 딥러닝을 이용한 모형이 회귀분석을 통해 구축된 모형보다 정확도가 높은 것으로 나타났다(표4-17참조). 특히, 보행사고율을 추정하는 경우에는 딥러닝을 이용한 모형의 R^2 가 0.42~0.60 정도로 현저히 높게 나타나 회귀분석으로는 설명이 어려운 다양한 독립변수들과 보행사고의 관계도 더 잘 설명할 수 있는 것으로 나타났다.

¹⁶⁾ 딥러닝에서는 활용된 독립변수가 종속변수와 관계가 낮은 경우에는 모형을 학습하는 과정에서 가중치가 작아지게 된다.

[표 4-17] 딥러닝과 회귀분석을 이용하여 예측한 보행사고 결과 비교

종속변수		모형	R ²		RMSE		MAE	
			Test	차이	Test	차이	Test	차이
보행사고 발생건수	전체	딥러닝	0.87	+0.19	12.02	-6.51	9.50	-3.95
		회귀분석	0.68		18.53		13.45	
	중상	딥러닝	0.72	+0.08	6.10	-1.45	4.78	-0.80
		회귀분석	0.64		7.55		5.58	
	경상	딥러닝	0.84	+0.22	7.81	-3.52	5.87	-2.39
		회귀분석	0.62		11.33		8.26	
보행사고율	전체	딥러닝	0.57	+0.33	2.25	-0.54	1.76	-0.39
		회귀분석	0.24		2.79		2.15	
	중상	딥러닝	0.42	+0.16	1.09	-0.12	0.85	-0.09
		회귀분석	0.26		1.21		0.94	
	경상	딥러닝	0.60	+0.40	1.33	-0.36	1.05	-0.22
		회귀분석	0.20		1.69		1.27	

3_장래 보행환경·행태 변화에 따른 보행사고 변화 예측

본 연구에서 구축한 보행사고 추정 모형을 이용하여 장래 보행환경 변화에 따른 보행사고 건수를 추정해 보았다. 모형의 구조는 앞에서 구축한 보행사고 발생 건수 추정 모형과 동일하게 은닉층 3개와 은닉층별 노드의 개수(64,32,16)로 설정하였고, 2019년 300개 역세권의 모든 데이터 셋을 이용하여 모형을 학습하였다.

장래 보행환경 변화를 설정하기 위해 스마트폰 이용의 변화가 보행환경에 영향을 주는 것으로 가정하였다. 즉, 본 연구에서 이용한 스마트폰 이용 데이터 중 최근 이용이 급격히 증가하며 보행환경에 영향을 주는 키보드와 배달 앱을 선정하고, 해당 앱의 사용이 증가하면 보행환경이 변화하는 것으로 설정하였다. 또한, 보행행태에 영향을 미치는 동영상, 음악 관련 앱의 이용자 수 변화도 장래 보행환경을 변화시키는 것으로 가정하였다. 선정된 4개 앱의 입력데이터 값을 각각 3배, 5배로 증가시키고 다른 입력변수는 동일하게 유지한 8개의 시나리오를 대상으로 보행사고 발생 건수의 변화를 추정하였다.

장래 보행사고 발생 건수 추정 결과를 살펴보면, 해당 앱의 이용이 증가할수록 보행사고가 증가하는 것으로 나타났다¹⁷⁾. 키보드 앱의 경우 현재보다 이용빈도가 3배, 5배 증가할 때 서울시에서 발생하는 보행사고가 각각 31.7%, 54.2% 증가하는 것으로 나타났다고, 배달 앱의 경우에도 20.0%, 41.7% 증가하여, 해당 앱들의 사용 증가가 보행환경을 악화시키고 보행사고를 유발할 수 있는 것으로 추정되었다(그림 4-25] 참조). 보행행태에 영향을 주는 동영상과 음악 앱 사용의 증가도 보행사고에 영향을 주었다. 특히, 보행 중 스마트폰 몰입을 유발하는 동영상 이용의 영향이 음악 이용보다 더 크게 나타났다(그림 4-26] 참조).

본 연구에서 추정한 장래 보행사고 발생 건수는 해당 앱 사용량의 영향만을 반영하였기 때문에 실제 보행사고 변화와는 다를 수 있다. 특히, 보행사고의 가장 큰 증가를 유발하는 키보드의 경우, 모형 학습 시 공유 키보드 활성화 초기인 2019년 데이터를 이용하여 키보드의 잦은 보도 이용 등 초기의 제도적 미비가 그대로 반영된 측면이 있다. 그럼에도 추정된 결과를 통해 제도적 보완, 이용자 의식 개선 등의 필요성을

¹⁷⁾ 2019년 서울시에서 발생한 전체 보행사고 건수는 9,896건이지만, 본 연구에서는 분석 단위인 역세권의 중복 지역이 존재하여 총 15,604건의 보행사고를 현재 기준으로 하였다.

