자유와 활력이 넘치는
파워풀 대구

FIX All on AI

Future Innovation tech eXpo Conference 2025

# AI와 디지털 트윈으로 진화하는 자율 주행 시뮬레이션

조윤희 / 모라이

1. MORAI Introduction
2. E2E 자율주행 관련 Project
3. NVIDIA COSMOS
4. MORDA 기반 Synthetic Fusion

# MORAI Glance

**2018**
Founded In

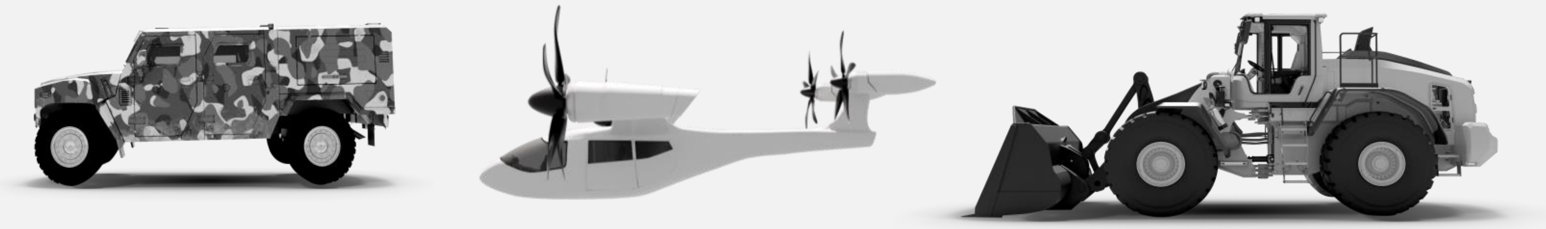
**370억원**
누적투자금액

**60+**
Members

**80% >**
Developers

## Our mission.

Empowering customers to
**pioneer the future of mobility**
with **innovative simulation technology**



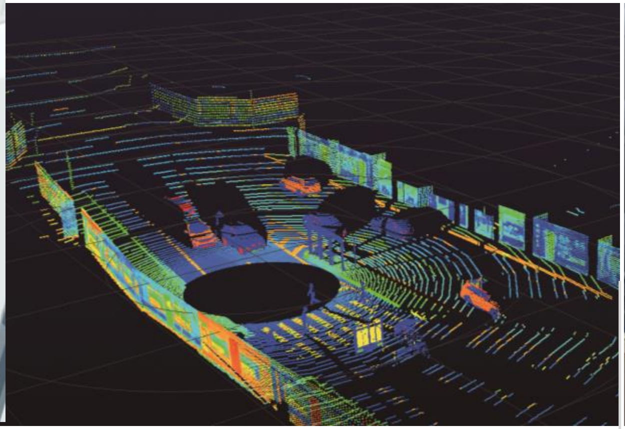
# MORAI 전문 기술 및 영역

**차세대 모빌리티를 지원하는 국내 유일의 Full Stack 자율주행 시뮬레이션 솔루션**

- 관제부터 자율차 수집 데이터, 시나리오 에디팅까지 할 수 있는 전체 파이프라인 보유
- 자율주행자동차, UAM, 무인 로봇, 무인 선박 등 차세대 핵심 모빌리티 및 무인이동체 전반의 시뮬레이션 지원

**95% 수준의 디지털 트윈 자동화 기술**

- 실제 정밀도로지도 데이터를 활용, 10cm 수준의 정밀도 환경을 디지털 트윈으로 구현
- 서울, 판교, 대구 및 라스베이거스, 싱가포르, 샌프란시스코, 시애틀 등 전 세계 30여 도시 구축



**다양한 종류의 차량 및 센서 모델링 및 학습 데이터 제공**

- 카메라, LiDAR, GPS, 레이더, IMU를 포함한 다양한 센서 모델을 통해 차량 동작의 정확한 시뮬레이션 제공
- 센서 데이터 자동 레이블링 및 가상 데이터셋 구축 제공. 날씨 및 조도 제어 기능과 함께 자동 주석 기능을 통해 사용자 맞춤형 데이터셋 자동 생성

**사실적인 시나리오로 자율주행 차량의 효과적인 성능 테스트 지원**

- 자율차가 수집한 현실 데이터를 활용하여 실제 교통 흐름을 모델링하고, 자율차가 다양한 상황에서 어떻게 행동해야 하는지 시뮬레이션
- 자율차 운행 시 예측하지 못한 상황에 대비하고, 자율주행 시나리오를 실제 도로 환경에서 효과적으로 테스트할 수 있도록 지원

# 주요 핵심 프로젝트

## Autonomous Vehicle

Validate Edge Case Scenario & Integrate SIL/HIL Platform

HYUNDAI AutoEver · a2z AUTONOMOUS · 42dot · NAVER LABS

## ADAS

Verification and Validation for ADAS Algorithm

HYUNDAI · HYUNDAI MOBIS · HYUNDAI AutoEver · Mando · JIAT 자동차융합기술원 · SAMSUNG ELECTRONICS

## Traffic Control

Autonomous Vehicle Control & Monitoring using Simulation

성남시 Seongnam City · 대구광역시 · SEOUL METROPOLITAN GOVERNMENT · 부산광역시 BUSAN METROPOLITAN CITY

## UAM

Urban VTOL Simulation & UAM Corridor Verification

HYUNDAI MOTOR GROUP · KAC KOREA AIRPORTS CORPORATION · KARI 한국항공우주연구원 · NSR National Security Research Institute

## Logistic

Logistics Center & Large-Scale Transportation Simulation

SAMSUNG HEAVY INDUSTRIES · NAVER LABS · SAMSUNG ENGINEERING

## Robotics

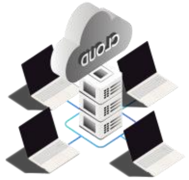
Autonomous Robot Service

neuromeka · KIRIA 한국로봇산업진흥원 · WeGo · WOOWABROS.

