Transactions of KSAE, Vol. 32, No. 3, pp.309-317 (March, 2024)

Check for updates

Copyright © 2024 KSAE / 220-08
pISSN 1225-6382 / eISSN 2234-0149
DOI http://dx.doi.org/10.7467/KSAE.2024.32.3.309

시나리오 데이터베이스와 생성형 모델을 이용한 상세 시나리오의 파라미터 선정과 선행차량 추종 및 끼어들기 시나리오 생성으로의 적용

정 선 암[1)] · 이 세 민[2)] · 정 영 훈[1)] · 김 승 환[1)] · 송 봉 섭[*2)]

아주대학교 D.N.A 플러스융합학과[1)] · 아주대학교 기계공학과[2)]

Parameter Selection of Safety-Critical Scenarios Based on Scenario Database and Generative Models and Its Application to Generation of Car-Following and Cut-in Scenarios

Seonarm Jeong[1)] · Semin Lee[2)] · Younghun Jeong[1)] · Seunghwan Kim[1)] · Bongsob Song[*2)]

[1)]Department of Data, Networks, and AI, Ajou University, Gyeonggi 16499, Korea
[2)]Department of Mechanical Engineering, Ajou University, Gyeonggi 16499, Korea
(Received 18 December 2023 / Revised 10 January 2024 / Accepted 11 January 2024)

Abstract : The selection method of concrete scenarios based on scenario database and generative models is proposed in this paper. Two sampling methods are used in selecting appropriate parameters along a performance measure. First, the parameter space is extracted from scenario DB, and a set of parameters are selected via random sampling. Once the set of concrete scenarios are simulated, their distribution is analyzed with respect to a specific measure. The second method is based on parameter generative models. Simulated scenarios are used to train a surrogate model, which is a multi-layer perceptron model. Then, a generative model is designed to search for the desired parameter based on the surrogate model. Thus, the second one can be used to compensate for the imbalance in randomly sampled concrete scenarios. Finally, the proposed selection method is more efficient than random sampling from the viewpoint of the distribution of two different measures.

Key words : Concrete scenario(상세 시나리오), Parameter space(파라미터 공간), Surrogate model(대리 모델), Gradient descent(경사 하강법), Generative model(생성형 모델), Multi-layer perceptron(다층 퍼셉트론), Autonomous driving scenario(자율주행 시나리오)

1. 서 론

전세계적으로 로봇택시와 같은 자율주행 서비스를 선보이면서 혁신적인 기술의 상용화 가능성을 지속적으로 보여주고 있다. 하지만 실제 자율주행모드에서 교통 사고가 발생하면서 자율주행의 안전성을 검증하기 위한 체계적 방법에 대한 관심이 동시에 크게 증가하고 있다. 일반적으로 통계적 또는 실 도로 주행 테스트(Field Operation Test, FOT) 기반의 검증 방법은 천문학적인 시간과 데이터가 필요하다고 알려져 있다. 이러한 한계를 보완하기 위해 시나리오 기반(Scenario-based) 검증 방법론이 제시되고 있다.[1)] 더 나아가 표준화된 시나리오 기반 검증방법이 논의되고 있으며 관련 툴체인(Toolchain)간의 호환성을 제고하기 위해 ASAM OpenScenario와 OpenDrive 같이 표준 시나리오 및 도로환경의 정의가 제안되고 있다.[2)]

이러한 표준화 노력에도 불구하고 시나리오를 구성하는 파라미터 공간의 차원(Dimension) 수나 파라미터 영역(Range) 내에서의 파라미터를 샘플링하는 방법에 따라 상세 시나리오의 개수가 기하급수적으로 증가(Combinatorial explosion)하는 문제가 존재한다.[3,4)] 예를 들어, Fig. 1(a)에서 보는 바와 같이 선행차량 추종(Car-following) 시나리오에서 파라미터 공간을 자율차량(HAV)과 선행차량

*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.

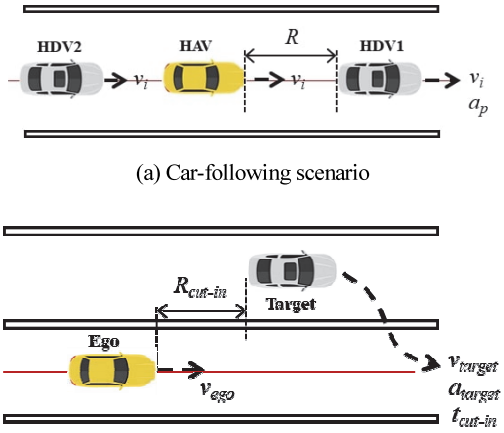
(HDV1)과의 상대거리(R), 선행차량의 감속도(a_p), 세 차량의 시작 속도(v_i)를 시나리오 파라미터로 고려한다면 상세 시나리오 개수는 다음과 같이 표현할 수 있다.[3)]

$$S^N = S^3 = R \times a_p \times v_i \tag{1}$$

여기서 S는 각 시나리오 파라미터 구간 내 샘플링개수를 뜻하며, N은 파라미터 공간의 차원 수를 뜻한다. 즉, Fig. 2에서 보는 바와 같이 각 시나리오 파라미터에 대해서 40개씩 등간격의 N-wise 샘플링을 한다면 40^3 또는 64,000개의 상세 시나리오가 생성된다.[5,6)]

또한 Fig. 1(b)와 같이 끼어들기(Cut-in) 시나리오의 경우, Euro NCAP에서 제안하는 파라미터 공간인 자 차량(Ego)의 속도(v_{ego})와 끼어들기 차량(Target)의 속도(v_{target}), 끼어들기 수행시간($t_{cut\text{-}in}$), 끼어들기 동안의 가속도(a_{target}), 그리고 끼어들기 시점 두 차량 간 상대거리($R_{cut\text{-}in}$)를 정의한다면 시나리오 개수는 다음과 같다.[7,8)]

$$S^N = S^5 = v_{ego} \times v_{target} \times t_{cut\text{-}in} \times a_{target} \times R_{cut\text{-}in} \tag{2}$$



(a) Car-following scenario

(b) Cut-in scenario

Fig. 1 Definition of scenario and parameter space

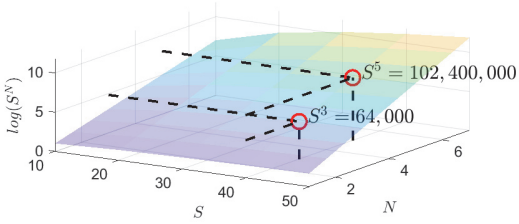


Fig. 2 The number of concrete scenarios with respect to dimension of parameter space(N) and the number of sampling(S)

Fig. 2에서 보는 바와 같이 각 시나리오 파라미터에 대해서 40개씩 등간격으로 샘플링을 한다면 40^5 또는 1.024×10^8개의 상세 시나리오가 생성된다. 위 두 시나리오에서 살펴본 바와 같이 시나리오 파라미터의 차원 수와 샘플링하는 개수에 따라 기하급수적으로 상세 시나리오의 개수가 증가하는 것을 확인할 수 있다.

