

딥러닝을 활용한 차량대기길이 추정모형 개발

Development of Vehicle Queue Length Estimation Model Using Deep Learning

이용주* · 황재성** · 김수희*** · 이철기****

* 주저자 : 아주대학교 교통연구센터 연구원
 ** 공저자 : 아주대학교 건설교통공학과 석·박사 통합과정
 *** 공저자 : 한국도로공사 도로교통연구원 책임연구원
 **** 교신저자 : 아주대학교 교통시스템공학과 교수

Lee Yong-Ju* · Hwang Jae-Seong** · Kim Soo-Hee*** · Lee Choul-Ki****

* Dept. of Transportation Research Institute, Univ. of Ajou
 ** Dept. of Construction and Transportation Eng., Univ. of Ajou
 *** Dept. of Research Institute, Korea Expressway Corporation
 **** Dept. of Transportation System Eng., Univ. of Ajou
 † Corresponding author : Choul-Ki Lee, cklee@ajou.ac.kr

Vol.17 No.2(2018)

April, 2018

pp.39~57

ISSN 1738-0774(Print)

ISSN 2384-1729(On-line)

<https://doi.org/10.12815/kits>

2018.17.2.39

Received 21 February 2018

Revised 20 March 2018

Accepted 2 April 2018

© 2018. The Korea Institute of Intelligent Transport Systems. All rights reserved.

요약

본 연구는 교통운영 개선에 필요한 빅데이터 및 인공지능 모델 개발의 일환으로서, 도시부의 링크통행시간 및 통과교통량 등 가용 데이터 등을 이용하여 교통변수로 활용도가 높은 차량대기길이와의 관계를 딥러닝(Deep Learning)을 통해 학습하고 추정하는 인공지능 모델을 구축하는 것을 목표로 하였다. 차량대기길이 추정모형은 데이터 분석결과를 토대로 하여 우선 차량대기길이의 링크 초과여부를 분류한 후 링크 초과 및 링크 미초과 상황에서의 차량대기길이 추정하는 3개의 모형으로 모델링하였다. 딥러닝 모형은 텐서플로우로 구현하였으며, 모든 모형은 DNN 구조로서 은닉층과 노드 개수를 다양화하여 학습 및 테스트 후 최소 오차를 나타내는 네트워크 구조를 선정하였다. 차량대기길이 링크 초과여부 분류 모형은 약 98%의 정확도를 나타냈으며, 미초과 모형은 15% 미만, 초과 모형은 5% 미만의 오차를 각각 나타내었다. 링크별 평균 오차는 12%로 도출되었다. 이를 기존 검지기 데이터 기반의 방식과 비교한 결과 오차가 약 39% 감소된 것으로 분석되었다.

핵심어 : 딥러닝, 빅데이터, 차량대기길이, 링크통행시간

ABSTRACT

The purpose of this study was to construct an artificial intelligence model that learns and estimates the relationship between vehicle queue length and link travel time in urban areas. The vehicle queue length estimation model is modeled by three models. First of all, classify whether vehicle queue is a link overflow and estimate the vehicle queue length in the link overflow and non-overflow situations. Deep learning model is implemented as Tensorflow. All models are based DNN structure, and network structure which shows minimum error after learning and testing is selected by diversifying hidden layer and node number. The accuracy of the vehicle queue link overflow classification model was 98%, and the error of the vehicle queue estimation model in case of non-overflow and overflow situation was less than 15% and less than 5%, respectively. The average error per link was about 12%. Compared with the detecting data-based method, the error was reduced by about 39%.

Key words : Deep Learning, Big Data, Vehicle Queue Length, Link Travel Time

I. 서론

1. 배경 및 목적

전 세계적으로 제4차 산업혁명이 진행 중인 가운데 교통 분야도 이러한 변화의 물결 속에 빅데이터 및 인공지능을 활용한 교통정책과 서비스 개발 요구가 증대되고 있다. 이에 맞춰 새로운 서비스를 정의하고 서비스 구현을 위해 새롭게 수집·저장해야 할 데이터들을 규정하며 빅데이터를 처리하는 인공지능 모델의 가능성을 사전에 검토하는 연구가 시급하다고 할 수 있다.

이러한 요구에 부응하여 본 연구의 방향을 거시적으로는 교통운영 개선에 필요한 빅데이터 및 인공지능 모델 개발로 규정하고, 미시적으로는 도시부를 대상으로 현재 가용할 수 있는 지능형교통체계(ITS) 데이터를 비롯하여 CCTV, 드론 등을 이용하여 조사 가능한 교통운영 데이터들을 검토하여 데이터들 간의 관계를 알아보고, 딥러닝(Deep Learning) 기법을 적용하여 이러한 관계를 학습, 추정하는 인공지능 모델을 구축하는 것으로 규정하였다.

이에 본 연구에서는 도시부 교통운영 개선을 위해 신호제어 분야 등에서 교통변수로서 중요도는 높으나 현 기술수준으로 인해 활용도가 떨어지는 차량대기길이를 추정하는 모델을 개발하고자 하였다. 기존의 검지기 데이터를 이용한 방식에서 벗어나 현재의 도시부 ITS 운영체계 등을 고려하여 링크통행시간, 통과교통량 등을 활용하여 차량대기길이를 추정하며, 다양한 링크 및 기하구조에서 발생하는 링크통행시간과 차량대기길이 간의 다양한 분포를 학습하고 이를 분류·추정하는 인공지능의 딥러닝 모델을 개발하는 데 본 연구의 목적을 두었다.

2. 연구 방법

본 연구는 크게 5단계로 이루어진다. 첫째 연구방향 정립, 둘째 관련 연구 및 이론 고찰, 셋째 데이터 구축 및 분석, 넷째 차량대기길이 추정모형 개발, 마지막으로 차량대기길이 추정모형 평가 순으로 수행된다.

각 절차별 연구내용을 정리하면, 첫째 연구방향 정립에서는 연구의 성격, 목적 및 범위 등을 명확히 설정하고, 둘째 관련 연구 및 이론 고찰에서는 기존의 차량대기길이 관련 연구동향과 딥러닝 관련 이론을 검토한다. 셋째 데이터 구축 및 분석에서는 가용 가능성, 조사 가능성, 실시간성 등을 고려하여 차량대기길이 추정에 필요한 교통운영 데이터들을 선정한 후, 데이터 추출을 위해 대상 네트워크에 대한 도로조건, 교통조건, 신호조건 등을 조사하여 교통시물레이션을 통해 해당 교통운영 데이터들을 구축하고 데이터 간의 관계를 분석한다. 넷째 차량대기길이 추정모형 개발에서는 데이터 간의 관계 분석결과를 토대로 차량대기길이 추정 인공지능 모형을 모델링하고 인공지능 모형의 학습 및 검증을 수행한다. 다섯째 차량대기길이 추정모형 평가에서는 개발된 인공지능 모형을 기존의 검지기 기반 방식과 비교·평가하여 개발모형의 우수성을 입증한다.

본 연구에서는 인공지능 모델을 구축하기 위해 최신 딥러닝 기법을 적용하며, DNN(Deep Neural Network) 구조를 구현한다. 딥러닝 프레임워크는 구글의 오픈소스 라이브러리인 텐서플로우(TensorFlow)를 사용한다.

II. 이론적 고찰

1. 차량대기행렬 관련 연구

1) 차량도착분포 기반 연구

Wardrop(1952)은 교차로에 접근하는 차량이 포아송분포를 따르며 이에 맞는 지체모형을 개발하였다. Winsten(1956), Dunne(1967)은 차량이 이항분포를 따라 대기행렬을 형성한다고 보고 지체모형을 개발하였다.

Newell(1965)는 차량의 차두시간 분포를 이동된 음지수분포를 이용하여 나타냈고, 근포화과 과포화 상태에서의 지체함수를 제시하였다. 특히, May(1965)와 Allsop(1972)은 차량의 도착을 균일하게 가정하여 지체시간 함수를 개발하였는데, 계산의 간편성으로 인해 현재에도 많은 분야에서 이용되고 있으나, 기본적으로 차량의 도착을 균일하게 본다는 가정은 오차를 포함하므로 신뢰성이 떨어진다.

김대웅 외(1998)는 차량 도착특성을 고려한 접근로의 대기행렬예측모형을 개발, 차량의 도착순서와 차두 시간과의 관계를 해석하고 시간 경과에 따른 정지차량의 패턴으로 지수모형을 제시하였다.

