

위험 시나리오 내 교통약자를 위한 어텐션 기반 합성곱 신경망을 이용한 충돌 예측 네트워크 개발 기초연구

김승환 · 송봉섭*

아주대학교 D.N.A.플러스융합학과

A Basic Study on Collision Prediction Based on Attention-Based Convolutional Neural Network for Vulnerable Road Users in Safety-Critical Scenarios

Seunghwan Kim · Bongsob Song*

Department of Data, Networks, and AI, Ajou University, Gyeonggi 16499, Korea
(Received 7 November 2025 / Revised 28 November 2025 / Accepted 28 November 2025)

Abstract : In this study, collision prediction networks for vulnerable road users (VRUs) are proposed and evaluated using simulations of safety-critical scenarios. While several networks for collision prediction are found in existing literature, predicting the potential collision of relatively small and occluded VRU with high dynamics might be limited. To enhance collision prediction performance, an attention-based convolutional neural network (ACNN) is proposed with collision probability (CP), so-called CP-ACNN. More specifically, simplified bird eye's view (SBEV) images including trajectory prediction as well as CP involving a body-fixed coordinate are generated based on simulation and real-driving datasets. It subsequently introduces a soft attention mechanism with four-region partitioning (front/rear-left/right) for more adaptive learning of spatial importance. Based on attention-based CNN, both feature extraction and classification are performed to predict collision for VRUs. Finally, evaluation results demonstrate that CP-ACNN consistently outperforms the other two networks in terms of accuracy and false positive rate.

Key words : Collision prediction(충돌 예측), Safety-critical scenario(위험 시나리오), Collision probability(충돌 확률), Attention-based(어텐션 기반), Convolutional neural network(합성곱 신경망), Vulnerable road user(교통약자)

Nomenclature

x	: relative longitudinal position	h	: flattened hidden feature vector of each area
y	: relative lateral position	\mathcal{W}	: attention weight
v_x	: relative longitudinal velocity	b	: bias
v_y	: relative lateral velocity	s	: attention score for each area
a_x	: relative longitudinal acceleration	α	: attention weight vector
θ	: heading angle to surrounding vehicle	c	: context vector
t_1	: collision prediction start time	\mathcal{L}_{CE}	: cross-entropy loss
t_c	: prediction time	\mathcal{L}_{att}	: attention entropy loss
t_2	: collision occurrence time	λ	: weighting factor for attention entropy loss
t_d	: collision prediction time, $t_2 - t_c$	CP	: collision probability

*Corresponding author, E-mail: bsong@ajou.ac.kr*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.

1. 서론

최근 수년 간 도로 교통 사망자 수는 꾸준히 감소하고 있으며, 이는 도로의 안전성 향상과 더불어 차량 자체의 충돌 예방 기술이 발전한 결과로 평가된다.¹⁾ 예를 들어 전방 충돌 경고(Forward collision warning), 자동 긴급 제동 시스템(Automatic emergency braking)과 같이 첨단 운전 지원 시스템(Advanced Driver Assistance System, ADAS)이 차량에 탑재됨에 따라 위험한 교통사고에서 운전자를 보호하는 결과를 보여주고 있다.²⁾ 하지만 도로 안전 연례 보고서에 따르면, 교통 시스템 내에서 상대적으로 보호 장치가 부족한 교통약자(Vulnerable Road User, VRU), 예를 들어 보행자, 자전거, 전동킥보드(E-scooter) 이용자 등은 차량과의 충돌 시 심각한 상해나 사망으로 이어질 확률이 높고, 전체 교통 사망자의 23%를 차지할 정도로 심각한 위협에 노출되어 있다.³⁾ 이러한 문제를 해결하기 위해서는 ADAS 또는 자율주행 차량이 사전에 충돌 위험을 감지하고 예방할 수 있는 기술적 대응이 필요하다.

최근 연구들은 다양한 시나리오 기반 테스트와 검증을 통해 자율주행차의 운용 안전을 보장하고자 하며, VRU의 비정형 행동을 예측하는 딥러닝 기반 접근 방식이 활발히 연구되고 있다.⁴⁾ 실제 도로 환경에서는 교통약자의 행동 패턴이 매우 다양하고 불규칙적이며, 날씨, 조명, 차선 조건, 보행자 밀도 등 수많은 외적 요인의 영향을 받는다.⁵⁾ 이러한 복잡성을 반영하기 위해 시나리오 기반의 데이터셋을 활용한 딥러닝 모델의 훈련과 상황별 성능 평가는 자율주행차의 운용 안전성을 향상시키는 데 중요한 역할을 하고 있다.

자율주행 시스템은 복잡한 작동 환경과 다양한 비정형 상황에 노출되며, 이로 인해 기존의 규칙 기반 검증 방식만으로는 충분한 안전성을 보장하기 어렵다. 이에 따라 최근에는 보다 정교한 시나리오 기반 안전성 평가 접근법이 자율주행 분야에서 주목받고 있다.⁶⁾ 예를 들어, ISO 34502는 자율주행 시스템의 안전성을 평가하기 위한 시나리오 기반 프레임워크를 제안하며, 시나리오 구조화 및 데이터 기반 검증의 중요성을 강조한다.⁷⁾ 특히 PEGASUS 프로젝트에서는 시나리오를 기능(Functional) 시나리오, 논리(Logical) 시나리오, 상세(Concrete) 시나리오 수준으로 구분하고 다양한 수준의 구분을 통해 실제 테스트에 사용할 수 있는 상세 시나리오를 구성하는 절차를 제안한다.⁸⁾ 또한 데이터베이스를 이용하는 시나리오 접근법으로 Testing-based와 Falsification-based 시나리오의 적용 가능성, 구체성 확보 측면에서 우수한 성능을 보고하고 있다.⁹⁾

자율주행 시스템의 판단 정확도를 높이기 위해 입력 표현 방식 또한 발전해 왔다. 원시 센서(Raw sensor) 데이터 접근 방법은 주변 환경에 대한 이용 가능한 데이터를 모두 포함하므로, 다양한 센서 신호에서 유용한 특징을 추출할 수 있는 장점이 있다. 그러나 입력 데이터를 처리하는 데 많은 계산 리소스를 요구하므로, 자율주행 차량의 온보드(Embedded) 구현에는 현실적인 제약이 존재한다. 이에 따라 시각적으로 직관적이며 정형화된 공간 표현 방식인 BEV(Bird's Eye View) 기반 표현이 널리 활용되고 있으며, 더 나아가 운행 환경의 중요한 객체만을 간결하게 나타내는 SBEV(Simplified Bird's Eye View)가 기존 예측 분류 연구에서 계속 사용되어 왔다.¹⁰⁾ SBEV는 입력 데이터의 복잡도를 크게 낮추면서도 의미 있는 정보를 유지할 수 있어 CNN(Convolution Neural Network)을 효율적인 입력으로 활용하는 데 적합하다. 하지만 SBEV 방식도 운전 환경 내 정적 및 동적 객체 상태를 추정하는 인식 모듈의 한계를 상속하고 있다는 점이 단점으로 지적된다.¹¹⁾

