

생성형 AI 기반 고위험 시나리오 생성을 통한 자율주행 레이싱 시스템 안전성 평가 프레임워크

Generative AI-based Safety Evaluation Framework for Autonomous Racing System via Safety-critical Scenario Generation

이 현 빈¹, 이 상 현¹, 양 형 준¹, 권 철 현^{1*}

(Hyeonbin Lee¹, Sanghyeon Lee¹, Hyeongjoon Yang¹, and Cheolhyeon Kwon^{1*})

¹Department of Mechanical Engineering, Ulsan National Institute of Science and Technology

Abstract: This paper introduces a generative artificial intelligence (AI)-based safety evaluation framework for autonomous racing systems, focusing on efficiently searching safety-critical racing scenarios utilizing domain knowledge, optimization, and machine learning. The proposed framework consists of three main phases: 1) dataset generation, 2) conditional variational auto-encoder (CVAE) model training, and 3) safety-critical scenario generation and evaluation. In the first phase, the dynamic scenario generation is automatically processed by leveraging ontological domain knowledge and genetic algorithm to efficiently establish a potentially safety-critical driving dataset. In the second phase, we train the CVAE network with the driving dataset generated from the first phase, allowing for diverse and realistic variations in driving scenarios. In the final phase, safety-critical scenarios are generated through the trained CVAE network by adversarially varying the scenarios. Experimental results show that the proposed framework identifies various safety-critical scenarios in different racing conditions, exhibiting its effectiveness for safety evaluation of autonomous racing systems.

Keywords: safety evaluation, safety-critical scenario, autonomous driving system, generative ai, autonomous racing

1. 서론

최근 Indy Autonomous Challenge [1], Abu Dhabi Autonomous Racing League (A2RL) [2]와 같은 자율주행 레이싱 경기들이 개최되면서 자율주행 레이싱 기술 또한 세계적으로 발전하고 있으며, 자율주행 레이싱 기술의 발전에 따라 레이싱 환경에서의 자율주행 시스템 안전성 평가 기술의 필요성 또한 증가하고 있다. 자율주행 레이싱 시스템의 안전성을 평가하기 위해서는 고위험 시나리오(safety-critical scenario)를 생성하여 자율주행 시스템이 대응하지 못하는 코너 케이스(corner case)를 찾아내는 것이 중요하다. 하지만 이러한 고위험 시나리오는 실제 주행에서 자주 발생하지 않기 때문에 실제 주행 데이터에 존재하지 않는 시나리오에 미리 대비하는 것이 매우 힘들다. 더욱이, 자율주행 시스템의 안전성 평가를 위해서는 실제 도로환경과 유사한 환경이 구현되어 있는 high-fidelity simulator를 통한 테스트가 필요하다. 하지만 이러한 시뮬레이터들은 상대적으로 높은 연산량을 필요로 하여 연산 비용이 큰 시나리오를 임의로 설정하여 자율주행 시스템을 테스트하고 코너 케이스를 찾는 것은 불필요한 시뮬레이션 시간 및 자원을 증가시켜 매우 비효율적이다. 따라서 코너 케이스가 발생할 확률이 높은 고위험 시나리오를 기반으로 자율주행 시스템의 안전성을 평가함으로써 불필요한 시뮬레

이션 시간을 줄일 필요성이 있다. 그러나, 이러한 고위험 시나리오를 찾는 것은 일반적으로 매우 어려운 문제이며, 이에 따라 고위험 시나리오 생성에 관한 연구가 활발히 진행되고 있다. 고위험 시나리오 생성 방법론은 data-driven generation, adversarial generation, knowledge-based generation 방식으로 분류될 수 있다[3]. 본 논문에서는 각 방법론을 비교하여 표 1과 같이 작성하였다.

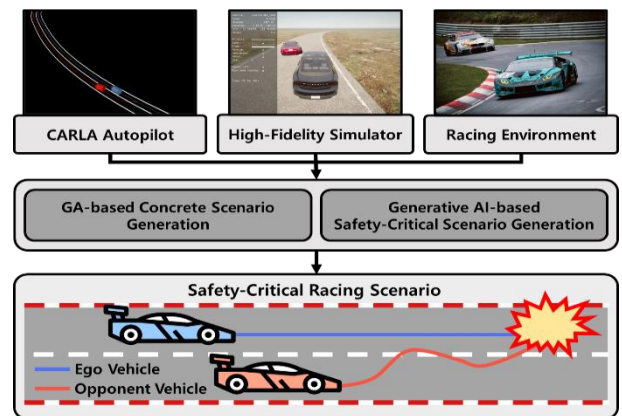


그림 1. 생성형 AI 기반 고위험 시나리오 생성 개념도.
Fig. 1. Concept of safety-critical scenario generation.

* Corresponding Author

Manuscript received February 24, 2025; revised March 18, 2025; accepted April 7, 2025

이현빈: 울산과학기술원 기계공학과 대학원생(bin2915@unist.ac.kr, ORCID[®] 0009-0001-8559-1529)

이상현: 울산과학기술원 기계공학과 대학원생(lsh0527@unist.ac.kr, ORCID[®] 0009-0006-8189-5206)

양형준: 울산과학기술원 기계공학과 대학원생(shineejuon@unist.ac.kr, ORCID[®] 0009-0009-7526-3132)

권철현: 울산과학기술원 기계공학과 교수(kwonc@unist.ac.kr, ORCID[®] 0000-0002-6084-8236)

※ 본 연구는 한국연구재단과 국토교통부/국토교통과학기술진흥원의 지원을 받아 연구되었음(RS-2024-00342930, RS-2024-00405707).

표 1. 고위험 시나리오 생성 방법론 비교.

Table 1. Comparison of methodologies for safety-critical scenario generation.

