



인공지능을 활용한 교통류(Traffic Flow) 예측

장 요약*

Abstract

최근 우리의 일상은 스마트폰과 인터넷의 조합으로 인해 한없이 편리해졌고, 이는 우리의 이동성에도 큰 변화를 가져왔다. 가본 적도 없고 생소한 장소를 손쉽게 검색해서 통행할 수 있고, 더 나아가 아직 일어나지도 않은 가깝거나 먼 미래의 통행도 미리 경험해 볼 수 있다. 이 모든 것들이 전문가들에 의해 고안된 잘 만들어진 ‘경험 교과서’를 인공지능이 잘 학습한 덕분인데, 이 ‘경험 교과서’를 잘 들여다보면 비단 발달한 인공지능 능력 뿐만은 아닌 듯하다. ‘경험 교과서’를 잘 집필할 수 있도록 바른 데이터와 양질의 자료를 수집하고 관리하기 위한 노력도 한몫하였고, 그사이에 눈부시게 발달된 하드웨어 성능 또한 일조한 것으로 보인다. 본고에서는 교통류 예측 분야에서 ‘경험 교과서’에 필수적으로 요구되는 주요 구성 요소들과 예측을 위한 최근까지의 흐름과 전개 방법, 인공지능의 발달로 달라진 변화내용들을 다루었다.

I. 서론

휴대전화 내비게이션 앱 등을 이용해서 길 찾기 서비스를 경험해 본 적이 있는 사람들은 알 것이다. 휴대전화 내비게이션 서비스가 가본 적도 없는 길들을 이리저리 잘도 찾아줄 뿐만 아니라, 내비게이션 앱에서 이야기하는 도착예정시간도 거의 정확하다. 최근 내비게이션 앱에서는 “미래운행정보” 정보 등도 함께 제공하는데, 현재 통행시간이 아닌 장래 특정 시점에 출발했을 때 예상되는 목적지까지의 예상 통행시간을 계산하여 제공한다. 현재 시점 도착예정시간은 수집된 도로 교통상황을 바탕으로 계산된다 치더라도, 그 누구도 겪어

* 국토연구원, 국토모니터링연구센터, 국토데이터랩 팀장, ycanns@krihs.re.kr

본 적 없고 알 수 없는 미래 도로 교통상황은 어떻게 계산된 것일까? 정답부터 이야기하자면, ‘경험’에 의한 예측이다. 그리고, 그 경험의 주체는 사람이 아닌 인공지능(Artificial intelligent, A.I.)이다. 대동여지도를 만든 김정호 선생님도 아니고 제아무리 많이 경험해 봤다고 한들, 전국 방방곡곡의 모든 도로를 전체 시간대에 걸쳐 경험한 이가 누가 있을까? 이 경험에 의한 예측은 전문가에 의해 잘 설계된 인공지능이 시간과 공간의 특징에 맞도록 적절한 학습을 한 뒤, 학습된 결과를 바탕으로 미래 도로 교통상황을 예측하고 이에 부합하는 통행시간을 제시하게 되는 논리이다. 즉, 경험적으로 가장 근접한 미래 도로 교통상황을 예측할 수 있는 규칙은 사람(전문가)이 만들고, 그 규칙을 잘 학습한 인공지능은 거의 모든 시간대에 걸쳐 전국에 교통상황을 예측한다. 본고에서는 어떤 데이터와 전문적인 학습 규칙을 바탕으로 교통상황이 학습되고, 신뢰도 높은 예측에 이르기까지 진행되는 일반적인 절차 등은 어떻게 이루어지는지를 이야기하고자 한다.

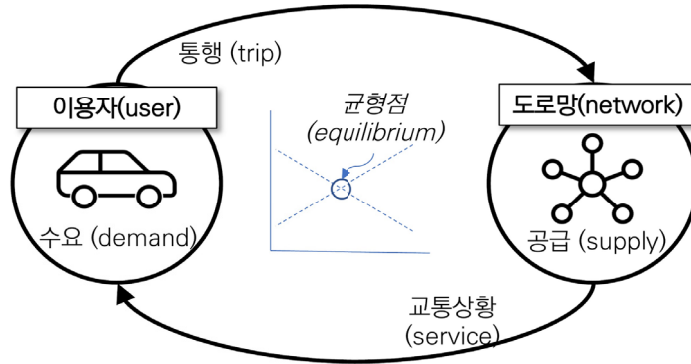
II. 교통의 기본적인 개념: 통행(Trip)

일반적으로 운전자(user)의 통행은 도로 서비스 상황에 따라 출발시간, 이동 경로 등 목적지까지의 가장 이상적인 조합이 고려되어 결정되고, 도로 기하구조가 복합적으로 이루어진 도로망(network)은 운전자 수요(demand) 정도에 따라서 혼잡 또는 원활 상태로 도로 서비스 상황이 반영된다. 다시 말해, 도로망 상황에 따라 교통 서비스(service)는 다양하게 변화하고, 해당 서비스의 소비자인 운전자(demand)는 다양하게 변화하는 교통 서비스의 공급상태에 따라 최적의 균형(equilibrium)의 조건에서 통행이 이루어지게 된다([그림 1] 참조). 수요와 공급 간의 균형에 따라 통행이 성사되기 때문에, 시시각각 변화하는 이러한 복합적 상호작용을 고려하여 장래 교통상황 예측을 한다는 것은 여간 쉽지 않은 일이다. 하지만, 현재 도로망과 이용자 간 균형에 대한 적절한 이해는 현재 교통상황 파악뿐만 아니라, 장래 교통상황 예측과 예측에 따른 대응에 이르기까지 그 이점이 다양하므로 간과해서는 안 된다.

통상, 이 ‘균형’을 이해하기 위한 정량적 척도로 ‘시간’ 가치를 사용한다. 도로망 혼잡도와 목적지까지의 접근성, 통행시간 변화(delay) 등을 종합적으로 고려하여 이루어지는 일련의 통행 의사결정 과정 등을 운전자 편익에 따른 지불의사(willingness-to-pay)로 환산하여 계산한다. 이는 곧, 통행시간 가치(value of time, VOT)로 환산되어 도로 혼잡통행료 징

수¹⁾, 고속도로 통행료 징수, 적정 대중교통 요금 결정 등 다양한 정책 의사결정 과정에 폭넓게 활용된다.