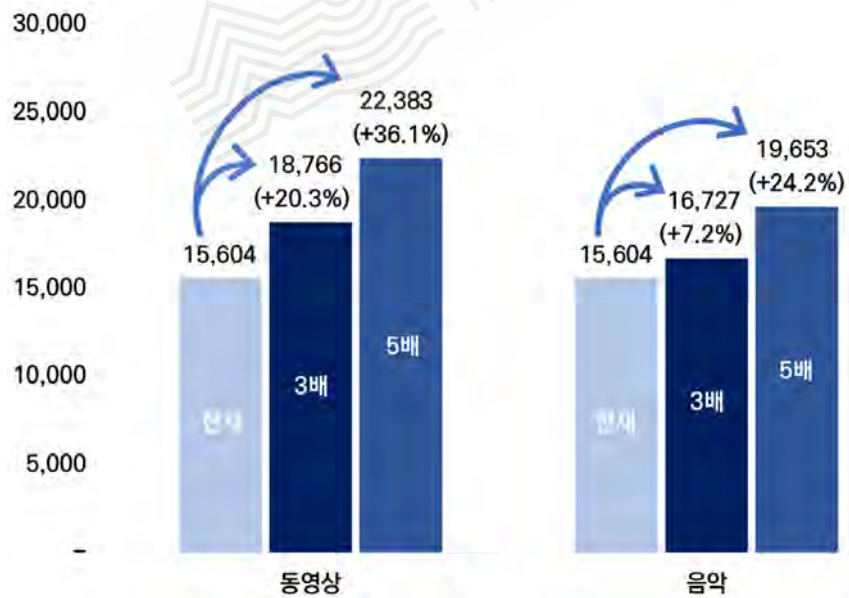
도출할 수 있고 다양한 보행친화 정책 수립의 기초 자료로 활용될 수 있을 것으로 기대된다.

(단위: 건)



[그림 4-25] 키보드, 배달 앱 사용 증가에 따른 보행사고 건수 추정

(단위: 건)



[그림 4-26] 동영상, 음악 앱 사용 증가에 따른 보행사고 건수 추정

4_딥러닝 분석의 시사점 및 향후 과제

본 연구에서는 최근 급변하고 있는 보행환경을 반영하기 위하여 다양한 데이터를 수집하여 보행사고 빅데이터를 구축하고, 이를 딥러닝을 활용해 분석하였다. 스마트폰 이용정보 등 새로운 데이터를 반영할수록 보행사고 변화를 보다 정확히 추정할 수 있었고, 특히 딥러닝을 적용하여 기존 회귀분석 등에서는 설명이 어려운 보행사고율도 높은 정확도로 추정할 수 있는 것을 확인하였다. 분석과정을 통해 도출된 시사점과 향후 과제는 다음과 같다.

우선, 보행 관련 데이터의 지속적인 수집이 필요하다. 본 연구에서는 보행사고 빅데이터를 구축하기 위해 다양한 분야의 데이터를 조사하였는데, 각 데이터의 구축시기가 상이하거나 단기간 구축 후 수집이 중단되는 경우가 많아 분석에 활용되지 못한 데이터가 많았다. 특히 인구, 자동차 등록 대수 같은 일부 사회경제지표를 제외한 데이터는 현재까지의 축적 기간이 길지 않아 장기적인 추세를 분석하는 데 한계가 있었고, 중간에 누락된 경우도 많아 데이터 전처리에 상당한 시간이 소요되었다. 본 연구를 통해 최근 보행행태와 보행환경이 변화하고 있고, 이러한 변화가 보행사고에 영향을 미치는 것을 확인한 만큼, 수집이 필요한 보행 관련 데이터를 선정하고 이를 지속적으로 수집하는 노력이 필요할 것이다.

또한, 데이터 수집의 공간적 단위가 상이하거나 모호한 경우가 많아 정비가 필요하다. 데이터의 공간적 단위가 상이한 경우, 이를 분석 단위별로 전환하는 과정이 필요하고, 이때 불필요한 가정이 개입되어 데이터가 왜곡될 가능성이 있다. 특히, 본 연구에서 이용한 스마트폰 이용 데이터의 경우 전국 단위 결과로만 제시되고 있어, 자치구 또는 역세권별로 세분화하는 데 어려움이 있었다. 본 연구에서 제시한 바와 같이 보행환경과 보행사고의 추세가 지역별로 다르게 나타나고 있기 때문에, 데이터의 공간적 범위를 보다 세분화할 필요가 있을 것이다.

