ARTICLE

베이지안 확률 추론 기반의 자율주행차 사고 계층분석 및 사고 심각도 모형 연구

이송하[1] · 박준영[2*]
[1]한양대학교 스마트시티공학과 석박사통합과정, [2]한양대학교 교통·물류공학과, 스마트시티공학과 부교수

A Study on Autonomous Vehicle Crash Hierarchy Analysis and Severity Model Based on Bayesian Probabilistic Inference

LEE, Songha[1] · PARK, Juneyoung[2*]

[1]The Ph.D. Course, Smart City Engineering, Hanyang University, Ansan 15588, Korea
[2]Associate Professor, Department of Transportation & Logistics Engineering, Smart City Engineering, Hanyang University, Ansan 15588, Korea

*Corresponding author: juneyoung@hanyang.ac.kr

Abstract

With the advancement of Advanced Driver Assistance System technology, autonomous driving technology is get a attention as a key component of smart transportation systems. However, verification of safety about autonomous vehicle is required due to the continuous traffic crash involving autonomous vehicles. In particular, to ensure the safety of vulnerable road users such as pedestrians as well as autonomous vehicle on handicap sections where autonomous mode is unexpectedly disengaged and the driver's response time is insufficient, this requires a technical/institutional approach and a continuous understanding for autonomous vehicle crashes. Therefore, this study used autonomous vehicle crash record data to extract impact factors from the autonomous vehicle collision reports provided by the California Department of Motor Vehicles, and used latent class analysis and ordered probit model with Bayesian probability inference to deal with the traffic crash sample. As a result, traffic crash involving vulnerable road users such as pedestrians and cyclists, scooter, mainly occur at bus stops or bicycle lanes where road users cross, and the traffic crash severity increases when light condition is dark street. On the other hand, for the traffic crash which is not involve vulnerable road users, the crash classes were divided and the traffic crash severity increased depending on weather conditions, autonomous vehicle movement, day and night, etc. When considering unobserved heterogeneity in geometry and weather conditions, it was found that the the variance size was smaller and the confidence interval was narrower based on the Bayesian ordered probit model with random effects unlike fixed effects. Specifically, the traffic crash severity increased mainly in movements including turns around signalized intersections and for the other crash, which does not include the vulnerable road users. This means that intersections require an exploring of vehicle movements due to the cross to the various road users, complexity of traffic flow, and unpredictability of vehicles and pedestrians. By developing a detailed autonomous vehicle safety evaluation scenario on these study, it is expected that more efficient and reasonable safety evaluation of level 4 to level 5 autonomous driving technology will be possible.

Keywords: autonomous vehicle collision report, bayesian latent class, random effect, traffic crash severity, text mining

J. Korean Soc. Transp.
Vol.42, No.1, pp.77-93, February 2024
https://doi.org/10.7470/jkst.2024.42.1.077

pISSN : 1229-1366
eISSN : 2234-4217

ARTICLE HISTORY
Received: 4 December 2023
Revised: 6 December 2023
Accepted: 15 January 2024

Copyright ©
Korean Society of Transportation

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

초록

첨단 운전자지원 시스템 기술 고도화에 따라 자율 주행 기술은 미래의 스마트 교통 시스템 주요 구성 요소로 주목받고 있지만 꾸준히 발생하는 자율주행차의 사고로 인해 안전성에 대한 체계적인 검증이 요구되고 있다. 특히, 예상치 못하게 제어권이 해제되어 운전자의 대응 시간이 충분하지 않은 상황적 범위라고 칭할 수 있는 핸디캡 구간이 반복적으로 발생하는 도시부 도로에서 자율주행차뿐만 아니라 보행자와 같은 취약한 도로 이용자들의 안전성을 확보하기 위해서는 기술적/ 제도적 접근과 자율 차 사고 원인에 대한 지속적인 이해가 필요하다. 따라서 본 연구는 캘리포니아 자동차관리국에서 제공하는 자율 차 사고 보고서 중 세부 영향 요인과 변수를 추출하기 위해 사고 세부 기록 데이터를 이용하였고 사고 표본 보완을 위해 베이지안 확률추론 기반의 잠재 계층분석과 순서형 프로빗 모형을 통해 계층별 사고 심각도와 설명 변수 간 상관성을 설명하였다. 그 결과 보행자, 자전거 이용자와 같은 취약한 도로 이용자와 관련된 사고는 주로 도로 이용자들이 교차가 발생하는 버스정류장 또는 자전거 도로가 있는 공간에서 발생하고 어두운 조명시설일 때 사고 심각도가 증가하였다. 반면에 나머지 사고의 경우, 기상 조건, 자율 차의 움직임, 주야 등에 따라 계층이 세분화되고 사고 심각도가 증가하였다. 특히, 그리고 기하구조와 기상 조건에 대한 이질성을 고려하였을 때, 고정 효과에 비해 임의 효과를 이용한 베이지안 순서형 프로빗 모형 경우, 분산의 크기가 작아져 신뢰구간 범위가 좁아진 것을 확인할 수 있었다. 이러한 연구 결과를 바탕으로 세분화된 자율 차 안전성 평가시나리오를 구성하면 Level 4-Level 5 수준의 자율 주행 기술에 대한 보다 효율적이고 합리적인 안전성 평가가 가능할 것이라고 기대한다.

주요어: 자율주행차 사고 보고서, 베이지안 잠재 계층분석, 임의 효과, 사고 심각도, 텍스트마이닝

서론

첨단 운전자 지원시스템(ADAS, Advanced Driver Assistance System) 기술이 고도화되면서 자율주행차(또는 자율 차)는 4차 산업혁명의 핵심 패러다임이자 스마트 교통 시스템의 주축으로 주목받고 있다. 자율주행차의 도입은 도로 안전을 크게 향상할 수 있는 잠재력을 가지고 있고 향상된 안전성을 기반으로 불필요한 지체가 감소하여 도로 및 교차로의 용량이 증대할 것으로 기대하여 GM, Google 등 주요 차량 제조업체에서는 다양한 교통환경을 대응할 수 있는 자율주행 시스템을 상용화하고자 노력하고 있다(Morando et al., 2018; Hoogendoorn et al., 2014). 또한, 자율주행 정의와 관련 교통법규를 제정하거나 상암, 판교 등 지역에서 시범운행을 수행하는 등 자율협력 주행 시대에 대응하기 위한 선제적 조치를 하고 있다.

하지만 각광받는 기대에 비해 자율주행차의 안전사고는 꾸준히 발생하여 여전히 안전성에 대한 의문점이 제기되고 있다. 대표적으로 미국 Uber사의 자율주행 3단계 택시가 도로를 횡단하는 보행자를 감지하였으나 차량이 스스로 대응하지 못하고 운전자에게 제어권을 이양하는 과정에서 충돌이 발생하였다(KICT, 2020). 이처럼 예상치 못하게 제어권이 해제되어 운전자의 대응 시간이 충분하지 않은 상황적 범위인 자율주행 핸디캡 구간에서 자율주행차의 주행 안정성을 확보하기 위해서는 기술적/ 제도적 접근과 자율주행차 사고 원인에 대한 지속적인 이해가 필요하다(Jeon and Kim, 2021). 센서 기반의 차량 안전성을 평가하기 위해서는 정량적인 연구가 필요하고 이를 기반으로 문제를 해결하기 위한 접근 방식이 필요하다(Sohrabi et al., 2021). 현재까지는 부족한 자율 차 안전성 평가 데이터 구축을 위해 충격 실험을 통해 차량의 종합적인 안전성을 평가하는 자동차 안전도 평가(New Car Assessment System, NCAP)를 수행하거나 미국 캘리포니아 주와 같이 차량 제조업체에 '자율 주행 제어권 이양 보고서(Disengagement Report)'와 자율 차 사고 발생 관련 보고서(Autonomous Vehicle Collision Reports) 제출을 요구하여 실증 데이터베이스를 구축하고 있다.