그림 1 운전자(user 또는 demand)와 도로망(network 또는 supply)



출처: 저자 작성

Ⅲ. 교통류(Traffic flow) 모니터링과 예측의 중요성

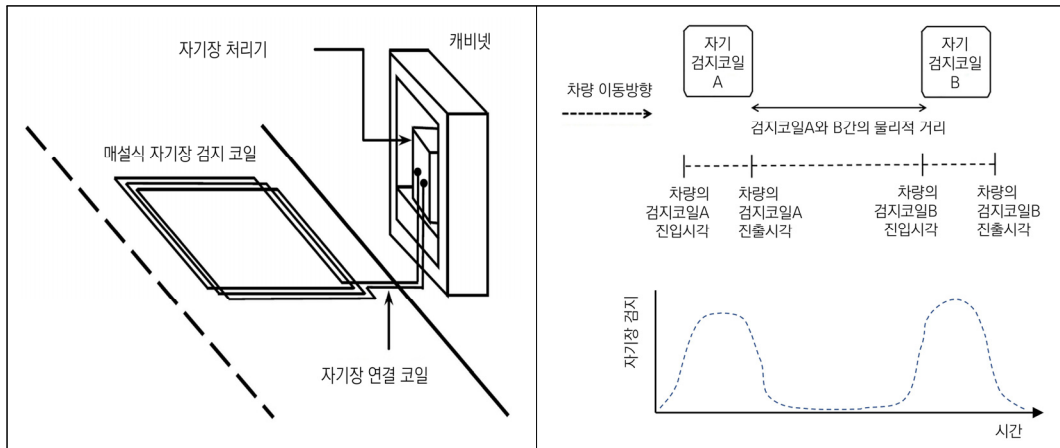
저마다의 목적에 의해 발생한 통행의 집합이 도로상에서 같은 방향으로 움직이는 흐름을 교통류(traffic flow)라 할 수 있는데, 도로를 이용하는 교통류의 움직임을 마치 수로관²⁾ 등을 이동하는 물결 흐름(flow)에 비유하기도 한다. 이러한 교통류의 흐름을 적절하게 이해하고 주기적으로 모니터링(monitoring)하는 것은 매우 중요한 교통정보로 활용되고 있다. 이는 현재의 교통운영(traffic operation)뿐 아니라, 도시계획 및 토지변화에 따른 장래 수요 예측 등에도 중요한 정보가 될 수 있기 때문이다. 교통운영에서는 실시간 교통류 상황을 모니터링하여 도로를 이용하는 이용자들에게 도로 혼잡상황이나 돌발상황 등에 대한 정보를 제공하여 합리적인 의사결정을 할 수 있도록 돕는가 하면, 도시계획에서는 도시를 통행하는 교통량을 고려하여 장래 신규시설 입지 적절성을 판단하는 데 활용하기도 한다. 교통사고 등 돌발상황 발생 예측을 통해 도로전광표지(가변정보판, variable message sign, VMS)에 표출하거나, 라디오 또는 문자 메시지 등으로 미리 알림이 가능하다. 또한, 위험도

1) 서울시 남산 1, 3호 터널 통과 차량에 대하여 혼잡통행료가 징수 운영 중이다

2) 도로에는 신호, 도로 폭, 차선 수, 설계속도, 도로 길이 등 물리적으로 고려해볼 수 있는 용량(capacity)의 요소들이 있다.

예측을 통해 고속도로 순찰차량(patrol car) 등을 위험도가 높은 장소와 시간에 배치하여 운전자들의 주의 상기와 위급상황 발생 시 즉각 대응할 수 있도록 예비할 수도 있어 도로망의 효율적인 운영에 특히 중요한 정보가 된다. 이러한 교통정보는, 개인정보 노출의 우려 때문에 개별차량에 대한 식별과 취득 정보를 수시로 모니터링하기 보다는 도로를 이용하는 교통류의 움직임을 다양한 방법으로 데이터화하여 수집하고 활용하고 있다.

그림 2 전통적인 매설방식 교통류 검지 방법 예시



출처: (왼) Lamas-Seco외, 2016의 그림 1, p.3을 이용하여 저자 한글 번역,(오) Mocholi Belenguer외, 2019의 그림2, p.4를 이용하여 저자 수정 및 한글 번역

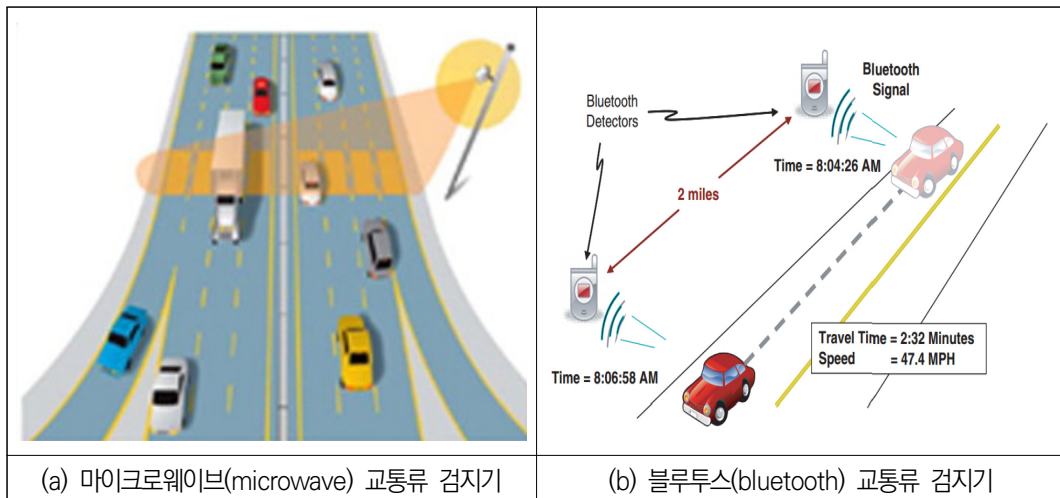
IV. 교통 데이터

전통적 교통류 관측은 도로에 매설된 전자기식 검지기 등을 통해 수집되었다. [그림 2]는 매설식 자기장 검지(inductive loop detector) 코일을 통한 차량 검지기의 예시인데, 작동 원리는 간단하다. 도로 노면에 사각형 등의 형태로 홈을 만든 뒤, 그 안에 자기장 코일 등을 매설한다. 매설된 코일은 도로 바깥쪽에 자기장 처리기 등을 설치하여 연결하고, 지속해서 일정량의 전류를 공급한다. 전도체의 성격을 가진 차량이 매설된 코일 위를 통과하는 경우 코일에 흐르는 전류에 변화가 발생하게 되는데, 이러한 전류의 변화를 감지해서 차량 수와 차종 등을 기록하는 형태이다. 도로를 이용하는 차량의 속도 검지를 위해서는, 유사한 검지 코일을 1m 등 인접한 노면에 수직으로 평행 매설하고 이 위를 차량이 통과하게 되면 전류의 변화를 연달아 감지하도록 구성한다. 차량이 두 코일을 통과하는 데 걸린 시간의