상세 시나리오 개수가 급격히 증가하는 문제(Explosion of concrete scenario)를 해결하기 위해 상세 시나리오를 선정(Selection of concrete scenario)하는 방법에 대한 연구가 활발히 진행되고 있다.[6,9)] 무작위로 샘플링 하는 랜덤 샘플링이 일반적이며, 이의 한계성을 다음 문제정의에서 논의하고자 한다. 다음으로 상세 시나리오의 파라미터의 모든 조합을 이용하는 대신 파라미터의 조합(2개 혹은 3개의 묶음)을 이용해서 상세 시나리오의 개수를 줄이는 Pair-wise 기법이 제안되고 있다.[10)] 이는 단순히 시나리오의 개수만 줄일 수 있을 뿐 원하는 시나리오나 중요도가 높은 시나리오가 누락될 수 있다. 또한 시나리오의 파라미터 Sweeping 기법을 통하여 판단 알고리즘이 취약한 평가 시나리오를 찾을 수 있는 방법이 제안되었지만 식 (2)와 같은 고차원 시나리오에 대해서는 다루고 있지 않다.[11)] 최근에는 기계학습이나 네트워크를 이용하는 방법이 제안되기도 한다. 그 예로 기계학습 기반의 분류 기법인 Support Vector Machine(SVM)을 사용하여 자동긴급제동(Automatic emergency brake)시스템을 테스트할 수 있는 추돌(Frontal collision) 시나리오를 선정하는 방법이 제안되었다.[4)] 또한 차량 추종 시나리오에 대한 위험도 성능 지표를 예측할 수 있는 Multi-Layer Perceptron(MLP) 기반 대리모델을 통해 위험도가 낮은 시나리오에서 높은 시나리오로 변하는 경계 값을 설정하여 운영설계영역(Operational design domain)을 탐색하는 연구를 찾아볼 수 있다.[3)]

더 나아가, 선정된 시나리오가 실제 도로에서 발생하는 시나리오인지 아닌지를 판단하는 정합성(Fidelity)을 고려해야 한다.[9)] 이를 위해 상세 시나리오의 파라미터 영역(Range)을 설정하기 위해 데이터 기반(Data-driven) 접근 방법이 연구되고 있다. 일본의 SAKURA 프로젝트에서는 고속도로에서 취득한 운전자의 주행데이터와 카메라를 통해 취득한 데이터를 활용하여 감속 및 끼어들기 시나리오의 상태변수에 대한 예상 범위를 선정하고 있다.[12)] 더 나아가 선형 회귀 기법을 통해 각 변수 간의 상관관계를 분석(Correlational analysis)하여 각 변수들이 갖는 값들의 조합에 따라 생성된 상세 시나리오가 실 도로에서 발생할 빈도 및 확률을 계산하였다. 또한 실 도로 주행 데이터로부터 취득한 파라미터들의 확률분포를 기반으로 한 몬테카를로 기법이 널리 사용되고 있다. 하지

만 실 도로환경에서는 위험한 상황이 거의 발생하지 않기 때문에 중요도 샘플링(Importance sampling)을 통해 생성한 위험한 확률 분포를 접목한 몬테카를로 기법으로 위험한 상세 시나리오를 생성한다.[13-15] 하지만, 이러한 접근방법은 대량의 실 도로 주행 데이터의 취득과 후처리 작업이 선행되어야 한다.

시나리오 파라미터의 선정뿐만 아니라 좀 더 다양하고 복잡한 교통환경을 고려하기 위해서 유럽 PEGASUS 프로젝트에서는 다양한 자료로부터(Knowledge-based) 생성된 테스트 시나리오를 저장하고 연관된 테스트 시나리오를 추출할 수 있는 데이터 베이스 기반 검증을 제안하고 있다.[16] 또한 시나리오 기반 검증 관련 문헌조사에 따르면 시나리오 데이터베이스(DB)를 기반으로 다양한 상세 시나리오를 선정하는 연구가 활발하게 진행되고 있다.[6] 예를 들어 ASAM OpenScenario에서 제공하는 기본(Logical) 시나리오인 끼어들기 시나리오 파일을 생성한 다음, 데이터베이스로부터 추출한 파라미터 값을 이용하여 상세 시나리오를 생성하였다.[17] 또한 생성된 상세 시나리오의 파라미터값이 실제 차량에서 나올 수 있는 궤적과 일치하는 것을 입증함으로써 데이터 베이스 기반으로 상세 시나리오를 선정하는 것이 중요함을 나타내고 있다.

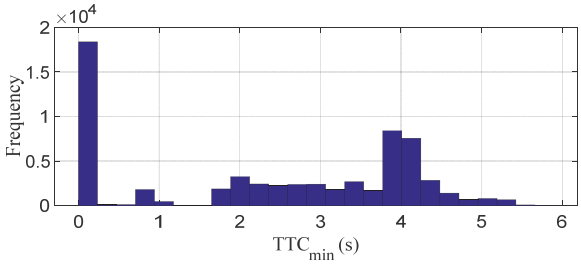
본 논문에서는 시나리오 데이터베이스를 제안하고 이를 기반으로 상세 시나리오를 선정하는 방법을 제안한다. 좀 더 구체적으로 우선 실 도로 주행 데이터(FOT)기반으로 설정한 파라미터 공간의 적절성을 확인한다. 다음으로는 주어진 파라미터 공간 내에서 랜덤 샘플링을 한 후 선정된 성능 지표에 따라 분포도를 분석한다. 이후 분포도 관점에서 추가해야 하는 상세 시나리오를 파라미터 생성형(Generative) 모델을 기반으로 상세 시나리오를 탐색한다. 이를 통하여 랜덤으로 선정한 시나리오와 생성형 모델을 기반으로 추가된 시나리오의 분포도를 비교하여 시나리오 증강(Scenario augmentation)의 효과를 보이고자 한다.