자율주행 시스템의 판단 기능 중 충돌 예측 기능은 충돌 부위를 정량적으로 분류하고 충돌 이전 단계에서 이를 인지해야 한다. 충돌 예측 기술은 크게 모델 기반과 데이터 기반으로 나뉜다. 모델 기반 알고리즘은 물리적 가정으로 해석 가능성이 높지만 복잡한 환경에서 성능이 제한되는 한계가 있으며, 데이터 기반 알고리즘은 실도로 주행에서 발생하는 노이즈에 민감하다고 알려져 있다.¹²⁾ 이러한 한계를 보완하기 위해 CP-CNN 알고리즘은 모델 기반 충돌 확률(Collision Probability, CP) 알고리즘을 통해 표현된 위험도를 CNN으로 학습 및 보정하여 물리 모델의 한계를 완화하였다.^{13,14)} 그러나 전통적인 Baseline CNN은 국소 패턴에 치우쳐 전역적, 공간적 관계(Spatial relationship)에 대한 모델링에 어려움이 있으며, 입력 변동(Adversarial perturbation)에 민감하다는 구조적 제약을 갖는다.^{15,16)} 이를 완화하기 위해 다양한 어텐션(Attention) 기반 접근이 제안되어 왔다. 채널 간 상호 연관성을 학습하여 특징 재보정을 수행하는 Squeeze-and-excitation(SE) 메커니즘을 제안하였으며, 선택적으로 전역 정보를 활용할 수 있도록 하였다.¹⁷⁾ 또한 채널 및 공간 차원에서 불필요한 특징을 억제하는 CBAM(Convolutional Block Attention Module)을 제안하여 정보 집중도를 향상시켰다.¹⁸⁾ 셀프 어텐션(Self-attention) 메커니즘은 특징 맵의 위치 간 상관관계를 학습하여 전역 정보와 국소 정보를 효과적으로 유지함으로써 Baseline CNN 대비 성능 향상을 입증하였다.¹⁹⁾

2. 문제 정의

Daimler-Benz의 연구 결과에 따르면, 충돌 0.5 s 전 경고를 통해 약 60 %의 충돌을 방지할 수 있고, 1.5 s 전 경고 시 약 90 %까지 사고 예방이 가능하다고 보고되었다.²⁰⁾ 또한 Euro NCAP AEB VRU test protocol v3.03에서는 충돌 1.5 s 전 경고를 기준으로 회피 가능성을 평가하고, 이 시점을 경고의 최소 요구 조건으로 규정하고 있다.²¹⁾ 이에 본 연구에서는 Fig. 1과 같이 위험(Safety-critical) 시나리오에서의 충돌 판단 시점을 정의하였다. 먼저, 위험 시나리오의 시작 시점은 보행자가 자차의 주행 경로 상, 즉 자차의 주행차선 내에 존재하는 시점으로 정의하였다. 자차와 상대 객체 간의 Time-To-Collision(TTC)이 1.5초가 되는 시점을 t_1 , 실제 충돌이 발생하는 시점을 t_2 , 처음 충돌을 판단하는 시점을 t_c 로 정의하였다. 또한 충돌 예측 시간 t_d 는 $t_2 - t_c$ 로 정의하였으며, 충돌 시점을 얼마나 빠르게 예측하였는지 나타내는 예측 평가 지표로 사용하였다. 본 연구에서는 다양한 차대교통약자(Car-to-VRU) 시나리오를 대상으로 보행자, 자전거 이용자 등 소형 객체에 대한 충돌 예측 성능을 확인하고자 한다.

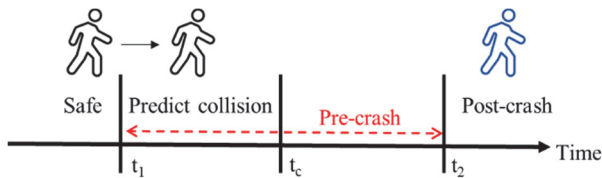


Fig. 1 Definition of collision prediction for pre-crash scenario

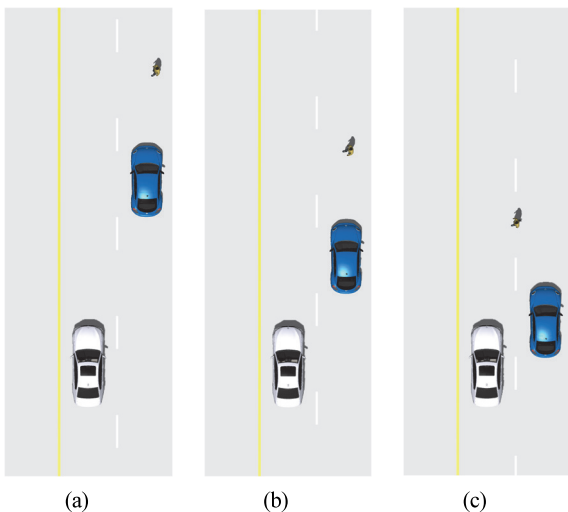


Fig. 2 Pedestrian crossing scenario with occlusion (a) $t = 9.61$ s, (b) $t = 10.21$ s, (c) $t = 10.77$ s

좀 더 구체적으로 Fig. 2(a)는 자차가 50 km/h로 주행 중일 때, 오른쪽 정지 차량에 의해 보행자가 가려진 상태에서 5 km/h로 횡단을 하고 있는 대표적인 AEB Pedestrian을 위한 테스트 시나리오를 보여주고 있다.²¹⁾ Fig. 2(b)는 자차의 전방 레이더 센서가 보행자를 처음 인식하는 시점으로, 아직 보행자가 자차 주행 경로 상에 있지 않아 충돌이 예측되지 않은 시점을 보여주고 있다. Fig. 2(c)에서는 보행자가 민첩한 행동(High dynamics)을 할 수 있는 구간으로 보행자가 횡단할 수도 있고 기다릴 수도 있는 상황이다. 만약 보행자가 기다린다면 안전(Safe)한 상황으로 판단해야 하며, 횡단한다면 0.33초(즉, TTC = 0.33) 내에 차량을 감속해서 보행자를 보호해야 하는 상황(Unsafe)에 직면한다. Fig. 2(c)에서 차량 주행 경로까지 도달하는 시간은 0.1초 이내로 5샘플(50 Hz 샘플링 기준)에 판단해야 하며, 가림(Occlusion) 현상으로 인해 보행자 인식 시간도 0.56초로 짧은 순간에 충돌 위험 판단이 이루어져야 한다.