자율주행차는 다양한 교통 시스템 구성 요소들이 복합적으로 상호작용하고 있는 도심부 도로에서 주로 시범 운행

하고 있어 보행자와 같이 취약한 도로 이용자(VRU, Vulnerable Road User) 안전에 대한 자율주행차의 예상 영향을 평가하고 이해의 필요성이 증가하고 있다. 예를 들어, 교통 신호로 통행우선권을 구분하는 상황과는 달리 비신호 교차로와 같이 손짓 등 비언어적인 의사소통을 빈번하게 발생하는 상황과 같이 운전자는 보행자와 같은 모든 도로 이용자들의 관점을 고려하여 의사결정 프로세스를 탐색하고 정의하는 것이 중요하다(Morris et al., 2021; Hagenzieker et al., 2020). 하지만 자율 차 사고 데이터를 활용한 다수의 연구에서는 부족한 사고 표본으로 인해 자율 차와 일반 차의 상호작용만을 고려하고 사고 영향 요인을 식별하여 분석 결과를 일반화하였다(Ashraf et al., 2021; Das et al., 2020). 또한, 보행자, 자전거 등 취약한 도로 이용자(VRU, Vulnerable Road Users) 계층과 관련된 데이터 표본을 확충하기 위해 설문조사, 시뮬레이션 등으로 대체하여 완벽하게 현실을 고증할 수 없다는 한계가 있다(Deb et al., 2017; Millard-Ball, 2018).

따라서, 본 연구는 2019년부터 2023년 8월까지 발생한 자율주행차 사고 중 취약한 도로 이용자가 관련된 사고와 그 외의 사고를 구분하고 잠재 계층분석을 통해 각 유형에 내재되어 있는 사고 패턴과 사고 심각도에 영향을 미치는 요인들을 확인하고자 하였다.

기존 문헌 고찰

1. 자율주행차 사고 보고서 기반 안전성 평가 관련 연구

Lee et al.(2023a)는 캘리포니아 차량 관리국에서 제공하는 자율 차 사고 165건을 대상으로 도심부의 사고시나리오 6개를 도출하였는데 이때 데이터 마이닝 기법 중 하나인 연관규칙을 이용하여 차량의 충돌 전 움직임과 날씨, 조명, 위치 등 다양한 사고 요인을 조합하였다. Ashraf et al.(2021)은 2016년부터 2020년 2월까지 발생한 198건의 자율주행차 관련 충돌사고 중 사고 심각도, 자율 주행 모드, 충돌 유형, 충돌 전 자율주행차 움직임을 기준으로 사고 패턴을 도출하였고 교차로에서 자율 차가 정차하고 있을 때 후미 추돌 유형의 사고의 빈도가 가장 높았다고 하였다. Das et al.(2020)은 베이지안 잠재 계층분석을 활용하여 자율주행차 사고에 내재된 주요 특성을 파악하고자 하였는데 주로 차량의 움직임이 회전 동작을 하거나 조명인 없는 도로에서 후미 추돌 또는 다중 차량이 포함된 사고가 발생할 가능성이 높았다. 또한, 비정형 데이터인 사고 세부 기록 내용을 이용해서 자율주행 모드 상황과 수동 주행 상황에 관련된 사고 특성을 이해하고자 하였다. Lee et al.(2023b)는 WordStat를 이용하여 사고 세부 내용에 포함된 사고 상황과 다양한 요인 추출 및 도로 기하구조 및 주변 교통환경 데이터 수집을 통해 자율주행차 도입에 따른 도심부 교통사고 변화를 확인하고자 하였다. Kutela et al.(2022)는 2017년에서 2020년까지의 자율주행차 사고 데이터를 수집하여 취약한 도로 이용자계층을 대상으로 사고 예측모형을 개발했는데 이때 취약한 도로 이용자는 보행자, 자전거 또는 스쿠터 이용자를 기준으로 하였고 전체 자율주행차 사고 중 해당 사고를 구분하기 위해 텍스트 마이닝 기법을 이용하였다. Novat et al.(2023)은 자율 차와 일반 차량간의 다양한 충돌 유형에 영향을 미치는 변수를 검토하기 위해 베이지안 네트워크를 이용하였고 그 결과, 자율 차와 일반 차 사이의 인지 반응시간 차이로 인한 혼합교통류 환경에서는 후미 추돌이 증가할 가능성이 높아서 이와 관련된 안전관리 전략을 사전에 수립할 필요가 있다고 제언하였다.

2. 교통사고 세부 기록을 활용한 교통 안전성 평가 관련 연구

교통사고 조사관 또는 관련 담당자가 사고 상황에 대한 자세한 기록을 작성하는 사고 세부 기록은 교통사고 심각도 및 관련 인과관계 요인 식별에 활용될 수 있다. Kwayu et al.(2021)은 구조 주제 모델링(STM, Structure Topic Modelling) 접근 방식과 네트워크 토폴로지(NT, Netwrok Topology) 분석을 사용하여 미국 미시간 주의 사망사고 내 충돌 유형별 요인간 관계를 확인하였다. Kwayu et al.(2020)은 사고 세부 기록을 활용하여 신호 교차로에서 운전

자의 위험한 행동에 대한 의미론적 분석 및 분류를 수행하였고 제안된 알고리즘은 상당한 정확도로 '신호 위반', '끼어들기 금지 위반'의 위험한 행동을 식별하였다. Arteaga et al.(2020)은 LIME(Local Interpretable Model-Agnostic Descriptions)을 활용하여 호주 퀸즈랜드 내 대형차량 사고의 인과관계 요인을 식별하였을 때 '측면충돌', '오토바이','손상 위치:차량 전면부' 등의 단어가 강한 연관성을 가지고 있다고 밝혔다. Kutela et al.(2022)는 VRU 관련 사고를 분류하기 위해 R 패키지 Quanteda 및 igraph를 사용하여 VRU 및 비 VRU 관련 충돌에 대해 가장 자주 발생하는 키워드 및 커뮤니티의 패턴을 시각화하고 텍스트 네트워크 분석을 수행하였다.

3. 도시부 교통사고 심각도 모형 관련 연구

도로 사용자 중 자동차 운전자보다 VRU는 사고 발생 시, 상처를 입을 가능성이 높아서 이러한 사고의 특성을 해석하고 사고 심각도를 줄이는 것이 교통안전 분야에서 핵심적인 연구 분야로 자리 잡고 있다. Yuan et al.(2022b)는 2021년 중국에서 발생한 차량과 VRU 간의 실제 교통 충돌 378건을 조사하여 부상 심각도에 영향을 미치는 인적, 차량 및 도로 요인간 관계를 해석하였을 때 VRU 사고를 예방하기 위해서는 자율주행차 Lidar 기술을 활용하여 위험 행동을 식별할 필요성이 있다고 하였다. 또한, 사고 심각도 추정시 관찰되지 않은 특성이 존재할 수 있어서 특정 변수의 확률분포에 따라 필연적으로 발생할 가능성도 있어서 관찰되지 않은 특성과의 상관관계를 경험적으로 탐색하기 위해서는 임의 효과와 임의 매개변수 설정을 통해 비관측 이질성을 고려할 필요가 있고 이에 관한 심층적인 연구가 필요하다(Yuan et al., 2022a).

4. 기존 연구들과의 차별성

기존의 연구들은 모든 도로 이용자들의 세부 특성을 고려하지 않고 자율 차 사고 요인의 영향력을 확인하기 위해 군집분석이나 연관 분석 등 단일적인 방법론을 이용하였다. 특히, 연관 분석은 대용량의 데이터에서 변수간 관계를 탐색하는 방법이지만 소량의 자율 차 사고 데이터만으로 인과관계를 일반화하기에는 한계가 있다. 또한, 자율주행차 사고 보고서 내 사고 조사 항목만을 이용하여 영향 변수 간 상호작용을 식별하기에는 어려움이 있어서 사고 세부 기록을 통해 구체적인 영향 요인을 추출할 필요성이 있다. 따라서 본 연구는 사고 캘리포니아 자동차관리국에서 제공하는 자율 차 사고 보고서를 기반으로 텍스트 분석을 통해 VRU 포함을 식별하였고 잠재 계층분석, 사고 심각도 모형을 개발함으로써 심각도 수준에 따라 유형별 사고 특성을 탐색하였고 부족한 사고 표본 수의 유효성을 확보하기 위해 베이지안 확률추론을 활용하였다.

분석 방법론

1. 연구 흐름도

본 연구는 자율 차 실사고 데이터 중 도로 이용자 계층을 기준으로 VRU 사고와 일반사고(수동 차량과 자율 차간 사고)를 구분하여 계층별 사고 심각도에 영향을 주는 요인을 분석하고자 하였다. 먼저, 비정형 데이터인 사고 세부 기록을 활용하여 VRU와 관련된 사고, 그 외의 사고를 구분하고 도로 이용자 계층별 사고에 내재 되어 있는 특성을 탐색하였다. 두 번째는 베이지안 확률추론을 기반으로 잠재된 세부 계층을 나누고 계층별 사고 심각도와 관련된 변수들의 영향력을 해석하고자 하였다(Figure 1).