방법론	특징	충실도	효율성	제어가능성
Data-driven generation	실 주행 데이터를 기반으로 실제 환경에 충실한 고위험 시나리오를 생성하여 충실도가 높음.	높음	낮음	낮음
Adversarial generation	자율주행 시스템에 능동적으로 위협을 가하는 방식을 통해 시나리오를 생성함으로써 고위험 시나리오를 효율적으로 생성할 수 있음.	낮음	높음	낮음
Knowledge-based generation	Knowledge 기반으로 시나리오 구성요소를 적절히 조합하거나 제한조건 적용을 통해 위험한 시나리오를 생성하여 안전성 평가 대상에 적합한 시나리오 생성이 가능함.	낮음	낮음	높음

Data-driven generation 방식에서는 실제 차량의 주행 데이터를 기반으로 시나리오를 생성한다. 차량 실 주행 데이터 중 고위험 시나리오를 추출하여 이를 기반으로 자율주행 시스템 평가에 사용할 수 있다[4]. 실제로는 고위험 시나리오가 발생할 확률이 매우 낮아 최근에는 이러한 실 주행 데이터를 통해 생성형 AI를 기반으로 데이터를 증강하여 비교적 적은 양의 주행 데이터로부터 많은 양의 다양한 고위험 시나리오를 생성하는 연구가 활발하게 진행되고 있다. [5]에서는 UMTRI (University of Michigan Transportation Research Institute)에서 수집한 주행 데이터셋(driving dataset) [6]을 활용하여 VAE (Variational Auto-Encoder) [7], GAN (Generative Adversarial Nets) [8]을 기반으로 고위험 시나리오를 생성하였으며, TrafficSim [9]에서는 ATG4D [10] 주행 데이터셋을 활용하여 CVAE (Conditional Variational Auto-Encoder) [11]를 기반으로 고위험 시나리오를 생성하였다.

Adversarial generation 방식에서는 보다 더 효율적인 시나리오 생성을 위해 자율주행 시스템에 능동적으로 위협을 가하는 시나리오를 생성한다. Learning-to-collide [12]에서는 강화학습을 통해 자율주행 차량의 주위에 있는 에이전트가 차량에 위협을 가하는 주행을 하도록 학습시키는 방식으로 고위험 시나리오를 생성하였다.

Knowledge-based generation의 경우에는 knowledge를 기반으로 생성된 시나리오에 제한조건을 적용하거나 시나리오 생성 모델이 시나리오를 생성할 때 knowledge를 기반으로 위험한 시나리오가 생성되도록 설정하여 고위험 시나리오를 생성한다. 그 중 특히, Ontology [13] 기반 시나리오 생성 방식은 knowledge를 기반으로 도로 환경, 차량, 보행자, 정적 장애물과 같은 객체의 종류, 크기, 위치, 속도 등 시나리오 구성 요소들을 적절히 조합하여 유용한 시나리오를 구성할 수 있다.

Data-driven generation 방식은 실 주행 데이터를 활용함으로써 실제 환경에서 발생할 수 있는 상황과 유사한 시나리오를 생성하여 현실과 이질감이 적은 시나리오를 생성할 수 있어 충실도(fidelity)가 높다는 장점이 있지만 현실에서 위험한 시나리오는 매우 적은 비율로 존재하며 평가하고자 하는 시나리오가 데이터에 존재하지 않을 수 있어 한계가 있다. Adversarial generation 방식은 시나리오의 자율주행 차량에 위협을 가하는 시나리오를 생성할 수 있어 효율적으로 다수의 위험한 시나리오를 생성할 수 있다는 장점이 있으나 대부분

AI 기반으로 학습한 모델을 통해 시나리오를 생성하는 등 실 주행 데이터를 기반으로 하지 않아 현실적인 환경을 잘 반영하지 못한다는 단점이 있다. Knowledge-based generation 방식은 knowledge를 기반으로 평가하고자 하는 시나리오를 자유롭게 생성할 수 있어 시나리오 제어가능성이 높다는 장점이 있지만 고위험 시나리오의 시나리오 파라미터를 적절히 설정하여 위험한 상황을 형성하는 데에 많은 시간이 소요될 수 있어 비효율적이라는 단점이 있다.

각 생성 방식의 단점을 보완하고자 STRIVE [14], AdvDiffuser [15]에서는 data-driven 및 adversarial generation 방식을 융합하여 시나리오를 생성하였으며, 각각 CVAE와 diffusion model [16]을 기반으로 NuScenes [17] 주행 데이터셋을 활용하여 학습된 생성형 AI를 통해 도로의 에이전트가 자율주행 차량에 위협을 가하는 주행을 하도록 적대적 최적화(adversarial optimization)를 진행함으로써 고위험 시나리오를 생성하였다. 또한, [18]에서는 STRIVE를 보완하고자 diffusion model을 적용하여 시나리오 다양성을 향상시켰다. 이와 같이 data-driven 및 adversarial generation 방식을 융합하면 실 데이터 기반으로 적대적 최적화를 통해 시나리오를 생성하여 현실성 있는 고위험 시나리오를 효율적으로 생성할 수 있다. 하지만 이러한 방식은 여전히 실 주행 데이터에 의존하여 안전성 평가 대상에 적합한 시나리오를 생성하기 어려울 수 있어 한계점이 존재한다.

이러한 연구 동향에 착안하여, 본 논문에서는 기존 고위험 시나리오 생성 방식의 한계를 극복하고자 그림 1과 같이 knowledge-based generation, adversarial generation 방식을 융합하여 고위험 시나리오를 생성하고 이를 기반으로 high-fidelity 시뮬레이터를 통해 자율주행 레이싱 시스템의 안전성을 평가하는 프레임워크를 제안한다. 본 논문에서 제안하는 프레임워크에서는 먼저 ontology 방법론 기반으로 구조화된 시나리오 파라미터 공간을 도출하여 논리 시나리오(logical scenario)를 구성한다. 이러한 논리 시나리오는 도로 환경, 각 차량의 위치, 속도 등 시나리오를 구성 요소를 필요에 따라 조합할 수 있도록 하여 시나리오 제어가능성을 확보할 수 있다. 구성된 논리 시나리오를 기반으로 GA (Genetic Algorithm)을 적용하여 시나리오 파라미터 값이 할당된 상세 시나리오(concrete scenario)를 생성하고, 해당 상세 시나리오를 자율주행 시스템에 적용하여 high-fidelity 시뮬레이션을 진행함으로써 시계열 차량 주행 데이터로 구성된 동적 시나리오(dynamic scenario)를