위 예제에서 본 바와 같이 교통약자(VRU)에 대한 충돌 예측을 좀 더 빠르고 짧은 순간에 정확하게 판단할 수 있는 네트워크 모델이 필요하다. 이를 위하여 우선 차량 간 충돌을 예측하는 알고리즘¹³⁾을 교통약자에 대한 충돌 예측 알고리즘으로 적용하고, 이를 개선시키기 위해 객체 간 상대적 중요도와 공간적 상호 관계(즉, 충돌 위험성)를 명시적으로 반영하는 Attention 메커니즘을 결합한 CP-ACNN(CP and Attention-based CNN)을 비교 평가하고자 한다.

3. 시나리오 기반 데이터셋 구성

3.1 시나리오 선정




본 연구를 위하여 우선 가상 데이터셋을 구성하였으며, 시뮬레이션을 위한 시나리오 선정이 필수적이다. 따라서 차량과 VRU 간 실제 교통 상황에서 발생한 충돌 시나리오를 중심으로 선정된 시나리오셋(A set of scenarios)을 구성하고 이를 시나리오 카탈로그(Scenario catalog)라고 명명한다. 참고로 실제 교통 사고 시나리오를 구분하기 위해서 유럽 교통 사고 데이터 표준인 IGLAD의 사고유형 번호를 이용하여 구분하였다.²²⁾

보행자(Pedestrian) 관련 사고 시나리오는 국내 교통사고 분석 시스템(TAAS) 통계로부터 11건, 유럽 PROSPECT에서 3건을 선정하여 총 14건으로 구성하였다.^{23,24)} 자전거(Cyclist) 사고 시나리오는 TAAS에서 14건, PROSPECT에서 3건을 반영하여 총 17건으로 구성하였다. 전동킥보드(E-scooter) 관련 사고 시나리오는 삼성교통안전문화연구소에서 사고 통계 분석한 9건, TITAN 데이터베이스(2018 ~ 2020)로부터 2건 추가하여 총 11건을 구성하였

Table 1 Number of VRU scenarios to virtual dataset

Scenario	Road type		Total
	Non-junction (NJ)	Junction (JN)	
Car-to-pedestrian	9	5	14
Car-to-cyclist	6	11	17
Car-to-e-scooter	2	9	11

Table 2 Occlusion scenarios in the scenario catalog

Scenario	IGLAD	Road type
Pedestrian crossing - from the left with obstacle		Non-junction
Pedestrian crossing - from the right while passing		Non-junction
E-scooter crossing - from the left with obstacle		Non-junction

다.^{25,26)} Table 1에서는 VRU 유형에 따라서 보행자, 자전거, 전동킥보드 사고 시나리오에 대해 비교차로(Non-junction, NJ)와 교차로(Junction, JN)로 구분하여 정리하였다. 결론적으로 시나리오 카탈로그는 총 42건의 시나리오를 고려하였으며, 이 중 Table 2와 같이 가림(Occlusion)을 포함하는 3건의 시나리오가 포함되었다.

3.2 복합 및 학습 데이터셋

위에서 선정된 시나리오에 대하여 시뮬레이션을 통하여 가상 데이터셋을 생성하고 네트워크 학습 데이터를 선정하는 과정을 설명하고자 한다.

3.2.1 복합 데이터셋

충돌 예측 알고리즘 개발을 위해서는 사고 상황을 포함하는 데이터가 필수적이다. 그러나 실제 교통 사고 관련 데이터는 대부분 블랙박스 영상 정도이며, 실제 VRU에 대한 위치나 속도 정보를 획득하는 것은 매우 드물고 충돌 관련 공개 데이터셋도 찾아볼 수 없다. 따라서 본 연구에서는 차량 시뮬레이터 중 하나인 IPG사의 CarMaker를 활용하여 사고 시나리오를 포함하는 가상 데이터를 생성하였다.²⁷⁾ Table 3에 나타난 바와 같이, 시나리오 카탈로그(총 42건의 Logical 시나리오)를 기반으로 총 31,847건의 상세(Concrete) 시나리오에 대해 시뮬레이션을 수행함으로써 가상 데이터를 구축하였다.

Table 3 A summary of training and test data

	Training	Test	Sum
Virtual dataset (# of concrete scenario)	10,382	21,465	31,847
Real driving dataset (# of snippet)	368	158	526
Sum	10,750	21,623	32,373

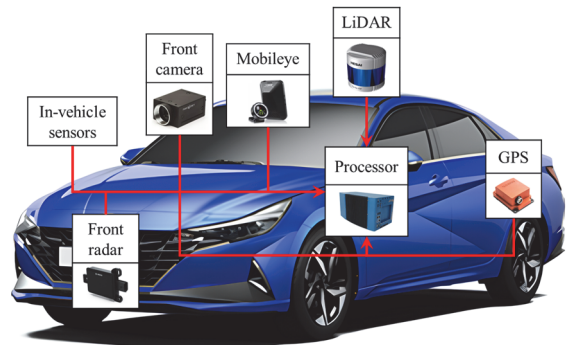


Fig. 3 Sensor configuration of the experimental vehicle

추가적으로 가상 데이터만을 사용할 경우, 실제 도로 주행 환경에서 발생하는 센서 노이즈 및 교통 혼잡 상황을 충분히 반영하지 못해 오판단(False alarm)이 발생할 우려가 있다. 이를 보완하기 위해 실도로 주행 데이터를 추가로 사용하였다. 해당 데이터는 실제 충돌 사고를 포함하고 있지 않기에 오판단에 대한 성능 제고를 위한 데이터라고 할 수 있다. 실도로 주행 데이터는 Fig. 3과 같이 전방 레이더(Field of view ± 10 deg) 및 비전센서(FOV ± 50 deg)가 장착된 실험 차량을 이용하여 취득하였다. 추가적으로 센서 융합 알고리즘을 거쳐서 주변 객체에 대한 트랙(Track)정보를 생성한다. 교통약자가 포함된 주행 경로는 약 5.44 Km이며, 주행 시간은 약 0.4 hr에 해당한다. 가상 데이터는 약 20초 정도에 해당하여 실도로 주행 데이터도 20초 길이를 Snippet이라고 정의하고 총 526개의 Snippet을 사용하였다. 가상 데이터 중 충돌 발생 여부는 Pre-crash 시나리오에서 충돌 1.5초 전부터 충돌이 발생한 시점(t_2)까지 기준으로 “Unsafe”로 라벨링하였다. 반면 실도로 주행 데이터는 사고가 발생하지 않는 주행 데이터므로 전체를 “Safe”로 라벨링하였다. 가상 데이터와 수집된 실도로 주행 데이터를 포함하는 복합 데이터셋을 Table 3에 요약하였다.