2) 검지기 데이터 기반 연구

Kang(2005)은 링크의 검지기에서 수집되는 개별차량의 점유율을 기반으로 대기행렬 길이 추정 알고리즘을 개발하였다. 기존 COSMOS 시스템에서는 평균차량길이와 점유시간을 이용하여 속도를 산출하고 다시 정체도로 환산 및 임계정체도(Thdoc)를 기반으로 최종적으로 대기행렬 길이를 산정한다. 이에 평균차량길이 적용에 따른 오차를 해소하기 위해 개별차량의 점유율을 이용한 연구이다.

Moon(2005)는 속도 산정 시 평균차량길이를 가정함으로써 발생한 오차와 함께 하류부로 갈수록 정체도가 높아진다는 가정이 실제와는 다른 경우를 지적하며, 기존 실시간 신호제어시스템 COSMOS의 한계를 분석하고 검지기 평균 점유시간 정보를 기반으로 대기행렬 길이를 예측하는 새로운 모형을 제시하였다.

3) 차량 궤적 데이터 기반 연구

Lee et al.(2016)는 신호제어를 위한 기초변수 산출의 일환으로 서울시에서 실시간으로 수집하는 택시 디지털운행기록장치(DTG) 데이터를 이용하여 대기행렬길이를 산출하였다. 택시 DTG 데이터를 기반으로 개별차량 시공도를 추출하고 정지·출발 시의 충격과 추세선이 서로 만나는 지점을 대기행렬 길이로 추정하였다.

2. 딥러닝 이론

딥러닝은 인공신경망(Artificial Neural Network) 이론이 발전된 형태의 인공지능이다. 다층 퍼셉트론을 통한 비선형 문제의 해결, 역전파(Back Propagation) 알고리즘을 통한 학습, 복잡한 문제 해결을 위한 신경망의 층수 확대, 최적화 과정에서의 학습 효율성을 개선한 경사하강법(Gradient Descent Algorithm)의 고도화, 경사감소소멸(Vanishing Gradient Descent) 문제에 대응한 활성화 함수(Activation Function)의 개선, 과적합(Overfitting)을 막기 위한 규제화(Regulation) 기법 개발 등 신경망 학습알고리즘의 발전이 이루어지고 있다.

딥러닝 이론은 인간의 뇌를 모델로 하여 여러 정보를 처리하는 알고리즘을 일컫는 인공신경망에 기반을 둔다. 기존 인공신경망 이론의 다층 퍼셉트론 구조가 적용되고 피드포워드(Feedforward)를 통한 예측과 역전파 알고리즘을 통해 가중치를 업데이트하는 학습이 수행된다. 이 때 예측치와 실제값의 오차를 최소화하는 가중치를 찾는 최적화 알고리즘은 경사하강법이 주로 적용되며, 경사하강법으로는 Batch Gradient Descent, Stochastic Gradient Descent(SGD), Momentum, AdaGrad, RMSProp, ADAM(Adaptive Moment Estimation) 등이 있다. 활성화 함수(Activation Function)는 은닉층의 노드에서 데이터와 가중치를 선형적으로 곱하여 모두 합산한 값에 임계치를 적용해 활성화 정도를 부여함으로써 실질적으로 의미 있는 데이터와 의미 없는 데이터를 필터링하는 역할을 수행하는데, 주요 활성화 함수에는 Sigmoid, tanh, ReLU 등이 있다.

딥러닝 모델에는 심층신경망(Deep Neural Network, DNN), 합성곱 신경망(Convolutional Neural Network,

CNN), 순환신경망(Recurrent Neural Network, RNN), LSTM(Long Short Term Memory) 등이 있으며, 본 연구에 적용된 DNN은 입력층(input layer)과 출력층(output layer) 사이에 복수개의 은닉층(hidden layer)들로 이뤄진 인공신경망을 의미하며, 가장 기본적인 딥러닝 구조이다.

본 연구에 사용한 텐서플로우(TensorFlow)는 기계 학습과 딥러닝을 위해 구글에서 만든 오픈소스 소프트웨어 라이브러리로서, Data Flow Graph를 사용하여 수치 연산을 한다. 데이터 플로우 그래프에서는 수학 계산과 데이터의 흐름을 노드(Node)와 엣지(Edge)를 사용한 방향 그래프로 표현한다. 노드는 수학적 계산, 데이터 입·출력, 그리고 데이터의 읽기·저장 등의 작업을 수행한다. 엣지는 노드들 간 데이터의 입출력 관계를 나타내는데, 동적 사이즈의 다차원 데이터 배열(텐서, tensor)을 옮기는 역할을 수행한다.

3. 기존 연구와의 차별성

차량대기길이 관련 기존 선행연구는 3가지 패러다임으로 변화되어 왔다. 1950~60년대에 수행된 통계학적 접근방식의 차량도착특성에 따른 확률분포모형 개발, 1990년대 실시간 신호제어시스템 개발과 병행된 검지기 데이터 기반 연구, 그리고 2010년대에 수행된 차량궤적데이터와 충격과 적용 연구 등이다. 확률분포모형은 다양한 차량도착특성에 따라 특정 모형으로 일반화하는 데 한계가 있으며, 검지기 데이터 기반 방식은 접근로별 교통상황별로 운영자의 적정 임계정체도 설정이 필요하고 루프검지기의 유지관리 문제로 운용상의 한계에 직면하였다. 최근의 차량궤적데이터 이용 및 충격과 적용 방식은 충분한 차량궤적데이터가 요구되며 데이터 수집율과 교통상황에 따른 보정계수 적용이 필요한 점이 단점으로 지적된다.

이에 따라 본 연구는 기존 선행연구의 고찰결과를 토대로 다음과 같은 차별화를 시도하였다.

첫째, 다양한 차량도착특성과 이에 따른 차량대기길이 확률분포를 통합적으로 수용할 수 있는 통합모형을 추구한다. 둘째, 임계정체도 등 중요 파라미터를 사람이 학습하여 설정하는 어려움이 없도록 한다. 즉 안정적인 지속적인 운영을 위해서는 기계가 학습하여 자동적으로 차량대기길이 모형을 최적화 하도록 한다. 셋째, 기존 보다 차량대기길이의 정확도를 높여 궁극적으로는 차량대기길이의 교통운영기법 적용에 있어 효용성을 제고할 수 있어야 한다. 넷째, 대기행렬 검지기과 같이 상시적인 교통수집장비 설치·운영을 통한 데이터 구득을 요구하지 않고, 현재 운영 중인 지능형교통체계 등에서 생성되는 가용 가능한 데이터를 적용한다.

Ⅲ. 데이터 구축 및 분석

1. 데이터 구축

차량대기길이를 추정하는 딥러닝 모형 개발에 필수적인 학습데이터 구축을 위해 Target Variable인 차량대기길이와 관련 있는 교통변수들을 조사하여 가용성, 조사 가능성, 실시간성 등을 고려한 입력변수(Input Variable)를 선정한다. 그리고 다양한 교통상황과 도로조건을 반영할 수 있도록 네트워크 단위의 공간적 범위를 선정하고 해당 네트워크 내의 도로조건, 신호조건 및 소통상태 현황을 조사한다. VISSIM을 이용한 시뮬레이션을 통해 해당 시간대의 교통상황을 유사하게 구현하며, 도출된 시뮬레이션 결과를 데이터로 구축한다.

1) 차량대기길이 관련 변수 선정

대기행렬길이는 신호교차로의 용량과 관련이 있기 때문에 도로용량편람에서 규정하고 있는 신호교차로 용량분석에 필요한 도로조건, 교통조건, 신호조건 등의 입력변수를 검토하였다. 해당 변수들 중에서 실시간

으로 파악이 어려운 변수를 제외하고 링크길이, 차로수와 신호조건 등 정적 변수들을 채택하였다.

도로관리청에서 운영 중인 지능형교통체계(ITS)에서는 링크통행시간, 링크통행속도, 통과교통량 등의 소통 상태 결과에 해당하는 변수를 실시간으로 수집·가공하여 도로이용자에게 제공하고 있다.

본 연구에서는 데이터 가용성, 조사 가능성, 실시간성 등을 고려하여 수요와 용량의 실시간 변화를 도출하는 원인변수를 제외한 ITS에서 소통상태지표로 수집·활용되는 결과변수를 이용하여 서로 간의 관계를 알아보는 데 중점을 둔다. 즉 교통의 3대 변수인 교통량, 속도, 밀도 간의 관계와 같이 다양한 도로조건별로 차량 대기길이와 링크통행시간, 통과교통량 등간의 관계를 학습하여 대기차량길이를 추정하며, 이에 맞게 데이터 구축대상 변수를 다음과 같이 설정하였다.