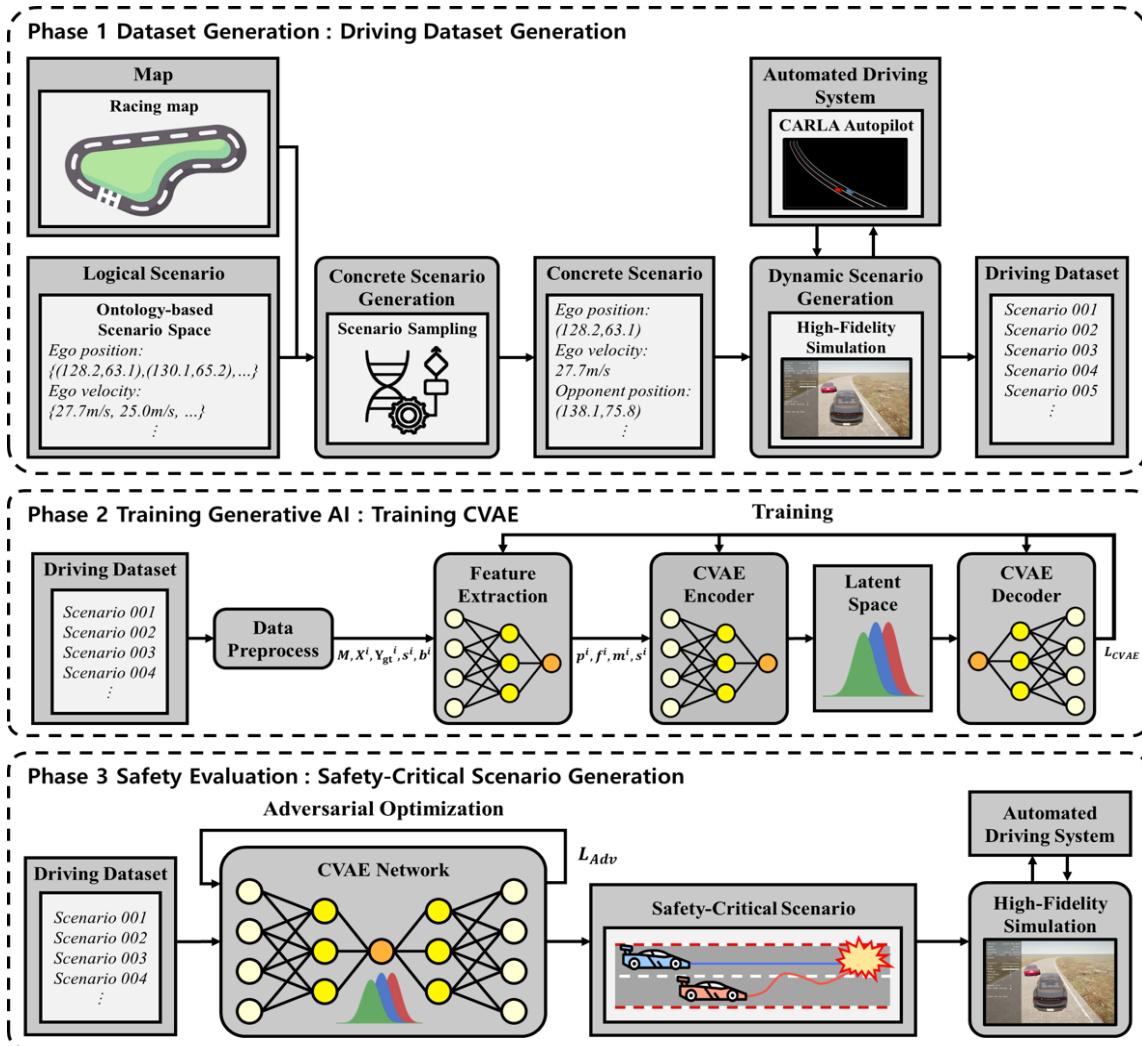


그림 2. 생성형 AI 기반 고위험 시나리오 생성을 통한 자율주행 레이싱 시스템 안전성 평가 프레임워크 개요도.

Fig. 2. Schematic of generative AI-based safety evaluation framework via safety-critical scenario generation.

생성한다. 이러한 high-fidelity 시뮬레이션을 통해 실 환경과 유사한 환경에서 생성한 데이터셋을 확보함으로써 충실도를 확보할 수 있다. 생성된 동적 시나리오는 수집되어 주행 데이터셋을 구축한다. 구축된 주행 데이터셋을 기반으로 생성형 AI 모델을 학습한 후, 적대적 최적화를 통해 고위험 시나리오를 생성한다. 이러한 과정을 통해 최적화와 AI를 기반으로 위험한 시나리오를 능동적으로 생성하여 효율성 또한 확보할 수 있다. 또한 최종적으로 생성된 고위험 시나리오를 기반으로 high-fidelity 시뮬레이션을 통해 실제 환경을 모사하여 자율주행 시스템을 평가한다.

본 논문에서는 아래와 같은 세 가지 연구 목표를 설정하였다.

1. 자율주행 레이싱 시스템 평가 프레임워크 구현: High-fidelity 시뮬레이션을 기반으로 한 자율주행 레이싱 시스템 안전성 평가 프레임워크를 구현하였다. 본 프레임워크는 domain knowledge, GA기반 최적화, 그리고 생성형 AI를 융합하여 고위험 시나리오를 효율적으로 탐색할 수 있도록 설계되었다.

2. 효율적인 주행 데이터셋 생성: 상세 시나리오는 domain knowledge를 기반으로 최적화를 통해 효율적으로 생성된다.

Domain knowledge를 ontology적으로 구조화하여 상세 시나리오의 생성 공간을 정의하였다. 이후, 이 생성 공간 내에서 GA를 적용하여 잠재적으로 위험한 주행 데이터셋을 효과적으로 생성하였다.

3. 생성형 AI기반 고위험 시나리오 탐색: 고위험 시나리오는 생성형 AI를 활용하여 실제 시뮬레이션과 유사한 결과를 생성함으로써 신속하게 탐색된다. 시뮬레이션에서 획득한 비교적 적은 양의 주행 데이터를 이용하여 CVAE 네트워크를 학습시켰으며, 이를 통해 이후의 동적 궤적을 생성할 수 있도록 하였다. 특히, 고위험 시나리오를 탐색할 때, 상대 차량의 궤적을 현실적이면서도 자율주행 차량에 위협을 가하는 시나리오를 생성하도록 적대적 최적화를 수행하였다.

이후의 섹션은 다음과 같은 내용으로 구성된다. 섹션 II에서는 본 논문에서 제안하는 AI 기반 자율주행 시스템 안전성 평가 프레임워크의 상세한 구조 및 적용 방법론 등을 기술하였다. 섹션 III에서는 제안하는 프레임워크를 적용하여 고위험 시나리오에 대한 자율주행 시스템의 안전성을 평가하고 생성한 고위험 시나리오를 분류하여 자율주행 시스템이 대응하거나 대응하지 못하는 시나리오를 통계를 기반

알고리즘 1. GA 기반 상세 시나리오 생성 알고리즘.
Algorithm 1. GA-based concrete scenario generation.