2. 문제 정의

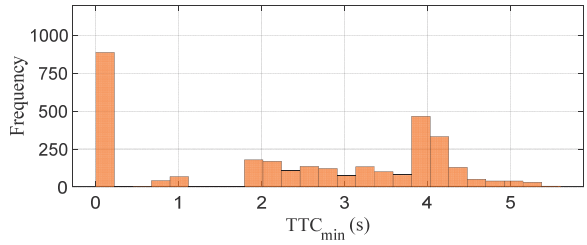
시나리오를 정의하는 파라미터의 차원이 높아질수록 상세 시나리오가 기하급수적으로 증가하는 문제를 Fig. 2에서 도식적으로 보여주고 있다. 따라서 모든 상세 시나리오를 구현하기보다는 상세 시나리오를 선정하는 과정이 필요하다. 가장 일반적인 방법으로 랜덤 샘플링이 있을 수 있다. 예를 들어 Table 1의 차량 추종 시나리오의 각 파라미터(R, a_p, v_i)인 최솟값, 최댓값 사이를 동일한 간격으로 40개씩 샘플링한 결과인 64,000개의 상세 시나리오를 생성 후, 그 중에서 3,200개의 상세 시나리오를 무작위로 선택하고 기존 64,000개의 시나리오와 샘플링된 시나리오를 비교하고자 한다. 비교를 위해 성능 지표가 필요하며 가장 보편적인 위험도 성능 지표는 Time-To-Collision(TTC)의 최솟값(TTC_{min})을 기준으로 Fig. 3에서 각각 보여주고 있다. 상세 시나리오의 분포가 형상적으로 비슷하며 두 시나리오셋 모두 TTC_{min}가 0~2초 사이에 상대적으로 상세 시나리오가 부족함을 보여주고 있다. 즉, 파라미터 공간에서 균일한 간격으로 시나리오 파라

Table 1 An example of parameter space for car-following and cut-in scenario

Param.	Description	Unit	Min	Max
R	Relative distance or range	m	25	64
a_p	Acceleration of preceding vehicle	g	-0.74	-0.35
v_i	Initial velocity of the i-th car	m/s	15	34.5
v_{ego}	Initial velocity of ego vehicle	km/h	30	110
v_{target}	Initial velocity of target vehicle	km/h	30	110
$t_{cut\text{-}in}$	Duration of cut-in	s	1	5
a_{target}	Acceleration of target vehicle during cut-in	m/s^2	-8	0
$R_{cut\text{-}in}$	Trigger distance of cut-in	m	2	50

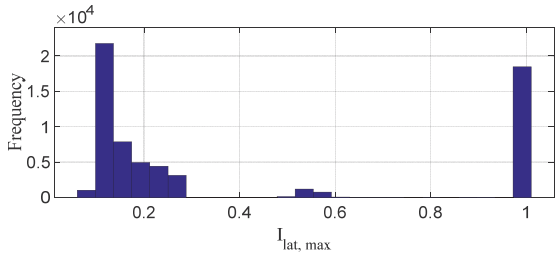


(a) N-wise sampling

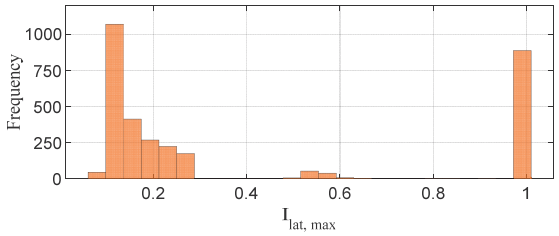


(b) Random sampling

Fig. 3 Distribution comparison of car-following scenarios with respect to minimum TTC



(a) N-wise sampling



(b) Random sampling

Fig. 4 Distribution comparison of car-following scenarios with respect to maximum lateral collision index

미터를 선정하였음에도 불구하고, 특정 지표의 분포가 불균형한 것을 볼 수 있다.

동일한 시나리오에 대하여 다른 위험도 성능 지표를 선택하여 분포도를 비교하고자 한다. 위험도 판단 방법 중 하나인 충돌 위험 인덱스(Lateral collision index)를 선정하였으며, 한 개의 시계열 시나리오에 대하여 최댓값($I_{lat,\ max}$)을 위험 지표로 선정하였다.[18] Fig. 4에도 64,000개의 상세 시나리오와 무작위로 선정된 3,200개의 상세 시나리오가 비슷한 경향을 보여주고 있으며, $I_{lat,\ max}$ 관점에서 약 0.3~0.5, 0.7~0.9 사이에 상대적으로 부족함을 보이고 분포의 불균형을 이루고 있음을 보여주고 있다.

3. 상세 시나리오 선정 방법

상세 시나리오가 파라미터 공간 차원이 높아질수록 기하급수적으로 증가하고 주어진 파라미터 공간 내에서의 일정한 간격의 샘플링(N-wise sampling)이나 랜덤 샘플링의 경우 모두 분포의 불균형 현상을 있을 수 있음을 Fig. 3과 Fig. 4에서 보여주고 있다. 문제를 보완하기 위하여 상세 시나리오를 선정하는 과정을 제안하고자 한다. 이러한 과정은 시나리오를 생성하는 프로세스에 기반하여 데이터베이스와 생성형 모델을 이용하여 최종적인 상세 시나리오를 선정하는 과정으로 구성되어 있다.