## Defense

Test MUM-T Scenario & Validate Autonomous Defense System

대한민국육군 Republic of Korea Army · LIG Nex1 · Hanwha Defense · KATECH 한국자동차연구원



## Data Generation

Generate Massive Virtual Data for AI Training

SAMSUNG · SAMSUNG ENGINEERING · NIA · Hanwha Systems

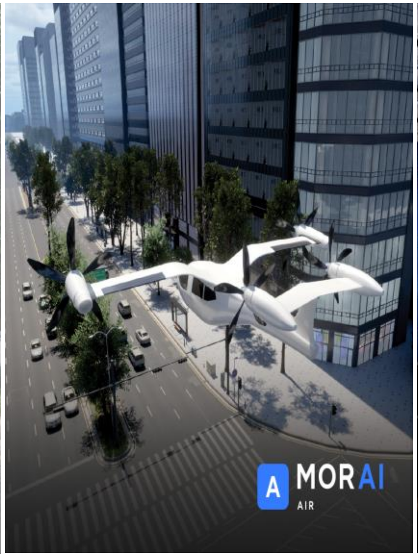
## Education

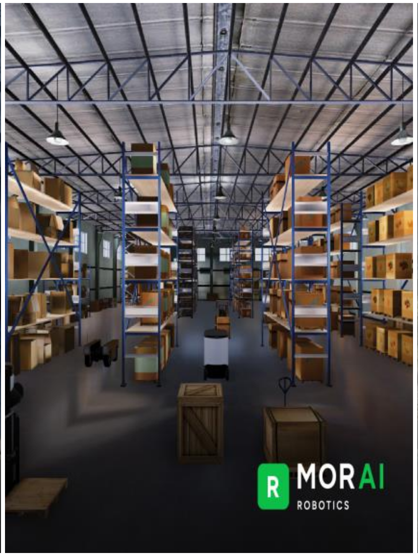
Provide Education Contents on Autonomous Driving & Host Competitions

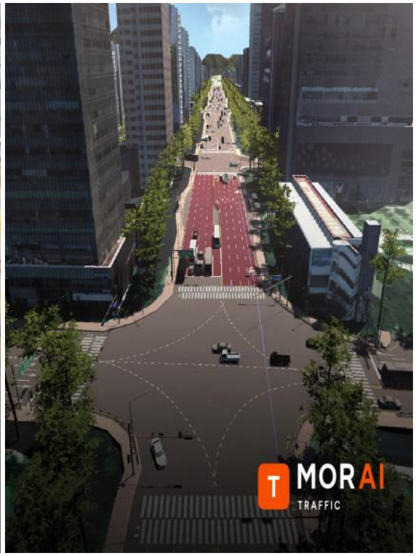
Carnegie Mellon University · UNIVERSITY OF MICHIGAN · 아주대학교 AJOU UNIVERSITY · 울산대학교 UNIVERSITY OF ULSAN

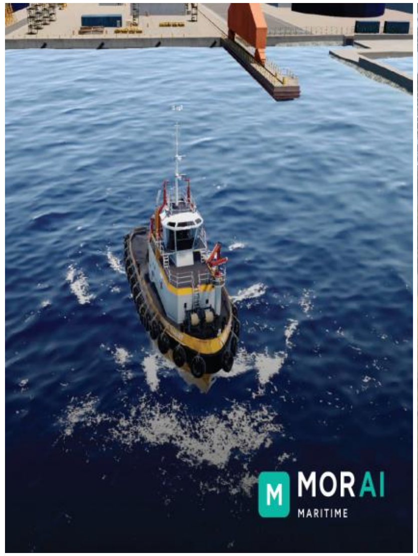
# 모빌리티 시뮬레이션 및 관제, 디지털 트윈

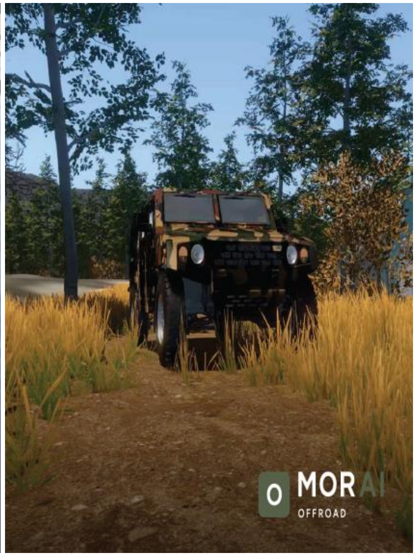
- 자율주행차에서부터 UAM, 무인 로봇, 무인 선박, 국방, 교통 관제 등 다양한 분야에 적용