Algorithm 1 GA-based Concrete Scenario Generation

Input: Set of ego vehicle positions (e_i, e_g) and opponent positions (o_i, o_g)

Output: Set of valid concrete scenarios satisfying the constraints

Initialize population of (e_i, e_g, o_i, o_g)

while Stored solutions < desired number **do**

for each candidate (e_i, e_g, o_i, o_g) **do**

 Compute distances:
 $d_e \leftarrow \|e_g - e_i\|$, $d_o \leftarrow \|o_g - o_i\|$
 $d_{eo} \leftarrow \|o_i - e_i\|$

 Compute lane density ρ

 Update maximum lane density θ_ρ

if ($\theta_{d,min} < d_e < \theta_{d,max}$ **and** $\theta_{d,min} < d_o < \theta_{d,max}$ **and** $\theta_{d_{eo},min} < d_{eo} < \theta_{d_{eo},max}$ **and** $\rho > \theta_\rho$) **then**

 Set fitness: $fitness \leftarrow -1$

else

 Set fitness: $fitness \leftarrow \frac{1}{\rho}$

if $fitness > 0$ **then**

 Store (e_i, e_g, o_i, o_g) as a valid solution

end if

end if

end for

end while

Return: Stored solutions

으로 분석하였다. 섹션 IV에서는 최종적으로 본 논문에서 제안하는 프레임워크의 유효성과 결론, 그리고 향후 연구의 방향성을 기술하였다.

II. AI 기반 자율주행 시스템 안전성 평가 프레임워크

본 논문에서 제안하는 자율주행 레이싱 시스템 안전성 평가 프레임워크는 그림 2의 개요도와 같이 구성되며, 주행 데이터 셋 생성 단계, 생성형 AI 모델 학습 단계, 생성형 AI 기반 고위험 시나리오 생성 및 자율주행 시스템 안전성 평가 단계, 총 세가지 단계로 구분할 수 있다. 먼저 주행 데이터 셋 생성 단계에서는 GA를 기반으로 다양한 상세 시나리오를 생성하고 생성된 상세 시나리오를 기반으로 high-fidelity 시뮬레이터를 통해 주행 데이터셋을 생성한다. 생성형 AI 학습 단계에서는 생성형 AI 모델인 CVAE를 기반으로 주행 데이터 셋을 학습한다. 다음으로, 학습된 CVAE 모델을 기반으로 적대적 최적화 과정을 통해 고위험 시나리오를 생성한다. 생성된 고위험 시나리오는 시뮬레이터에서 다시 구현하여 위협 여부를 확인할 수 있다.

1. 주행 데이터 셋 생성

주행 데이터 셋 생성 단계에서는 high-fidelity 시뮬레이터 상에서 자율주행 시스템을 기반으로 주행 시뮬레이션을 진행하여 주행 데이터 셋을 수집할 수 있는 데이터 셋 생성 자동화 프레임워크를 구축하였다. 먼저 시뮬레이션 상에 레이싱 시뮬레이션 환경과 map을 구축한다. 다음으로 ontology 방법론을 기반으로 domain knowledge를 적용하여 논리 시나리오, 즉 시나리오 파라미터 공간을 생성하고, 해당 논리 시나리오를 기반으로 최적화를 통해 상세 시나리오를 생성한다. 본 논문에서 제안한 프레임워크에서는 주행 데이터 셋을 기반으로 생성형 AI 모델을 학습하여 고위험 시나리오를 생성

표 2. 상세 시나리오 생성을 위한 GA 파라미터.
Table 2. GA parameters for generating concrete scenario.

Variables	Notations	Value
e_i	Ego vehicle spawn position	
e_g	Ego vehicle goal position	
o_i	Opponent vehicle spawn position	
o_g	Opponent vehicle goal position	
d_e	Arc length of ego vehicle path	
d_o	Arc length of opponent vehicle path	
d_{eo}	Arc length from ego vehicle start position to opponent vehicle start position	
ρ	Density assigned to each lane	
θ_ρ	Max. Density assigned to each lane	
$\theta_{d,min}$	Min. Distance of vehicle path	500 (m)
$\theta_{d,max}$	Max. Distance of vehicle path	1000 (m)
$\theta_{d_{eo},min}$	Min. Relative distance between vehicles start positions	10, 15, 20, 25 (m)
$\theta_{d_{eo},max}$	Max. Relative distance between vehicles start positions	15, 20, 25, 30 (m)

하기 때문에 주행 데이터 셋을 생성할 때 다양한 고위험 시나리오가 생성될 수 있는 적절한 상세 시나리오를 설정하는 것이 매우 중요하다. 따라서, 본 논문에서는 적절한 상세 시나리오를 설정하기 위해 GA를 적용하였다. GA를 기반으로 다양한 고위험 시나리오가 생성될 수 있는 상세 시나리오를 설정하기 위해서는 적절한 fitness function을 설계해야 한다. 본 논문에서 설계한 fitness function은 ego vehicle과 opponent vehicle의 spawn position과 goal position 사이의 거리와 ego vehicle과 opponent vehicle의 spawn position 간의 거리가 너무 짧거나 길지 않도록 설정하여 차량 사이의 상호작용이 자주 발생할 수 있도록 설계하였다. 또한, ego vehicle의 spawn position을 기준으로 map 상의 특정 지점에 시나리오가 밀집되지 않도록 설계하였다. 해당 fitness function을 기반으로 최종적으로 원하는 수만큼의 솔루션을 탐색하였을 경우, GA는 종료된다. 구체적인 GA 기반 상세 시나리오 생성 알고리즘

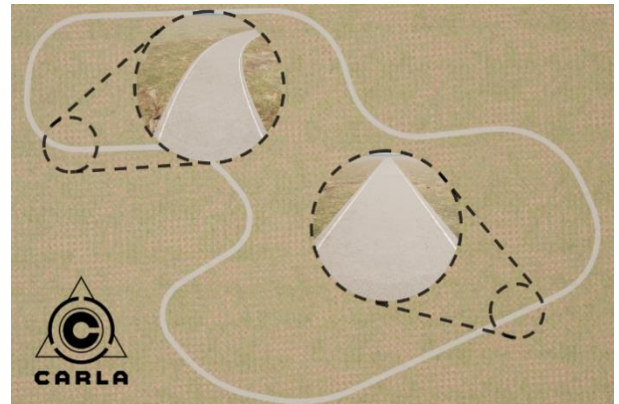


그림 3. CARLA 시뮬레이션 상에 구축한 레이싱 환경.
Fig. 3 Racing environment generated in CARLA simulator.