3.1 위험 시나리오 생성 프로세스

위험(Safety-critical) 시나리오 생성 프로세스는 크게 여섯 단계로 구분할 수 있다.[6] 첫 번째 단계는 Fig. 5의 시나리오 생성(Scenario generation)을 위한 전 단계로 시나리오를 정의하고, 교통사고 데이터나 주행데이터를 수집하고 분석하는 과정을 포함하고 있다.[4,19] 두 번째 단계인 시나리오 생성 단계는 선정된 기본(Logical) 시나리오를 시뮬레이션 환경에서 구현하기 위한 다양한 형태의 파일로 변환한다. 좀 더 구체적으로 본 연구에서는 상세 시나리오 시뮬레이션을 위해서 IPG CarMaker를 선정하였으며,[20] 시뮬레이션이 가능한 기본 시나리오 파일을 생성하였다. 추가적으로 파라미터 공간을 뜻하는 시나리오 파라미터 파일은 CSV 형식으로 생성하였다. 또한 주행 도로 환경이 파일로 생성되며, 주행 데이터는 시계열 형태의 파일로 데이터 저장소(Data repository)에 보관된다. 세 번째 단계인 시나리오 데이터베이스는 데이터 저장소 내의 다양한 형태의 파일을 탐색하고 추출하기 위하여 비관계형(NoSQL) 데이터베이스를 적용하였다. 즉, 데이터 저장소 파일정보를 JSON 형식으로 다큐먼트(Document)를 생성하고 Mongo DB기반으로 시나리오 데이터베이스를 구성하였다. Fig. 5의 상세 시나리오 선정

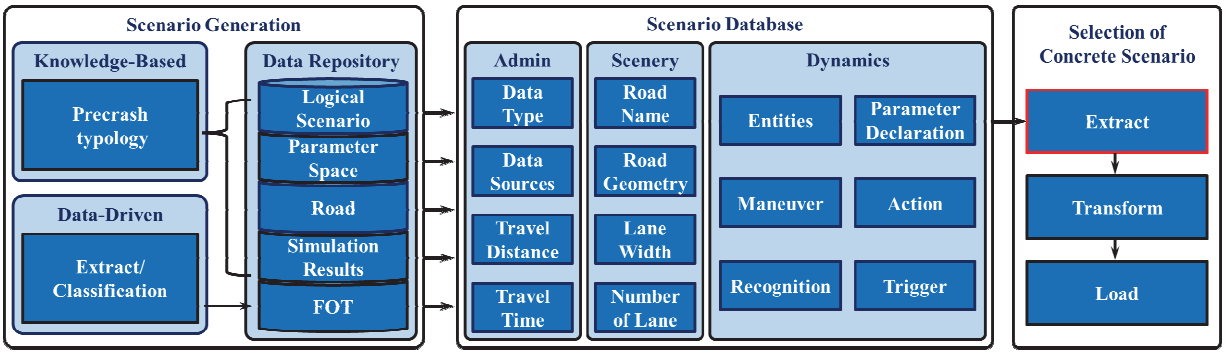


Fig. 5 Generation and selection process of safety-critical scenarios

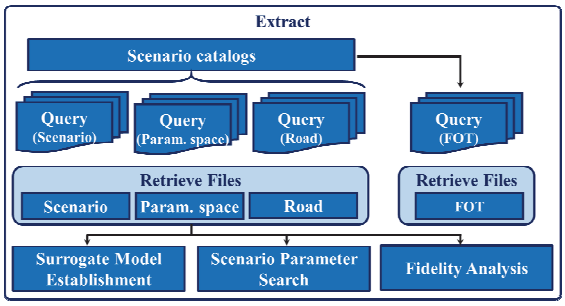


Fig. 6 Schematics of Extract process in selection of concrete scenarios

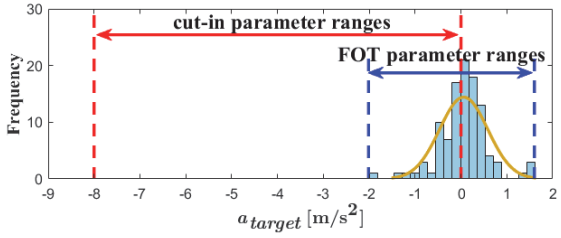
(Selection of concrete scenario)은 시나리오 생성 프로세서의 네 번째 단계에 해당이 된다. 상세 시나리오 선정 과정은 다시 세 단계로 구분된다. Fig. 6에서 볼 수 있듯이, 우선 추출(Extract) 단계에서는 원하는 기본(Logical) 시나리오를 쿼리(Query)를 통하여 추출하고, 시나리오 파라미터를 샘플링하는 과정, 주행 도로 파일을 쿼리를 통해서 탐색하는 과정, 마지막으로 실 도로 주행데이터를 기반으로 파라미터 영역(Range)의 적절성(Fidelity)을 분석하는 과정을 모두 포함한다. 본 논문에서는 시나리오 파라미터를 샘플링하는 과정을 집중적으로 설명하고 있다.

다음으로 변환(Transform) 단계에서는 추출된 상세 시나리오에 대하여 선정된 도로(Road) 정보와 실 도로 주행 데이터 정보를 모두 종합적으로 시뮬레이션을 할 수 있는 파일로 변환하는 단계이다. 마지막으로 탑재(Load) 단계에서는 시뮬레이션의 환경에 따라 시뮬레이션 자동화 또는 테스트 자동화(Test automation)를 하기 위한 파일을 생성하고 시뮬레이터에 탑재할 수 있는 결과물을 생성하는 단계이다. 이를 통해서 다섯 번째 단계인 시뮬레이션 수행과정과 여섯 번째 단계인 평가과정으로 진행이 된다.

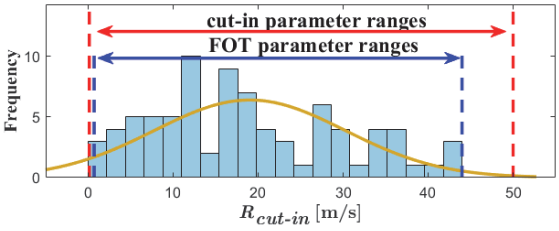
3.2 시나리오 데이터베이스

시나리오의 구성요소는 Fig. 5의 데이터 저장소(Data Repository)에서 보여주는 바와 같이 크게 기본(Logical) 시나리오 파일, 파라미터 공간을 저장한 시나리오 파라미터 파일, 도로 파일, 실 도로 주행 테스트(FOT) 데이터로 구성되어 있다. 이러한 다양한 형태의 파일을 탐색하여 추출하기 위해서는 데이터 스키마(Schema)가 필요하며 이를 기반으로 시나리오 데이터베이스와 데이터 저장소 내의 파일과의 NoSQL 데이터베이스를 구축하였다. Fig. 5의 시나리오 데이터베이스에서 보는 바와 같이 ASAM OpenScenario와 OpenLabel의 정의를 확대하고 Admin, Scenery, Environment, Dynamics로 구분하여 시나리오 데이터베이스를 구축하였다.[2,21,22] Environment의 상세 시나리오에 대한 시뮬레이션이 본 논문에서는 고려되지 않았기에 관련 내용이 표시되지는 않았다.