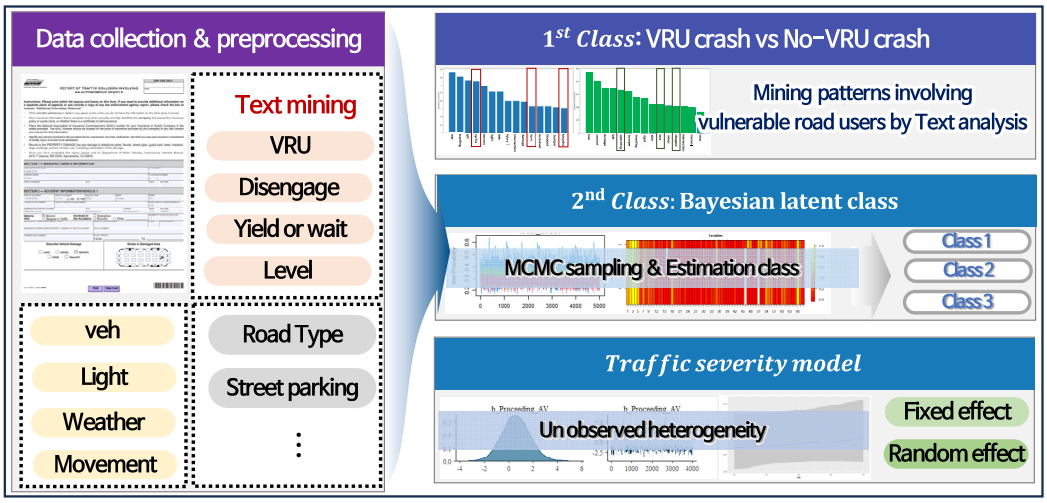

Figure 1. Research flow chart

2. 활용 데이터

본 연구는 활용한 캘리포니아 자동차관리국(DMV, Department of Motor Vehicles)에서 제공하는 개방형 데이터인 자율주행차 사고 데이터 중에서 2019년-2023년 8월까지 수집된 522건의 사고 데이터를 이용하였다. 캘리포니아 자동차관리국은 자율주행차 운행을 위한 체계적인 프레임워크 마련을 위해 차량 제조업체에 자율주행차 평가 및 상용화의 권한을 부여하고 사고 발생 시 10일 이내 보고서를 작성하여 제출하도록 규정하고 있다. 해당 보고서는 그림 사고 발생 위치, 자율주행차 또는 일반차량 움직임, 충돌 유형, 자율주행 모드 유무에 대한 정보를 제공하고 있다(Figure 2).

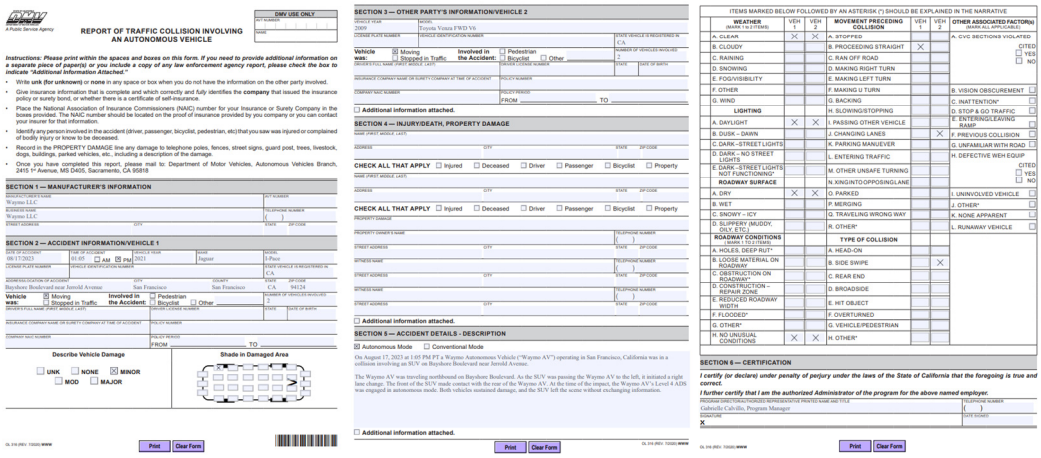

Figure 2. Example of report of traffic collision involving an autonomous vehicle

또한, 자율 차 사고 발생 시 주변 환경 특성의 영향력을 확인하기 위해 Google Earth를 통해 도로 유형, 차선 수, 노상 주차 유무, 버스정류장 유무, 자전거 도로 유무, 버스 또는 택시 전용 도로 유무, 회전 차로 유무, 중앙분리대 유무에 대한 공간정보를 수집하였다. 그리고 VRU 사고와 일반사고를 구분하기 위해 자율 차 사고 세부 기록 데이터를 이용하였고 VRU 포함 유무를 포함해서 총 4가지 설명 변수를 추가하였고, 사고 세부 기록 내용 데이터의 전처리 과정을 Figure 3에 나타냈다. 사고 기록에서 VRU와 관련된 사고를 추출하기 위해 관련 연구 문헌들을 참고하여 보행

자, 자전거/오토바이 이용자와 관련이 있거나 충돌이 발생했다고 명시된 어구(involving pedestrian/cyclist/scooter, by pedestrian/bike) 포함 여부에 따라 VRU사고를 추출하였다(Morris et al., 2021; Olszewski et al., 2019). 본 연구는 도심부 도로를 대상으로 분석하기 위해 사고 발생 장소가 고속도로 또는 간선도로에서 발생한 16건의 사고를 제외하고 텍스트 분석을 통해 추출된 VRU와 관련된 사고는 67건, 그 외의 사고는 439건의 데이터를 이용하였고, 사고 심각도 모형에 활용된 변수는 Table 1과 같다. Table 1의 AV는 자율주행차를 의미하고 MV는 자율주행차가 아닌 상대 일반차량(수동 주행 차량)을 의미한다.

Crash narrative	Yielding or waiting O/X(Blue word)	Disengagement O/X(Orange word)	VRU O/X (Red word)	Autonomous driving level (Green word)
On April 26, 2022 at 1:19 PM PST a Waymo Autonomous Vehicle ("Waymo AV") operating in San Fra ncisco, California was in a collision involving a cyclist on Folsom Street at Mabini Street. While compl eting a right turn from Folsom Street onto Mabini Street in manual mode, the test driver slowed to yi eld to a cyclist approaching from the right. The cyclist made contact with the right front fender of the Waymo AV after the Waymo had come to a stop. At the time of the impact, the Waymo AV's Level 4 ADS was not engaged and a test driver was operating the Waymo AV in manual mode. The Waymo AV sustained minor damage.	1	0	1	4
At about 11:35 am our vehicle (car 1) was waiting behind a USPS truck to make an unprotected left t urn, meanwhile one other vehicle (car 2) waited behind our vehicle. Another non involved vehicle noti ces the queue of vehicles and passes on the right in the bike lane. As the USPS truck begins to make the turn, car 1 begins to move forward at 7 MPH then slowed to 3 MPH when yielding to the vehicle passing in the bike lane. Car 2 directly behind us strikes our passenger rear bumper. The safety drive r then disengaged the vehicle (car 1) and moved into a nearby parking lot to exchange information.	1	1	0	4

Figure 3. Process of creating explanatory variables from crash narrative

Table 1. Summary statistics of the independent variables

Variable	Description	Frequency	Proportion (%)
Severity	No injury	47	9.28
	Minor	370	73.21
	Mod	76	15.01
	Major	13	2.50
Autonomous mode	Autonomous	270	53.35%
	Conventional	236	46.65%
Number of vehicle	1 vehicle	59	11.66
	2 vehicle	443	87.54
	3 vehicle	4	0.08
Weather	Clear	440	86.95
	Cloudy	44	8.69
	Rain	19	3.75
	Fog	2	0.6
Light	Daylight	348	68.77
	Dark streetlight	146	28.85
	Dust dawn	10	1.97
	Dark nostreetlight	2	0.39
Surface	Dry	471	93.08
	Wet	24	4.74
	Other	11	2.17
AV_behavior	Backing_AV	21	4.15
	Changing lanes_AV	8	1.58
	Entering_AV	5	0.98
	Left turn_AV	25	4.94
	Merging_AV	2	0.39
	Parked_AV	4	0.79
	Parking-manuvering_AV	17	3.35
	Passing_AV	2	0.39
	Proceeding_AV	138	27.27
	Right turn_AV	27	5.33
	Slowing_AV/Stopping_AV	45	8.89
	Stopped_AV	204	40.31
	U turn_AV	3	0.59
	Other_AV	4	0.79