구조는 알고리즘 1과 같다. 이 때, 상세 시나리오 구성을 위해 사용되는 파라미터는 표 2와 같으며, 시나리오에 따라 변하는 각 차량의 spawn, goal position, 경로 길이, 차량 사이의 거리, GA 솔루션을 찾을 때마다 갱신되는 spawn position의 밀도와 같은 변수로 구성되어 있다. 또한, 경로 길이 및 차량 사이의 거리와 같은 파라미터가 정의되는 범위는 각 변수의 최소, 최대값을 통해 결정된다. 이러한 파라미터를 통해 고위험 시나리오가 생성될 가능성이 높은 다양한 상세 시나리오를 생성할 수 있다. 이렇게 생성된 상세 시나리오를 기반으로 high fidelity 시뮬레이션을 진행하여 자율주행 시스템의 동적 시나리오를 생성할 수 있으며 이러한 동적 시나리오 데이터를 수집하여 주행 데이터셋을 구축한다.

2. 생성형 AI 학습

생성형 AI 학습 단계에서는 이전 단계에서 생성한 주행 데이터셋을 기반으로 생성형 AI 모델을 학습시킨다. 본 논문에서 생성형 AI 모델은 실제 도심로 주행 데이터 셋인 NuScenes [17] 데이터셋을 기반으로 고위험 시나리오를 생성하는 STRIVE [14]의 CVAE 네트워크를 적용하였다. CVAE 네트워크에서 학습이 될 수 있도록 전 처리 과정을 거친 주행 데이터 셋은 feature extraction 과정을 통해 인코더 학습에 적절한 feature로 추출된다. 해당 feature는 인코더를 통해 latent space를 생성하고, 디코더는 생성된 latent space로부터 trajectory를 재생성한다.

가장 먼저, feature extraction 과정에서는 map, trajectory, one-hot encoding 벡터 형태인 semantic class, 그리고 각 agent의 dimension 정보를 포함하고 있는 bounding box를 기반으로 ego vehicle과 opponent vehicle의 feature들이 추출된다. 과거의 trajectory와 미래의 trajectory는 각각 4-layer MLP(Multi-Layer Perceptron)를 거쳐 feature vector로 추출되고, 각 위치별 map 데이터 또한 CNN(Convolutional Neural Network)을 거쳐 map feature로 추출된다. 이렇게 추출된 feature들을 기반으로 CVAE의 prior 및 posterior 네트워크는 GNN(Graph Neural Network) 형태의 scene graph로 구축된다. Prior 네트워크와 posterior 네트워크는 비슷한 형태이지만 posterior 네트워크는 future trajectory의 feature를 추가적으로 포함하고 있다. Scene graph의 node는 ego vehicle과 opponent vehicle의 feature집합으로 이루어져 있으며, 차량 간의 관계성을 나타내는 edge feature는 차량의 feature와 차량 간의 상대적 위치, heading 정보를 이용하여 계산된다. 각 노드들의 output은 가우시안 분포(gaussian distribution) 파라미터, 즉 평균과 표준편차의 형태로 계산되며 이러한 파라미터들을 기반으로 latent space가 형성된다. 디코더 네트워크는 한 스텝의 future trajectory를 순차적으로 생성한다. 먼저 latent variable이 prior 또는 posterior latent space로부터 샘플링되며 매 t 타임스텝마다 GNN은 t-1, 즉, 한 스텝 과거의 trajectory feature와 map feature를 기반으로 가속도와 각가속도를 생성한다. 생성된 가속도와 각가속도는 bicycle model을 통해 t 스텝의 상태로 계산되며, 계산된 상태를 기반으로 past feature와 map feature가 업데이트된다. 디코더 네트워크는 이러한 방식을 반복하며 trajectory를 순차적으로 생성한다.

모델 학습 과정에서 네트워크의 파라미터들은 loss function을 기반으로 ADAM optimizer를 이용하여 업데이트된다.

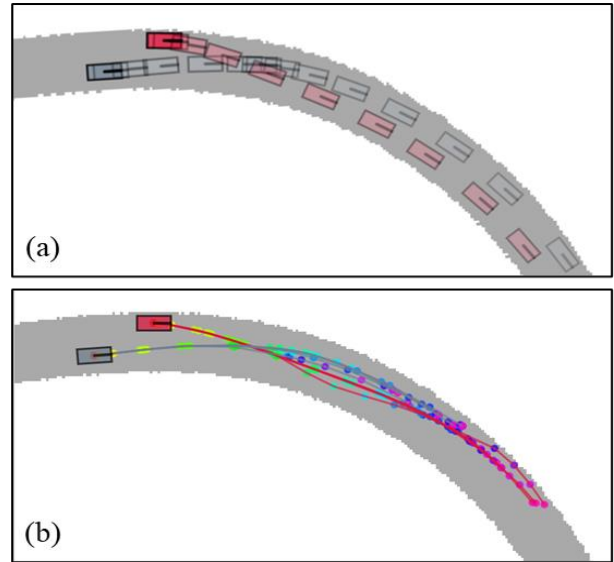


그림 4. 생성형AI 기반 고위험 시나리오 생성 결과.

Fig. 4. Dynamic scenario generated based on GA and high-fidelity simulation (a). Generated safety-critical scenario by adversarially varying dynamic scenario with generative AI (b).

CVAE의 loss function은 세개의 term으로 구성되어 있다. 먼저 reconstruction loss는 posterior 네트워크로부터 생성된 상태가 ground truth에 가까울수록 작은 값을 가지며 posterior 네트워크가 잘 학습되도록 한다. KL (Kullback-Leibler) loss는 prior와 posterior latent space 사이의 KL divergence를 통해 계산되며, prior 네트워크가 posterior네트워크와 유사하게 학습할 수 있도록 한다. Collision loss의 경우 각 차량 사이의 충돌 정도와 agent들과 non-drivable area와의 충돌 정도를 계산하여 최대한 충돌이 발생하지 않도록 하여 타당한 trajectory를 생성할 수 있다.