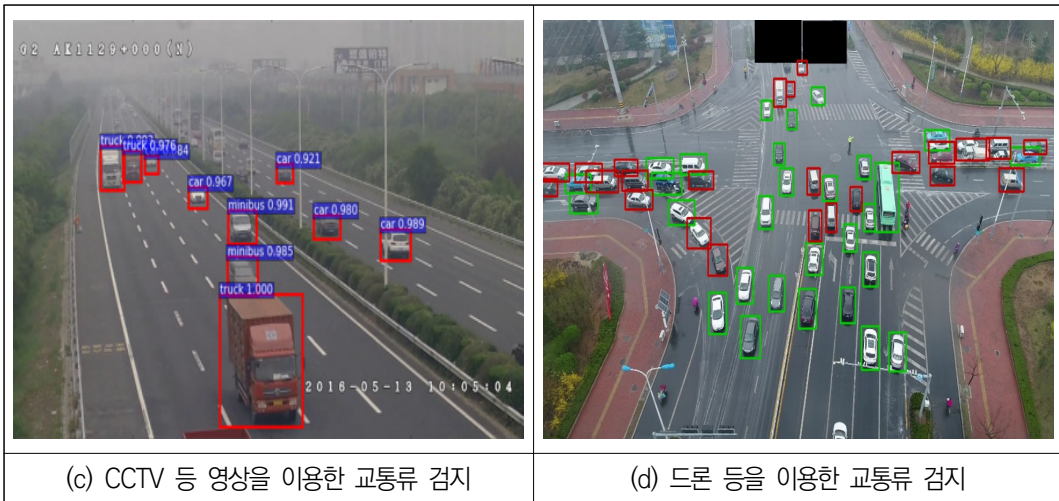
차와 두 코일 간의 물리적 거리를 환산하면, 속도(즉, 거리/시간, 거리를 시간으로 나눈 값)가 산출되는 형태이다. 통상적으로(30초 등) 일정 간격을 두고 교통량과 속도 정보를 지속해서 수집한다.

최근에는 블루투스 등 센서를 이용한 익명화된 검지기, 마이크로웨이브 등을 이용한 교통류 검지기에서 폐쇄회로TV(CCTV) 등을 활용한 영상검지기에도 그 검지 방식을 확대해 나가고 있다. 특히 CCTV 등 영상장치를 활용한 교통류 검지 방법은, 딥러닝 등 인공지능 분야 기술력과 하드웨어 성능이 최근 급성장하면서 그 활용도가 높아지고 있는데, 기존 인프라를 심분 활용하면서 교통류 검지 사각지대에 속한 지방부 및 소로 등에도 활용이 가능하다는 강점이 있다([그림 3]참조). 영상장치의 활용도는 더 나아가, 드론 등 무인기(Unmanned aerial vehicles, UAV)에도 그 범위를 확장하였는데, 체코의 한 기업인 DataFromSky³⁾는 이러한 트렌드에 맞추어 교통류뿐만 아니라 보행자, 자전거 등을 한 번에 수집할 수 있는 서비스를 제시하여, 그 활용도를 회전교차로의 주행궤적(trajecory) 및 안전성 분석, 주차장 주차시간 계산, 대규모집회 군집 인원 집계 등 다양하게 확장하여 운영 중이다([그림 4] 참조). 해당 기업의 홍보영상은 인공지능을 재미있게 활용한 사례로 <부도 1>에 첨부된 QR 코드를 통해 확인할 수 있다.

그림 3 최근 활용되고 있는 다양한 교통류 검지기 예시



3) DataFromSky(<https://datafromsky.com/>, 2022년 5월 26일 검색)



출처: (a) Manaa 외, 2015, (b) Dabiri and Kevin, 2018, (c) Chen 외, 2018, (d) Wu 외, 2019

그림 4 CCTV, 드론 등을 이용한 영상기기와 인공지능이 접목된 교통류 검지 예시

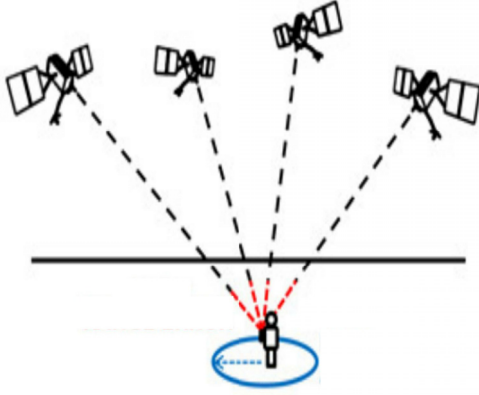
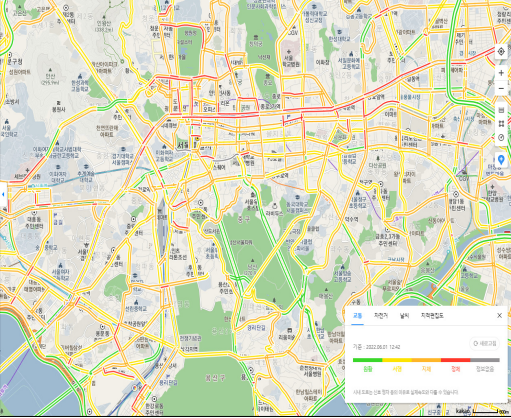
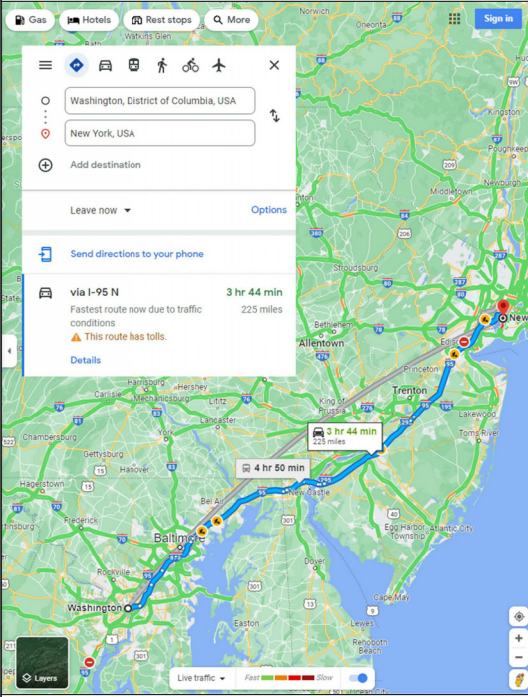


출처: DataFromSky 홈페이지 및 홍보 유튜브 영상⁴⁾

스마트폰 보급량이 늘고, 위치기반 서비스가 늘어나 위치정보에 대한 사용자들의 거부감이 경감되면서 스마트폰 이용자들의 이용행태 또한 자연스럽게 데이터화 되고 있는데, 그 대표적인 예가 스마트폰 내비게이션 앱이다. 스마트폰 이용자의 익명 위치정보(Global positioning system, GPS)를 일종의 관측 표본화(observation sampling)하여, 이용자의 접속 시간에 따른 위치정보 변화를 활용한 교통상황 정보를 내비게이션을 비롯한 지도 서비스에 실시간