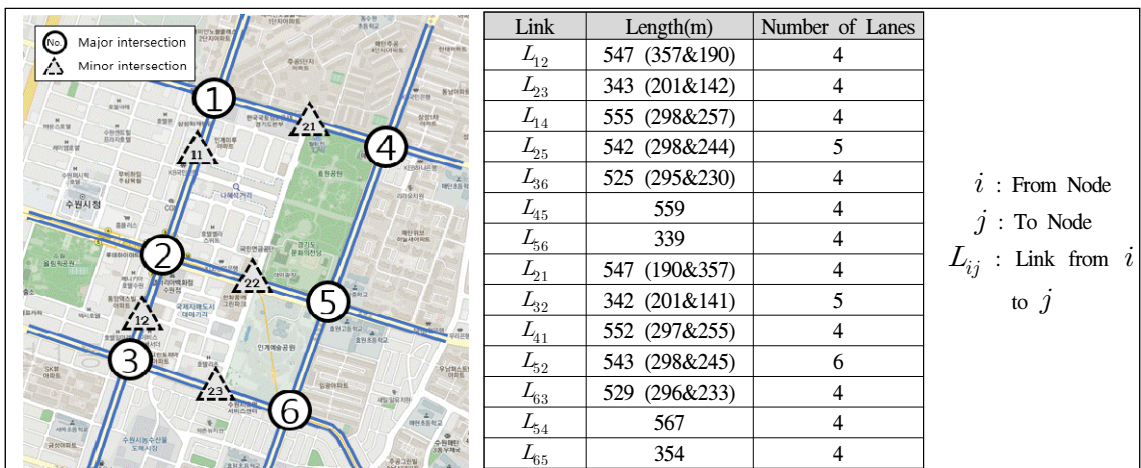
<Table 1> Variables Target of Data Construction

Purpose	Variable Personality	Variable (9unit)
Classification	Road Condition	Link Number, Link Length(m), Number of Lanes(unit)
	Signal Condition	Cycle Length(sec), Green Time(sec), Offset(sec)
Classification + Regression	Traffic Condition	Link Travel Time(sec), Passing Volume(veh), Queue(m)

2) 대상 네트워크 현황

모형의 일반화를 위해 비포화에서부터 과포화까지 다양한 교통상황이 발생하는 교차로를 대상으로 하였다. 차량도착특성은 상·하류부 교차로의 상황, 연동 등에 의해 영향을 받으므로 이를 반영하기 위해 독립교차로나 축이 아닌 네트워크를 대상으로 하였으며, 또한 본 연구에서는 실제 통행시간 데이터를 사용하지 않으나 향후 개발모형의 현실 적용이 용이할 수 있도록 DSRC RSE 등 교통수집장치가 설치되어 통행시간 정보가 생성되는 링크를 대상으로 하였다. 이에 따라 수원시 내 상습적으로 교통혼잡이 발생하는 CBD지역이며, DSRC를 통해 통행시간 정보가 생성되는 인계동 지역의 네트워크를 선정하였다.

선정된 네트워크는 6개의 중요교차로(CI)로 구성된 3X2 구조이며, 그 외에도 네트워크 내에는 5개의 비중요교차로가 위치해 있어 총 11개의 교차로가 포함되어 있다. 네트워크 구조 및 링크번호체계, 링크별 길이와 차로수 현황은 다음과 같다.



<Fig. 1> Road Network Structure

시간대는 다양한 교통상황을 반영하기 위해 오후첨두 시간대를 대상으로 하여 실제 교차로별 신호DB 및 교통여건을 반영하였고, 해당 시간대의 ITS 통행속도를 토대로 유사한 소통상황을 VISSIM으로 구현하였다.

3) 데이터 수집환경

미시적 교통시뮬레이터인 VISSIM에 조사자료를 바탕으로 대상 네트워크를 구축하고 신호자료를 입력하며, 통행속도를 실제 조사한 값과 비슷하도록 교통량, 회전비율을 조절하여 분석시간대의 소통상황을 구현하였다. 특히 17~19시 시간대에서 발생한 비포화에서부터 과포화까지의 다양한 교통상황을 구현하였다.

구간통행시간은 중요교차로 구간별 Vehicle Travel Time Measurement를 이용하여 수집하였고, 차량대기길이는 Queue Counter를 설정하여 수집하였다. 데이터 수집주기는 공통 신호주기인 3분으로 설정하였다.

VISSIM은 다양한 확률 모형이 적용되어 있어 Random Seed를 달리하면 매 시뮬레이션 마다 확률적인 차량행태가 반영되어 매번 다양한 교통상황이 도출된다. 이를 이용하여 2개월가량의 데이터를 저장하는 개념으로 총 40회의 시뮬레이션을 통해 데이터를 저장하였다. 시뮬레이션 분석 기간은 네트워크 균형상태(Equilibrium)를 고려하여 20분을 제외하고 2시간 동안의 데이터를 수집하였다.

<Table 2> Data Collection Environment

Classification	Setting Environment
Analysis Time	2hours (17:00 ~ 19:00)
Simulation Count	40times (2 month assumption)
Vehicle Type Ratio	Car 90%, Heavy Vehicle 10% (Bus 8%, Freight Car 2%)
Data Collection Cycle	3minute (Apply common signal cycle in network)

4) 데이터 구축 결과

총 40회 시뮬레이션을 통해서 시뮬레이션을 통해 동일 환경에서의 통행시간에 따른 다양한 차량대기길이 데이터를 수집하였다. 링크별로 1,600개씩 전체 14개 링크에 총 22,400개의 데이터를 구축하였으며, 링크별 통행시간, 통과교통량, 차량대기길이 발생범위는 다음과 같다.

<Table 3> Collected Data Statistics

Link	The Number of Data(EA)	Travel Tim(s)			Passing Volume(veh)			Queue Length(m)		
		Min	Max	Average	Min	Max	Average	Min	Max	Average
L_{12}	1,600	46.9	199.4	69.33	56	124	95	37.0	528.8	149.4
L_{21}	1,600	50.4	360.4	171.8	53	129	89	104.6	991.2	452.8
L_{23}	1,600	51.4	239.1	119.7	36	113	88	76.1	951.8	450.7
L_{32}	1,600	42.1	125.3	61.87	52	107	77	28.5	195.3	69.4
L_{14}	1,600	58.7	379.1	164.8	48	104	75	48.0	1003.8	347.6
L_{41}	1,600	92.6	530.4	260.2	42	84	60	118.3	987.8	432.5
L_{25}	1,600	88.0	272.9	137.2	62	105	85	84.5	604.4	206.3
L_{52}	1,600	83.8	171.5	119.3	37	89	66	23.4	167.8	54.9
L_{36}	1,600	63.8	299.9	112.3	35	90	70	69.9	589.5	178.2
L_{63}	1,600	48.0	142.3	74.3	42	98	71	57.5	407.4	164.6
L_{45}	1,600	40.7	348.1	156.7	57	111	86	28.8	632.5	395.7
L_{54}	1,600	49.8	372.6	91.5	73	126	106	78.2	565.3	312.6
L_{56}	1,600	30.1	121.3	55.0	70	136	100	29.6	418.5	204.9
L_{65}	1,600	34.6	192.5	83.8	47	89	73	59.3	418.4	219.1

2. 데이터 분석

데이터 분석의 목적은 DNN 모형에 입력변수로 설정할 변수를 찾기 위함으로, 표적변수로 설정된 최대 차량대기길이와 링크통행시간, 통과교통량 간의 상관관계를 분석한다. 또한 차량대기길이 해당 링크를 초과하여 상류부까지 형성될 때 상류부 링크 통행시간과의 상관관계를 분석한다.

1) 통행시간 및 차량대기길이 현황

오후 침두시간대를 대상으로 시뮬레이션을 수행한 만큼 많은 링크에서 비포화 상태에서 전이를 거쳐 과포화 상황까지 다양한 교통상황이 나타났으며, 이에 따라 링크 전반의 통행시간도 최소 30~90초, 1주기(180초)미만의 통행시간에서부터 최대 300~500초, 2~3주기 가량의 통행시간까지 발생하였다.