3.2.2 학습 데이터의 균형

학습 데이터셋 구성은 크게 두 가지 관점을 고려하여 수행하였다. 우선 Table 1에 요약된 바와 같이 42개의 시

나리오가 모두 포함되도록 Scenario-balanced 기준을 적용하여 복합(Mixed) 데이터셋에서 학습 데이터를 선정하였다. 가상 데이터와 실도로 주행 데이터를 시나리오 단위로 분류하였으며, 전체 31,847건의 상세 시나리오 중 교통약자 유형별 비율을 조정하였다. 구체적으로 자전거(Cyclist) 사고 시나리오는 32%에서 12%로, 전동킥보드(E-scooter) 사고 시나리오는 30%에서 15%로, 보행자(Pedestrian) 사고 시나리오는 28%에서 16%로 학습 데이터셋을 선정하였다. 또한, 선정된 상세 시나리오에 대하여 충돌 여부에 따라 안전 데이터는 0.5초, 충돌 데이터는 0.05초 간격으로 샘플링하여 SBEV 이미지를 생성하였다. 이를 통해 총 261,615개의 학습데이터(SBEV 기준)를 선정하였다.

다음으로 네트워크 출력의 균형(Class-balanced)을 고려하여 학습데이터를 선정하였다. 충돌은 전방위에서 발생 가능하고 충돌의 유무(Unsafe/safe)뿐만 아니라 Fig. 5와 같이 12개의 충돌 부위(Collision mode)를 고려한 Class-balanced 학습 데이터셋을 구성하였다.²⁸⁾ 좀 더 구체적으로 위에서 고려된 Scenario-balanced 학습데이터에 대해서 각 충돌 부위별로 6,000개의 학습 데이터를 랜덤하게 선정하여 총 72,000개의 충돌 데이터(SBEV 데이터 기준)를 선정하였다. 최종적으로 Table 4와 같이 충돌(Unsafe) 및 안전(Safe) 비율을 51%와 49%로 구성하여 시나리오(Scenario-balanced)와 출력 관점의 균형(Class-balanced)을 동시에 고려하였다.

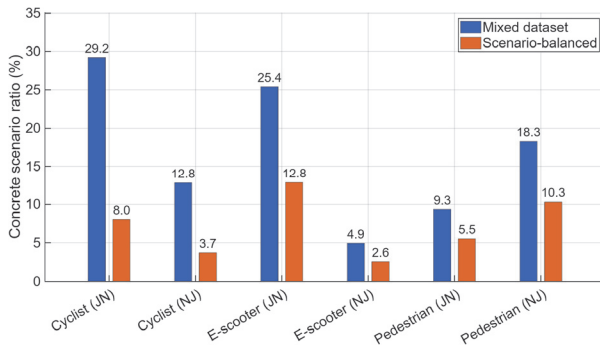


Fig. 4 Distribution of training data with respect to VRU and road geometry

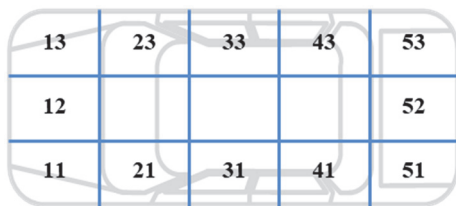


Fig. 5 Definition of impact point sections

Table 4 Number of samples in training dataset

	Unsafe	Safe	Total
Scenario-balanced	193,740	67,875	261,615
Class-balanced	72,000	67,875	139,875

4. 교통약자 충돌 예측 알고리즘

교통약자에 대한 충돌 예측을 위해 제안된 알고리즘은 추상화(Abstraction), 특징 추출(Feature extractor), 그리고 충돌 예측 분류(Decision)의 세 단계로 구성된다. Fig. 6은 알고리즘의 주요 처리 절차를 단계별로 나타낸다. 보다 구체적으로는, VRU에 대한 미래 궤적과 충돌 확률(CP)을 기반으로 SBEV 이미지를 생성한 뒤, 어텐션 메커니즘이 결합된 CNN 구조를 활용하여 해당 SBEV로부터 주요 특징 추출(Feature extraction)을 수행한다. 마지막으로, 충돌 위험을 구별(Classification)하는 단계를 수행한다.

4.1 추상화

제안된 알고리즘은 레이더와 비전 센서의 트랙(Track) 정보 입력으로 센서 융합(Sensor fusion)을 통해 다음과 같은 주변 객체의 트랙 정보를 생성한다.

$$\mathbf{x}_n = [x, y, v_x, v_y, \theta, a_x]^T \in \mathbb{R}^6 \quad (1)$$

다음으로 등가속도(Constant acceleration, CA) 모델 기반 칼만 필터(Kalman filter)를 사용하여 주변 객체에 대한 상대적인(Body-fixed coordinate) 미래 위치를 예측하고, 이를 기반으로 충돌 확률(CP)을 계산한다.¹³⁾ 다음 단계로 센서 융합 결과, 예측된 미래 위치, 충돌 확률 정보를 종합하여 SBEV 이미지로 추상화(Abstraction)한다.

네트워크의 입력인 SBEV는 RGB채널의 형태로 구성되며, 비전 센서에서 제공되는 차선 관련 정보인 정적(Static) 계층 정보, 센서 융합에서 제공되는 위치 및 속도 정보인 동적(Dynamic) 계층 정보, 그리고 미래 위치(Trajectory prediction) 및 충돌 확률(CP) 정보인 메타(Meta) 계층 정보로 구성된다. Fig. 7(a)에서 보는 바와 같이 Fig. 2(c)의 보행자와 정지 차량에 대한 예측 궤적(5샘플, 0.2초)을 박스 형태로 표현하고 있으며, 충돌 확률(CP)을 R 채널로 나타낸다. 충돌 확률은 자차 및 상대 객체 간 미래 위치 확률 분포가 일정 충돌 영역 D에 속할 확률을 의미하며 다음과 같이 계산된다.

$$CP(t) = \max_{i=1, \dots, N} \text{Prob}(x_{t+iT} \in D) \quad (2)$$

여기서 i 는 미래 예측 단계, N 은 총 예측 단계 수를 나타

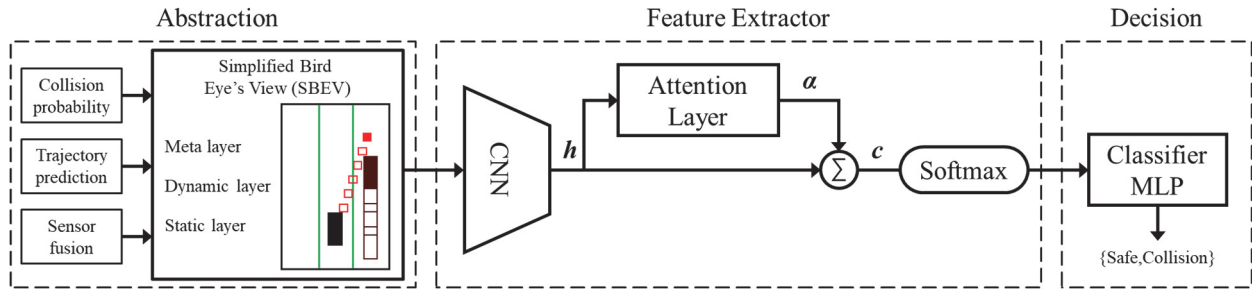


Fig. 6 Overview of processing steps and model architecture

낸다. 상대 상태 벡터는 $x_{t+iT} = [\tilde{p}_x \tilde{p}_y]^T \in R^2$ 로 정의되며, 자차량 기준 좌표계의 종·횡방향 상대 위치를 포함하며 확률 밀도 함수는 다음과 같다.