Table 1. Summary statistics of the independent variables (continued)

Variable	Description	Frequency	Proportion (%)
MV_behavior	Backing_MV	24	4.74
	Changing lanes_MV	30	5.92
	Entering_MV	8	1.58
	Left turn_MV	25	4.94
	Parked_MV	37	7.31
	Parking-manuvering_MV	8	1.58
	Passing_MV	19	3.75
	Proceeding_MV	213	42.09
	Ran off road_MV	20	3.95
	Right turn_MV	30	5.93
	Slowing/Stopping_MV	12	2.37
	Stopped_MV	18	3.55
	Xing into opposing lane_MV	7	1.38
	Other_MV	45	8.89
	None_MV	10	1.97
Crash type	Rear end	159	31.42
	Sideswipe	92	18.18
	Head on	44	8.69
	Broadside	30	5.92
	Hit object	28	5.53
	Other	151	29.84
	None	2	0.39
Night or day	Day	350	69.17
	Night	256	30.83
Time of the week	Weekday	383	75.69
	Weekend	123	24.31
Involving an AV's yielding or waiting	Yes	109	21.54
	No	397	78.45
Involving a VRU	Yes	67	13.24
	No	439	86.76
Involving a disengagement for autonomous mode	Yes (Disengagement)	85	16.79
	No	421	83.20
Level	Level 2	4	0.79
	Level 3	252	45.84
	Level 4	242	47.82
	Level 5	8	5.53
Road type	Signalized intersection	225	44.46
	Unsignalized intersection	135	26.67
	Local street	145	28.65
	Other	1	0.19
Number of lane	1 lane	74	14.62
	2 lane	144	28.45
	3 lane	87	17.19
	4 lane and more	182	39.91
Including street parking	Yes	358	70.75
	No	148	29.25
Including bus stop	Yes	176	34.78
	No	330	65.21
Including bike lane	Yes	282	55.73
	No	224	44.27
Including divided median	Yes	93	18.38
	No	413	81.62
Including unprotected turning lane	Yes	121	23.91
	No	385	76.09

note: reference coding scheme, in which the design variables (or dummy variables) take the value of 0 (reference category) or 1, was used to represent the various categories of the explanatory factors

3. 텍스트 마이닝(Text mining)

비정형 데이터에 나타나는 단어 또는 문구의 빈도를 계산하는 방법을 TF(Term Frequency) 방법이라고 하고, 극단적 빈도에 따라 중요도가 달라지는 것을 방지하기 위해 역 문서 빈도(문서의 빈도에 특정 식을 취함)를 사용하여 단어마다 중요도를 가중치로 주는 TF-IDF(Term Frequency-Inverse Document Frequency)를 활용하여 사고 세부 기록 내 VRU 사고와 일반사고의 특징을 구분하고자 하였다. 이때 TF-IDF의 수식은 Equation 1-Equation 3을 통해 주요 사고 특성의 빈도를 계산할 수 있다.

$$TF(t,d) = \frac{Number\,of\,word\,'t'\ on\,docs\,'d'}{Number\,of\,all\ word\ on\,docs\,'d'} \tag{1}$$

$$IDF(t,D) = \log\left(\frac{Number\,of\,all\,docs}{Number\,of\,docs\ that\ cludes\,word\,'t'}\right) \tag{2}$$

$$TF-IDF(t,d,D) = TF(t,d) \times IDF(t,D) \tag{3}$$

4. 베이지안 잠재 계층분석(Bayesian latent class analysis)

다른 여러 가지 분포를 포함하는 유한 혼합 모형에 속하는 잠재 계층 분석(LCA, Latent Class Analysis)은 이질적인 충돌 데이터로부터 동질적인 하위 클래스를 식별하여 하위 클래스 내 동질성과 하위 클래스 간의 이질성을 극대화하는 확률론적 모델 기반 군집분석 방법이고 수식은 Equation 4-Equation 7과 같다(Chen et al., 2020; Ahlmann-Eltze et al., 2018; Das et al., 2020; Li et al., 2018). Equation 4와 Equation 5의 α, β는 외부적으로 해석되어 모델 희소성을 관리하는 매개변수이고 Equation 4는 Dirichlet 분포에 따라 군집 크기를 정의하고 Equation 5에 따라 각 데이터 z를 잠재 클래스에 할당하고 Equation 7은 개인 i(확률 벡터 $U_{j,k}$에 따른 다항 분포에서 클래스 k에 속함)의 응답을 규정한다. 모델의 결합 분포 방정식은 Equation 8과 같다.

$$\lambda \mid \alpha \sim Dirichlet(\alpha) \tag{4}$$

$$z_i \mid \lambda \sim Multinomial(\lambda) \tag{5}$$

$$U_{j,k} \mid \beta \sim Dirchlet(\beta) \tag{6}$$

$$X_{i,j} \mid U_j, z = k \sim Multinomial(U_{j,k}) \tag{7}$$

$$p(\lambda, z, U, X \mid \alpha, \beta) = p(\lambda \mid \alpha)\prod_{i=1}^{I} p(z_i \mid \lambda)\prod_{j=1}^{J}\prod_{k=1}^{K} p(U_{j,k} \mid \beta)\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} p(X_{i,j} \mid U_{j,k}) \tag{8}$$

베이지안 확률추론은 데이터의 사전분포 때문에 특정 변수의 출현 정도인 사후분포를 추정하는 확인하는 방법으로 사후 확률분포를 추정하는 방법에는 EM 알고리즘, Gibbs 샘플링, 변형된 베이즈 정리 기반 근사화 방법이 있다(White and Murphy, 2014). 본 연구에서는 마르코프 연쇄 체인에 기반한 확률분포로부터 원하는 분포의 표본을 추출하는 MCMC(Markov Chain Monte Carlo) 샘플링 알고리즘 중 Gibbs 알고리즘을 활용하였다. 영향 요인의 수가 많고 복잡할수록 완전 조건부 분포를 구하기 어려운데 이를 해결하기 위해 고차원의 분포를 저차원으로 바꿀 수 있는 Gibbs 알고리즘을 이용하였고 샘플링 과정은 Figure 4와 같다.

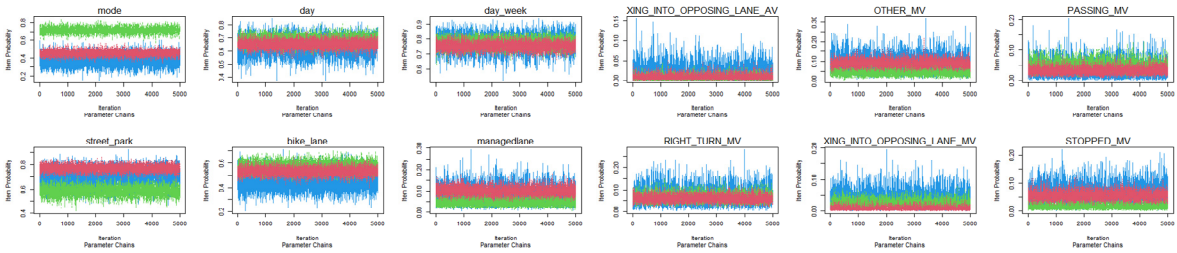

Figure 4. Process of estimation posterior class distribution using Gibbs sampling

5. (베이지안) 순서형 프로빗 모형

사고 심각도와 같은 순위가 있는 자료를 설명하기 위해서는 주로 순서형 프로빗 모형을 사용하여 사고 등급에 따라 해석이 가능한 모형을 적용하고 있다(Lee et al., 2020; Yuan et al., 2022a). X_i는 설명 변수의 관측값이고, ϵ_i는 정규분포를 따르는 오차항을 의미하고 β_i는 설명 변수의 계수이며 사고 심각도가 Y_i일 때, 관찰 i가 j번째 부상 심각도일 확률은 Equation 9-Equation 11로 표현할 수 있다. 본 연구에서는 VRU와 관련된 사고와 그 외의 사고 중 잠재 계층별로 사고 심각도 모형을 개발하였고 독립변수는 앞서 언급한 자율주행차 실사고 보고서 자료를 기반으로 정의하였다. 종속 변수인 부상 심각도는 대물 피해, 경상(Minor), 중상(Mod), 치명상(Major)으로 범위를 구분하였는데 VRU와 관련된 사고의 경우, 치명적인 사고의 표본이 부족하여 중상과 같은 범주로 구분하여 Y_i는 0, 1, 2로 범주를 설정하였고, 그 외의 사고의 Y_i는 0, 1, 2, 3의 범위를 가지도록 표현하였다.