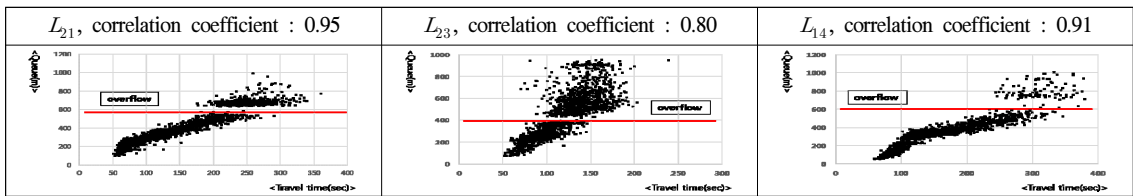
주기당 통과교통량은 통행시간 분포 및 교통상황 전이 추세와는 다르게 일정한 분포를 나타내고 있다. 그 이유는 할당된 녹색신호시간에 통과하는 차량대수는 용량으로 한정될 수밖에 없기 때문이다.

링크별 차량대기길이는 통행시간과 유사한 추세를 나타내고, 통행시간 발생 범위와 같이 비포화 및 과포화 상태를 보이며, 차량대기길이 또한 최소 30m, 최대 1km를 초과하고 있다.

2) 변수 간 상관관계 분석

① 링크통행시간과 차량대기길이 상관분석

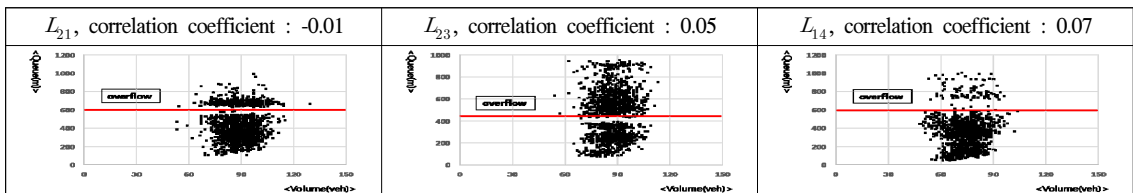
링크별 통행시간과 차량대기길이 간의 관계를 알아보기 위하여 상관분석을 수행하였으며, 링크별 상관계수는 다음과 같이 높게 나타났다.



(Fig. 2) Relationship Between Link Travel Time and Vehicle Queue Length

② 통과교통량과 차량대기길이 상관분석

링크별 통과교통량과 차량대기길이 간의 관계를 알아보기 위하여 상관분석을 수행하였으며, 링크별 상관계수는 다음과 같이 낮게 나타났다.

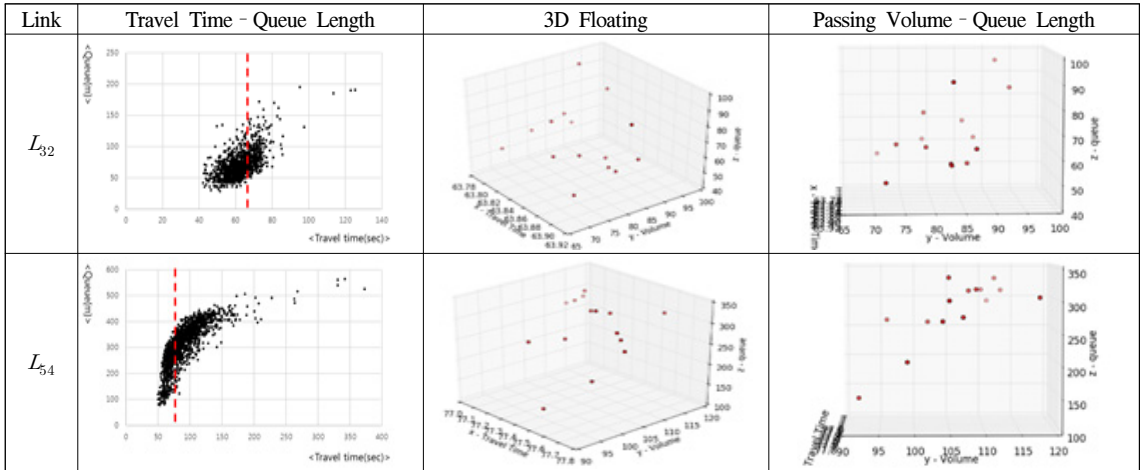


(Fig. 3) Relationship Between Passing Volume and Vehicle Queue Length

③ 통과교통량의 차량대기길이 분류 기여도 분석

앞서 링크통행시간은 차량대기길이와 비교적 높은 상관관계를 나타내고 있으나 일부 링크의 경우 특정

링크통행시간에 차량대기길이가 넓게 분포되어 있어 링크통행시간을 이용한 차량대기길이 추정 시 많은 오차를 수반하게 된다. 오차를 최소화하기 위해서는 분류 및 회귀에 기여할 수 있는 변수가 추가적으로 필요하다. 이에 차량대기길이와는 직접적인 상관관계가 낮은 통과교통량이 특정 링크통행시간에 넓게 분포된 차량대기길이를 분류하는 데 있어 기여할 수 있는지에 대한 분석을 수행하였다. 다음은 2개 링크를 대상으로 특정 링크통행시간에 넓게 분포된 차량대기길이 단면을 확대하여 링크통행시간과 차량대기길이, 통과교통량의 3차원 그래프로 분석한 그림이다. 단면을 확대하여 3차원으로 분석해 보니 이러한 특정 상황에서는 통과교통량이 차량대기길이와 상관관계가 비교적 높아 분류 및 회귀에 기여할 수 있을 것으로 판단하였다.

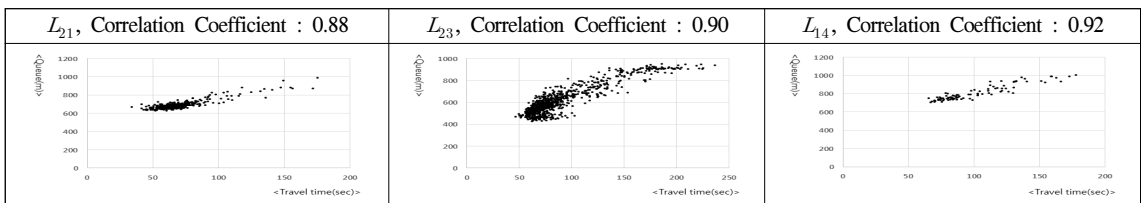


<Fig. 4> Analysis of Classification Contribution to Vehicle Queue Length of Passing Traffic

3) 링크 초과 상황에서의 변수 간 상관관계 분석

① 상류부 통행시간과 차량대기길이 상관분석

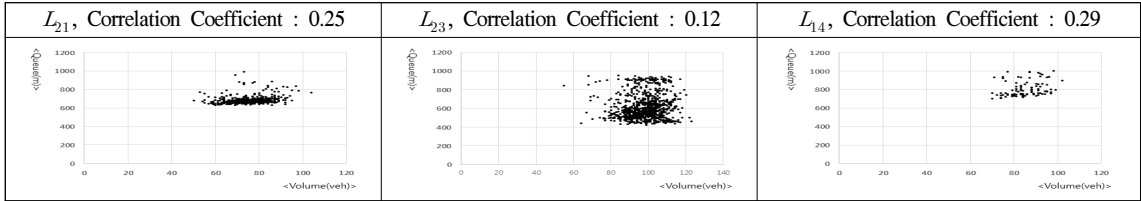
차량대기길이가 링크를 초과하는 상황에서는 당시의 차량대기길이는 해당 링크의 통행시간과 관계성이 떨어진다. 이에 따라 차량대기길이가 해당 링크를 초과하여 상류부 링크에도 영향을 주기 때문에 상류부 링크의 통행시간과의 상관관계를 분석하였다.



<Fig. 5> Relationship Between Link Travel Time and Vehicle Queue Length in a Queue Overflow Situation

② 상류부 통과교통량과 차량대기길이 상관분석

상류부 링크의 통과교통량은 통행시간과 다르게 링크를 초과하는 차량대기길이와는 낮은 상관관계를 보였으며, 이는 통과교통량과 차량대기길이와의 상관분석과 유사한 결과를 보이고 있다고 할 수 있다. 다만, 특정 통행시간에 넓게 분포된 차량대기길이를 분류할 수 있는 변수로서의 역할을 수행하여 오차를 줄이는데 기여할 수 있을 것으로 예상된다.