3. 생성형 AI 기반 고위험 시나리오 생성

생성형 AI 기반 고위험 시나리오 생성 단계에서는 기존에 학습한 CVAE 네트워크를 기반으로 적대적 최적화 과정을 통해 고위험 시나리오를 생성하고 최종적으로 생성된 시나리오를 시뮬레이터를 통해 평가한다. 해당 과정에서는 주행 데이터셋과 이전 단계에서 학습한 CVAE 네트워크를 기반으로 future trajectory를 생성하고, objective function을 통해 latent variable을 최적화한다. 그 과정에서 생성되는 latent variable과 trajectory를 기반으로 loss function이 계산되며, 이를 통해 realistic하고 ego vehicle에 위협적인 trajectory를 생성할 수 있다. 또한, ego vehicle의 경우 latent space로부터 생성한 trajectory를 ground truth trajectory와 매칭하는 과정을 통해 실제 trajectory와 유사한 trajectory를 생성하도록 최적화된다. 이러한 과정을 통해 시뮬레이션 기반 주행 데이터 셋으로부터 생성형 AI를 통해 다양한 고위험 시나리오를 생성할 수 있다. 이 때 생성되는 고위험 시나리오는 각 차량의 시계열 상태 정보로 구성되어 있는 동적 시나리오이며 해당 정보를 기반으로 opponent vehicle이 trajectory를 추종하도록 제어하여 high-fidelity 시뮬레이터 상에 재현할 수 있다. 이 때, 생성된 ego vehicle의 trajectory는 AI를 통해 생성되어 물리적 환경을 잘 반영하지 못할 수 있다. 따라서 ego vehicle의 경우 고위험

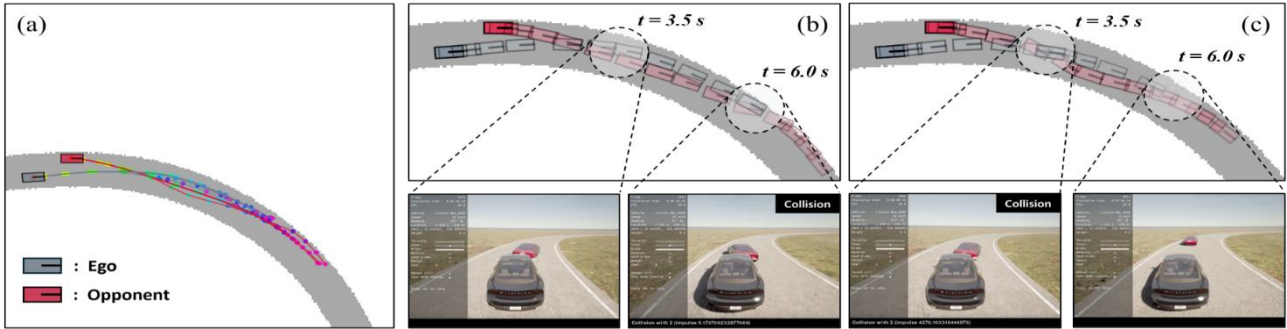


그림 5. 생성형 AI 기반으로 생성된 다양한 고위험 시나리오와 이를 반영한 CARLA 기반 자율주행 시스템 평가.

Fig. 5. Various safety-critical scenarios generated based on generative AI using a single dynamic scenario (a). ADS safety evaluation in CARLA simulation based on one (b) and another (c) example case of safety-critical scenarios from (a).

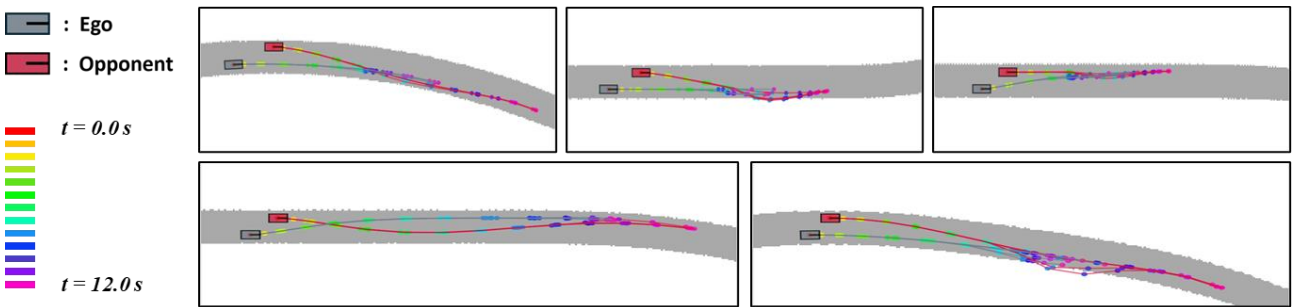


그림 6. 생성형 AI 기반 적대적 최적화를 통한 다양한 고위험 시나리오 생성 결과.

Fig. 6. Various safety-critical scenarios generated based on generative AI and adversarial optimization.

시나리오의 기존 동적 시나리오가 생성되었던 상세 시나리오의 정보를 적용하여 실제로 자율주행 시스템이 고위험 시나리오와 동일하게 충돌하는지, 또는 자율주행 시스템이 대응하여 충돌을 방지하는지 판단하여 자율주행 시스템을 평가한다.

최종적으로 high-fidelity 시뮬레이터를 통해 생성된 고위험 시나리오를 기반으로 자율주행 시스템의 안전성을 평가할 수 있다.

III. 실험 및 결과

본 논문에서는 보다 더 현실적인 주행 데이터 셋을 수집하고자 CARLA 시뮬레이터[19]를 사용하였다. CARLA 시뮬레이터는 Unreal Engine 기반의 high-fidelity 시뮬레이터로 최근 자율주행 연구에 활발히 사용되고 있다[20]. CARLA 시뮬레이션 상에 그림 3과 같이 레이싱 환경을 구축하였으며 Ontology 기반의 시나리오 생성을 위해 OpenStreetMap [21] 기반의 map을 구축하였다. 또한, CARLA 시뮬레이션 상에 시나리오를 구현하고 데이터를 수집하기 위해 ROS2 (Robot Operating System 2)를 기반으로 통신 환경을 구축하였다. 자율주행 시스템은 CARLA-ROS Bridge [22]에서 제공하는 CARLA AD Agent를 기반으로 구현되었다. 구축한 CARLA 레이싱 환경에서 map의 lane 정보를 기반으로 논리 시나리오를 구성하였으며 주요 파라미터는 표 2와 같다. 차량 간 거리 정보의 최소값과 최대값, $\{\theta_{d_{eo},min} \theta_{d_{eo},max}\}$ 을 Set 1: {10m, 15m}, Set 2: {15m, 20m}, Set 3: {20m, 25m}, Set 4: {25m, 30m}로 설정하여 총 4종류의 논리 시나리오를 기반으로 알고리즘 1

의 GA 알고리즘을 적용하여 CARLA 시뮬레이터 상에 각 논리 시나리오 별 100개의 레이싱 시나리오를 구현하였다. 이를 바탕으로 총 400개의 레이싱 주행 시나리오를 수집하여 주행 데이터 셋을 구축하였다. CARLA AD Agent와 opponent vehicle의 주행 특성을 학습하고자 구축한 데이터셋을 기반으로 CVAE 네트워크를 학습하였다. 학습된 CVAE를 기반으로 주행 시나리오에 대해 적대적 최적화를 진행하였으며, 그림 4 (a)와 같은 하나의 주행 시나리오에서 적대적 최적화를 진행하여 그림 4 (b)와 같이 다양한 고위험 시나리오를 생성하였다. 생성된 고위험 시나리오는 그림 5와 같이 CARLA 시뮬레이션을 통해 CARLA AD Agent를 기반으로 opponent vehicle의 동적 시나리오를 재현하여 해당 시나리오를 평가하였다. 그림 5 (b)와 (c)는 같은 기존 주행 데이터로부터 적대적 최적화되었으나 서로 다른 경로로 진행이 되며, 충돌하는 시점 또한 다르다. 이러한 시나리오는 실제로 CARLA를 기반으로 재현해 보았을 때, 고 위험 시나리오가 생성된 대로 각각 6초, 3.5초에 충돌이 발생하며 적대적 최적화를 통해 생성된 대로 시나리오가 재현이 되는 것을 확인할 수 있었다. 이러한 시나리오 외에 총 400개의 주행 시나리오를 바탕으로 그림 6과 같이 다양한 고위험 시나리오를 생성하고 CARLA 시뮬레이션을 통해 평가하였다.