(a) Snapshot of cut-in scenario in FOT data



(b) Distribution of a_{target}



(c) Distribution of $R_{cut\text{-}in}$

Fig. 7 Example of FOT cut-in scenario and distribution of parameters

다음으로 시나리오 데이터베이스를 기반으로 선행차량 추종 및 끼어들기 시나리오를 시뮬레이션하기 위한 추출과정을 설명하고자 한다. 시나리오 파일과 그에 해당하는 파라미터 파일의 정보를 쿼리(Query)를 통하여 탐색한다. 예를 들어서 Fig. 3과 Fig. 4에서 보는 바와 같이 시뮬레이션을 수행하기 위한 선행차량 추종 시나리오 파일, 64,000개의 시나리오 파라미터 파일, 직선주행을 위한 도로 파일이 탐색될 수 있다. 이후 64,000개의 시뮬레이션이 자동으로 수행될 수 있는 테스트 자동화 기능과 연동이 되어 시뮬레이션을 수행하고 시나리오 별 위험도(TTC_{min}, $I_{lat,max}$)를 구할 수 있다.

마지막으로 실 도로 주행데이터를 기반으로 파라미터 영역의 적절성을 살펴보는 예제를 설명하고자 한다. 시나리오 데이터베이스의 탐색을 통하여 경부 고속도로에서 2시간 40분동안 취득한 데이터 내 113건의 끼어들기(Cut-in) 시나리오가 있음을 확인할 수 있다. Fig. 7(a)는 그 중 하나의 시나리오를 보여주고 있으며, Fig. 7(b)와 (c)는 113건에 대한 파라미터 공간의 분포를 보여주고 있다. Fig. 7(b)와 (c)의 적색 구간은 Table 1에서 설정한 파라미터 공간이며, 청색 구간은 FOT의 파라미터 공간을 의미한다. Table 1에서 설정한 v_{ego}, v_{target}, $t_{cut\text{-}in}$, $R_{cut\text{-}in}$의 파라미터 공간은 FOT 데이터의 파라미터 공간을 모두 포함하는 것으로 그 적절성을 확인했으며, 그 예시로 Fig. 7(c)와 같이 $R_{cut\text{-}in}$에 대해서 FOT 데이터의 파라미터 공간을 포함하는 것을 나타내고 있다. 하지만 Fig. 7(b)의 a_{target}은 일부 포함되지 않는 것을 볼 수 있는데, Table 1에서 설정한 a_{target}의 범위는 추돌이 발생할 수 있는 전방 차량이 감속하는 상황만을 고려한 범위이기 때문이다.[4] 따라서 끼어들기 도중 가속하는 상황은 설정한 범위에 포함되지 않는다.

3.3 시나리오 파라미터 선정

3.3.1 시나리오 대리 모델

상세 시나리오의 파라미터를 선정하기 위해서 우선적으로 시나리오 대리 모델(Surrogate model)을 개발하고자 한다. 여기서 대리 모델이라 함은 주어진 성능지표에 대해서 시뮬레이션의 결과를 예측할 수 있는 네트워크 기반 모델이다. 즉, Fig. 3(b)와 Fig. 4(b) 같이 무작위로 선정된 상세 시나리오를 시뮬레이션을 수행하고 해당되는 위험도(TTC_{min}, $I_{lat,max}$)를 Ground Truth(GT)로 선정하여 대리 모델을 학습하고, 이를 기반으로 시나리오 파라미터 공간 영역 내에 있는 모든 파라미터에 대해서 상세 시나리오의 위험도를 예측하기 위한 모델이다.

문헌에서 제시하는 Multi-Layer Perceptron(MLP)기반 대리 모델을 활용하여 대리 모델의 구조를 Fig. 8과 같이 5개의 Hidden layer로 구성하였다.[3] 대리 모델의 입력은 시나리오 파라미터($p_j \in \mathbb{R}^N$)이며 출력은 위험도($\hat{y}_k$)이다.

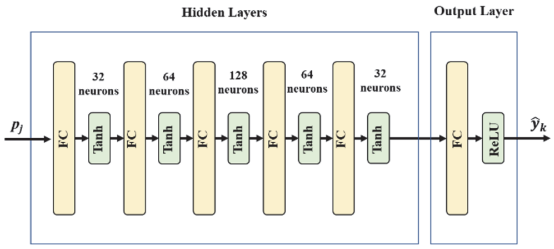


Fig. 8 Structure of surrogate model

여기서 아래 첨자 j는 선행차량 추종 시나리오나 끼어들기 시나리오를 구분하며 아래 첨자 k는 TTC_{min}와 $I_{lat,max}$와 같은 위험도 인자를 구분한다. 더 구체적으로 선행차량 추종 시나리오에 대한 파라미터는 식 (1)과 Table 1에, 끼어들기에 대해서는 식 (2)와 Table 1에 정의되어 있다.