## Drive

Drive의 실제와 동일한 시뮬레이션 환경, 센서 및 차량 모델을 통해 가상으로 자율주행 차량 검증

## AIR

Air의 실제와 동일한 비행 환경을 통해 드론, UAM 및 고정익 항공기 등 기체의 비행 시스템 검증

## Robotics

실내외 시뮬레이션 환경, 센서 및 차량 모델을 통해 가상으로 AMR 검증

## Traffic

현실과 동일한 교통환경을 Digital Twin 기반 관제 시스템으로 표현하여 정밀한 교통관제 지원

## Maritime

Maritime은 자율 운항 선박 개발자들이 가상 환경에서 알고리즘을 개발 및 테스트

## Offroad

Offroad는 전시 상황 시뮬레이션 환경 제공

# XIL Environments

SIL / MIL
(Software-In-the Loop)
(Model-In-the Loop )

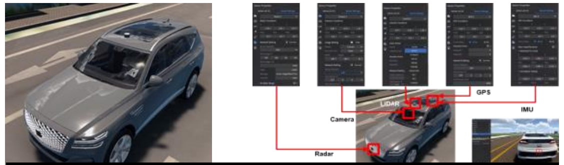
HIL
(Hardware-In-the Loop)

VIL
(Mixed Reality)
(Vehicle-In-the Loop)

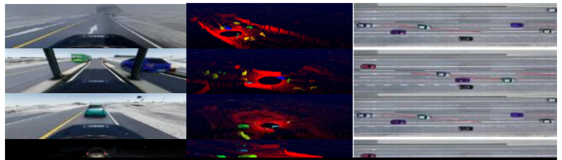
DIL
(Driver-In-the Loop)

Dynamics

Traffic



Sensors



Dataset

| XIL | SIL / MIL | HIL | VIL | DIL | Field Of Test |
|---|---|---|---|---|---|
| Driver | Virtual | Virtual | Virtual (ADAS/AD) | Real | Real |
| Systems | Virtual | Real | Real | Virtual / Real | Real |
| Dynamics | Virtual | Virtual / Real | Real | Virtual / Real | Real |
| Traffic | Virtual | Virtual / Real | Virtual / Real | Virtual / Real | Real |
| Sensors | Virtual | Virtual / Real | Virtual / Real | Virtual / Real | Real |

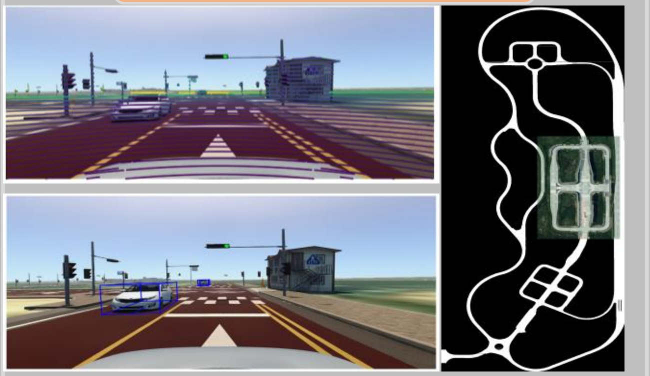
# 현대자동차 End to End 자율주행 경진대회

## 프로젝트 정보

| 프로젝트 명 | 현대자동차 End to End 자율주행 경진대회 |
|---|---|
| 목적 | MORAI Virtual Simulation 기반 AI End to End 개발에 대한 대학 연구 집중 유도 및 인재 육성을 위함 |
| 프로젝트 기간 | 2024년 10월 ~ 2025년 9월 |
| 고객 | 현대자동차 자율주행챌린지 운영사무국 |
| 대회 일정 | 2/18 ~ 19 예선 1차 → 3/ 예선 2차 → 5/ 본선 1차 → 9/ 본선 2차 |
| 제공 내용 | (아래 참조) |