$$\begin{aligned} & \text{Prob}(x_{t+iT} \in D) \\ &= \iint_{(\tilde{p}_x, \tilde{p}_y) \in D} p_{t+T}(\tilde{p}_x, \tilde{p}_y | Y_t) d\tilde{p}_x d\tilde{p}_y \end{aligned} \quad (3)$$

Y_t 는 관측된 모든 정보를 뜻하며, $p_{t+T}(\tilde{p}_x, \tilde{p}_y | Y_t)$ 는 $[\tilde{p}_x \tilde{p}_y]^T$ 의 확률 분포인 확률 밀도 함수(PDF)이다. 확률 분포 적분은 영역 D 의 x 축 및 y 축 방향의 경계 값에서 평가된 해당 누적분포함수(Cumulative distribution function, CDF)를 이용한다. 식 (3)에서 확률 분포 적분은 각 축의 1차원 CDF로 분리하여 충돌 확률(CP)를 다음과 같이 근사하여 계산하였다. 본 연구에서는 칼만 필터 기반 예측 단계에서 상대 위치 오차가 약한 상관 구조를 갖는 2차원 Gaussian으로 모델링된다는 조건부 독립 가정(Conditional independence assumption) 하에 식(3)의 2차원 적분을 각 축의 1차원 CDF 곱으로 근사하였다.

$$\begin{aligned} \text{Prob}(x_{t+iT} \in D) &= [F_X(x_{max}) - F_X(x_{min})] \\ &\quad \times [F_Y(y_{max}) - F_Y(y_{min})] \end{aligned} \quad (4)$$

여기서 F 는 각 축에 대한 정규 누적 분포 함수이며, 최종적으로 CP는 식(2)와 같이 각 예측 단계에서 계산된 확률 중 최댓값으로 정의된다.

Fig. 7(b)와 (c)에서는 차선 정보와 자차 정보를 통합하여 통합적으로 구성된 SBEV를 보여주고 있다. 비전 센서에서 추출된 차선 정보는 도로 형상을 간결하게 표현하기 위해 3차 다항식 형태로 모델링되며, 다음과 같이 나타낸다.

$$X = C_3 \cdot Z^3 + C_2 \cdot Z^2 + C_1 \cdot Z + C_0 \quad (5)$$

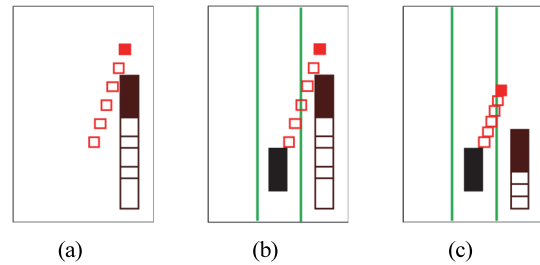


Fig. 7 Examples of meta layer and SBEV for pedestrian crossing scenario (a) Meta layer, (b) SBEV at $t = 10.21$ s, (c) SBEV at $t = 10.77$ s

여기서 X 는 차선까지의 횡방향 위치, Z 는 종방향 위치, C_0 는 차선 위치, C_1 은 차선의 기울기, C_2 은 차선의 곡률, C_3 은 차선의 곡률 변화율을 의미한다. SBEV 이미지는 (3, 201x101) 크기로 구성되었으며, 네트워크 입력 전 단계에서 공간적 정렬과 연산 안정성을 위해 패딩(Padding) (11, 7)을 적용한다. 이를 통해 최종 입력 크기는 (3, 208, 112)로 확장되며, 알고리즘 입력으로 사용된다.

4.2 특징 추출 및 충돌예측

본 장에서는 특징 추출(Feature extractor)의 성능 향상을 위해 영역 기반 공간 어텐션(Spatial attention) 메커니즘을 적용한다. 이 메커니즘을 사용하는 목적은 VRU 객체의 미래 궤적에 가장 큰 영향을 미치는 영역을 식별하고, 해당 영역에 선택적으로 집중하는 것이다. 특징 추출은 기존 문헌에서 사용되었던 알고리즘¹³⁾과 동일한 구조인 3개의 연속된 합성곱(Convolution) 계층과 Max pooling 연산으로 구성되며, 입력 이미지로부터 Feature map을 생성한다. SBEV의 픽셀 정보는 공간적으로 희소(Sparse)해서 매우 깊은 CNN 구조를 사용할 경우 성능 제고 효과가 크지 않다. 본 연구에서는 합성곱 계층 수를 늘리기 않고 공간 어텐션(Spatial attention) 메커니즘의 효과를 살펴보고자 한다.

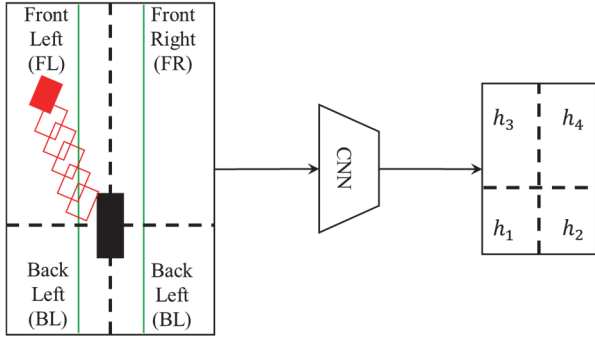


Fig. 8 Four areas of input representation in the feature map

추출된 Feature map은 Fig. 8과 같이 차량 중심 좌표를 기준으로 전·후방 및 좌·우 축을 따라 네 개의 영역, 즉 Front-left(FL), Front-right(FR), Back-left(BL), Back-right(BR)으로 분할된다. 본 연구의 4영역 분할 구조는 차선 변경 예측 알고리즘에서 제안된 구조를 기반으로 한다.¹¹⁾ 네 영역이 주행 위험과 객체 행동을 구분하는 가장 의미있는 공간 단위를 제시하고 있으며, BEV 입력의 Sparsity 특성으로 인해 과도하게 복잡한 영역 분할은 성능 향상에 기여하지 않는다고 보고하였다. 이러한 기존 연구 결과를 바탕으로 본 연구에서도 동일한 네 영역 기반 어텐션(Attention) 구조를 채택하였다. 각 영역 Feature map은 평탄화(Flatten)를 통해 특징 벡터 h_1, h_2, h_3, h_4 로 표현된다. Flatten된 특징 벡터는 어텐션 레이어(Attention layer)를 통해 가중치가 계산되며, 충돌 가능성이 높은 영역에 더 큰 중요도를 부여하는 학습 과정을 수행한다. 어텐션(Attention) 가중치는 선형 변환(Linear layer)을 통해 다음과 같이 계산된다.