4) DataFromSky 홍보 유튜브 영상(https://www.youtube.com/watch?v=xwQky6aM_lo, 2022년 5월 26일 검색)

그림 5 스마트폰 서비스를 활용한 교통상황 서비스 예시

	
(a) GPS 원리 예시	(b) GPS 정보를 활용한 교통상황 지도 서비스
	<pre> <DistanceMatrixResponse> <status>OK</status> <origin_address>Washington, DC, USA</origin_address> <destination_address>New York, NY, USA</destination_address> <row> <element> <status>OK</status> <duration> <value>14078</value> <text>3 hours 55 mins</text> </duration> <distance> <value>367654</value> <text>228 mi</text> </distance> </element> </row> </DistanceMatrixResponse> </pre>
(c) Google Map교통상황 및 경로탐색 서비스	(d) Google Map API 제공예시(XML 방식)

출처: (a) Kumaraige, 2018의 그림 2, p.23을 수정, (b) 카카오지도의 서울시 일대 교통상황 지도 서비스 예시⁵⁾, (c) Google Map⁶⁾ (d) Google Map API⁷⁾

5) 카카오지도(<https://map.kakao.com>, 2022년 6월 1일 검색)

6) 구글맵(Google Map, <https://www.google.com/map>, 2022년 6월 1일 검색)

7) Google Map API, distance matrix

으로 제공 중이다. 또한, 교통 데이터에 관한 관심과 필요성이 증대되면서, 내비게이션 지도 서비스 기업들은 응용 프로그래밍 인터페이스(application programming interface, API) 방식을 통해 누구나 사용할 수 있도록 외부에 공개하기 시작하였고, 이는 새로운 교통 데이터로 자리매김하게 되었다. 스마트폰 서비스를 활용한 교통상황 서비스 예시를 도식화한 [그림 5]는 지도 서비스에 교통 혼잡 상황을 단계별 색으로 표현하는가 하면(b), 원하는 출발지에서 목적지까지의 경로와 거리 정보를 지도상에 안내해주고 예상 시간 등을 제공해주기도 한다(c). 하지만, 이러한 서비스가 API를 통해 데이터로 제공(d)되기 시작하면서, 전 세계 대부분 지역의 통행시간과 경로 정보 등에 손쉽게 접근할 수 있게 되었다.

V. 일반적인 교통류 예측 과정

인공지능을 이용한 교통류 예측의 일반적인 과정을 [그림 6]에 도식화하였다. 인공지능을 이용한 예측이라 하더라도, 예측 모형만 바뀐 형태가 대부분이기 때문에 전통적인 예측 방법론에서 크게 벗어나지는 않는다.

제일 먼저, 교통류 예측에 활용되는(적합한) 「데이터 수집」 단계가 필요하다. 고전적으로는 매설식 차량 검지기 등이 주를 이루었으나, 최근 발달 된 응용 기술과 방법론 고도화로 소셜미디어 등 다양한 정보들이 데이터로 활용되기도 한다(Ni 외. 2014, Essien 외. 2021, Yao and Qian, 2021).

가장 고도화된 기술과 숙련도가 요구되는 다음 단계는 바로 「데이터 정제」 단계이다. 최근 급부상한 인공지능 분위기에 따라 인공지능 학습에 필요한 고품질 (빅)데이터에 관한 관심이 뜨거워졌고, 바른 출력값을 위한 올바른 입력값(Garbage-in-garbage-out, GIGO)의 중요성이 주목받기 시작했다. 최근에는 데이터과학자(Data Scientist)라는 신종 업종이 세계적으로 가장 촉망받는 직업 중 하나로 급부상할 정도로 그 열기가 뜨겁다.⁸⁾ 최종 활용 목적에 따라 데이터 정제와 모형의 선택이 달라지기는 하나, 이 단계에서 수집된 데이터를 정제하고 향후 사용될 입력변수들을 선별하는 상관도분석 단계 등도 필수적으로 요구된다. 데이터가 지나치게 방대한 경우에는 랜덤 샘플링기법 등으로 접근하는 것도 방법이나, 대

(<https://developers.google.com/maps/documentation/distance-matrix/start#xml>, 2022년 6월 1일 검색)

8) 하버드 비즈니스 리뷰, "Data scientist: the sexiest job of the 21st century", 2012년 10월 호

(<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>, 2022년 5월 23일 접속)

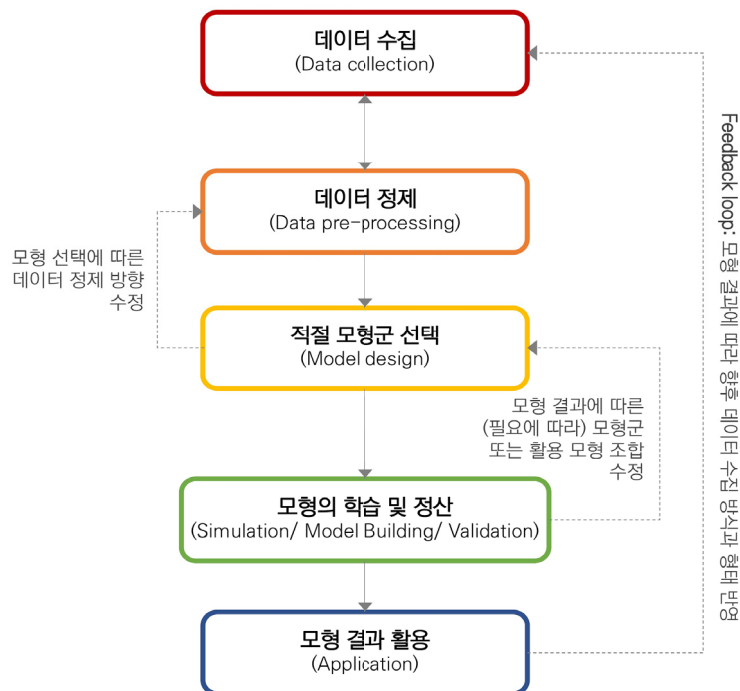
부분 경우는 데이터가 많지 않은 경우가 문제가 된다. 교통류 예측 문제에 따라 다르겠지만, 예측하고자 하는 시간 범위에 통상 3배에서 4배 정도의 과거 데이터가 존재했을 때 과소적합(underestimation) 또는 과대적합(overestimation)의 오류에 빠질 위험이 줄어들 수 있다. 예를 들어, 현시점을 기준으로 1시간 후의 교통류를 예측하고자 한다면, 같은 위치에서 관측된 3~4시간 이전부터의 연속적인 데이터가 최소한으로 요구되는 셈이다. 하지만 통상적으로는, 3시간 이전부터의 데이터뿐만 아니라, 1주일 전에서부터 한 달 전까지 같은 요일 데이터(lookback data) 구성이 예측력 높은 모형 구축을 위해 필수적으로 요구된다. 수집된 데이터에 비해 유실된 정보가 지나치게 많을 경우, 보간법(interpolation) 등의 해결책도 가능하겠지만, 이런 경우 데이터 분포가 편중되는 단점이 있어서 데이터 특성값(eigen value)이나 푸리에변환(Fourier transform) 등의 단계를 통해 데이터 특성이 드러나도록 전처리하는 것도 방법이다. 데이터의 범주가 다양하고 그 위계가 지나치게 복잡한 경우에는 다차원 축소 알고리즘 등의 단계를 거쳐 계산에 요구되는 부담을 덜어주는 방법도 널리 쓰인다(Pearson, 1901, Van der Maaten and Hinton, 2008). 데이터 축소알고리즘은 다소 생소할 수도 있으니, 추가적인 경험이 필요한 경우에는 <부도 1>에 첨부된 시뮬레이션 시각화 도구를 활용하여 참고하면 좋다.