다음은 모델의 학습 과정 및 결과이다. 선행차량 추종 시나리오 같은 경우, 총 64,000개의 상세 시나리오에서 랜덤으로 추출한 3,200개의 상세 시나리오를 학습 데이터로 사용하였다. 끼어들기 같은 경우, 추종 시나리오와 유사한 개수의 상세 시나리오를 생성하기 위해 각 파라미터(v_{ego}, v_{target}, $t_{cut\text{-}in}$, a_{target}, $R_{cut\text{-}in}$)별로 9개씩 샘플링하여 총 59,049(9^5)개의 상세 시나리오를 생성하였으며, 끼어들기 거동이 일어나기 위해 상대 차량 속도(v_{target})보다 자차량 속도(v_{ego})가 더 큰 상세 시나리오를 추출하여 총 26,244개의 상세 시나리오를 전체 시나리오셋으로 설정하였다. 그 중, 랜덤으로 3,000개의 상세 시나리오를 추출하여 학습 데이터에 사용하고, 모델의 정확성 평가를 위해 다음과 같이 MSE(Mean-squared-error)를 사용하였다.

$$MSE = \frac{1}{M}\sum_{l=1}^{M}\left(\hat{y}_{k,l} - y_{k,l}\right)^2 \tag{3}$$

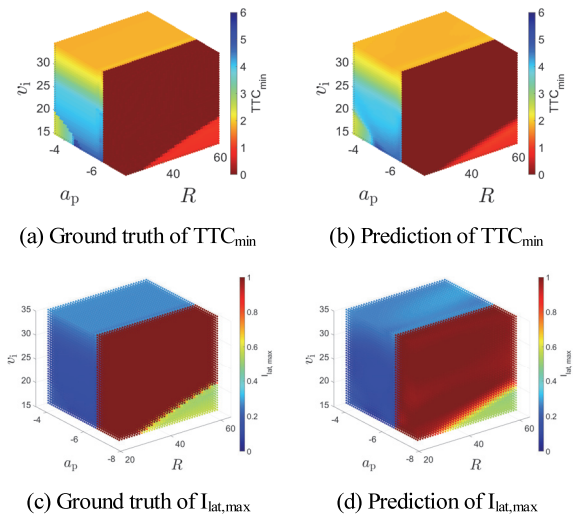
여기서 M은 평가에 사용된 상세 시나리오의 개수, $\hat{y}_k$은 모델에서 예측한 위험도, y_k는 시뮬레이션을 통해 구한 위험도이다.

Table 2는 선행차량 추종(Car-following) 및 끼어들기(Cut-in) 시나리오 64,000개와 26,244개에 대한 위험도 예측 결과를 MSE지표로 각각 보여주고 있으며 추종 시나리오의 경우 기존 연구에서 제시된 TTC_{min} 기준 약 0.01보다 적은 오차로 학습이 되었음을 확인할 수 있다.[3] 하지만 끼어들기 시나리오의 경우 추종 시나리오보다 MSE 기준 오차가 10배 이상 크지만 학습 횟수(Epoch)를 증가하여도 MSE가 감소하지 않아 최종적으로 Epoch 500, Batch size 32로 학습하였다.

Fig. 9와 Fig. 10은 선행차량 추종 시나리오와 끼어들기 시나리오에 대해서 파라미터 3개를 선정하여 시뮬레이션 결과를 통해 구한 위험도와 모델을 통해 나온 위험

Table 2 Mean-squared-error(MSE) of the threat assessment index

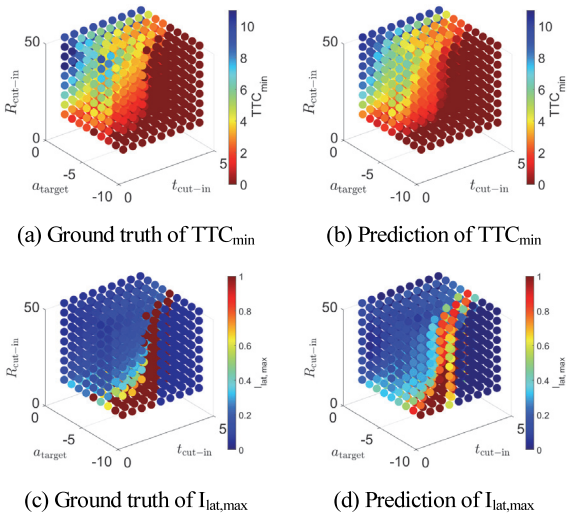
Scenario (# of scenario)	TTC_{min} (s)		$I_{lat,max}$	
	Train (3,200)	Validation (64,000)	Train (3,000)	Validation (26,244)
Car-following	0.003	0.005	0.0004	0.0005
Cut-in	0.020	0.055	0.0051	0.0065



(a) Ground truth of TTC_{min} (b) Prediction of TTC_{min}

(c) Ground truth of $I_{lat,max}$ (d) Prediction of $I_{lat,max}$

Fig. 9 Comparison between ground truth and output of surrogate model with respect to TTC_{min} and $I_{lat,max}$ in car-following scenario



(a) Ground truth of TTC_{min} (b) Prediction of TTC_{min}

(c) Ground truth of $I_{lat,max}$ (d) Prediction of $I_{lat,max}$

Fig. 10 Comparison between ground truth and output of surrogate model with respect to TTC_{min} and $I_{lat,\ max}$ in cut-in scenario

도를 크기에 따라 색을 나타낸 그림이다. 위험도의 크기에 따라 색으로 나타내었고, 각 시나리오 별 TTC_{min}, $I_{lat,\ max}$ 모두 비슷한 형색을 나타낸다.

더 나아가, 두 그림의 비교를 통해서 상대적으로 끼어들기 시나리오에 대한 데이터의 연속성(Continuity)이 부족함을 확인할 수 있으며, 이러한 이유로 끼어들기 시나리오에 대한 대리모델 성능이 상대적으로 낮은 것으로 예상된다.

3.3.2 시나리오 파라미터 탐색

시나리오 파라미터 탐색은 문제 정의에서 언급했던 랜덤 샘플링의 위험도 성능 분포 불균형 문제가 발생한다. 따라서 경사 하강법 기법 중 Adam 알고리즘[3)]을 이용하여 위험도가 높은, 즉 TTC_{min}이 작은 방향으로 상세 시나리오의 파라미터를 탐색한다. 더 나아가 파라미터 탐색 과정 중 설정한 TTC_{min} 범위에 도달하면 상세 시나리오 파라미터를 샘플링한다.