### 제공 내용

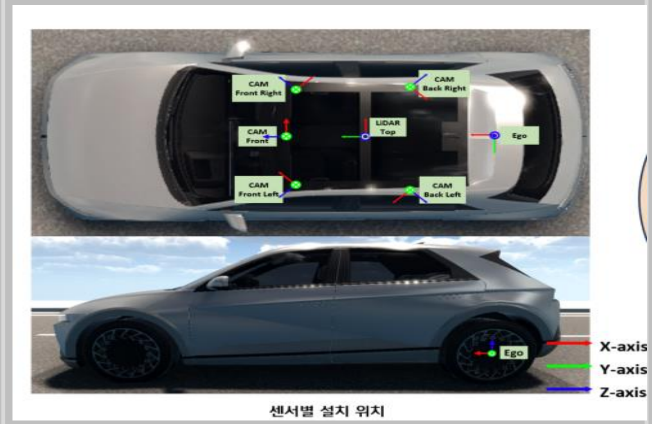
**Dataset 제공**



Sensor 및 GT 데이터 예시

- Map / Sensor / GT Data

**센서 구성**

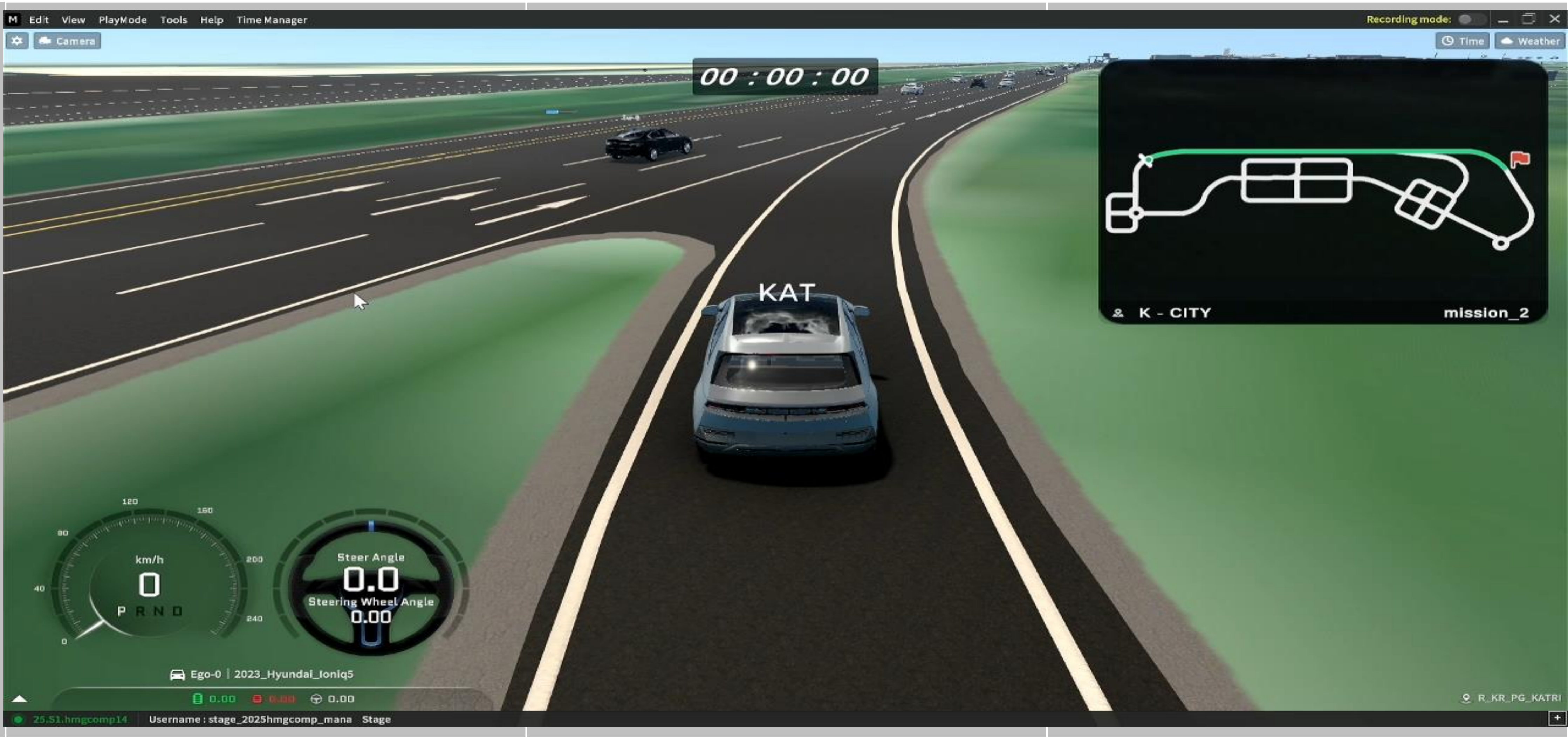



- 카메라5대/Lidar1대/GNSS/IMU

**대회 Mission 가상 제공**


- 신호 인지 / 차량 & 보행자 회피 / 톨게이트 / 차선 변경 / 특수 상황

# 현대자동차 End to End 자율주행 경진대회

| 프로젝트 정보 | |
|---|---|
| 대회 동영상<br>(도심 고속도로 주행영상 ) |  |

# AI 학습을 위한 Synthetic Dataset Generation

ASIL D READY
Functional Safety
www.sgs-tuev-saar.com

CERTIFICATE NO FS/71/220/23/1039 PAGE 1/1
ZERTIFIKAT NR.: SEITE(N)

| LICENCE HOLDER / GENEHMIGUNGSINHABER | MANUFACTURER / HERSTELLER |
| --- | --- |
| MORAI Inc.<br>12th Fl, V-PLEX, 501, Teheran-ro,<br>Gangnam-gu,<br>Seoul,<br>Republic of Korea | MORAI Inc.<br>4th Fl, L7 Gangnam Tower, 415,<br>Teheran-ro, Gangnam-gu,<br>Seoul, 06160,<br>Republic of Korea |

| PROJECT NO/-ID / PROJEKT-NR/-ID | LICENSED TEST MARK / GENEHMIGTES PRÜFZEICHEN | CERT. REPORT NO. / ZERTIFIKATSBERICHT NR. |
| --- | --- | --- |
| T2U1-AU01 | ASIL D READY Functional Safety | T2U10002 is an integral part of this certificate. ist ein integraler Bestandteil dieses Zertifikats. |

**Certified product(s)** *Zertifizierte(s) Produkt(e)*: Camera Sensor Module

Version 230105.S4.ISO7

**Tested according to** *Geprüft nach*: ISO 26262-8:2018, clause 11.4.4 and 11.4.9

**Technical Data and Parameter** *Technische Daten und Parameter*: The Camera Sensor Module is suitable to be part of a simulation analysis tool for validating the safety goals of an item up to ASIL D, according to ISO 26262.