딥러닝 적용을 통해 다음과 같은 모형 구축의 시사점이 도출되었다. 딥러닝 특성상 이용하는 데이터가 증가하거나 모형의 구조를 복잡하게 하는 경우 모형의 성능이 증가할 것으로 예상되었다. 하지만, 너무 많은 데이터를 이용하거나 모형의 구조가 지나치게 복잡해지는 경우 오히려 모형의 성능이 저하되는 경우가 발생하였다. 특히, 스마트폰 이용 데이터 중 보행환경과 상관이 적은 346개의 애플리케이션 데이터를 모두 적용하는 경우 추정 결과가 왜곡되는 현상이 발생하였다. 이러한 현상은 구축된 모형이

훈련 데이터에만 과대적합(Overfitting)하여 발생하는 것으로 추정되며, 이를 방지하기 위해 입력 데이터 범위와 모형 구조에 대한 분석가의 판단과 경험이 필요한 것으로 생각된다.

또한, 데이터의 공간적, 시간적 특성을 반영할 수 있는 발전된 딥러닝 알고리즘의 적용도 고려하여야 한다. 본 연구에서는 딥러닝 알고리즘 중 인공신경망 알고리즘을 이용하여 해당 역세권의 보행사고를 설명하기 위해 해당 연도, 해당 지역의 특성만을 반영하였다. 하지만, 시간적 특성을 반영하기 위한 알고리즘(RNN(Recurrent Neural Network)), LSTM(Long Short Term Memory Network)이나 공간적 특성을 반영하는 알고리즘(CNN(Convolutional Neural Network), GCN(Graph Convolutional Network))이 지속적으로 개발되고 있어 이를 적용하면 보다 정확한 추정이 가능할 것으로 기대된다.



05

맺음말



05. 맺음말

최근 보행행태가 변화하면서 보행사고의 발생 추이도 변화하고 있다. 지난 13년간 서울시에서 발생한 13.7만 건의 보행사고를 분석한 결과, 심각도가 높은 사망, 중상사고는 감소하고 있지만, 경상사고 발생은 증가하고 있는 것으로 나타났다. 보행사고의 피해보행자 중 20대 보행자의 사고 발생률이 증가하고 있어 인구지표의 변화만으로는 설명할 수 없는 연령대별 특성도 영향을 미치는 것을 알 수 있었다. 또한, 보행사고의 변화 추세는 지역별로도 상이하게 나타나고 있어, 보행사고를 설명하기 위해서는 지역 특성을 고려해야 함을 알 수 있었다.

특히 보행 중 스마트폰 이용은 보행행태를 크게 변화시키고 있다. 서울시민 1,000명을 대상으로 온라인 조사를 시행한 결과, 보행자의 대부분이 보행 중 스마트폰을 이용하고 있었고, 주의 산만을 유발하는 ‘게임’, ‘동영상’ 등의 보행 중 이용 비율도 상당히 높은 것으로 나타났다. 대부분의 보행자들이 다른 보행자의 스마트폰 사용으로 인해 불편함을 경험하고 있어, 스마트폰을 이용하는 보행행태가 보행사고를 유발하는 원인으로 작용할 수 있음을 알 수 있었다.

변화된 보행행태와 보행환경을 반영하여 보행사고를 설명하기 위하여 보행사고 데이터, 교통, 사회·경제 데이터, 스마트폰 이용 데이터를 포함하는 보행사고 빅데이터를 구축하고 딥러닝을 이용하여 분석하였다. 보행행태를 대변하는 스마트폰 데이터를 반영하는 경우 모형의 성능이 향상되는 것을 확인할 수 있었다. 또한, 딥러닝을 적용하여 빅데이터를 분석하는 경우, 기존의 회귀분석보다 모형의 정확도나 데이터 활용도에서 우수한 점을 알 수 있었다.

본 연구를 통해 빅데이터 구축과 딥러닝 적용의 가능성을 확인하였다. 하지만, 현재의 보행사고 관련 데이터는 구축기간이 상대적으로 짧고 데이터 수집의 공간적 범위가 상이하여 활용에 제한이 있었다. 특히, 보행행태 반영을 위한 스마트폰 이용 데이터가 전국 단위로 구축되어 공간적 세부단위로 활용하기에 제약이 있었다. 특히, 스마트폰

이용 데이터의 활용 범위는 보행사고 분석 이외에도 다양할 것으로 예상되어 보다 세분화된 데이터를 구축하는 방안을 마련해야 할 것이다. 또한, 딥러닝의 장점을 극대화하기 위해 현재 개발되고 있는 다양한 최신 알고리즘에 대한 이해와 적용 경험이 필요하고, 적절한 모형의 구조를 구축하기 위한 분석가의 경험 축적도 필요할 것이다.