Fig. 11은 그 예시이며, 랜덤으로 생성된 초기 상세 시나리오의 파라미터에서 경사 하강법을 통해 TTC_{min}의 값이 작은 방향으로 탐색한다. 이때 TTC_{min}의 설정한 값($TTC_{min,\ th}$ = 1.4s)에 도달하게 된다면, 그 때의 파라미터(R = 28.02 m, a_p = -0.62 g, v_i = 32.01 m/s) 근처 범위($TTC_{min,\ th} - \varepsilon < \hat{y}_k < TTC_{min,\ th} + \varepsilon$, ε = 0.1s) 내에서 샘플링을 진행하고 탐색은 종료되며, 이러한 과정을 반복하여 상세 시나리오를 추출하게 된다. $I_{lat,\ max}$의 경우 TTC_{min}와 반대로 위험도가 높을수록 값이 크기 때문에 $I_{lat,\ max}^{-1}$이 작은 방향으로 파라미터를 탐색하고 범위를 ε= 0.1로 설정하였다.

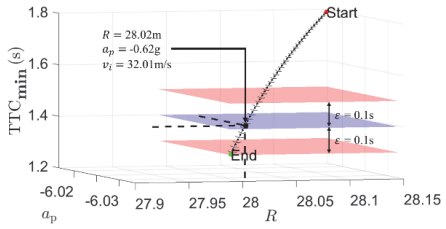


Fig. 11 Schematics of parameter generative model based on surrogate model

4. 선행차량 추종 및 끼어들기 시나리오 생성

Fig. 3(b)와 Fig. 4(b)에서 보인 바와 같이 위험도 관점에서 상세 시나리오 분포의 불균형을 해결하기 위해 시나리오 파라미터 탐색을 통한 상세 시나리오를 추가적으로 생성하고자 한다. 먼저 선행차량 추종 및 끼어들기 시나리오 각각 전체 시나리오에서 대리 모델의 학습에 사용된 랜덤으로 추출한 3,200개, 3,000 개의 상세 시나리오를 이용한다. 그 후에 선행차량 추종 시나리오 같은 경우, 랜덤으로 추출한 3,200개의 상세 시나리오의 결과와 탐색을 통해 추출한 상세 시나리오를 추가한 결과를 보여주고, 끼어들기 시나리오의 경우, 랜덤으로 추출한 3,000개의 상세 시나리오의 결과와 탐색을 통해 추출한 상세 시나리오를 추가한 결과를 보여주고자 한다.

4.1 선행차량 추종 시나리오에 대한 파라미터 선정

선행차량 추종 시나리오의 경우, Fig. 12에서 볼 수 있

듯이, 랜덤으로 추출한 3,200개의 상세 시나리오셋은 TTC_{min} 관점에서 1~2초 사이에 분포가 없는 것을 확인할 수 있다. 따라서 경사 하강법을 통해 파라미터를 탐색할 때 TTC_{min} 값이 1.4s 근처의 파라미터만 샘플링을 진행하였으며, 최종적으로 300개의 상세 시나리오를 추출하였다. 그 후, 랜덤으로 추출한 3,200개의 상세 시나리오의 결과와 탐색을 통해 추출한 300개의 상세 시나리오를 합산한 결과를 Fig. 12에서 볼 수 있으며, 기존 랜덤 샘플링 결과에 TTC_{min}이 1.4s 부근의 상세 시나리오가 추가되어 분포해 있는 것을 확인할 수 있다. 이에 따라 Fig. 13에서 $I_{lat,\ max}$에 대한 결과도 볼 수 있으며, $I_{lat,\ max}$ 값이 0.4인 곳에 분포가 추가되어 그려졌다.

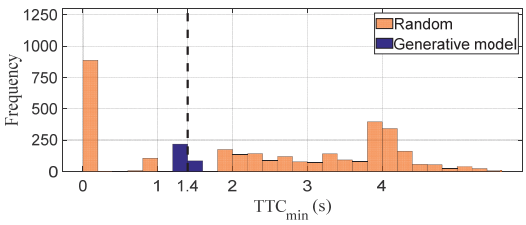


Fig. 12 Distribution of car-following scenarios with respect to TTC_{min}

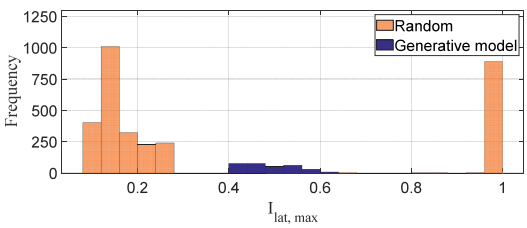


Fig. 13 Distribution of car-following scenarios with respect to $I_{lat,\ max}$

4.2 끼어들기 시나리오의 파라미터 선정

끼어들기 시나리오에 대해서도 선행차량 추종 시나리오와 동일하게 탐색을 진행하였다. 하지만 $I_{lat,\ max}$ 값을 기준으로 시나리오 탐색한다는 것이 다르다. Fig. 14에서 볼 수 있듯이, TTC_{min} 관점에서 부족한 상세 시나리오는 5~10 s이며, 이는 안전한 시나리오로 판단하여 $I_{lat,\ max}$ 관점에서 부족한 상세 시나리오를 생성하고자 하였다. Fig. 15를 보면, $I_{lat,\ max}$ 값이 0.6~0.9 사이에 있는 부분이 부족하여 경사 하강법을 통해 $I_{lat,\ max}$ 값이 0.7 근처의 파라미터만 샘플링을 진행하였다. 앞에서 언급했듯이, 랜덤으로 추출한 3,000개의 상세 시나리오와 탐색을 통하여 추출한 300개의 상세 시나리오를 합한 결과의 분포를 Fig. 15에서 보여주고 있으며, $I_{lat,\ max}$값이 0.6~0.8사이의 분포가 추가된 것을 확인할 수 있다.