$$s_i = \mathcal{W}_i h_i + b_i, \quad i = 1, 2, 3, 4 \quad (6)$$

여기서 $\mathcal{W}_i \in \mathbb{R}^{1 \times d}$ 와 $b_i \in \mathbb{R}$ 는 학습 가능한 파라미터이며, 각 영역에 대해 독립적으로 존재한다. 계산된 스칼라 Score s_i 는 Softmax 함수를 통해 정규화 되어 영역별 어텐션(Attention) 가중치 α_i 로 변환된다.

$$\alpha_i = \frac{e^{s_i}}{\sum_{j=1}^4 e^{s_j}}, \quad \sum_{i=1}^4 \alpha_i = 1 \quad (7)$$

최종적으로 어텐션 기반 특징 벡터 c 는 영역별 가중합(Weighted sum)으로 표현된다. 아래와 같이 계산된 c 는 이후 Decision인 분류기(Classifier)로 전달되며, 충돌 분류의 입력으로 사용된다.

$$c = \sum_{i=1}^4 \alpha_i h_i \quad (8)$$

영역 기반 어텐션(Spatial attention) 구조는 이미지 내 모든 영역을 동일 처리하는 기존 CNN 구조와 달리 중요 영역에 선택적으로 집중하게 한다.

마지막으로 Fig. 6의 Decision은 Feature extractor로부터 전달된 통합 특징 벡터 c 를 입력으로 받아 충돌판별하는 클래스 분류기(Classifier) 역할을 수행한다. 분류기는 두 개의 완전연결층(Fully connected layer)과 비선형 활성화 함수 ReLU, 그리고 드롭아웃(Dropout)으로 구성되어 있으며, 이는 과적합을 방지하고 일반화 성능 향상을 위한 구조이다. 최종 출력층은 Softmax 함수를 통해 각 클래스별 확률을 계산한다.

5. 실험 결과

5.1 학습 환경

성능 평가를 위해 Table 3과 같이 가상 및 실도로 주행 데이터를 사용하였고, Table 4와 같이 학습 데이터셋을 구성하였다. 해당 알고리즘은 Python 3.9 기반 파이토치(PyTorch)로 구현하였다. 학습 및 평가는 Intel Core i9-12900KF (3.19 GHz) CPU와 NVIDIA RTX 3090 Ti GPU 시스템에서 수행되었다. 학습은 Mini-batch 방식을 적용하였으며, 배치 크기는 128로 설정하였다. 최적화 알고리즘은 Adam을 사용하였고, 학습률(Learning rate)은 0.001로 설정하였다. 전체 학습 횟수는 100 epochs로 수행하였으며, 손실 함수로는 Cross-entropy loss를 사용하였다. 또한 어텐션(Attention) 모듈의 불확실성을 억제하기 위해 추가적인 Attention entropy loss를 도입하였다. 따라서 최종 손실 함수는 다음과 같이 정의된다.

$$\mathcal{L}_{total} = \mathcal{L}_{CE} + \lambda_{att} \mathcal{L}_{attn} \quad (9)$$

이때 \mathcal{L}_{attn} 은 어텐션(Attention) 가중치의 엔트로피 향으로, 알고리즘이 중요도가 높은 영역에 집중하도록 유도한다. 가중 계수 λ_{att} 는 하이퍼파라미터 탐색(Grid search)을 통해 최적값을 결정하였다.

5.2 평가 지표

충돌 예측 시스템의 성능을 정량적으로 평가하기 위해 세 가지 핵심 성능지표(Key Performance Index, KPI)를 정의하였다. KPI 1은 충돌 발생 전 예측 시점으로, 위험(Safety-critical) 시나리오에서 실제 충돌 발생 시점 t_i 를 기준으로 알고리즘이 최초 충돌 예측 시점 t_c 의 차이로

정의된다. 본 지표는 알고리즘이 충돌을 얼마나 조기에 판단하는지를 나타낸다.

$$t_d = t_i - t_c, \quad 1.5s > t_d \quad (10)$$

KPI 2는 False Positive Rate(FPR)로, 실제로 충돌이 발생하지 않았음에도 불구하고 알고리즘이 충돌로 잘못 판단한 비율을 의미한다. 마지막으로 KPI 3는 정확도(Accuracy)로, 전체 예측 결과 중 실제 라벨과 일치하는 비율을 나타내며 알고리즘의 전반적인 분류 성능을 평가하는 지표이다. 모든 평가지표는 알고리즘의 출력값이 충돌로 판정된 시점을 기준으로 산출하였다.

충돌 예측 결과는 실제 정답 라벨(Ground truth)과 알고리즘의 예측 조합에 따라 네 가지 유형으로 분류된다. True Positive (TP)는 충돌이 발생하는 시나리오에서 알고리즘이 Pre-crash 구간($t_1 - t_2$)에서 올바르게 예측한 경우를 의미한다. True Negative (TN)는 충돌이 발생하지 않는 시나리오 전 구간에서 일관되게 ‘Safe’로 판단한 경우이다. False Positive (FP)는 충돌이 발생하지 않음에도 알고리즘이 충돌로 잘못 판단한 경우이며, 이는 두 가지 상황으로 나뉜다. 하나는 Pre-crash 구간 이전($t < t_1$)에 너무 이른 시점에 충돌을 예측한 경우, 다른 하나는 충돌이 없는 시나리오에서 충돌 발생으로 오판단한 경우이다. 마지막으로 False Negative (FN)는 실제 충돌이 발생함에도 불구하고 Pre-crash 구간($t_1 \sim t_2$)동안 충돌을 탐지하지 못하고 ‘Safe’로 분류한 경우를 의미한다.

5.3 성능 평가

기존 문헌에서 제시된 CP-CNN¹³⁾, 어텐션(Attention)만을 포함한 ACNN, 그리고 CP와 어텐션(Attention)을 모두 포함한 제안 알고리즘 CP-ACNN에 대하여 비교 평가를 수행하였다. 우선 Table 5는 전체 가상 데이터에 대한 성능 비교 결과를 보여 주고 있다. 충돌 예측 시간 관점에서는 ACNN이 제일 높지만, 정확도나 FPR의 성능 지표 관점에서는 CP-ACNN이 가장 우수 우수함을 보여주고 있다. 이는 CP-ACNN이 모델 기반 충돌 확률 예측(CP)의 구조적 안정성과 어텐션(Attention) 기반 적응적 표현력(Adaptive representation)을 효과적으로 결합함으로써, 기존 알고리즘의 한계를 보완할 수 있는 가능성을 보여준다.