시나리오 생성 결과를 분석하기 위해 그림 7과 같이 시나리오 유형을 분류하였다. 일반적으로 고위험 시나리오를 생성 모델의 성능을 판단할 때, risk metric으로 충돌시나리오의 개수 또는 비율, 다른 차량과의 거리, 도로를 벗어난 정도 등의 metric을 기준으로 성능을 판단한다. 본 논문에서는 그림



그림 7. 충돌 유형 별 고위험 시나리오 유형 분류.

Fig. 7. Safety-critical scenario classification based on the collision cases.

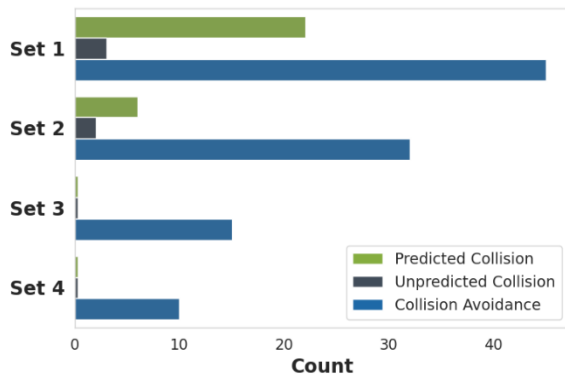


그림 8. 생성된 시나리오에 대한 평가 결과 통계.

Fig. 8. Collision statistics of the simulation based on generated safety-critical scenarios.

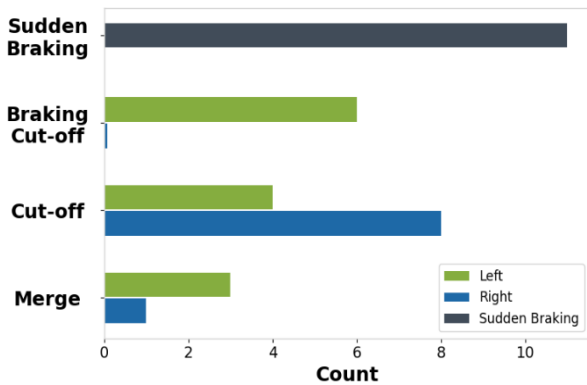


그림 9. 시나리오 유형에 따른 평가 결과 통계.

Fig. 9. Collision statistics according to the scenario class.

8, 그림 9와 같이 충돌 시나리오의 개수를 기준으로 결과를 분석하였다. 하지만 CVAE 네트워크를 학습할 때 ego vehicle의 ground truth 정보를 기반으로 학습되므로 시뮬레이션 결과가 CVAE 기반 적대적 최적화를 통해 생성된 고위험 시나리오와 상이할 수 있다. 따라서, 생성된 고위험 시나리오에서 생성한 시나리오대로 충돌이 발생하는 경우(predicted collision), 생성한 시나리오가 아니지만 충돌이 발생하는 경우(unpredicted collision), CARLA AD Agent가 위험한 시나리오에 대응하여 충돌을 회피한 경우, collision avoidance로 분류하였다. 또한, 시뮬레이션 특성 상 노이즈가 발생할 수 있으며, 여러 번 시뮬레이션 시행 시에 다른 결과로 연결될 수 있어 한 개의 고위험 시나리오에 대해 5번의 시뮬레이션을 시행하였다. 4가지 논리 시나리오를 기반으로 생성한 동적 시나리오 400 개를 통해 생성한 고위험 시나리오를 Set 1~Set 4으로 분류하여 그 결과를 그림 8과 같이 분석하였다. 차량간 거리를 가장 가깝게 설정한 논리 시나리오 Set 1에서 충돌이 발생하는 시나리오가

가장 많았으며 차량간 거리를 가장 멀게 설정한 Set 4에서 충돌이 발생하는 시나리오가 가장 적었음을 확인할 수 있다. 또한, Set 3, Set 4에서는 충돌이 발생하는 시나리오가 거의 없고 대응하는 시나리오만 생성이 된 것으로 보아 거리의 여유가 충분할 경우 위험한 상황이 발생하더라도 대응할 수 있음을 확인할 수 있다.

또한, 그림 9의 경우 시나리오 평가 결과를 유형 별로 분류한 결과이다. 급정지 시나리오가 가장 많이 발생하였으며, 우측에서 끼어들기 시나리오가 다음으로 많이 발생하였다. 전체적으로 좌우의 결과가 상이한 것은 도로의 형상과 관련이 있을 수 있으나 결과 데이터의 양이 많지 않아 더 많은 데이터를 생성하고 추가적으로 분석할 필요성이 있다.

IV. 결론

본 논문에서 제안하는 자율주행 레이싱 시스템 안정성 평가 프레임워크는 ontology 적인 domain knowledge와 GA기반 최적화, 생성형 AI를 통해 고위험 시나리오를 생성하고 물리적 환경을 잘 반영할 수 있는 high-fidelity 시뮬레이션을 통해 자율주행 시스템을 효과적으로 평가함으로써 효율적이고 신뢰도 있는 안전성 평가 프로세스가 가능하다.