〈Fig. 6〉 Relationship Between Passing Volume and Queue Length in a Queue Overflow Situation

4) 데이터 분석 결과

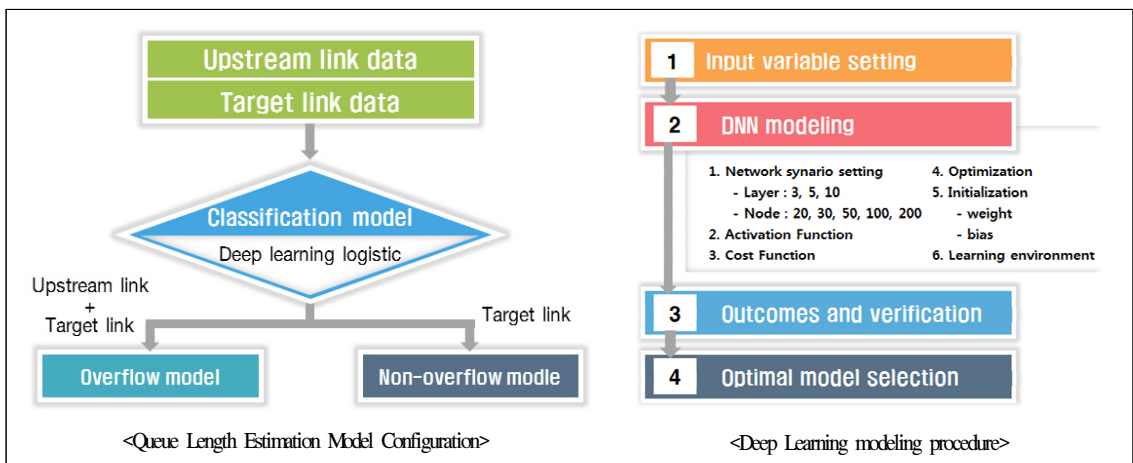
해당 링크 통행시간과 차량대기길이는 상관관계가 높고, 통과교통량은 차량대기길이와 상관관계가 낮다. 단, 특정 링크통행시간에 차량대기길이가 넓게 분포되어 통행시간에 따른 추정 오차가 클 것으로 예상되었는데 이 구간에 대해서는 해당 링크통행시간의 통과교통량이 차량대기길이를 분류할 수 있는 것으로 분석되었다.

차량대기길이가 해당 링크를 초과하여 상류부 링크까지 생성되는 상황에서는 해당 링크 통행시간은 상류부 링크까지 이어지는 차량대기길이를 설명하지 못하였다. 이때에는 상류부 링크 통행시간이 차량대기길이와 상관관계가 높은 것으로 분석되었다.

IV. 딥러닝 기반 차량대기길이 추정모형 개발

1. 개요

앞서 제시된 데이터 분석결과를 토대로 차량대기길이 추정 모형은 크게 3가지 인공지능 모형으로 구성된다. 차량대기길이의 링크 초과 여부를 추정하는 로지스틱 모형과 링크 미초과 차량대기길이 추정모형, 링크 초과 차량대기길이 추정모형이다. 각 딥러닝 모형에는 데이터 분석결과에서 제시한 바와 같이 각각의 모형에 적합한 회귀 및 분류 용도의 데이터가 입력변수로 설정되어 진다.



〈Fig. 7〉 Vehicle Queue Length Estimation Model Configuration and Modeling Procedure

2. 차량대기길이 링크 초과여부 분류모형

1) 로지스틱 회귀분석 고찰

로지스틱 회귀는 Cox(1958)에 제안한 확률 모델로서 독립 변수의 선형 결합을 이용하여 사건의 발생 가능성을 예측하는데 사용되는 통계 기법이다. 로지스틱 회귀 분석과는 다르게 종속 변수가 범주형 데이터를 대상으로 하며 입력 데이터가 주어졌을 때 해당 데이터의 결과가 특정 분류로 나뉘기 때문에 일종의 분류(Classification) 기법으로도 볼 수 있다.

독립 변수가 $[-\infty, \infty]$ 의 어느 숫자이든 상관없이 종속 변수가 항상 범위 $[0, 1]$ 사이에 있도록 한다. 이는 오즈비(odds ratio)를 로짓(logit) 변환을 수행함으로써 얻어진다. 오즈비(odds ratio)는 성공 확률이 실패 확률에 비해 몇 배 더 높은가를 나타내며 여기에 로그를 취하면 입력값의 범위가 $[-\infty, +\infty]$ 일 때 출력 값의 범위를 $[0, 1]$ 로 조정한다.

$$odds = \frac{p(y=1|x)}{1-p(y=1|x)} \tag{1}$$

$$\text{logit}(p) = \log \frac{p}{1-p} \tag{2}$$

$$y_i = \ln\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + e_i, e_i \sim \text{Normal}(0, \sigma^2) \tag{3}$$

$$p_i = \Pr(Y=1|x) = \frac{e^{(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}}{1 + e^{(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}} + e_i^* = \frac{1}{1 + e^{-(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}} + e_i^* \tag{4}$$

2) 딥러닝 기반 로지스틱 회귀 구현 기법

로지스틱 회귀의 가설함수를 선형회귀의 비용함수¹⁾에 바로 대입하는 경우에는 울퉁불퉁하여 기울기가 0이 되는 지점이 다수 발생하는 non-convex 함수가 되고, 이로 인해 전역최솟값(global minima)이 아닌 지역최솟값(local minima)에 수렴하는 문제가 발생하게 된다.

이에 신경망 이론에서는 비용함수를 convex 함수가 되도록 조정하여 경사하강법(Gradient descent algorithm)이 원활히 작동될 수 있도록 이 문제를 해결한다. 실제 값이 1인 경우와 0인 경우 각각에 로그를 취함으로써 convex 함수 형태를 만들어 비용을 최소화 할 수 있도록 만든 것이다

$$H(x) = \frac{1}{1 + e^{-(Wx+b)}} \tag{5}$$

$$\text{cost}(W) = \frac{1}{m} \sum c(H(x), y) \tag{6}$$

$$c(H(x), y) = \begin{cases} -\log(H(x)) & : y = 1 \\ -\log(1-H(x)) & : y = 0 \end{cases} \tag{7}$$

$$c(H(x), y) = -y \log(H(x)) - (1-y) \log(1-H(x)) \tag{8}$$

여기서, $H(x) : \text{Hypothesis of Logistic Function}$

1) 비용함수(Cost Function or Loss Function)란 딥러닝 모형을 통한 추정값과 실제값의 차이 즉 오차를 나타내는 함수를 의미한다.

3. 딥러닝 기반 차량대기길이 링크 초과여부 추정 모형

1) 입력변수 설정

차량대기길이와 통행시간, 통과교통량 간의 상관관계 및 특히, 차량대기길이의 링크 초과 시 상류부 링크와 높은 상관관계를 고려하여 입력변수로는 해당링크와 더불어 상류부 링크를 포함한 총 10개로 설정하였다.

〈Table 4〉 Input Variables of the Model to Classify whether Vehicle Queue is a Link Overflow

Link	Input Variable
Upstream Link	Link Number(1), Number of Lanes(2), Length of Link(3), Travel Time(4), Passing Volume(5)
Target Link	Link Number(6), Number of Lanes(7), Length of Link(8), Travel Time(9), Passing Volume(10)

2) DNN 모델링

신경망의 네트워크는 은닉층(Hidden Layer) 개수를 각각 3, 5개와 은닉층의 노드 개수를 각각 20, 50개로 설정하여 학습을 수행한다. 활성화 함수는 최근에 가장 많이 쓰이고 성능이 우수한 Lelu를 적용한다. 비유함수는 앞서 로지스틱 회귀분석의 딥러닝 구현 기법을 고찰한 바와 같이 다음과 같이 설정한다.

$$c(H(x), y) = -y \log(H(x)) - (1-y) \log(1-H(x)) \tag{9}$$

최적화를 위한 경사하강법은 최근 가장 성능이 좋은 ADAM을 적용한다. 학습을 통해 최적화될 weight와 bias의 초기화 방법은 각각 tf.Zeros, tf.random_normal를 적용한다. 학습(Training) 및 검증(Test)은 홀드아웃(holdout) 기법을 적용한다. 레이블된 데이터셋을 학습 데이터(80%)와 테스트 데이터(20%)로 나눈 후, 학습 데이터로 모델을 학습시키고 테스트 데이터의 레이블을 예측하여 정확도를 구하는 방식으로서 과적합 방지 및 모델의 일반화 검증 방법이다.

3) 학습 및 검증을 통한 최적모형 선정

DNN 모델링 시나리오에 따라 은닉층 개수와 노드 개수로 구분된 각 케이스별로 10회씩 학습 및 테스트를 시행하여 평균오차가 가장 적은 DNN 구조를 우선 선정한 후 해당 구조 내에서 오차가 최소인 순번의 모형을 최적모형으로 선정하였다.