The certificate is based on voluntary tests. The compliance of the certified product against the requirements of above listed functional safety standards was evaluated. Any changes to the design, components or processing may require repetition of some parts of the certification to retain the certification. All applicable requirements of the testing and certification regulations of SGS-TÜV Saar GmbH have to be complied, see www.sgs-tuv-saar.com/tcr-muc and www.sgs-tuv-saar.com/gtc-muc.

Certification Body
for Functional Safety &
Cyber Security
SGS-TÜV Saar GmbH
*Zertifizierungsstelle für Funktionale Sicherheit & Cyber Sicherheit*

Reference to SGS Certification Database

Munich, Feb 24, 2023

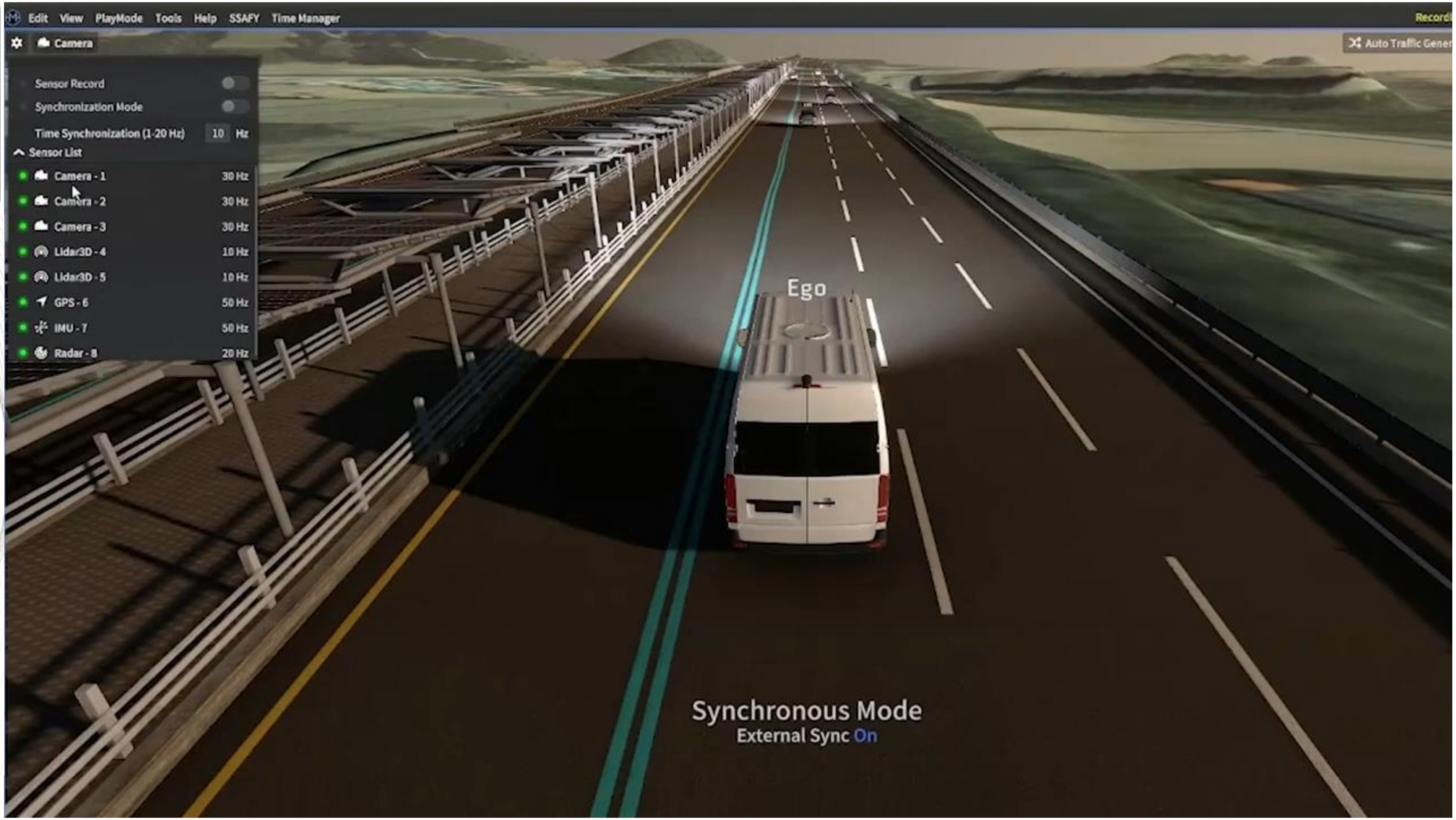
C. Schürmann

Cornelius Schürmann

SGS-TÜV Saar GmbH, Hofmannstr. 50, 81379 München, Deutschland / Germany

Website: www.sgs-tuv-saar.com
E-Mail: fs@sgs.com

CERTIFICATE



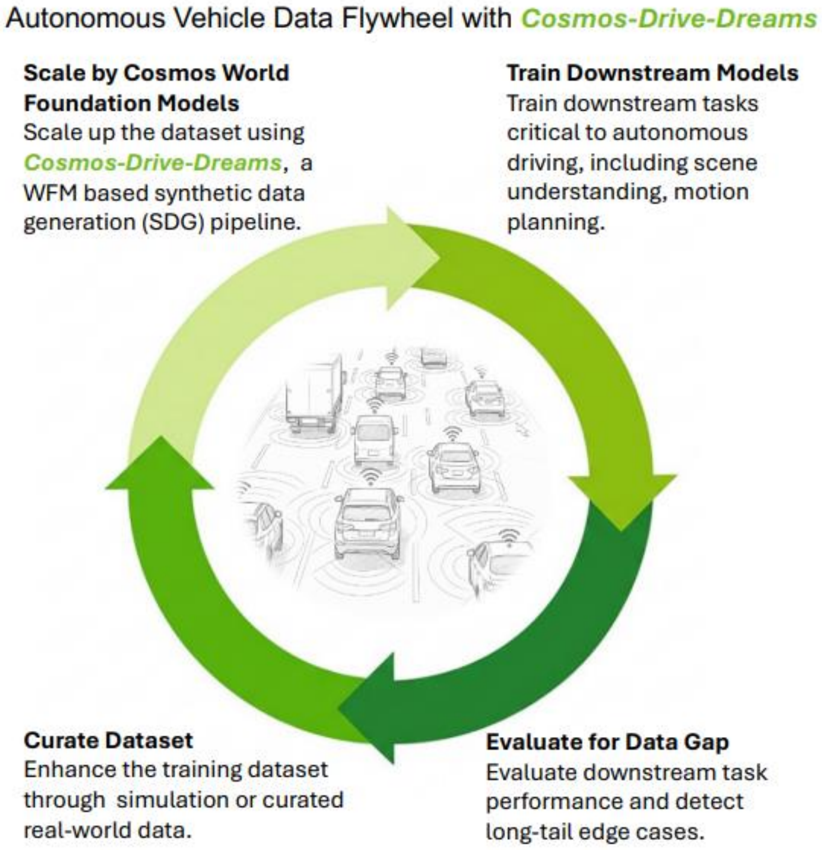
# NVIDIA COSMOS

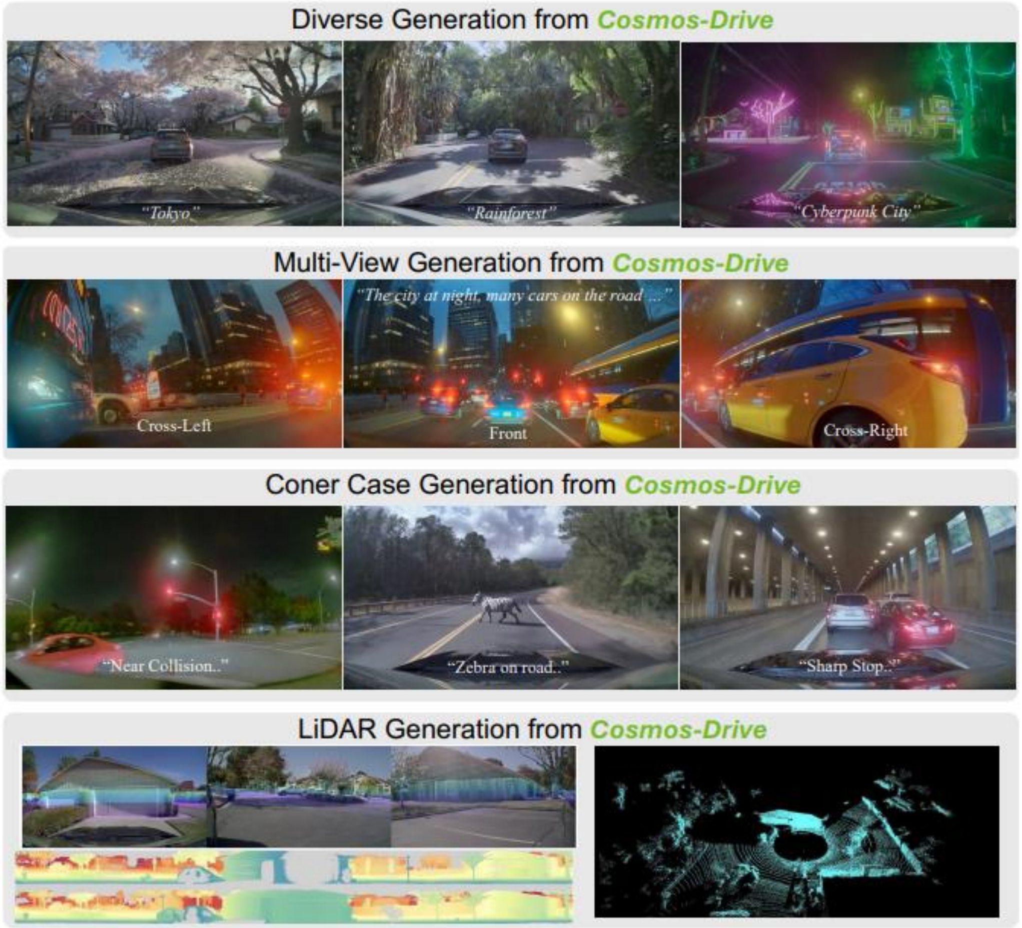
- A world foundation model platform for physical AI
- COSMOS WFM(World Foundation Model)
  - 2000만 시간(약 2283년) 분량의 영상데이터 학습
  - 뛰어난 예측 능력 ~ 사고력
    - Physical AI를 위한 Platform으로 사용 가능