다음으로 가상 데이터 중 가림(Occlusion) 상황이 포함된 사고 시나리오 3건에 대해 충돌 예측 성능을 비교하였다. Table 6에서 보는 바와 같이, 기존 CP-CNN 대비 ACNN과 CP-ACNN은 정확도(Accuracy) 측면에서 각각 30.07%, 8.68% 수준의 성능 향상을 보였다. 또한 제안한 CP-ACNN은 기존 알고리즘 대비 오탐율(FPR)

Table 5 Performance comparison of collision prediction with respect to virtual dataset

		ACC (%)	FPR (%)	Avg t_d (s)
VRU	CP-CNN	87.13	21.95	1.05
	ACNN	81.55	29.07	1.23
	CP-ACNN	89.14	18.51	1.08

Table 6 Performance comparison of collision prediction based on simulation data (only occlusion scenarios)

		ACC (%)	FPR (%)	Avg t_d (s)
VRU	CP-CNN	51.44	79.15	1.12
	ACNN	81.51	30.39	1.10
	CP-ACNN	60.12	66.09	1.09

Table 7 Performance comparison of collision prediction based on experimental data

		ACC (%)	FPR (%)	Avg t_d (s)
VRU	CP-CNN	93.83	6.17	1.09
	ACNN	90.91	9.09	1.01
	CP-ACNN	98.48	1.51	1.07

을 79.15%에서 30.39%로 13.06% 상대 감소하였다. 이러한 정량적 개선 결과는 어텐션(Attention) 구조가 이미지 내 소형 객체의 가림(Occlusion) 현상에 더 민감하게 반응하여 위험도를 빠르고 안정적으로 인지할 수 있는 성능적 이점을 갖는다.

제안된 네트워크의 효과를 분석하기 위해서 Fig. 2에서 제시한 보행자 시나리오에 대하여 세 가지 네트워크의 Feature map을 시각화한 결과를 Fig. 9에서 보여주고자 한다. 해당 가림 기반 시나리오는 Pre-crash 시나리오이며, 충돌 회피를 위한 시간이 매우 짧기 때문에 충돌 예측 평가 항목 중 KPI 1과 같이 충돌을 얼마나 빠르게 판단하는지가 핵심 성능 요소이다. Fig. 9(a)는 Fig. 2(b)와 동일 시점의 SBEV를 보여주고 있으며, 각 알고리즘의 Feature activation을 보여준다. 색상은 네트워크가 주목한 영역을 의미하며, 빨간색에 가까울수록 높은 활성화(High activation)로 해당 영역을 충돌 판단에 주요 근거로 사용했음을 의미한다. 반면, 파란색 영역은 낮은 활성화(Low activation)로 분류 결정에 기여도가 낮은 부분이다. 차차 레이더 센서가 보행자를 최초로 인지한 시점인 10.21초는 Fig. 9(b)-(d)에서 보는 바와 같이 세 알고리즘 모두 Safe로 판단하였다. 보행자의 차차 주행 차선 진입 시점인

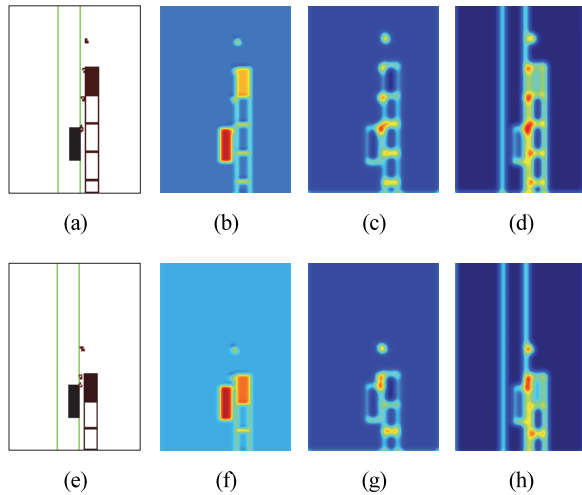


Fig. 9 SBEV, visualization results, and collision prediction to pedestrian crossing scenario (a) SBEV at $t = 10.21$, (b) CP-CNN Pred: Safe, (c) ACNN Pred: Safe, (d) CP-ACNN Pred: Safe, (e) SBEV at $t = 10.77$, (f) CP-CNN Pred: Safe, (g) ACNN Pred: Collision, (h) CP-ACNN Pred: Collision

10.77초는 Fig. 9(f)-(h)에서 확인할 수 있듯이 CP-CNN에서는 충돌을 예측하지 못했지만 ACNN과 CP-ACNN은 보행자 영역에서 높은 활성도를 보이며 충돌을 예측하였다. 특히, ACNN은 10.71초부터 충돌을 예측하였으며 보행자를 인지한 시점부터 0.5초 만에 판단을 수행하였다. 더 나아가 CP-ACNN은 10.31초 시점부터 충돌을 예측하였으며, 보행자를 인지 이후 0.1초 만에 충돌을 판단하였다. 이는 어텐션(Attention) 메커니즘이 가림 기반 시나리오에서 기존 CP-CNN 대비 Pre-crash 상황을 더 빠르고 정확하게 인식함을 확인할 수 있다.

Table 7은 센서 노이즈와 복잡한 객체 구성(High complexity)을 포함한 실도로 주행 데이터에 대한 테스트 결과를 나타낸다. CP-ACNN은 기존 알고리즘과 유사한 평균 판단 시점(Avg t_d)을 유지하면서 FPR이 5% 감소하고 Accuracy가 5% 향상되었다. 이는 어텐션(Attention) 기반 구조가 실제 환경에서도 충돌 예측의 신뢰성과 안전성 향상을 보여준다.

6. 결론

본 연구는 자율주행 차량 환경에서 교통약자(VRU)와의 충돌 예측 알고리즘을 제안하였다. 기존 CP-CNN은 입력 이미지 내 소형 객체에 대해 지역적 위험도를 구분하지 못하고 전체 영역에 동일한 가중치를 부여함으로써 가림(Occlusion) 상황이나 복잡한 교통 환경에서 충돌 위험 예측 성능 저하가 있었다. 해당 문제 개선을 위하여 객체의 상대적 중요도와 공간적 위험 분포를 반영할 수

있도록 영역 기반 어텐션(Spatial attention) 메커니즘 결합의 CP-ACNN(CP-Attention-based CNN)을 제안하였다. 가상 및 실도로 주행 데이터의 실험 결과, CP-ACNN은 기존 CP-CNN 대비 평균 정확도(Accuracy) 향상과 오탐율(False Positive Rate, FPR) 감소 측면에서 우수한 성능을 보였으며, 특히 가림(Occlusion) 기반 Pre-crash 상황에서 더 빠른 예측 성능을 확인하였다.