〈Table 5〉 Errors by the DNN Structure depending on the Number of Hidden Layers and Nodes in the Vehicle Queue Link Overflow Classification Model

Hindden Layer		3		5	
Node		20	50	20	50
1	Training	0.0499	0.0512	0.0526	0.0522
	Test	0.0520	0.0569	0.0454	0.0490
(skip)					
8	Training	0.0506	0.0495	0.0514	0.0519
	Test	0.0505	0.0562	0.0494	0.0546
9	Training	0.0534	0.0509	0.0505	0.0511
	Test	0.0459	0.0504	0.0582	0.0556
10	Training	0.0510	0.0525	0.0506	0.0532
	Test	0.0506	0.0446	0.0513	0.0483
average	Training	0.0508	0.0513	0.0514	0.0522
	Test	0.0518	0.0515	0.0515	0.0511

차량대기길이 링크 초과여부 분류 모형의 정확도는 학습데이터 97.89%, 검증데이터 97.92%로 각각 나타났으며, 굉장히 높은 정확도를 나타내고 있다.

<Table 6> The Accuracy of the Vehicle Queue Link Overflow Classification Model

	The Number of Data	Actual		Estimation		Accuracy (%)
		Overflow	non-Overflow	Overflow	non-Overflow	
Training	17,900	1,469	16,431	1,481	16,419	97.89
Test	4,480	4,114	366	4,131	349	97.92

4. 차량대기길이 링크 미초과 추정모형

1) 입력변수 설정

차량대기길이가 링크를 초과하지 않을 시에는 해당 링크의 통행시간과 상관관계가 높았고, 단 통과교통량은 특정 통행시간대에 넓게 집중된 대기길이의 분류 기능을 수행할 수 있을 것으로 판단한다.

<Table 7> Input Variables of the Vehicle Queue Estimation Model in Case of Vehicle Queue non-Overflow Situation

Link	Input Variable
Target Link	Link Number(1), Number of Lanes(2), Length of Link(3), Travel Time(4), Passing Volume(5)

2) DNN 모델링

신경망의 네트워크는 은닉층(Hidden Layer) 개수를 각각 5, 10개와 은닉층의 노드 개수를 각각 30, 50, 100, 200개로 설정하여 학습을 수행한다. 활성화 함수는 최근에 가장 많이 쓰이고 성능이 우수한 Lelu를 적용한다. 비용함수는 추정값과 실제값의 차이가 클수록 더 큰 오차를 나타내는 평균제곱오차(Mean Squared Error, MSE)를 적용한다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \tag{10}$$

최적화를 위한 경사하강법은 최근 가장 성능이 좋은 ADAM을 적용한다. 학습을 통해 최적화될 weight와 bias의 초기화 방법은 모두 tf.random_normal을 적용한다. 학습(Training) 및 검증(Test)은 홀드아웃(Holdout) 기법을 적용한다. 레이블된 데이터셋을 학습 데이터(80%)와 테스트 데이터(20%)로 나눈 후, 학습 데이터로 모델을 학습시키고 테스트 데이터의 레이블을 예측하여 정확도를 구하는 방식으로서 과적합 방지 및 모델의 일반화 검증 방법이다. 오차 계산 및 최적화를 위한 데이터 전달방식은 일부 데이터를 전달하는 미니배치(Mini-batch) 방식을 적용하였으며, epoch 당 학습횟수는 전체 학습데이터를 미니배치로 나눈 값이 된다.

3) 학습 및 검증을 통한 최적모형 선정

DNN 모델링 시나리오에 따라 은닉층 개수와 노드 개수로 구분된 각 케이스별로 10회씩 학습 및 테스트를 시행하여 평균오차가 가장 적은 DNN 구조를 우선 선정한 후 해당 구조 내에서 오차가 최소인 순번의 모형을 최적모형으로 선정하였다.

〈Table 8〉 Errors by the DNN Structure depending on the Number of Hidden Layers and Nodes in the Vehicle Queue Estimation Model in Case of Vehicle Queue non-Overflow Situation

Hindden Layer		5				10			
Node		30	50	100	200	30	50	100	200
1	Training	1,448	1,431	1,367	1,346	1,414	1,415	1,391	1,384
	Test	1,489	1,484	1,403	1,527	1,552	1,449	1,438	1,332
2	Training	1,422	1,393	1,374	1,362	1,478	1,453	1,408	1,380
	Test	1,438	1,501	1,384	1,483	1,424	1,418	1,323	1,315
(skip)									
10	Training	1,427	1,380	1,391	1,343	1,429	1,432	1,369	1,367
	Test	1,697	1,477	1,456	1,479	1,385	1,479	1,368	1,503
Average	Training	1,444	1,421	1,377	1,382	1,446	1,407	1,380	1,374
	Test	1,484	1,434	1,453	1,509	1,479	1,481	1,563	1,459

차량대기길이 링크 미초과 모형의 추정 오차는 약 14%이며, 평균절대오차(Mean Absolute Error, MAE)와 평균절대백분율오차(Mean Absolute Percentage Error, MAPE)는 다음과 같다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \tag{11}$$

$$MAPE(\%) = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) * 100 \tag{12}$$

〈Table 9〉 The Error of the Vehicle Queue Estimation Model in Case of Vehicle Queue non-Overflow Situation

queue length (m)	Training		Test	
	MAE(m)	MAPE(%)	MAE(m)	MAPE(%)
0 ~ 100	12.65	18.41	13.40	19.20
100 ~ 200	24.33	17.02	26.26	18.51
200 ~ 300	34.63	14.26	32.65	13.47
300 ~ 400	29.31	8.51	29.44	8.56
400 ~ 500	36.06	8.20	32.93	7.53
500 ~ 600	30.81	5.74	33.69	6.24
Total	26.14	14.20	26.27	14.54

5. 차량대기길이 링크 초과 추정모형

1) 입력변수 설정

차량대기길이가 링크를 초과할 시의 추정모형은 앞서 차량대기길이 링크 초과여부 분류모형과 마찬가지로 관련 있는 링크의 정보 즉 상류부와 해당 링크의 총 10개로 설정하였다.

〈Table 10〉 Input Variables of the Vehicle Queue Estimation Model in Case of Vehicle Queue Overflow Situation

Link	Input Variable
Upstream Link	Link Number(1), Number of Lanes(2), Length of Link(3), Travel Time(4), Passing Volume(5)
Target Link	Link Number(6), Number of Lanes(7), Length of Link(8), Travel Time(9), Passing Volume(10)

2) DNN 모델링

신경망의 네트워크는 은닉층(Hidden Layer) 개수를 각각 3, 5개와 은닉층의 노드 개수를 각각 30, 50, 100, 200개로 설정하여 학습을 수행한다. 그 외 DNN 모델링은 앞서 차량대기길이 링크 미초과 추정모형의 설정 방법과 동일하다.

3) 학습 및 검증을 통한 최적모형 선정

DNN 모델링 시나리오에 따라 은닉층 개수와 노드 개수로 구분된 각 케이스별로 10회씩 학습 및 테스트를 시행하여 평균오차가 가장 적은 DNN 구조를 우선 선정한 후 해당 구조 내에서 오차가 최소인 순번의 모형을 최적모형으로 선정하였다.

〈Table 11〉 Errors by the DNN Structure depending on the Number of Hidden Layers and Nodes in the Vehicle Queue Estimation Model in Case of Vehicle Queue Overflow Situation

Hindden Layer		3				5			
Node		30	50	100	200	30	50	100	200
1	Training	2,024	1,846	1,516	1,618	2,055	1,667	1,570	1,700
	Test	1,553	2,265	1,930	1,685	1,785	1,455	1,785	2,598
8	Training	1,678	2,010	1,659	1,611	1,824	1,820	1,859	1,592
	Test	2,424	1,265	1,434	1,621	2,149	1,506	1,556	1,516
(생략)									
10	Training	2,031	1,818	1,706	1,608	2,016	1,921	1,930	1,612
	Test	1,749	1,963	1,744	1,739	2,318	2,721	1,379	1,821
Average	Training	2,196	1,840	1,646	1,619	2,013	1,829	1,717	1,679
	Test	2,247	1,921	1,722	1,744	2,020	1,984	1,775	1,784

차량대기길이 링크 초과 모형의 추정 오차는 약 5% 미만이며, MAE와 MAPE는 다음과 같다.