Source: https://youtu.be/9Uch931cDx8?feature=shared

# COSMOS-Drive-Dreams

- Scalable Synthetic Driving Data Generation with World Foundation Models
- COSMOS를 이용한 자율주행 SDG(Synthetic Generation Pipeline)

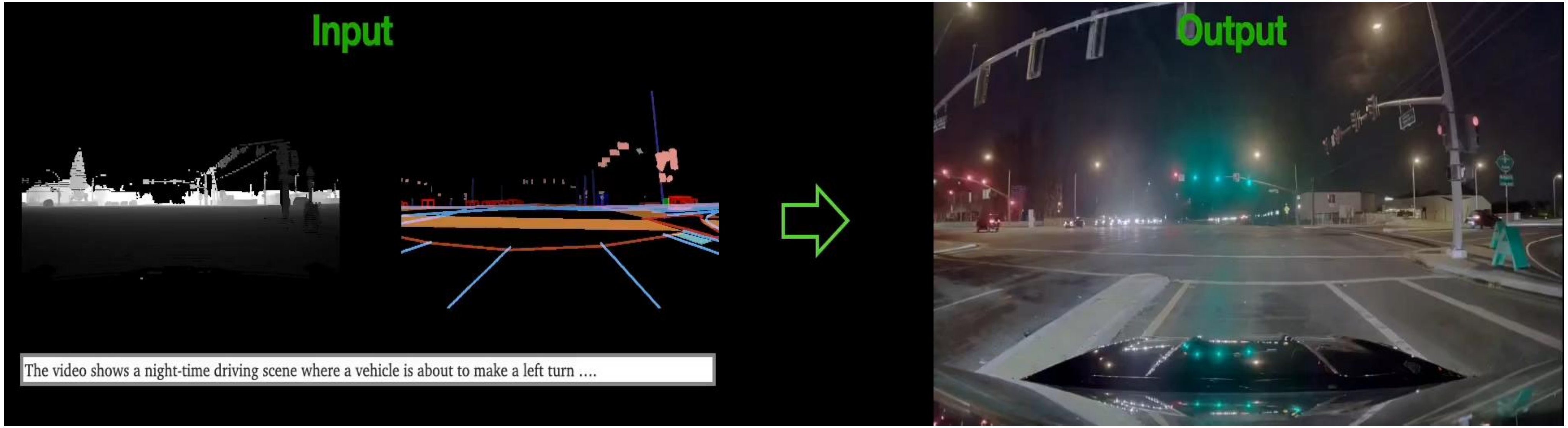



Source: https://github.com/nv-tlabs/Cosmos-Drive-Dreams

# COSMOS-Drive-Dreams

- Cosmos-Transfer1
- 자율주행 데이터 생성을 위한 COSMOS의 Downstream Model
  - Input / Output 모두 영상 포맷
  - Input: LiDAR(Depth map), HD Map(3D BBox + 3D Map feature), Text Prompt
  - Output: 합성 영상




✓ LiDAR PCD를 Depth map으로 변환하여 사용
✓ Image plane에 투영된 BBox 및 Map feature 사용
  ✓ 차로 중앙선, 정지선, 표지판, 신호등, 횡단보도 등 7가지
✓ Input 형태로 변환을 위한 toolkit 제공

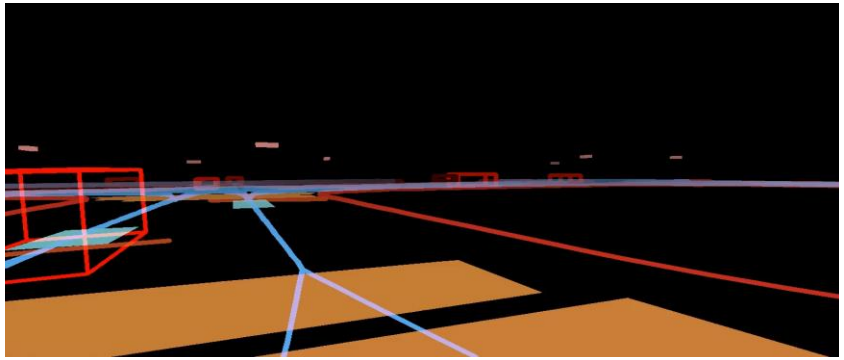
✓ 동일 Image plane에 transfer 된 Single-view  영상 출력
✓ Multi-view 출력은 Single-view 생성 후 별도 처리 필요
✓ Real-to-Real, Sim-to-Real 모두 가능