본 연구에는 몇 가지 한계가 존재한다. 첫째, 시물레이션 기반 가상 시나리오는 다양한 위험 상황을 효과적으로 생성할 수 있으나, 다중 VRU 상호작용, 복잡 교차로 구조, 군집 보행과 같은 다양한 실도로 상황을 완전히 반영하기 어렵다는 한계가 있다. 실도로 주행 데이터 또한 충돌 상황을 포함하지 않기 때문에 충돌 위험 예측의 전체 분포를 충분히 반영하지 못하는 제약이 존재한다. 둘째, 본 연구에서 활용한 SBEV는 입력 복잡도를 줄이고 실시간성을 확보할 수 있다는 장점이 있으나, 실제 도로 환경에서 나타나는 복잡한 배경 정보, VRU의 세부 형상 등이 충분히 반영되지 못한다. 또한 차선 정보를 3차 방향으로 단순화하고 객체를 확률 기반 점 집합으로 표현함에 따라 정교한 공간 정보가 일부 손실되는 구조적 한계가 있다.

향후 연구에서는 도시 교차로, 도시 혼잡 환경, 다중 VRU 상호작용 등의 복합 시나리오를 포함하는 실도로 주행 데이터를 활용하고자 한다. 더 나아가 BEV 변환 기법과 Transformer 기반 어텐션 구조를 접목하여 SBEV 표현의 단순화로 인해 손실되는 공간 정보를 보완하여 제안 알고리즘의 범용성과 실효성을 확보하고자 한다.

후 기

본 논문은 2025년도 산업통상자원부 및 한국산업기술 기획평가원(KEIT) 연구비 지원에 의한 연구임(과제번호: RS-2024-00448797, 과제명: 도심 자율주행을 위한 인공지능 기반 보행자·마이크로모빌리티 충돌회피 및 경감시스템 기술 개발)

References

- 1) World Health Organization, Global Status Report on Road Safety 2023, <https://www.who.int/teams/social-determinants-of-health/safety-and-mobility>, 2023.
- 2) M. Aleksa, A. Schaub, I. Erdelean, S. Wittmann, A. Soteropoulos and A. Fördös, "Impact Analysis of Advanced Driver Assistance Systems (ADAS) Regarding Road Safety – Computing Reduction Potentials," European Transport Research Review, Vol.16, No.1, Paper No.39, 2024.

- 3) International Transport Forum, Road Safety Annual Report 2020, https://www.oecd.org/en/publications/road-safety-annual-report-2020_f3e48023-en.html, 2020.
- 4) C. Zhang and C. Berger, "Pedestrian Behavior Prediction Using Deep Learning Methods for Urban Scenarios: A Review," *IEEE Transactions on Intelligent Transportation Systems*, Vol.24, No.10, pp.10279–10301, 2023.
- 5) A. Rasouli and J. K. Tsotsos, "Joint Attention in Driver-Pedestrian Interaction: From Theory to Practice," arXiv preprint arXiv:1802.02522, 2018.
- 6) J. Lee, U. Jung and B. Song, "Critical Scenario Generation for Collision Avoidance of Automated Vehicles Based on Traffic Accident Analysis and Machine Learning," *Transactions of KSAE*, Vol.28, No.11, pp.817–826, 2020.
- 7) ISO, ISO 34502 – Road Vehicles – Test Scenarios for Automated Driving Systems – Scenario Based Safety Evaluation Framework, 2022.
- 8) H. Winner, K. Lemmer, T. Form and J. Mazzega, "PEGASUS – First Steps for the Safe Introduction of Automated Driving," *Road Vehicle Automation 5*, pp.185–195, 2019.
- 9) D. Nalic, T. Mihaji, M. Baumler, M. Lehmann, A. Eichberger and S. Bernsteiner, "Scenario Based Testing of Automated Driving Systems: A Literature Survey," *FISITA Web Congress*, Vol.10, 2020.
- 10) S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings and A. Mouzakitis, "Deep Learning-Based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review," *IEEE Transactions on Intelligent Transportation Systems*, Vol.23, No.1, pp.33–47, 2022.
- 11) S. Mozaffari, E. Arnold, M. Dianati and S. Fallah, "Early Lane Change Prediction for Automated Driving Systems Using Multi-Task Attention-Based Convolutional Neural Networks," *IEEE Transactions on Intelligent Vehicles*, Vol.7, No.3, pp.758–770, 2022.
- 12) S. Lee, Y. Jeong and B. Song, "Multi-Task Prediction of Collision and Trajectories Based on Transformer Network for Safety-Critical Scenarios of Automated Vehicles," *Transactions of KSAE*, Vol.32, No.10, pp.843–852, 2024.
- 13) S. Lee, B. Song and J. Shin, "Collision Prediction in an Integrated Framework of Scenario-Based and Data-Driven Approaches," *IEEE Access*, Vol.12, pp.55234–55247, 2024.
- 14) J. Jansson, Collision Avoidance Theory with Application to Automotive Collision Mitigation, Ph.D. Dissertation, Linköping University, Sweden, 2005.
- 15) N. Kapila, J. Glattki and T. Rathi, "CNNtention: Can CNNs Do Better with Attention?," arXiv preprint arXiv:2412.11657, 2024.
- 16) G. Rangel, J. C. Cuevas-Tello, J. Nunez-Varela, C. Puente and A. G. Silva-Trujillo, "A Survey on Convolutional Neural Networks and Their Performance Limitations in Image Recognition Tasks," *Journal of Sensors*, 2024.
- 17) J. Hu, L. Shen and G. Sun, "Squeeze-and-Excitation Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.7132–7141, 2018.
- 18) S. Woo, J. Park, J. Y. Lee and I. S. Kweon, "CBAM: Convolutional Block Attention Module," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp.3–19, 2018.
- 19) I. Bello, B. Zoph, Q. Le, A. Vaswani and J. Shlens, "Attention Augmented Convolutional Networks," *Proceedings of the IEEE International Conference on Computer Vision*, pp.3285–3294, 2019.
- 20) N. T. S. Board, Special Investigation Report – Highway Vehicle and Infrastructure-Based Technology for the Prevention of Rear-End Collisions, <http://www.nts.gov>, 2001.
- 21) European Commission, European New Car Assessment Programme, <https://www.euroncap.com/en/for-engineers/protocols>, 2020.
- 22) IGLAD, Codebook, <https://www.iglad.org>, 2020.
- 23) Traffic Accident Analysis System, <https://taas.koroad.or.kr>, KoROAD, 2017.
- 24) J. Stoll, A. Schneider, M. Wisch, P. Seiniger and T. Schaller, "The Addressed VRU Scenarios within PROSPECT and Associated Test Catalogue," *Techn. Ber., PROSPECT Consortium*, 2016.
- 25) Samsung Fire, Safety Enforcements Considering E-Scooter's Braking Performance and Accident Characteristics, <https://www.samsungfire.com>, 2021.
- 26) Tennessee's Integrated Traffic Analysis Network, <https://titan.safety.tn.gov>, 2020.
- 27) IPG Automotive, CarMaker Reference Manual Version 9.0.2, IPG Automotive GmbH, 2020.
- 28) M. Wisch, A. Hellmann and M. Lerner, "Car-to-Car Accidents at Intersections in Europe and Identification of Use Cases for the Test and Assessment of Respective Active Vehicle Safety Systems," *Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles*, pp.10–13, 2019.