그리고 제안한 프레임워크는 임의로 설계한 환경에서 고위험 시나리오 생성이 가능하여 실제 데이터셋에 의존하지 않고 새로운 가상의 환경을 직접 만들어서 다양한 환경에서의 자율주행 시스템 안전성 평가가 가능하다. 본 논문에서는 주행 데이터셋이 많이 존재하지 않는 레이싱 환경을 시뮬레이션 상에 구현함으로써 새로운 주행 데이터셋을 구축하여 이를 기반으로 새로운 가상의 환경에서 고위험 시나리오를 생성하고 고위험 시나리오를 통해 자율주행 시스템을 평가하였다. 새로운 환경에서 고위험 시나리오를 생성함으로써 본 프레임워크의 시나리오 제이가능성을 입증하였으며 추후에 본 프레임워크의 충실도와 효율성을 증명하기 위해 data-driven 및 adversarial generation 방식을 통합한 시나리오 생성 프레임워크를 활용할 계획이다. Data-driven generation 방식의 경우 실제 데이터셋에 기반하는 반면 본 프레임워크는 데이터셋을 직접 생성하여 시나리오를 생성하여 사실 상 직접적인 비교는 불가능하다. 하지만 추후 실제 데이터셋이 존재하는 환경과 유사한 환경을 구축하여 해당 환경에서 제안한 프레임워크를 기반으로 시나리오를 생성하고 data-driven 및 adversarial generation 방식을 통합한 프레임워크를 통해 실제 데이터셋을 기반으로 시나리오를 생성할 예정이며, 각 시나리오 생성 과정 및 결과를 일정 시간 동안 탐색한 시나리오 개수, 종류 등 정량적 지표를 통해 비교하여 제안한 프레임워크의 충실도 및 효율성을 실험적으로 증명할 계획이다.

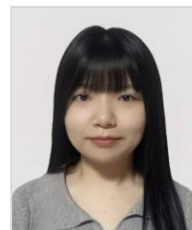
추후 연구로는 실제 환경을 보다 잘 반영한 고위험 시나리오를 생성하기 위해 실제 데이터를 활용한 시나리오 생성 프레임워크와 본 프레임워크의 간극을 줄이는 방향으로 연구를 진행할 예정이다. 이를 통해 제한한 프레임워크는 knowledge-based generation, data-driven generation, 그리고 adversarial generation 방법론들의 장점들을 고루 통합할 수 있을 것이라 기대한다.

REFERENCES

- [1] Indy Autonomous Challenge, <https://www.indyautonomouschallenge.com/>.
- [2] Abu Dhabi Autonomous Racing League (A2RL), <https://a2rl.io/>.
- [3] W. Ding, C. Xu, M. Arief, H. Lin, B. Li, and D. Zhao, "A survey on safety-critical driving scenario generation—A methodological perspective," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 6971–6988, Mar. 2023.
- [4] R. Van der Made, M. Tideman, U. Lages, R. Katz, and M. Spencer, "Automated generation of virtual driving scenarios from test drive data," *Proc. of the 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, National Highway Traffic Safety Administration, paper no. 15-0268, Gothenburg, Sweden, Jun. 2015.
- [5] W. Ding, W. Wang, and D. Zhao, "A new multi-vehicle trajectory generator to simulate vehicle-to-vehicle encounters," arXiv preprint arXiv:1809.05680, Sep. 2018.
- [6] W. Wang, C. Liu, and D. Zhao, "How much data are enough? A statistical approach with case study on longitudinal driving behavior," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 2, pp. 85–98, Jun. 2017.
- [7] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *Proc. of the International Conference on Learning Representations (ICLR)*, Banff, Canada, Dec. 2013.
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672–2680, 2014.
- [9] S. Suo, S. Regalado, S. Casas, and R. Urtasun, "TrafficSim: Learning to simulate realistic multi-agent behaviors," *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, pp. 10400–10409, Jun. 2021.
- [10] B. Yang, W. Luo, and R. Urtasun, "Pixor: Real-time 3D object detection from point clouds," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, pp. 7652–7660, Jun. 2018.
- [11] K. Sohn, H. Lee, and X. Yan, "Learning structured output representation using deep conditional generative models," *Advances in Neural Information Processing Systems*, vol. 28, pp. 3483–3491, 2015.
- [12] W. Ding, B. Chen, M. Xu, and D. Zhao, "Learning to collide: An adaptive safety-critical scenarios generating method," *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, pp. 2243–2250, Oct. 2020.
- [13] G. Bagnschik, T. Menzel, and M. Maurer, "Ontology based scene creation for the development of automated vehicles," *Proc. of the IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, pp. 1813–1820, Jun. 2018.
- [14] D. Rempe, J. Phillion, L. J. Guibas, S. Fidler, and O. Litany, "Generating useful accident-prone driving scenarios via a learned traffic prior," *Proc. of the IEEE/CVF Conference on Computer*

Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, pp. 17305–17315, Jun. 2022.

- [15] Y. Xie, X. Guo, C. Wang, K. Liu, and L. Chen, "AdvDiffuser: Generating adversarial safety-critical driving scenarios via guided diffusion," *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, UAE, pp. 9983–9989, Oct. 2024.
- [16] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," arXiv preprint arXiv:2010.02502, Oct. 2020.
- [17] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuScenes: A multimodal dataset for autonomous driving," *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 11621–11631, Jun. 2020.
- [18] M. Jeon, K. Cho, and H. Lee, "Driving scenario generation based on latent diffusion model for autonomous driving vehicles," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 31, no. 1, pp. 14–19, Jan. 2025.
- [19] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," *Proc. of the Conference on Robot Learning (CoRL)*, Mountain View, CA, USA, pp. 1–16, Oct. 2017.
- [20] H. Kim, H. Ra, A. Kim, and S. Kim, "A study on construction of transfer learning-based dataset for object detection using CARLA," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 30, no.2, pp. 175–182, Feb. 2024.
- [21] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," *IEEE Pervasive Computing*, vol. 7, no. 4, pp. 12–18, Dec. 2008.
- [22] CARLA ROS Bridge Documentation, <https://carla.readthedocs.io/projects/ros-bridge/en/latest/>.



이 현 빈

2022년 울산과학기술원 기계항공공학(학사). 2022년~현재 울산과학기술원 기계공학과 석박사통합과정 재학 중. 관심분야는 인공지능, 자율주행, 시스템 안전성 검증.



이 상 현

2023년 울산과학기술원 기계공학(학사). 2023년~현재 울산과학기술원 기계공학과 석박사통합과정 재학 중. 관심분야는 인공지능, 자율주행, 시스템 안전성 검증.



양 형 준

2025년 울산과학기술원 기계공학(학사). 2025년~현재 울산과학기술원 기계공학과 석박사통합과정 재학 중. 관심분야는 인공지능, 자율주행, 시스템 안전성 검증.



권철현

2010년 서울대학교 기계항공공학(학사).
2013년 Purdue University 항공우주공학
(공학석사). 2017년 Purdue University 항
공우주공학(공학박사). 2019년~현재 울
산과학기술원 기계공학과 교수. 관심분
야는 고신뢰 사이버 물리 시스템 제어.