$$Y_i = \beta_i X_i + \epsilon_i,\ i = 1 \cdots, n \tag{9}$$

$$Y_i = j, \qquad \text{if } u_{i,j-1} \le Y_i \le u_{i,j} \Leftrightarrow \begin{cases} 0 \text{ if } -\infty \le Y_i \le 0(PDO) \\ 1 \text{ if } 0 < Y_i \le 1(\text{min}or\,jury) \\ 2 \text{ if } 1 < Y_i \le 2(\text{mod}\,jury) \\ 3 \text{ if } 2 < Y_i \le 3(major\,jury) \end{cases} \tag{10}$$

$$\begin{aligned} P(y=0) &= \Phi(-\beta_i X_i) \\ P(y=1) &= \Phi(\mu_1 - \beta_i X_i) - \Phi(-\beta_i X_i) \\ P(y=2) &= \Phi(\mu_2 - \beta_i X_i) - \Phi(\mu_1 - \beta_i X_i) \end{aligned} \tag{11}$$

또한, 사고 심각도 모형 개발 시, 도로 구간 또는 기상 조건의 이질성(Heterogeneity)을 설명하기 위해 최대 우도추정법 기반 임의 효과를 적용하였고, 임의 효과는 독립변수의 계수 β가 임의 분포항(Random distributed term) φ_i이 있는 Equation 12와 같이 표현할 수 있다. Equation 12를 순서형 프로빗 모형에 적용한 사고 심각도별 발생확률은 Equation 13과 같고 정규분포 형태의 확률분포를 이용하여 모형의 계수를 추정하였다. 여기서 P_j는 자율 차 사고가 발생했을 때, 심각도 j를 초래하는 사고 발생확률, $f(\beta_i \mid \varphi)$는 정규분포 확률밀도함수와 관련된 매개변수의 벡터를 의미한다. 최대 우도추정법을 이용하면 무작위로 분포된 자율주행차 사고의 영향 변수 계수 β_i가 정규분포 밀도함수 f에 따라 $\beta_i \sim N(0, \sigma_\beta^2)$, $\varphi_i \sim N(0, \sigma^2)$를 따르는 사고 심각도별 발생확률이 추정된다.

$$\beta_i = \beta + \varphi_i \tag{12}$$

$$P_j = \int \frac{\exp(\beta_i X_{i,j})}{\sum_i \exp(\beta_i X_{i,j})} f(\beta_i \mid \varphi) d\beta \tag{13}$$

표본 크기는 회귀 모델의 매개변수 추정 정확도에 큰 영향을 미칠 수 있어서 일반적으로 계량경제학 분야에서 모수 추정에 있어 관찰된 데이터의 우도를 사용하여 미지의 사전분포 신뢰수준을 바탕으로 사후분포를 추정하는 베이지안 모델을 활용한다(Xie et al., 2009). 본 연구에서 제한된 자율 차 사고의 표본을 기반으로 제안한 모형의 신뢰도를 향상하기 위해 베이지안 확률추론을 이용하였고. 베이지안 순서형 프로빗 모형에 입력된 설명 변수는 일반 순서형 프로빗 모형에서 유의수준 0.1 이하로 통계적으로 유의한 설명 변수에 따른 사고 심각도별 발생확률을 추정하였다.

6. 베이지안 모형 선택

본 연구에서는 모형 적합도 평가지표로 DIC(Deviance Information Criterion) 지표를 활용하였고 DIC는 모형의 설명력과 복잡성에 대한 패널티의 합으로 추정되며 사후 분포로분터 얻은 매개변수의 로그 가능도와 실질적인 모수의 개수에 따라 편차의 사후 기대치가 달라진다. $\overline{D}$는 사후 기댓값으로 정의되는 모형의 설명력을 의미하고 PD는 모형의 복잡성으로 모형이 주어진 데이터의 특성을 잘 설명할수록 DIC 지표 값이 작아지고 Equation 14처럼 표현할 수 있다(Spiegelhalter et al., 2002).

$$DIC = \overline{D} + PD = D(\overline{\theta}) + 2PD \tag{14}$$

연구 결과

1. 텍스트 마이닝 기반 VRU 사고와 일반사고 특성 분석 결과

사고 세부 기록을 바탕으로 VRU에 관련된 사고와 그 외의 사고를 구분하기 위해 텍스트 마이닝 분석을 수행하였는데 차량 제조업체 이름(Cruise, Waymo, Zoox 등)과 불용어는 연구 시 분석 단어에서 제외하였다. 도로 이용자 계층별 사고 발생 시 상위 특징 15개를 추출하였는데 그 결과는 Figure 5와 같고 파란색 그래프가 VRU와 관련된 사고, 초록색 그래프는 그 외의 수동 자동차 이용자와 자율주행차 간 사고를 의미한다. VRU가 포함된 사고의 경우 주로 자전거를 이용하는 사람(cyclist, bicylist)이 도로 이용자의 단어 빈도가 높았고 교차하는 장소(intersection)에서 회전(turn)에 의한 움직임이 VRU 관련 사고에서 높은 빈도를 차지하고 있다. 반면에 VRU 관련 사고를 제외한 그 외의 사고는 자율 차가 시범운행(test) 중이거나 정차(stopped)와 관련된 움직임의 특성 빈도가 높았다는 것을 확인할 수 있었다. 그리고 두 가지 유형의 사고 모두 'Rear', 'left', 'right', 'side', 'bumper' 단어 빈도가 높았고 충돌 시 후미 범퍼, 좌측, 우측 순으로 충돌 위치를 식별할 수 있었다. 텍스트 마이닝을 활용한 다른 연구에서도 자율주행차의 사고 유형은 대부분 후방 좌측 범퍼와 후방 우측 범퍼가 파손과 관련된 주제라고 나타내고 있다(Lee et al., 2023b).

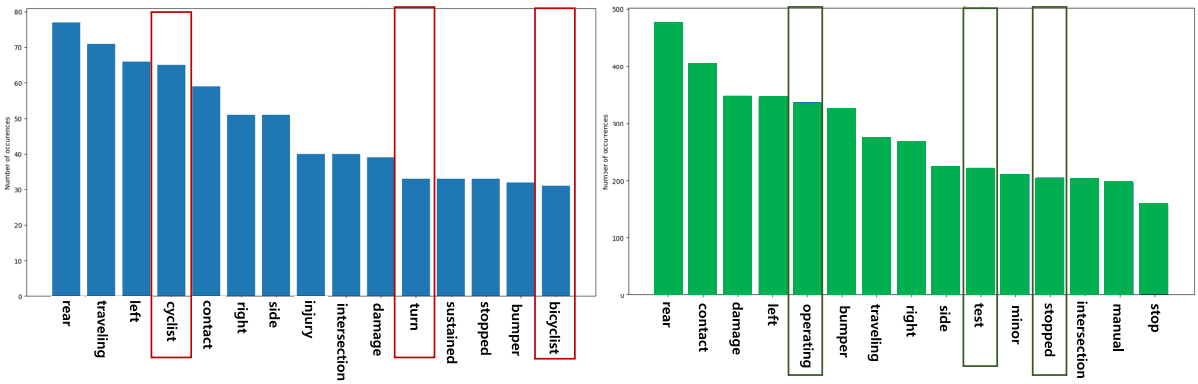

Figure 5. Results of top crash characteristics keyword with text mining for VRU crash and other crash

2. 베이지안 잠재 계층분석 기반 자율 차 사고 패턴 연구 결과

부족한 사고 표본을 보완하기 위해 도로 이용자 계층 사고별 베이지안 잠재 계층분석을 수행하였는데 VRU와 관련된 사고는 계층을 구분할수록 사후분포가 수렴하지 않아 계층이 1개인 경우보다 아닌 경우의 모형 적합도가 컸다(Table 2). 1개 계층인 VRU 사고는 주로 도로 교통이용자들이 교차하는 교차로에서 자율주행 모드가 해제되거나 운전자가 수동 운전하는 경우 서행, 회전과 같은 움직임이 있을 때 후미 추돌 또는 측면 충돌이 발생하여 경미한 사고 심각도를 가지게 된다. 특히, 자전거 이용자, 보행자가 많이 통행할 수 있는 자전거 이용 가능 도로, 버스정류장이 주변에 있으면 충돌이 발생할 빈도가 높다(Figure 6).