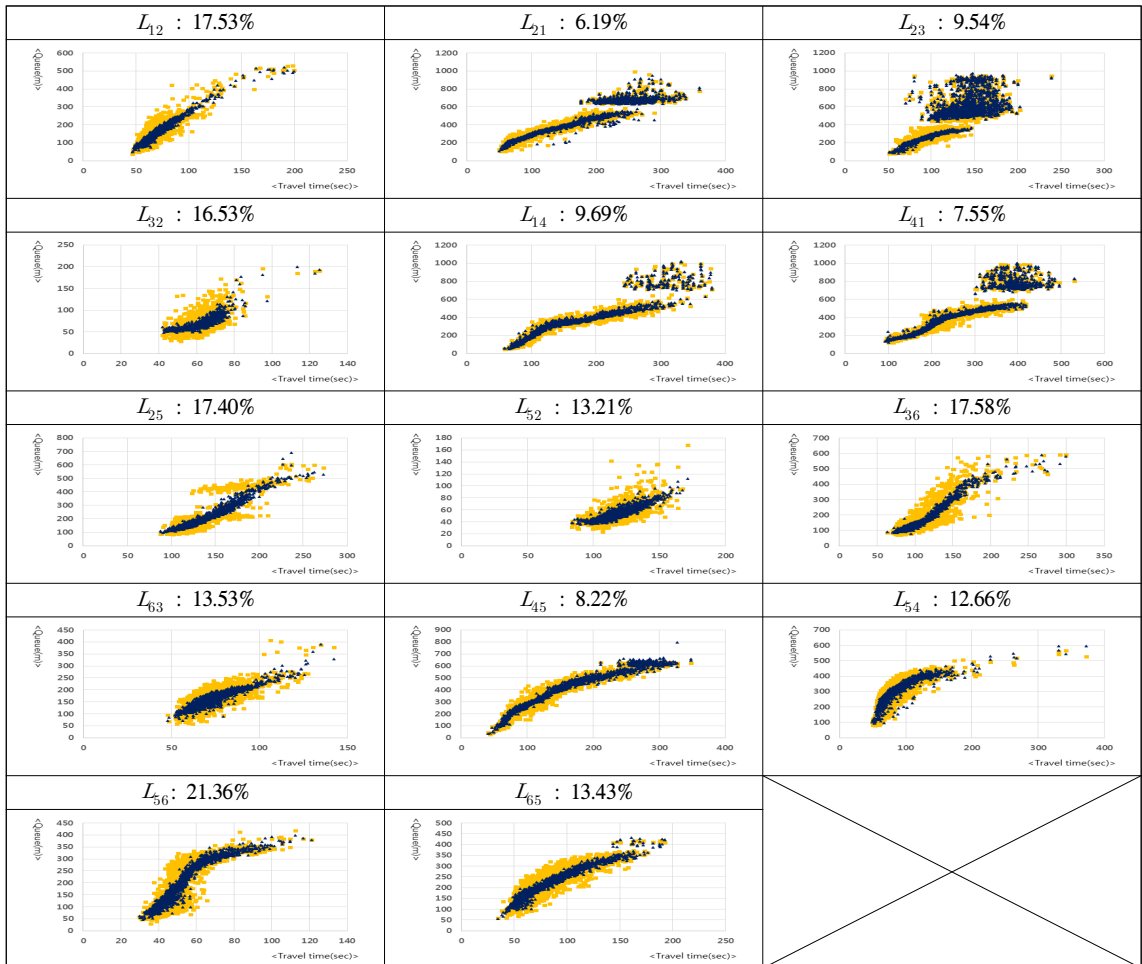
〈Table 12〉 The Error of the Vehicle Queue Estimation Model in Case of Vehicle Queue Overflow Situation

Queue Length (m)	Training		Test	
	MAE(m)	MAPE(%)	MAE(m)	MAPE(%)
400 ~ 500	50.56	10.89	44.32	9.62
500 ~ 600	28.89	5.25	27.49	5.01
600 ~ 700	24.23	3.71	24.87	3.81
700 ~ 800	28.47	3.82	24.22	3.25
800 ~ 900	36.57	4.32	33.02	3.95
900 ~	28.88	3.08	31.54	3.36
Average	29.63	4.69	28.24	4.53

6. 링크별 차량대기길이 추정 결과

네트워크 내에 있는 전체 14개 링크에 대한 링크별 차량대기길이 추정에 따른 오차(MAPE)는 평균 12.82%로 나타났다. 가장 낮은 오차는 6.19%, 가장 높은 오차는 21.36%가 도출되었다. 가장 높은 정확도를 나타낸 링크(L_{21})는 통행시간과 차량대기길이간의 선형 관계 및 링크 초과 차량대기길이에 대한 높은 추정 정확도에 기인한 것으로 분석되며, 가장 낮은 정확도를 나타낸 링크(L_{56})의 경우에는 링크통행시간 40~60초대에 넓게 분포된 차량대기길 이로 인해 타 링크에 비해 큰 오차가 발생한 것으로 분석된다.

차량대기길이가 링크를 초과한 상황이 나타난 4개 링크 모두 높은 정확도를 나타냈으며, 특정 통행시간대에 차량대기길이가 넓게 분포한 경우에는 정확도가 저하된 결과를 나타내고 있는 것을 알 수 있다.



〈Fig. 8〉 Plotting and MAPE of Vehicle Queue Length Estimation by Link

V. 차량대기길이 추정모형 평가

1. 평가 개요

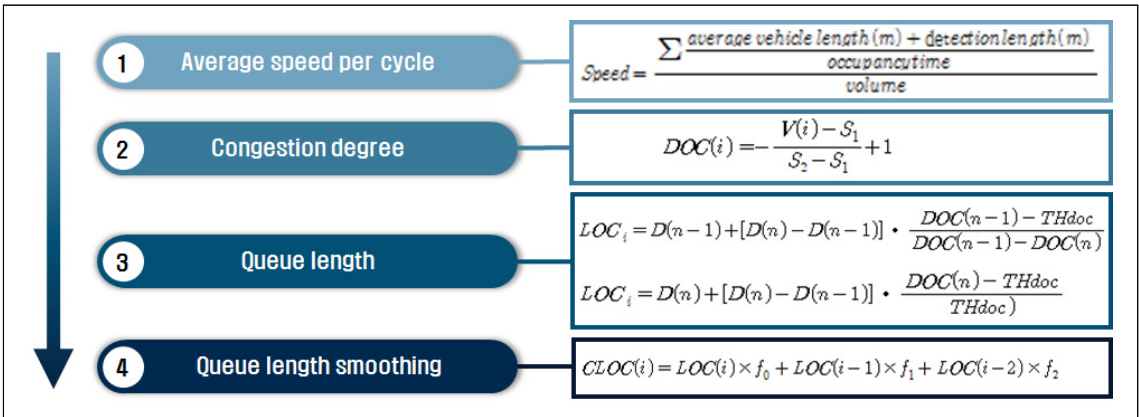
본 연구에서 개발한 딥러닝 기반의 차량대기길이 추정모형에 대한 평가는 실시간 신호제어시스템에 적용된 검지기 데이터 기반의 차량대기길이 추정 모형과의 비교 분석을 통해 수행하였다.

연구대상 네트워크 내에서 비포화부터 과포화 상황까지 다양한 교통상황을 나타낸 L_{21} 링크를 대상으로 VISSIM 시뮬레이션을 통해 평가를 수행하였다. 547m의 해당 링크에 대기행렬검지기를 3개 지점에 설치하고 분석기간 2시간(총 40주기) 동안의 주기별 평균속도를 구독하였다. 검지기별 주기별 평균속도를 정체도로 환산 후 차량대기길이를 산출하였다. 딥러닝 모형의 차량대기길이는 최적 모형에서의 L_{21} 링크 추정결과를 이용하였다. MAE, MAPE를 두 모형의 비교를 위한 평가지표로 선정하여, 그 결과를 비교하였다.

2. 검지기 데이터 기반 차량대기길이 산출

1) 검지기 데이터 기반 차량대기길이 산출 방법

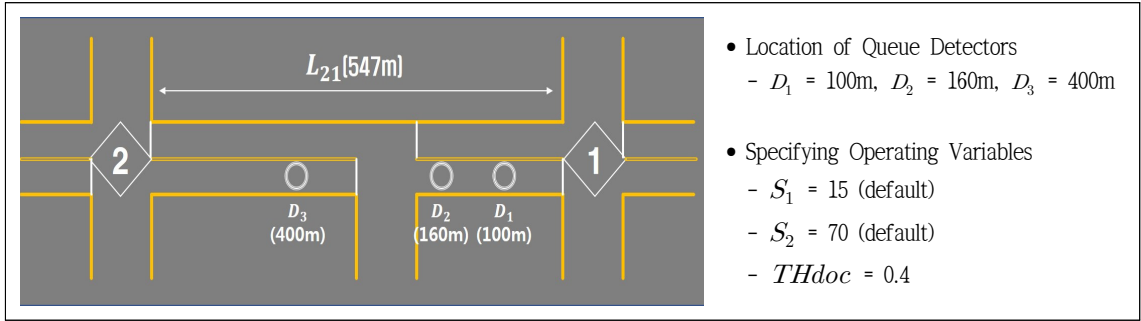
실시간 신호제어시스템에서는 대기행렬길이를 산정하기 위해 정체도를 산정하는데, 정체도는 밀도와 속도 관계식에서 밀도를 대체하는 개념이다. 각 대기행렬 검지기에서 수집되는 개별차량의 점유시간을 속도로 환산하여 주기별 평균속도를 산출한 후 주기별 평균속도를 이용하여 정체도를 구한다. 각 대기행렬 검지기별 정체도를 연결하는 직선을 이용하여 임계 정체도에 해당하는 지점까지를 정체 길이로 판단하고 이를 차량대기길이를 산출한다. 교통상황의 점진적인 변화 추세를 고려하여 해당 주기에 이전 2주기를 포함하여 총 3주기 동안의 차량대기길이를 평활화(Smoothing)하여 해당 주기의 최종 차량대기길이를 산정한다.