「적절 모형군 선택」 단계에서 필수적으로 이해하고 접근해야 하는 것이 바로 모형군 학습 방식이다. 미국 사람들이 모국어인 영어에 능하고 한국 사람이 한국어를 잘하듯이, 인공지능 모형들도 그 태생이 있어 강점과 약점에 따라 선별적으로 활용해야 한다. 교통류 예측에 주로 활용되는 인공지능 모형군은 대략 세 가지인데, 지도학습(supervised learning), 자가지도학습(self-supervised learning), 그리고 강화학습(reinforcement learning) 등이 그것이다. 지도학습은 회귀분석과 마찬가지로, 구하고자 하는 입력변수와 종속변수 간의 관계가 확실하고, 예측하고자 하는 시점에도 독립변수가 분명하게 존재하는 경우에 활용하는 모형군이다. 대표적인 모형으로는 랜덤포레스트(random forest)와 인공신경망(Artificial neural networks, ANNs) 등이 있다. 자가지도학습은 독립변수와 종속변수에 대한 구분이 모호한 경우 활용되는데, 입력데이터(독립변수) 분포를 바탕으로 종속변수의 특징을 유추해 내는 것이 특징인 모형이다. 교통류 예측 분야에서는 주로 시계열적 변화(variation)가 많거나 수집된 데이터에 유실 데이터가 많은 경우에 활용되는데, 대표적인 모형이 오토엔코더(AutoEncoder) 모형 등이 있다(Lv 외 2014). 마지막으로 강화학습은, 도로기하구조 변화에 따른 운전자 형태 예측 시뮬레이션 등에 활용되는 학습방법으로, 어떠한 행동주체가 되는 객체(agent)와 환경(environment) 간의 행동과 보상 관계를 규명하고 이를 기계가 학습함으로써, 점차 객체를 환경에 맞게 적응하도록 유도하는 것이 특징이다(Valkanis 외, 2020,

Nie, 외. 2020).

모형군 선택 이후부터는 요구되는 단계별 조건이 까다롭지 않다. 모델마다 필수적으로 요구되는 결정계수(hyper-parameter)들을 유전자알고리즘(genetic algorithm, GA) 등과 같은 휴리스틱 최적화 방법론을 이용하여 선택된 모형과 정제된 데이터를 학습을 진행한 뒤 예측도 높은 결과값을 도출하여 활용한다.

그림 6 인공지능을 이용한 교통류 예측의 일반적인 과정

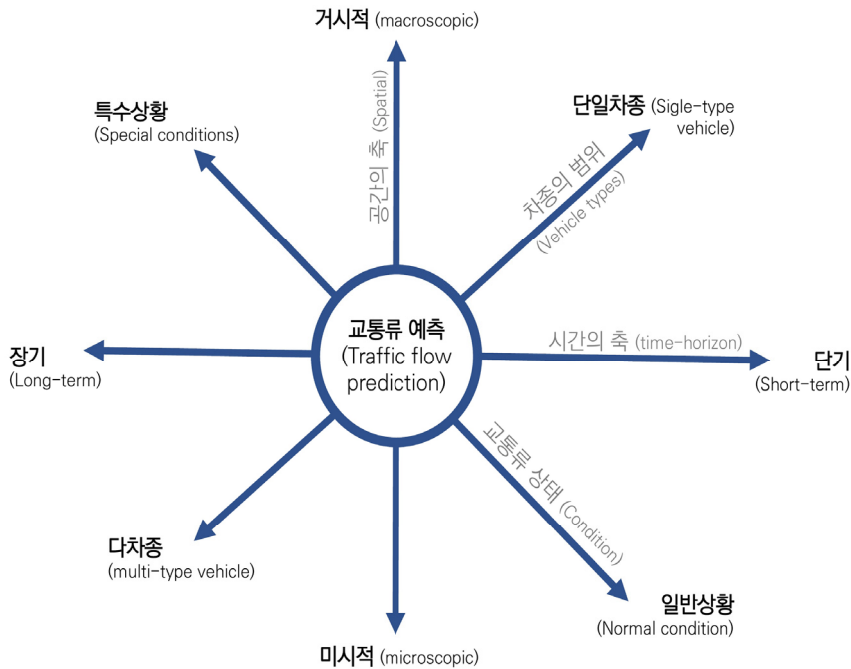


출처: 저자 작성

VI. 교통류 예측의 범위와 종류

교통류 예측은 예측하고자 하는 범위와 목적에 따라 다양하게 구별되고 그 특성이 결정된다. [그림 7]은 예측하고자 하는 범위와 목적에 따라 구별되는 교통류 예측 모형군의 구분을 도식화한 내용이다. 본고에서는 네 가지의 축으로 구별하여 기술하였지만, 시·공간의 축이 나머지 축보다도 더욱 그 중요성이 강한 것이 특징이다.

그림 7 교통류 예측의 범위와 특징 구분



출처: 저자 작성

가. 시간의 축(time-horizon)에 따른 구분

먼저, 예측하고자 하는 시간 축(time-horizon)에 따라서 단기(short-term)와 장기(long-term) 모형으로 구분이 되는데, 짧게는 현재를 기준으로 30초 후부터 30초에서 5분 내의 간격으로 대략 30분 후까지를 단기 모형군으로 취급한다. 대개 이런 모형군은 교통사고 발생 등과 같은 돌발 사고 예측에 주로 활용되는데, 예측되는 모형은 차량 내 시스템(in-vehicle system)에 탑재되어 운전자 사고 예측 알림 등에 활용되기도 한다. 30분 내외의 예측이 활용되는 경우에는 운전자보다는 교통운영관리자(traffic operator)가 활용하는 경우가 많은데, 돌발상황 또는 교통흐름 와해(breakdown) 등이 예상되는 경우, 해당 시간대와 지점 또는 구간에 도로순찰(patrol) 담당자나 경찰을 투입하여 교통류 내의 운전자들에게 경각심을 유도하기 위한 목적으로 주로 활용된다. 장기 모형군의 활용은 대략 30분 또는 1시간 단위로 현재를 기준으로 30분 이후부터 길게는 한 달 후까지 예측이 가능한 모형을 장기 모형군으로 취급한다. 장기모형은 그 예측의 범위가 상대적으로 길어서 예고가 가능한 특정 지역과 시기의 스포츠 행사나 공사현장 운영에 따른 교통류 영향 예측에 주로

활용된다.