Table 2. Result of bayesian latent class model fit for VRU crash and No-VRU crash

Number of cluster	VRU crash			No-VRU crash		
	AIC	BIC	DIC	AIC	BIC	DIC
Cluster 1	−2226.42	−2341.11	−2212.63	−15216.36	−15998.37	−15995.38
Cluster 2	−2233.32	−2358.65	−2210.56	−15226.08	−15829.60	−16110.28
Cluster 3	−2191.16	−2454.07	−2125.78	−14977.85	−16013.37	−16190.76
Cluster 4	−2201.31	−2481.78	−2127.92	−14977.38	−16002.74	−14858.15
Cluster 5	−2200.70	−2476.34	−2129.53	−15900.37	−15311.01	−15140.74
Cluster 6	−2202.71	−2478.89	−2131.32	−15245.66	−15895.96	−15188.67

Figure 6. Results of bayesian latent class analysis: VRU traffic crash characteristic

VRU가 포함되지 않은 그 외의 사고(No-VRU crash)는 계층의 수가 3개일 때 DIC의 값이 가장 작았으며 4개 계층부터는 DIC 값이 점차 커져서 잠재 계층을 3개로 설정하였다(Table 2). Figure 7처럼 계층 1(Cluster1)은 220개, 계층 2(Cluster2) 166개, 계층 3(Cluster 3)은 53개로 계층화되었다. 계층 1은 맑은 기상환경에서 Level 3 자율 차가 상대 차량 1대와 신호 교차로 부근에서 충돌이 발생하였는데 중앙분리대가 없고 자전거 이용 가능 도로, 노상 주차가 가능한 도로들이 위치하는 도로 기하학적 특징을 가지고 있다. 그리고 자율 차, 일반 차는 차량의 움직임이 있을 때 충돌하였고 이는 다른 자율 차 계층분석을 수행한 연구결과와 유사한 결과를 두고 있다(Das et al., 2020). 계층 2의 경우, 자율주행 모드로 교차로를 주행하는 Level 4의 자율 차가 정차하고 있는 상황에서 감속 또는 회전, 직진과

같은 움직임이 나타나는 일반 차와의 후미 추돌의 사고 유형의 계층이 생성되었다. 이는 후방 충돌과 관련된 3가지 차량 움직임은 자율주행 모드를 충돌의 원인 중 하나로 식별했고 혼합 교통 상황에서 후방 충돌의 위험이 높다(Ashraf et al., 2021). 계층 3은 비 또는 안개와 같은 기상 조건일 때 노면이 젖은 상황에서 사고가 발생한다는 특징을 확인했다(Figure 8).

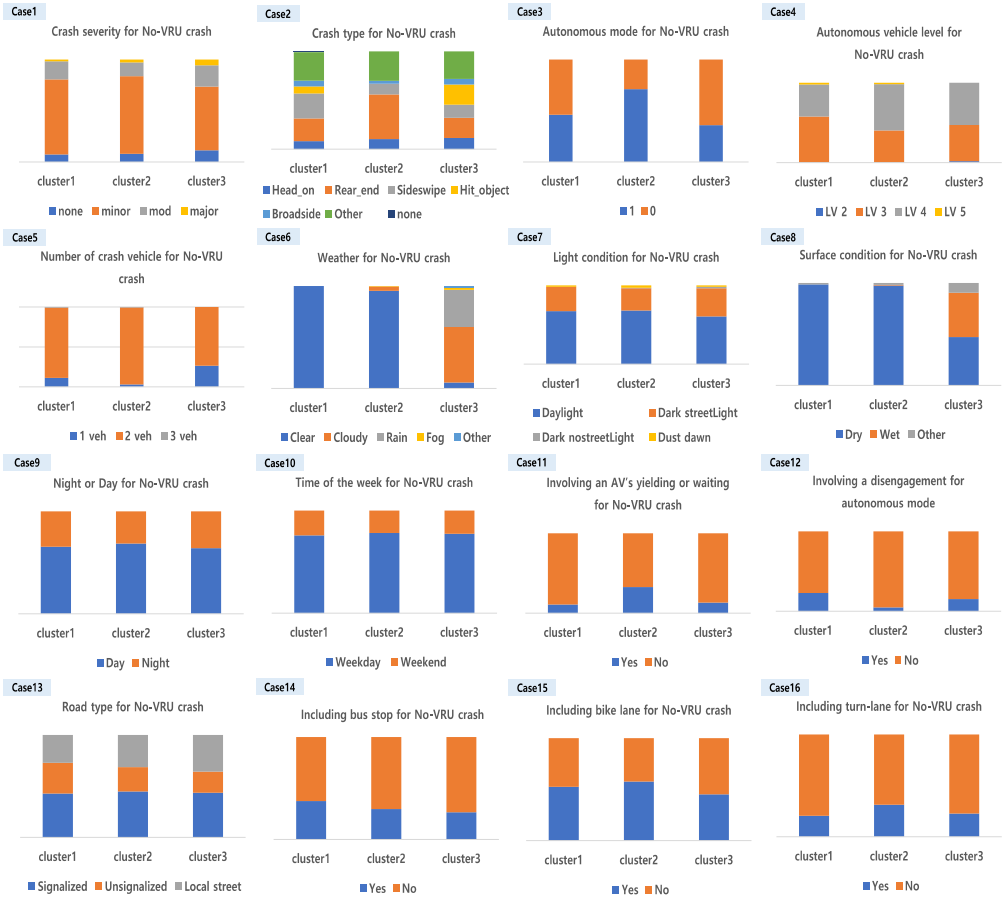

Figure 7. Results of bayesian latent class analysis: No-VRU traffic crash characteristic 1

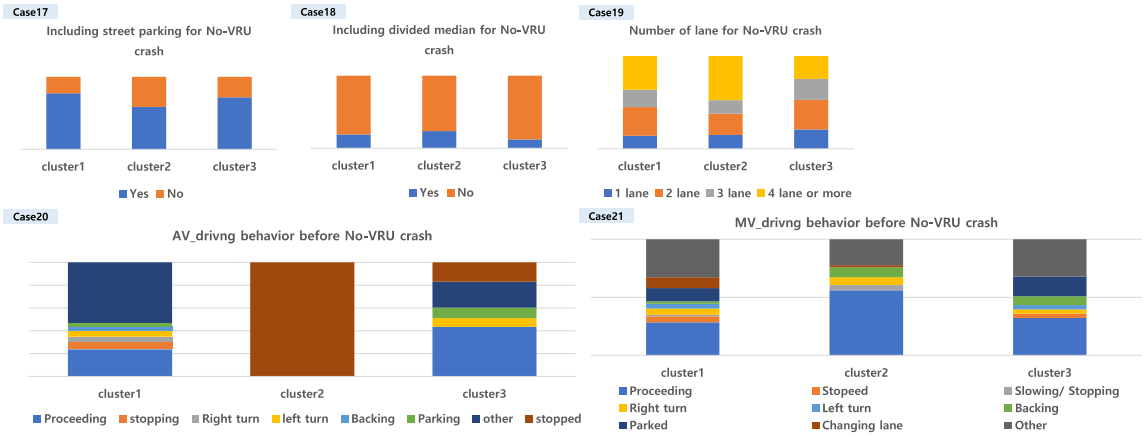

Figure 8. Results of bayesian latent class analysis: No-VRU traffic crash characteristic 2

3. 사고 심각도 모형 연구 결과

본 연구에서는 자율주행차 실사고 데이터를 기반으로 사고 심각도에 영향을 주는 요인을 분석하기 위해 순서형 프로빗 모형, 임의효과 기반 순서형 프로빗 모형을 적용하였고, 부족한 사고 표본으로 인한 모형 설명력을 향상하고자 베이지안 확률추론을 이용하였다.

Table 3과 같이 VRU와 관련된 사고 모형의 경우, 고정효과(fixed effect) 모형에 비해 임의 효과(random effect)가 있는 모형의 DIC 지표가 4정도 작아 임의 효과를 가정했을 때 유의미한 차이가 있음을 확인할 수 있다. VRU가 포함된 사고는 자전거 이용 가능 도로가 있는 경우, 신호 교차로 위치, 중앙 분리대가 없는 경우, 노상 주차 가능 도로, 도로가 어두운 경우(가로등 등화 상태) 사고 심각도가 증가하는 것으로 나타났다. 자율 차의 움직임의 경우 회전하거나 천천히 서행하는 움직임이 확인될 때 자율주행차의 사고 심각도가 높은 것을 확인하였다. 도시부 VRU 관련 사고에서 자전거 도로가 교차로 주변에 위치하면 교차 횡단하고자 하는 도로 이용자들이 늘어나 차량과 보행자 또는 자전거 이용자와의 충돌 빈도가 증가하는데 특히 어두운 거리일수록 사고 심각도가 증가하였다. 이런 열악한 조명 조건하에서도 자율 차의 감지 시스템을 개선하는 것이 필요하다(Salon et al., 2018; Asgarzadeh et al., 2018, Zhu and Meng, 2022).