〈Fig. 9〉 Vehicle Queue Length Calculation Process based on Detector Data

2) 대기행렬검지기 설치 및 임계정체도 설정

L_{21} 링크(길이 547m) 상에 대기행렬 검지기를 다음과 같이 3개 지점에 설치하고, 해당 링크의 차량대기길이와 검지기별 정체도 등을 고려하여 최적의 임계정체도(THdoc)를 0.4로 설정하였다.



<Fig. 10> Location of Queue Detectors

3) 차량대기길이 산출 결과

실시간 신호제어시스템의 검지기 데이터 기반 차량대기길이 산출방식을 연구대상 네트워크 내 L_{21} 링크에 적용한 결과 분석시간(2시간, 40주기) 동안 산출한 차량대기길이는 평균 16.2%의 오차(MAPE)를 나타냈다.

3. 평가 결과

본 연구에서 개발한 딥러닝 기반의 차량대기길이 추정모형을 L_{21} 링크에 적용한 결과 평균 9.83%의 오차를 나타낸 반면, 실시간 신호제어시스템에 적용된 대기행렬검지기 데이터를 이용한 차량대기길이는 평균적으로 16.2%의 오차를 나타냈다. 이에 본 연구에서 개발한 딥러닝 기반의 모형은 기존 실시간 신호제어시스템에 적용되었던 검지기 데이터 기반의 방식 보다 우수한 성능을 보이는 것으로 분석되었다.

<Table 13> Comparison of Vehicle Queue Length Error between Two Models

Model	MAE(m)	MAPE(%)	Remarks
Detecting Data Based Model	53.52	16.21	
Deep Learning Based Model	29.60	9.83	39% increase

VI. 결 론

본 연구는 교통운영 개선에 필요한 빅데이터 · 인공지능 모델 개발의 일환으로서, 도시부의 링크통행시간 및 통과교통량 등 가용 데이터 등을 이용하여 교통변수로 활용도가 높은 차량대기길이와의 관계를 딥러닝(Deep Learning)을 통해 학습하고 추정하는 인공지능 모델을 구축하는 것을 목표로 하였다.

링크별로 다양한 차량대기길이 분포를 수용하는 통합모형 개발을 위해 6개의 중요교차로를 포함한 총 11개의 교차로와 14개 링크(중요교차로 간을 연결)로 구성된 수원시 CBD 지역 내의 네트워크를 공간적 범위로 하여, 해당 네트워크를 VISSIM 시뮬레이션으로 구현하여 데이터를 구축하였다.

딥러닝 모형의 입력변수 설정을 위해 링크통행시간 및 통과교통량과 차량대기길이와의 상관분석을 수행하였다. 링크통행시간은 차량대기길이와 높은 상관관계가 있으며 통과교통량은 특정 링크통행시간에 넓게 분포된 차량대기길이를 분류하는 데 있어 일조하였다. 특히 차량대기길이가 해당 링크를 초과하여 상류부 링크까지 생성될 때에는 상류부 링크통행시간과 높은 상관관계가 분석되었다.

이에 따라서 차량대기길이 추정모형은 우선 차량대기길이의 링크 초과여부를 분류한 후 링크 초과 및 링크 미초과 상황에서 각각 차량대기길이를 추정하는 3개의 모형으로 모델링하였다. 모든 모형은 DNN 구조를 기반으로 하였으며, 입력변수 및 활성화 함수, 비용 함수, 가중치 초기화, 최적화 알고리즘, 학습환경 등을 설정하고 은닉층과 노드 개수를 다양화하여 학습 및 테스트 후 최소 오차를 나타내는 네트워크 구조를 선정하였다.

차량대기길이 링크 초과여부 분류 모형은 약 98%의 정확도를 나타냈으며, 링크 미초과 모형은 15% 미만, 링크 초과 모형은 5% 미만의 오차를 각각 나타내었다. 링크별 평균 오차는 약 12%로 도출되었다. 이를 기존 실시간 신호제어시스템에 적용된 대기행렬검지기 방식과 비교한 결과 오차가 감소된 것으로 분석되었다.

본 연구에서는 VISSIM 시뮬레이션을 통해 구축한 네트워크 데이터를 활용하여 다양한 링크에 적용할 수 있는 딥러닝 모형을 개발하였으나, 이러한 한계점을 극복하기 위해 실제 현장조사를 통한 차량대기길이와 링크통행시간, 통과교통량 등의 데이터를 구축하여 본 연구에서 제시한 딥러닝 기법의 검증과 동시에 다양한 도로로의 확대 가능방안에 대한 연구를 제안하는 바이다.

ACKNOWLEDGEMENTS

이 논문은 미래창조과학부 재원으로 경찰청과 치안과학기술연구개발사업단의 지원을 받아 수행된 치안과학기술연구개발사업임.(PA-A000001)

REFERENCES

- Allsop R. E.(1972), "Delay at a Fixed Time Traffic Signal-Theoretical Analyses," *Transportation Science*, vol. 6, no. 3, pp.260-285.
- Cox D. R.(1958), "The Regression Analysis of Binary Sequences," *Journal of the Royal Statistical Society, Series B(Methodological)*, vol. 20, no. 2, pp.215-242.
- Dunne M. C.(1967), "Traffic Delay at a Signalized Intersection with Binomial Arrivals," *Transportation Science*, vol. 1, pp.24-31.
- Hwang J. -S., Lee Y. -J. and Lee C. -K.(2017), "Development of Queue Length, Link Travel Time Estimation and Traffic Condition Decision Algorithm using Taxi GPS Data," *The Journal of the Korea Institute of Intelligent Transportation Systems*, vol. 16, no. 3, pp.59-72.
- Kang J. -H.(2005), "The Development of Traffic Queue Length Estimation Algorithm Using the Occupancy Rates," *The Journal of the Korea Institute of Intelligent Transportation Systems*, vol. 4, no. 2, pp.13-22.
- Kim D. -O. and Woo Y. -H.(1998), "Modelling for Queue Estimation on the Approach of Signalized Intersections," *Journal of the Korean Society of Civil Engineers*, vol. 18, no. 2-2, pp.159-171.
- Kingma D. P. and Ba J. L.(2014), "ADAM: A Method for Stochastic Optimization," arXiv:1412.6980.
- Lee Y. -J., Hwang J. -S. and Lee C. -K.(2016), "Study on Queue Length Estimation using GPS Trajectory Data," *The Journal of the Korea Institute of Intelligent Transportation Systems*, vol. 15, no. 3, pp.45-51.
- May A. D.(1965), "Traffic Flow Theory the Traffic Engineer's Challenge," *Proc. Inst. Traffic Eng.*, pp.290-303.

- Moon H. -R.(2005), "A development of Traffic Queue Length Measuring Algorithm Using Average Occupancy Time Based on COSMOS," *Korean Society of Civil Engineers Conference*, pp.4230-4234.
- Newell G. F.(1965), "Approximation Methods for Queues with Applications to the Fixed cycle Traffic Light," *SIMA Rev.*, vol. 7, no. 2, pp.223-240.
- Rosenblatt F.(1958), "The perceptron : A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp.386-408.
- Rumelhart D. E., Hinton G. E. and Williams R. J.(1986), "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 9, pp.533 - 536.
- Wardrop J. G.(1952), "Some Theoretical Aspects of Road Traffic Research," *Proc. Inst. Civ. Eng.* 1, pp.325-362.
- Winsten C. B.(1956), M. Beckmann and C. B. McGuire, "Studies in the Economics of Transportation," Yale University Press.