본 연구에서 제안한 빅데이터 구축과 딥러닝 적용 방안은 보행사고 이외의 다른 분야에도 적용될 수 있을 것으로 기대된다. 특히, 본 연구의 분석 주제였던 보행사고는 발생 가능성이 현저히 낮아 특정 모형을 통해 분석하기에는 한계가 있었다. 하지만, 본 연구에서 제안한 빅데이터 구축-딥러닝 적용 분석 방안은 최근 증가하는 새로운 데이터를 활용한 교통혼잡 예측, 도심의 적정 차량 대수 추정, 대중교통 수요 추정 등 다양한 분야에 활용이 가능할 것으로 기대된다.



참고문헌

- 곽호찬 외, 2014, “유전자 프로그래밍을 이용한 고속도로 사고예측모형”, 「대한교통학회지」, 32권 4호, pp.369-379.
- 류종득, 2018, “딥러닝을 이용한 고속도로 교통사고 예측모델 개발”, 「아주대학교」,
- 박철영, 이수기, 2016, “가로환경 특성이 보행자 교통사고에 미치는 영향 분석”, 「한국도시설계학회지」, 17권 3호, pp.105-121.
- 백태현 외, 2014, “간선도로 기능별 보행사고 심각도 분석과 모형 개발”, 「한국도로학회 논문집」, 16권 4호, pp.111-118.
- 이하식 외, 2019, “유동인구를 고려한 보행사고 노출률 추정”, 「대한교통학회지」, 37권 4호, pp.302-321.
- 서울연구원, 2019, 「‘걷는 도시, 서울’ 정책효과와 향후 정책방향」.

Abstract

Big Data Analysis of Pedestrian Accidents in Seoul Using Deep Learning

Youngjun Han · Hasik Lee

Due to the application of various pedestrian-friendly policies, the number of pedestrian accidents in Seoul has decreased. However, the use of smart phones has changed the walking behaviour, and has worsened the pedestrian environment due to disruption of delivery motorcycles or electric scooters. As a result, the decline in pedestrian accidents has abated, whereas the number of minor accidents has increased. A survey of walking behaviour among Seoul citizens indicated that 69.0% of pedestrians use smart phones while walking, and 44.9% and 23.6% of them watch a 'Video' or play a 'Game' respectively, which leads to vulnerable pedestrian behaviour. In order to reflect this change in pedestrian behaviour, we developed pedestrian big data including smart phone usage, and modelled pedestrian accidents using deep learning techniques. Compared with traditional linear regression, the developed model provides a better estimate of accident rate. We also predicted pedestrian accidents with respect to changes in future pedestrian environment, such as increasing the use of electric scooter. The framework of this study based on deep learning and big-data analysis may be applicable to other areas including prediction of traffic congestion or demand for public transit.

Contents

01 Introduction

- 1_Background and Purpose of the Study
- 2_Main Contents and Research Methods

02 Analysis of Pedestrian Accidents in Seoul

- 1_Pedestrian Accident Data
- 2_Analysis of Pedestrian Accidents
- 3_Implications of the Analysis

03 Analysis of Pedestrian Behaviour in Seoul through a Survey

- 1_Survey Overview
- 2_Smart Phone Usage Status of Pedestrians while Walking
- 3_Smart Phone Usage Status of Driver while Driving
- 4_Citizen's Perception on How to Control Smart Phone Usage while Walking

04 Estimation Pedestrian Accidents Using Big Data and Deep Learning

- 1_Building Big Data for Explaining Pedestrian Accidents
- 2_Building and Application of Deep Learning Model for Estimating Pedestrian Accidents
- 3_Estimation of Changes in Pedestrian Environment and Behavior in the Future
- 4_Implications and Future Study of Deep Learning Analysis

05 Conclusions

빅데이터와 딥러닝 활용한
서울시 보행사고 분석과 시사점

서울연 2019-BR-25

발행인 서왕진

발행일 2020년 7월 31일

발행처 서울연구원

ISBN 979-11-5700-524-6 93530 6,000원

06756 서울특별시 서초구 남부순환로 340길 57

이 출판물의 판권은 서울연구원에 속합니다.