Table 3. Results of the bayesian ordered probit model for VRU crash

	Ordered probit model with fixed effect			Ordered probit model with random effect		
	Mean	SD	Cred.interval	Mean	SD	Cred.interval
Involving bike_lane	0.412	0.584	(−0.690, 1.587)	0.293	0.359	(−0.210, 0.608)
Light_Dark streetlights	0.156	0.642	(−1.12, 1.401)	0.196	0.454	(−0.69, 1.06)
Signalized intersection	0.498	0.303	(0.017, 1.881)	0.213	0.371	(−0.52, 0.96)
Devided_median	−1.181	1.496	(−4.380, 0.9021)	−0.597	0.346	(−1.541, 0.101)
Turning_AV	0.730	0.489	(−0.211, 1.663)	0.934	0.538	(−0.12, 1.98)
Street_parking	0.244	0.701	(−0.870, 1.889)	0.468	0.476	(−0.42, 1.42)
Proceeding_AV	0.358	0.368	(−0.366, 1.059)	0.612	0.450	(−0.26, 1.55)
DIC	−71.361			−75.056		

VRU를 제외한 계층별 사고 심각도 모형의 결과는 Tables 4, 5와 같다. 계층 1(Cluster 1)과 계층 2(Cluster 2)의 경우, 고정 효과 모형에 비해 임의 효과를 가진 모형의 DIC 값의 차이가 8 이상이 나서 통계적으로 모형이 다르다고 할 수 있으나 계층 3(Cluster 3)의 DIC 값 차이가 4 미만으로 비관측 이질성에 대한 임의 효과 유무의 차이가 없다고 판단할 수 있다. 계층 1의 경우, 야간이거나, 중앙분리대가 설치되지 않거나 자율주행 모드가 해제되는 상황에서 사고 차량이 많을수록 사고 심각도가 증가하였다. 계층 2는 신호 교차로이면서 중앙분리대가 없고 버스와 같은 대중교통이 회전 움직임이 있을 때 사고 심각도가 증가하게 되는데 적은 사고 표본으로 인해 넓은 신뢰구간을 가지게 되어 음의 계수를 가질 수도 있다. 향후 더 많은 표본을 확보하여 신호 교차로와 같이 다양한 도로 사용자의 교차, 교통흐름의 복잡성으로 인한 예측 불확실성을 가지는 교차 공간을 중심으로 차량의 회전 움직임에 대한 이해가 필요하다(Chan, 2006; Wang et al., 2017). 계층 3의 경우, 사고 차량이 증가할수록 야간 상황이면 각각 122.5%, 101.3% 사고 심각도가 증가하고 자율 차가 우회전 시 일반 차가 직진 또는 정차하고 있는 상황에서 사고 심각도가 92.3%, 282.6% 사고 심각도가 증가한다. 우회전 차량과 통과 차량간 상충은 자주 발생하는데 이는 자율 차의 도입 시 도로 안전 분야에서 중요한 문제로 주목받고 있다(Zhu et al., 2022; Guo et al., 2020).

Table 4. Results of the bayesian ordered probit model for class 1 in No-VRU crash

	Ordered probit model with fixed effect			Ordered probit model with random effect		
	Mean	SD	Cred.interval	Mean	SD	Cred.interval
Cluster 1						
Number of vehicle	0.350	0.271	(−0.163, 0.893)	0.0314	0.244	(−0.451,0.523)
Night	0.561	0.277	(0.209, 0.941)	0.578	0.181	(0.231, 0.577)
Disengagement	0.289	0.213	(−0.131, 0.698)	0.196	0.193	(−0.199, 0.585)
Divided median	−0.391	0.223	(−0.829, 0.041)	−0.286	0.212	(−0.691, 0.134)
DIC	−168.306			−178.869		

Table 5. Results of the bayesian ordered probit model for class 2,3 in No-VRU crash

	Ordered probit model with fixed effect			Ordered probit model with random effect		
	Mean	SD	Cred.interval	Mean	SD	Cred.interval
Cluster 2						
Divided median	−0.371	0.255	(−0.861, 0.123)	−0.311	0.246	(−0.803, 0.161)
Transit	1.306	0.701	(−0.103, 2.645)	0.826	0.599	(−0.338, 1.987)
Other_unsafe_turning_MV	3.035	0.799	(1.579, 4.663)	1.796	0.584	(0.683, 2.955)
Signalized intersection	0.153	0.216	(−0.264, 0.563)	0.140	0.208	(−0.263, 0.544)
Unprotected turning lane	0.361	0.233	(−0.088, 0.824)	0.323	0.227	(−0.112, 0.754)
DIC	−121.924			−129.434		
Cluster 3						
Number of vehicle	1.225	0.661	(−0.063, 2.538)	0.175	0.462	(−0.721, 1.065)
Night	1.013	0.425	(0.201, 1.847)	0.581	0.366	(−0.116, 1.297)
Managed lane	1.328	0.756	(0.200, 1.845)	0.479	0.561	(−0.622, 1.563)
Right turn_AV	1.551	0.844	(−0.042, 3.234)	0.291	0.588	(−0.621, 1.563)
Other_MV	1.771	0.857	(0.145, 3.469)	0.598	0.584	(−0.547, 1.749)
Proceeding_MV	0.923	0.471	(0.013, 1.870)	0.563	0.384	(−0.200, 1.301)
Stopped_MV	2.826	1.016	(0.842, 4.826)	0.955	0.670	(−0.356, 2.278)
Sideswipe	0.923	0.658	(−0.336, 2.224)	0.791	0.516	(−0.201, 1.777)
DIC	−57.685			−53.866		

연구 결론

본 연구는 자율 차 실사고 데이터를 활용하여 자율 차 사고 심각도에 영향을 주는 영향을 분석하기 위해 캘리포니아 자동차관리국에서 제공하는 자율 차 사고 보고서 내 영향 변수 외에 도로 기하학적 요소, 비정형 데이터 내 잠재 설명 요인을 선정하고 도로 이용자 계층, 잠재 계층으로 구분하여 베이지안 확률추론 기반의 분석을 진행하였다.

분석 결과, VRU와 관련된 사고는 주로 자전거 이용자가 포함되어 있고 회전 움직임에 의해 충돌이 발생한다는 특징이 있고 다양한 구성 요소가 교차할 수 있는 교차로, 자전거 이용자, 자전거 이용 가능 도로, 버스정류장이 주변에 있다는 특징이 있다. 베이지안 기반 순서형 프로빗 모형 기반 사고 심각도 모형 결과, VRU 사고는 자전거 이용 가능 도로, 노상 주차, 어두운 상황에 대한 설명 변수는 사고 심각도에 양의 상관성을 가지고 있으며 일반적으로 교차로와 도로 구간, 특히 보행자가 많이 이용하는 도로 구간의 양호한 조명의 중요성이 있는 기존의 문헌들의 연구 결과와 같이 센서 융합형 자율 차가 어두운 조명 상태에서 물체를 감지하는 능력이 중요하고 이에 대한 차량 제조업체의 노력이 필요하다(Olszewski et al., 2019; Morris et al., 2021). 특히, 고정된 알고리즘 또는 규칙을 바탕으로 주행하는 차량과는 달리 정해져 있는 규칙 없이 통행하는 VRU의 예측할 수 없는 행동을 이해 및 예측하는 것이 필요하다. 또한, VRU를 포함하지 않는 사고는 자율 차의 움직임의 유무와 기상 상태 등에 따라 사고 패턴이 달라지고 사고 심각도와의 상관성을 확인하였다. 특히, 계층 1과 계층 2의 사고 특성에 경우 모형 적합도 결과 DIC값이 8이상 차이가 발생하였고 베이지안 기반 임의 효과를 가진 모형이 적합하다는 것을 의미한다. 이를 통해 고정 효과에 비해 임의 효과 모형의 신뢰구간 범위가 좁아져 과소 또는 과대 추정으로 인한 오차의 한계를 해결할 수 있을 것으로 기대된다. 마지막으로 심층적인 교통사고 메커니즘을 이해하기 위해서는 더 많은 사고 표본을 수집할 필요성이 있고 인구의 유동성을 판단할 수 있는 토지이용 데이터를 추가 활용하거나 자율 차 센서로부터 수집되는 데이터를 이용한다면 Level 4-Level 5 수준의 자율 주행 기술에 대한 보다 효율적이고 합리적인 테스트를 위해 안전 평가시나리오 구성 요소로 발전할 수 있을 것으로 기대한다.