나. 공간의 축(spatial-level)에 따른 구분

공사 현장(work zone) 작업에 따른 운전자 행태분석이나, 운전자 졸음운전 상태에서의 사고인지 반응과 이로 인한 교통류 속도 변화 예측과 같이 분석의 공적 범위를 미시적(microscopic)으로 볼 것인지 아니면, 서울 도심에 교통사고 발생 등으로 얼마만큼의 지체가 발생할지, 이번 추석 연휴에 어느 만큼의 귀경길 차량이 발생할 것인지 등 분석의 범위를 거시적(macroscopic)으로 볼 것인지에 따라 모형 선택과 시공간적 데이터 구성 또한 다르게 전개된다.

다. 교통류 상태(condition)와 차종의 범위에 따른 구분

앞서 언급한 시공간 범위에서 많은 부분이 중첩되기는 하나, 예측하고자 하는 교통류 상태가 출퇴근 등 일반적인 교통류 상태인지 아니면 교통사고나 폭설 등 특수상황에 따른 교통류 상태인지에 따라 시공간적 데이터 수집과 구성이 구별되어 전개된다. 이러한 특수성은 차종의 범위에 따른 구분에서도 유사하게 적용된다.

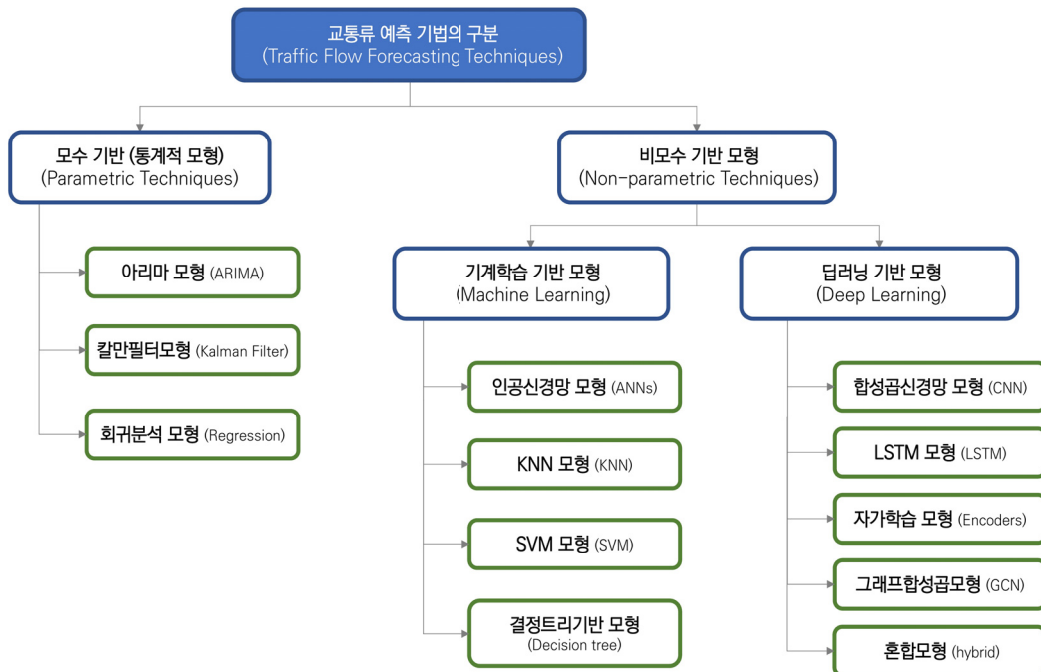
Ⅶ. 인공지능을 이용한 교통류 예측

전통적인 교통류 예측은 통계 기반 모델이 주를 이루었다. 하지만, 누적되는 데이터가 점차 방대해지고 이러한 방대한 데이터를 처리할 수 있는 기계학습에 관한 관심이 집중되면서, 다양한 기계학습 모형이 소개되었다. 기계학습 범주 중 하나인 딥러닝 모형도 최근 하드웨어 시장의 급성장과 함께 그 활용도가 지속해서 고조되고 있다.

모수모형이라고 불리는 Parametric 기반 모형들은 전통적인 통계적 모형으로, 시계열 특징이 강한 교통류 예측 범주에서는 ARIMA(Autoregressive integrated moving average) 모형 등이 가장 널리 활용되었다. ARIMA 모형의 예측력이 우수하기는 하나, 관측된 교통량 데이터의 시계열 관계가 복잡하거나 그 양이 방대한 경우 전수 데이터를 기반으로 학습이 이루어지는 ARIMA 모형은 그 효용성이 제한적인 것이 특징이다(Smith and Michael, 1997, Smith 외, 2002).

흔히들 이야기하는 인공지능 기반 모형을 비모수(Non-parametric)모형이라 할 수 있다. 딥러닝(Deep learning)은 기계학습(Machine learning)에서 파생된 학습 방법으로, 딥러닝 모형 대부분은 인공신경망(ANNs)에 그 모태를 두고 있다. 일반적인 인공신경망 모형은 입력층(input layer)-은닉층(hidden layer)-출력층(output layer)으로 이루어져 있고 각 층은 통상적으로 단일 층으로 구성된다. 전통적인 인공신경망 모형은 입력층과 은닉층 사이에 교신역할을 하는 전달계수(weight factor)를 조절하면서 결과값을 출력층으로 전달하는 형태로, 전달계수를 조절하는 은닉층의 역할이 상당히 중요한데, 딥러닝 모형은 본격적인 학습이 이루어지는 이 은닉층을 복수 개 이상으로 구성하여 기계의 학습을 심도(deep) 있게 할 수 있도록 고안하면서 붙여진 이름이다. 교통류 예측에 가장 널리 활용되는 모형으로는 결정트리기반의 랜덤포레스트 모형이 있는데, 블랙박스(black-box)로만 여겨지던 기계학습 범주에 독립변수의 민감도를 확인할 수 있는 통계수치를 제공하는 특징과 오차에도 높은 예측력(robust)으로 교통류 예측뿐만 아니라 기타 예측 문제에도 그 활용도가 매우 높다(Chang 외, 2020, Hou 외, 2017).