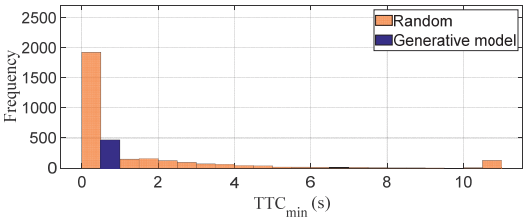


Fig. 14 Distribution of cut-in scenarios with respect to TTC_{min}

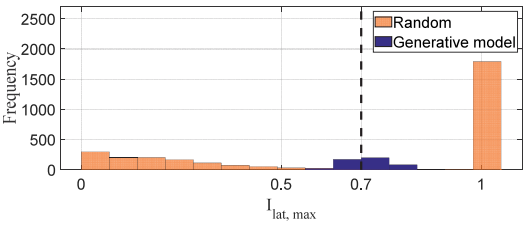


Fig. 15 Distribution of cut-in scenarios with respect to $I_{lat,\ max}$

5. 결 론

본 연구에서는 선행차량 추종 및 끼어들기 시나리오에 대한 파라미터 공간 내에서 특정 위험도 관점에서 상세 시나리오를 선정하는 방법을 제안하였다. 랜덤 샘플링에서 발생할 수 있는 시나리오의 불균형 문제를 보완하고자 대리 모델을 기반으로 경사 하강법을 이용해 원하는 범위의 성능 지표에 해당하는 상세 시나리오의 파라미터를 추출하는 생성형 모델을 이용하여 시나리오의 위험도 지표 관점에서 좀 더 균일한 분포를 보여주는 상세 시나리오 선정방법을 개발하였다. 향후 계획으로 시나리오 및 성능 평가 지표를 다양화하고, 이에 좀 더 정확한 생성형 모델을 고도화하고자 한다.

후 기

이 논문은 산업통상자원부가 지원한 ‘자동차산업 미래기술혁신을 위한 오픈플랫폼 생태계 구축’사업의 지원을 받아 수행된 연구 결과이다[과제명: 자동차산업 미래기술혁신을 위한 오픈플랫폼 생태계 구축 / 과제고유번호: P0018434].

References

1) N. Kalra and S. M. Paddock, “Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?,” Transportation Research Part A: Policy and Practice, Vol.94, pp.182-193, 2016.
2) ASAM, ASAM OpenX SIMULATION, https://www.

asam.net, 2023.

3) Y. Wang, R. Yu, S. Qiu, J. Sun and H. Farah, "Safety Performance Boundary Identification of Highly Automated Vehicles: A Surrogate Model-Based Gradient Descent Searching Approach," IEEE Transactions on Intelligent Transportation Systems, Vol.23, No.12, pp.23809-23820, 2022.
4) 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.
5) T. Ponn, D. Fratzke, C. Gnandt and M. Lienkamp, "Towards Certification of Autonomous Driving: Systematic Test Case Generation for a Comprehensive but Economically Feasible Assessment of Lane Keeping Assist Algorithms," Vehicle Technology and Intelligent Transport Systems International Conference Proceedings, pp.333-342, 2019.
6) S. Riedmaier, T. Ponn, D. Ludwig, B. Schick and F. Diermeyer, "Survey on Scenario-Based Safety Assessment of Automated Vehicles," IEEE Access, Vol.8, pp.87456-87477, 2020.
7) EuroNCAP, AEB Car-to-Car Test Protocol v3.0.3, https://cdn.euroncap.com/media/62794/euro-ncap-aeb-c2c-test-protocol-v303.pdf, 2023.
8) EuroNCAP, Assisted Driving - Highway Assist Systems v1.0, https://cdn.euroncap.com/media/58813/euro-ncap-ad-test-and-assessment-protocol-v10.pdf, 2023.
9) 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, pp.6971-6988, 2023.
10) F. Schuldt, A. Reschka and M. Maurer, "A Method for an Efficient, Systematic Test Case Generation for Advanced Driver Assistance Systems in Virtual Environments," Automotive Systems Engineering II, pp.147-175, 2018.
11) J. Jung and K. Lee, "Automatic Scenario Generation for Decision Algorithm Performance Evaluation of Autonomous Vehicle via Scenario Parameter Sweeping Method," Int. J. Automotive Technology, Vol.23, No.5, pp.1383-1391, 2022.
12) H. Nakamura, H. Muslim, R. Kato, S. Préfontaine-Watanabe, H. Nakamura, H. Kaneko, H. Imanaga, J. Antona-Makoshi, S. Kitajima, N. Uchida, E. Kitahara, K. Ozawa and S. Taniguchi, "Defining Reasonably Foreseeable Parameter Ranges Using Real-World Traffic Data for Scenario-Based Safety Assessment of Automated Vehicles," IEEE Access, Vol.10, pp.37743-37760, 2022.
13) D. Zhao, H. Lam, H. Peng, S. Bao, David J. LeBlanc, K. Nobukawa and Christopher S. Pan, "Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques," IEEE Transactions on Intelligent Transportation Systems, Vol.18, No.3, pp.595-607, 2017.
14) D. Zhao, X. Juang, H. Peng, H. Lam and D. J. LeBlanc, "Accelerated Evaluation of Automated Vehicles in Car Following Maneuvers," IEEE Transactions on Intelligent Transportation Systems, Vol.19, No.3, pp.733-744, 2018.
15) E. de Gelder and J. P. Paardekooper, "Assessment of Automated Driving Systems Using Real-Life Scenarios," IEEE Intelligent Vehicles Symposium (IV), pp.589-594, 2017.
16) A. Pütz, Z. Adrian, B. Julian and E. Lutz, "System Validation of Highly Automated Vehicles with a Database of Relevant Traffic Scenarios," ITS European Congress Proceedings, Vol.1, pp.1-8, 2017.
17) A. Tenbrock, A. König, T. Keutgens and H. Weber, "The ConScenD Dataset: Concrete Scenarios from the highD Dataset According to ALKS Regulation UNECE R157 in OpenX," IEEE Intelligent Vehicles Symposium Workshops, pp.174-181, 2021.
18) K. Kim, D. Kim and K. Yi, "Development of Lateral Collision Risk Index," KSAE Annual Conference Proceedings, pp.1376-1381, 2012.
19) W. G. Najm, J. D. Smith and M. Yanagisawa, "Pre-Crash Scenario Typology for Crash Avoidance Research," National Highway Traffic Safety Administration, No.DOT-VNTSC-NHTSA-06-02, 2007.
20) IPG Automotive, CarMaker Reference Manual Version 7.0.2, IPG Automotive GmbH, 2014.
21) ASAM, ASAM OpenLABEL User Guide, https://www.asam.net/standards/detail/openlabel, 2023.
22) G. Bagschik, T. Menzel and M. Maurer, "Ontology Based Scene Creation for the Development of Automated Vehicles," IEEE Intelligent Vehicles Symposium (IV), pp.1813-1820, 2018.