Funding

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00142565).

알림

본 논문은 대한교통학회 제89회 학술발표회(2023.10.12)에서 발표된 내용을 수정·보완하여 작성된 것입니다.

ORCID

LEE, Songha http://orcid.org/0000-0002-0134-4687
PARK, Juneyoung http://orcid.org/0000-0002-1598-3367

REFERENCES

Ahlmann-Eltze C., Yau C. (2018), MixDir: Scalable Bayesian Clustering for High-Dimensional Categorical Data, In 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 526-539, IEEE.

Arteaga C., Paz A., Park J. (2020), Injury Severity on Traffic Crashes: A Text Mining with an Interpretable Machine-Learning Approach, Safety Science, 132, 104988.

Asgarzadeh M., Fischer D., Verma S. K., Courtney T. K., Christiani D. C. (2018), The Impact of Weather, Road Surface, Time-of-Day, and Light Conditions on Severity of Bicycle-Motor Vehicle Crash Injuries, American Journal of Industrial Medicine, 61(7), 556-565.

Ashraf M. T., Dey K., Mishra S., Rahman M. T. (2021), Extracting Rules from Autonomous-Vehicle-Involved Crashes by Applying Decision Tree and Association Rule Methods, Transportation Research Record, 2675(11), 522-533.

Chan C. Y. (2006), Defining Safety Performance Measures of Driver-Assistance Systems for Intersection Left-Turn Conflicts, In 2006 IEEE Intelligent Vehicles Symposium, 25-30, IEEE.

Chen Y., Luo R., Yang H., King M., Shi Q. (2020), Applying Latent Class Analysis to Investigate Rural Highway Single-Vehicle Fatal Crashes in China, Accident Analysis & Prevention, 148, 105840.

Das S., Dutta A., Tsapakis I. (2020), Automated Vehicle Collisions in California: Applying Bayesian Latent Class Model, IATSS Research, 44(4), 300-308.

Deb S., Strawderman L., Carruth D. W., DuBien J., Smith B., Garrison T. M. (2017), Development and Validation of a Questionnaire to Assess Pedestrian Receptivity Toward Fully Autonomous Vehicles, Transportation Research Part C: Emerging Technologies, 84, 178-195.

Guo Y., Liu P., Wu Y., Chen J. (2020), Evaluating How Right-Turn Treatments Affect Right-Turn-on-Red Conflicts at Signalized Intersections, Journal of Transportation Safety & Security, 12(3), 419-440.

Hagenzieker M. P., Van Der Kint S., Vissers L., Van Schagen I. N. G., De Bruin J., Van Gent P., Commandeur J. J. (2020), Interactions Between Cyclists and Automated Vehicles: Results of a Photo Experiment, Journal of Transportation Safety & Security, 12(1), 94-115.

Hoogendoorn R., Van Arem B., Hoogendoorn S. (2014), Automated Driving, Traffic Flow Efficiency, and Human Factors, Transportation Research Record, 2422, 113-120.

Jeon H., Kim J. (2021), Analysis on Handicaps of Automated Vehicle and Their Causes Using IPA and FGI, The Journal of The Korea Institute of Intelligent Transport Systems, 20(3), 34-46.

KICT (2020), Improvement Road Infrastructures to Strengthen Driving Safety of Automated DrivingCar, The Second report, South Korea.

Kutela B., Das S., Dadashova B. (2022), Mining Patterns of Autonomous Vehicle Crashes Involving Vulnerable Road Users to Understand the Associated Factors, Accident Analysis & Prevention, 165, 106473.

Kwayu K. M., Kwigizile V., Lee K., Oh J. S. (2021), Discovering Latent Themes in Traffic Fatal Crash Narratives Using Text Mining Analytics and Network Topology, Accident Analysis & Prevention, 150, 105899.

Kwayu K. M., Kwigizile V., Zhang J., Oh J. S. (2020), Semantic N-gram Feature Analysis and Machine Learning: Based Classification of Drivers' Hazardous Actions at Signal-Controlled Intersections, Journal of Computing in Civil Engineering, 34(4), 04020015.

Lee H. J., Kang M. H., Song J. I., Hwang K. Y. (2023a), Development of Autonomous Vehicle Traffic Accident Scenarios in Urban Areas Based on Real-World Accident Data Using Association Rule Mining, J. Korean Soc. Transp., 41(3), Korean Society of Transportation, 375-393.

Lee S. Y., Han S. J., Chung Y. S. (2020), Injury Severity of Freeway Crashes in Rainy Weather Using an Ordered Probit Model, J. Transp. Res, 27(1), 1-11.

Lee S., Arvin R., Khattak A. J. (2023b), Advancing Investigation of Automated Vehicle Cashes Using Text Analytics of Crash Narratives and Bayesian Analysis, Accident Analysis & Prevention, 181, 106932.

Li Y., Lord-Bessen J., Shiyko M., Loeb R. (2018), Bayesian Latent Class Analysis Tutorial, Multivariate Behavioral Research, 53(3), 430-451.

Millard-Ball A. (2018), Pedestrians, Autonomous Vehicles, and Cities, Journal of Planning Education and Research, 38(1), 6-12.

Morando M. M., Tian Q., Truong L. T., Vu H. L. (2018), Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures, Journal of Advanced Transportation.

Morris A. P., Haworth N., Filtness A., Nguatem D. P. A., Brown L., Rakotonirainy A., Glaser S. (2021), Autonomous Vehicles and Vulnerable Road-Users—Important Considerations and Requirements Based on Crash Data from Two Countries, Behavioral Sciences, 11(7), 101.

Novat N., Kidando E., Kutela B., Kitali A. E. (2023), A Comparative Study of Collision Types Between Automated and Conventional Vehicles Using Bayesian Probabilistic Inferences, Journal of Safety Research, 84, 251-260.

Olszewski P., Szagała P., Rabczenko D., Zielińska A. (2019), Investigating Safety of Vulnerable Road Users in Selected EU Countries, Journal of Safety Research, 68, 49-57.

Salon Deborah, Andrew McIntyre (2018), Determinants of Pedestrian and Bicyclist Crash Severity By Party at Fault in San Francisco, CA, Accident Analysis & Prevention, 110, 149-160.

Sohrabi S., Khodadadi A., Mousavi S. M., Dadashova B., Lord D. (2021), Quantifying the Automated Vehicle Safety Performance: A Scoping Review of the Literature, Evaluation of Methods, and Directions for Future Research, Accident Analysis & Prevention, 152, 106003.

Spiegelhalter D. J., Best N. G., Carlin B. P., van der Linde A. (2002), Bayesian Measures of Model Complexity and Fit (with discussion), Journal of the Royal Statistical Society, Series B, 64, 583-639.

Wang X., Zhao D., Peng H., LeBlanc D. J. (2017), Analysis of Unprotected Intersection Left-Turn Conflicts Based on Naturalistic Driving Data, In 2017 IEEE Intelligent Vehicles Symposium (IV), 218-223, IEEE.

White A., Murphy T. B. (2014), BayesLCA: An R Package for Bayesian Latent Class Analysis, Journal of Statistical Software, 61, 1-28.

Xie Y., Zhang Y., Liang F. (2009), Crash Injury Severity Analysis Using Bayesian Ordered Probit Models, Journal of Transportation Engineering, 135(1), 18-25.

Yuan Q., Xu X., Wang T., Chen Y. (2022a), Investigating Safety and Liability of Autonomous Vehicles: Bayesian Random Parameter Ordered Probit Model Analysis, Journal of Intelligent and Connected Vehicles, 5(3), 199-205.

Yuan Q., Zhai X., Ji W., Yang T., Yu Y., Yu S. (2022b), Correlation Analysis on Accident Injury and Risky Behavior of Vulnerable Road Users Based on Bayesian General Ordinal Logit Model, Sustainability, 14(23), 16048.

Zhu H., Zhu S., Iryo-Asano M., Nakamura H. (2022), Investigating Driver Reactions to Movements of Autonomous Vehicle in Permitted Right Turn Through Driving Simulator Experiments, Transportation Research Part F: Traffic Psychology and Behaviour, 89, 385-398.

Zhu S., Meng Q. (2022), What Can We Learn from Autonomous Vehicle Collision Data on Crash Severity? A Cost-Sensitive CART Approach, Accident Analysis & Prevention, 174, 106769.