그림 8 교통류 예측 범위와 특징 구분



출처: Medina-Salgado 외, 2022 p.2의 그림1과, Smith and Michael, 1997, Smith 외, 2002,의 내용을 바탕으로 저자 각색

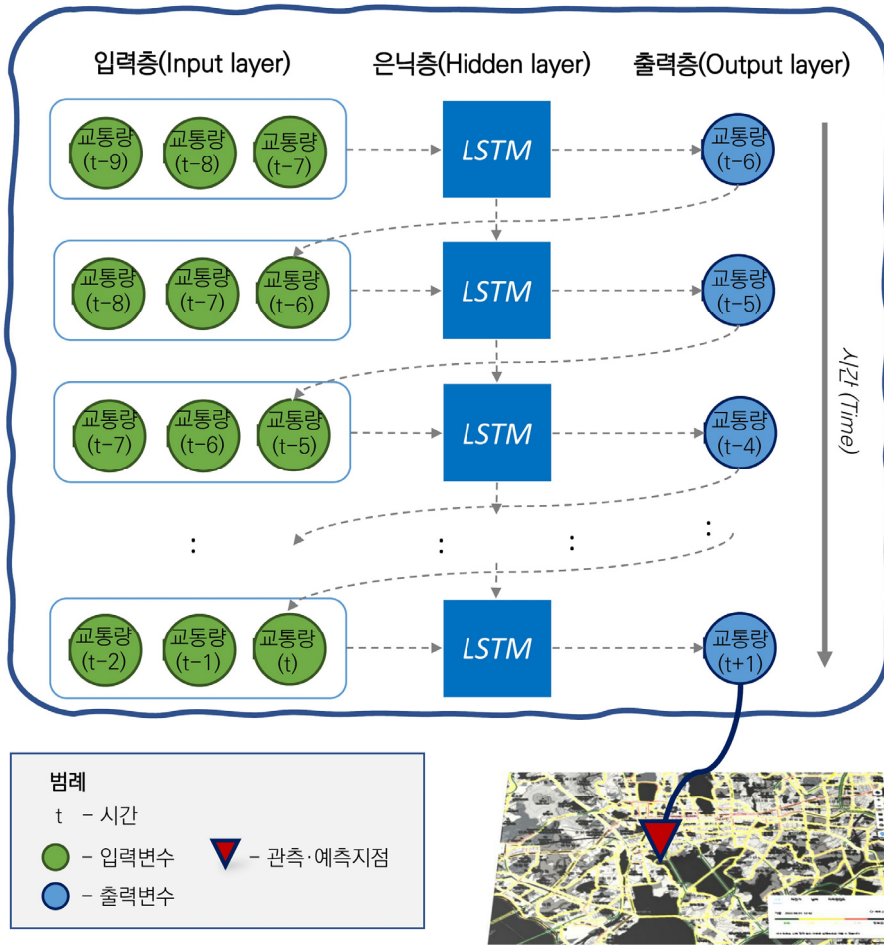
최근 다양한 시도가 이루어지고 있는 딥러닝 범주에서는 합성곱신경망(Convolutional neural network, CNN)이 대표적인데, 2012년 세계 이미지 분류 경진대회(ILSVRC⁹⁾)에서 토론토대학의 Geoffrey Hinton 교수가 이끄는 AlexNet 기반의 CNN이 우승하게 되면서 그 관심이 시작되었다.¹⁰⁾ 합성곱신경망 모형학습은, 입력데이터를 매트릭스 형태로 변환하여 초기 매트릭스에 분포된 데이터의 특성은 유지하면서 매트릭스의 크기를 순차적으로 줄여나가는, 이른바 특성맵(feature map)이라는 단계를 통해 데이터 변환을 거치는 것이 특징이다. 매트릭스 구성에 최적화된 이미지 데이터가 초기부터 그 효용성이 높았으나, 최근에는 이미지 데이터뿐만 아니라 각종 검지기에서 구득되는 교통량 정보를 단순히 학습하고 예측하는 형태로 그 활용도가 확장되는 추세이다. 합성곱의 형태는 아니지만, 반복적인 인공신경망 모형의 형태로 구성된 LSTM(Long Short-Term Memory)의 활용도가 시계열 특성이 강한 교통류 예측 분야에서 널리 활용되고 있다. 학습에 필요한 시계열 데이터를 일정 구간으로 나누어 중첩 학습하도록 하면서 계산 부담은 줄이고, 시계열 패턴 학습은 고수할 수 있도록 설계된 것이 특징이다([그림 9] 참조). 기존 모형의 응용 및 적용 시도가 다양해지면서, 예측력 제고를 위한 혼합모형(hybrid) 접근방법 등으로 새로운 시도가 제안되고 있다. 혼합 모형은 이중 모형을 부분적으로 결합한 형태로, 가장 널리 결합되는 형태는 합성곱신경망 특성맵 단계와 LSTM 시계열 학습방식을 융합하여 적용한 방식으로 활용되고 있는 형태가 대다수이다(Chen 외, 2018, Wu 외, 2019 등).

9) ILSVRC, ImageNet Large Scale Visual Recognition Challenge

(<https://image-net.org/challenges/LSVRC/#:~:text=The%20ImageNet%20Large%20Scale%20Visual,image%20classification%20at%20large%20scale,>, 2022년 5월 20일 검색)

10) 2010년부터 매해 개최되는 이 대회에서, AlexNet이라는 신경망 모형으로 16.4%의 오차율을 보이며 우승을 거머쥐게 되었는데, 직전년도 우승팀의 오차율이 28.2%였던 점을 고려하면 엄청난 향상이라 할 수 있다. 기존 중앙처리장치(Central processing unit, CPU) 대비 수백 배 이상의 동시 병렬계산이 가능한 그래픽처리장치(Graphical processing unit, GPU)의 효용성이 해를 거듭할수록 극대화되면서 방대한 데이터에서 오는 계산 부담도 자유로워졌다.

그림 9 교통량 예측을 위한 LSTM의 일반적인 모형 구성



출처: 저자 작성(지도: 카카오지도의 서울시 일대 교통상황 지도를 수정하여 삽입)

VIII. 결론

본고에서는 교통류 예측에 있어서 어떠한 고전적인 접근방법이 있었고, 최근 인공지능 발달로 인해 어떠한 움직임이 있는지 등을 살펴보았다. 수요(운전자)와 공급(도로망) 간의 균형에 따라 통행이 발생하게 되는 교통의 논리에 따라 도로 위를 통행하는 교통류는 지금, 이 순간도 시시각각 변화한다. 인간이 가진 지식과 시간으로는 이렇게 복잡하게 얽힌 인과관계 모두를 실시간으로 파악하고 더 나아가 미래 통행을 예측하기에는 물리적 한계가 분명하지만, 통행의 발생 논리를 면밀하게 파악할 수 있도록 다양한 방법으로 설계하고 조합

하는 학습구성(경험 교과서)에 따라 잘 학습된 인공지능은 인간의 물리적 한계를 극복할 수 있도록 새로운 편리를 제공하기에 이르렀다. 기존의 방법론에서 한계로 거론되었던 데이터 처리와 계산시간 등의 한계는 최근 GPU 등 폭발적인 하드웨어 성능이 개선되면서 상당수 극복되었고, 그 결과 다양한 기계학습과 딥러닝 등 인공지능 방법론이 빠르게 발전하고 있다. 하지만, 양질의 원시 데이터 수집과 효과적인 모델 전개를 위한 데이터 전처리에 요구되는 전문성과 중요성은, 아직도 이러한 인공지능 패러다임에서도 여전히 큰 숙제로 남아 있다. 이는 비단 교통류 예측 범주에 국한된다고 생각되지 않을 만큼, 전 분야에 공통적인 숙제일 것이다.

고도의 계산시간 등의 한계로 제한되어왔던 교통류 예측 분야에서도 CPU에서 GPU로의 하드웨어의 비약적인 전개가 지금의 인공지능 패러다임을 부추겼다면, 최근 소개되고 있는 양자컴퓨팅(Quantum computing)은 향후 또 어떠한 패러다임을 선도할 수 있을지 그리고 그 영향으로 우리의 삶은 얼마나 더 편리해질 수 있을지 기대된다.

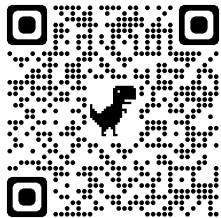
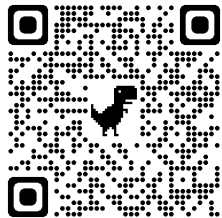
참고문헌

- Ni, Ming, Qing He, and Jing Gao. "Using social media to predict traffic flow under special event conditions." In The 93rd annual meeting of transportation research board, 2014.
- Essien, Aniekan, Ilias Petrounias, Pedro Sampaio, and Sandra Sampaio. "A deep-learning model for urban traffic flow prediction with traffic events mined from twitter." World Wide Web 24, no. 4(2021): 1345-1368.
- Yao, Weiran, and Sean Qian. "From Twitter to traffic predictor: Next-day morning traffic prediction using social media data." Transportation research part C: emerging technologies 124(2021): 102938.
- Pearson, Karl. "LIII. On lines and planes of closest fit to systems of points in space." The London, Edinburgh, and Dublin philosophical magazine and journal of science 2, no. 11(1901): 559-572.
- Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9, no. 11(2008).
- Lv, Yisheng, Yanjie Duan, Wenwen Kang, Zhengxi Li, and Fei-Yue Wang. "Traffic flow prediction with big data: a deep learning approach." IEEE Transactions on Intelligent Transportation Systems 16, no. 2(2014): 865-873.

- Valkanis, Anastasios, Georgia A. Beletsioti, Petros Nicopolitidis, Georgios Papadimitriou, and Emmanouel Varvarigos. "Reinforcement learning in traffic prediction of core optical networks using learning automata." In 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics(CCCI), pp. 1–6. IEEE, 2020.
- Nie, Laisen, Zhaolong Ning, Mohammad S. Obaidat, Balqies Sadoun, Huizhi Wang, Shengtao Li, Lei Guo, and Guoyin Wang. "A reinforcement learning-based network traffic prediction mechanism in intelligent internet of things." *IEEE Transactions on Industrial Informatics* 17, no. 3(2020): 2169–2180.
- Lamas-Seco, Jos J., Paula M. Castro, Adriana Dapena, and Francisco J. Vazquez-Araujo. "SiDIVS: Simple detection of inductive vehicle signatures with a multiplex resonant sensor." *Sensors* 16, no. 8(2016): 1309.
- Mochol Belenguer, Ferran, Antonio Mocholí Salcedo, Antonio Guill Ibañez, and Victor Milián Sánchez. "Advantages offered by the double magnetic loops versus the conventional single ones." *PloS one* 14, no. 2(2019): e0211626.
- Manaa, Karmel, Maram Rabe'e'a, and Loay Khalaf. "Traffic control by digital imaging cameras." In *Emerging Trends in Image Processing, Computer Vision and Pattern Recognition*, pp. 231–247. Morgan Kaufmann, 2015.
- Dabiri, Sina, and Kevin Heaslip. "Transport-domain applications of widely used data sources in the smart transportation: A survey." *arXiv preprint arXiv:1803.10902*(2018).
- Chen, Linkai, Feiyue Ye, Yaduan Ruan, Honghui Fan, and Qimei Chen. "An algorithm for highway vehicle detection based on convolutional neural network." *Eurasip Journal on Image and Video Processing* 2018, no. 1(2018): 1–7.
- Wu, Zhenyu, Karthik Suresh, Priya Narayanan, Hongyu Xu, Heesung Kwon, and Zhangyang Wang. "Delving into robust object detection from unmanned aerial vehicles: A deep nuisance disentanglement approach." In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1201–1210. 2019.
- Kumarage, Sakitha Pathmila. "Use of crowdsourced travel time data in traffic engineering applications." PhD diss., 2018.
- Medina-Salgado, Boris, Eddy S nchez-DelaCruz, Pilar Pozos-Parra, and Javier E. Sierra. "Urban traffic flow prediction techniques: A review." *Sustainable Computing: Informatics and Systems*(2022): 100739.
- Smith, Brian L., and Michael J. Demetsky. "Traffic flow forecasting: comparison of modeling approaches." *Journal of transportation engineering* 123, no. 4(1997): 261–266.

- Smith, Brian L., Billy M. Williams, and R. Keith Oswald. "Comparison of parametric and nonparametric models for traffic flow forecasting." *Transportation Research Part C: Emerging Technologies* 10, no. 4(2002): 303–321.
- Chang, Yohan, Nipjyoti Bharadwaj, Praveen Edara, and Carlos Sun. "Exploring Contributing Factors of Hazardous Events in Construction Zones Using Naturalistic Driving Study Data." *IEEE Transactions on Intelligent Vehicles* 5, no. 3(2020): 519–527.
- Hou, Yi, Praveen Edara, and Yohan Chang. "Road network state estimation using random forest ensemble learning." In *2017 IEEE 20th International Conference on Intelligent Transportation Systems(ITSC)*, pp. 1–6. IEEE, 2017.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25(2012).

부도 1 영상처리와 인공지능이 접목된 교통 감지 예시와 데이터 축소 알고리즘 QR코드

		Embedding Project: https://projector.tensorflow.org/(2022년 5월 20일 접속)	
--	--	---	---

출처: DataFromSky 홈페이지 및 홍보 유튜브 영상¹¹⁾, Embedding project

11) DataFromSky 유튜브 영상(https://www.youtube.com/watch?v=xwQky6aM_lo, 2022년 5월 26일 검색)