Source: https://github.com/nvidia-cosmos/cosmos-transfer1

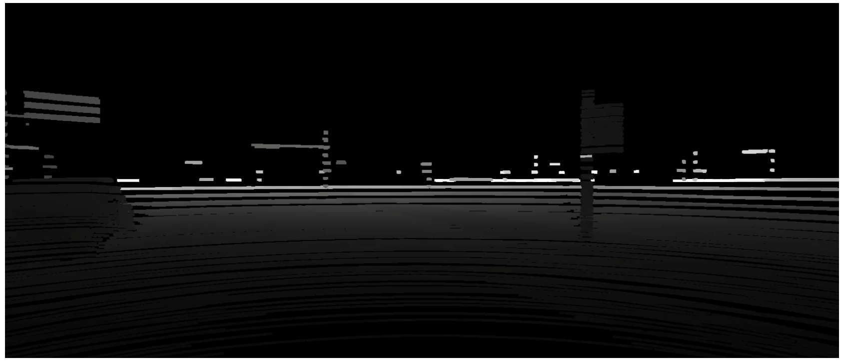
# COSMOS-Drive-Dreams

- Sim-to-Real Demo Using Cosmos-Transfer1
- Input: MORAI Data Clip, Output: Style-transferred MORAI

Converting



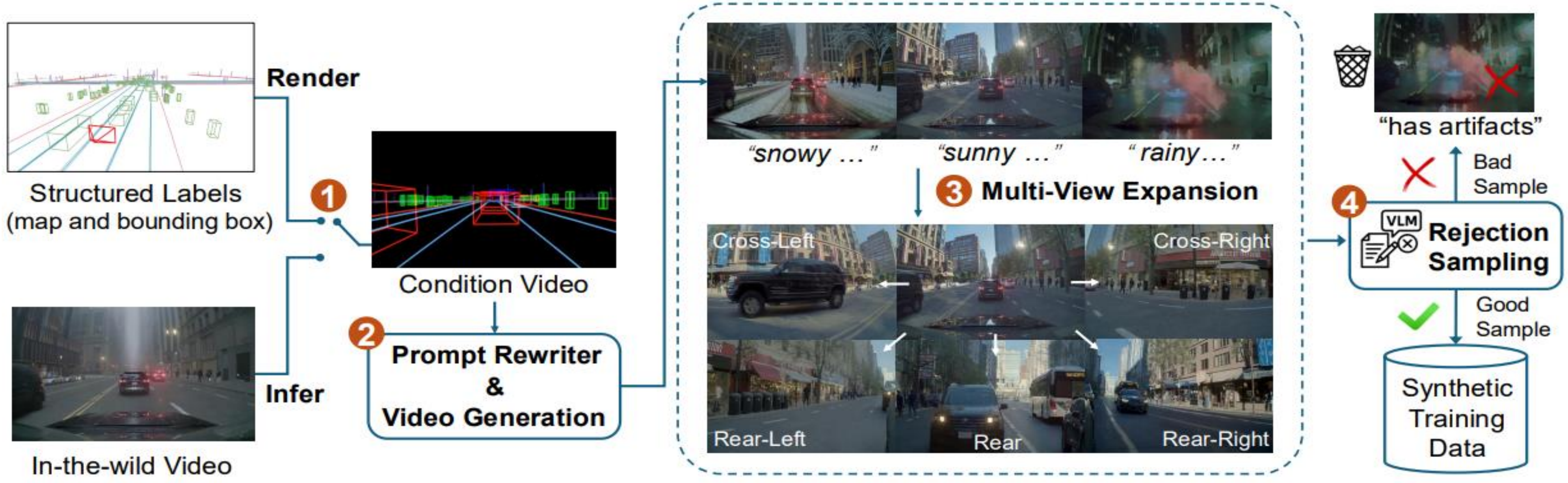
Cosmos-transfer1 Input



SDG by Cosmos-transfer1

# COSMOS-Drive-Dreams

- SDG Method and Validation Result




*Cosmos-Drive-Dreams* **SDG Pipeline**

| Dataset | Waymo | | | RDS-HQ (2k) | | | | RDS-HQ (20k) | | | | RDS-HQ (2k, MV) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LET-AP ↑ | All | Ext. Wea. | Night | All | Rainy | Foggy | Night | All | Rainy | Foggy | Night | All | Rainy | Foggy | Night |
| w/o SDG | 0.625 | 0.642 | 0.580 | 0.299 | 0.280 | 0.265 | 0.285 | 0.459 | 0.432 | 0.418 | 0.448 | 0.333 | 0.319 | 0.320 | 0.324 |
| $R_{s2r} = 0.5$ | **0.634** | **0.650** | **0.599** | 0.332 | 0.316 | 0.305 | 0.311 | 0.477 | 0.456 | 0.435 | 0.460 | **0.345** | **0.330** | **0.328** | **0.330** |
| $R_{s2r} = 1$ | 0.627 | **0.650** | 0.598 | **0.337** | **0.323** | **0.308** | **0.321** | **0.489** | **0.468** | **0.442** | **0.478** | 0.335 | 0.324 | 0.317 | **0.330** |

**Validation on Camera-based 3D Object Detection**

Source: https://github.com/nv-tlabs/Cosmos-Drive-Dreams

# COSMOS-Drive-Dreams

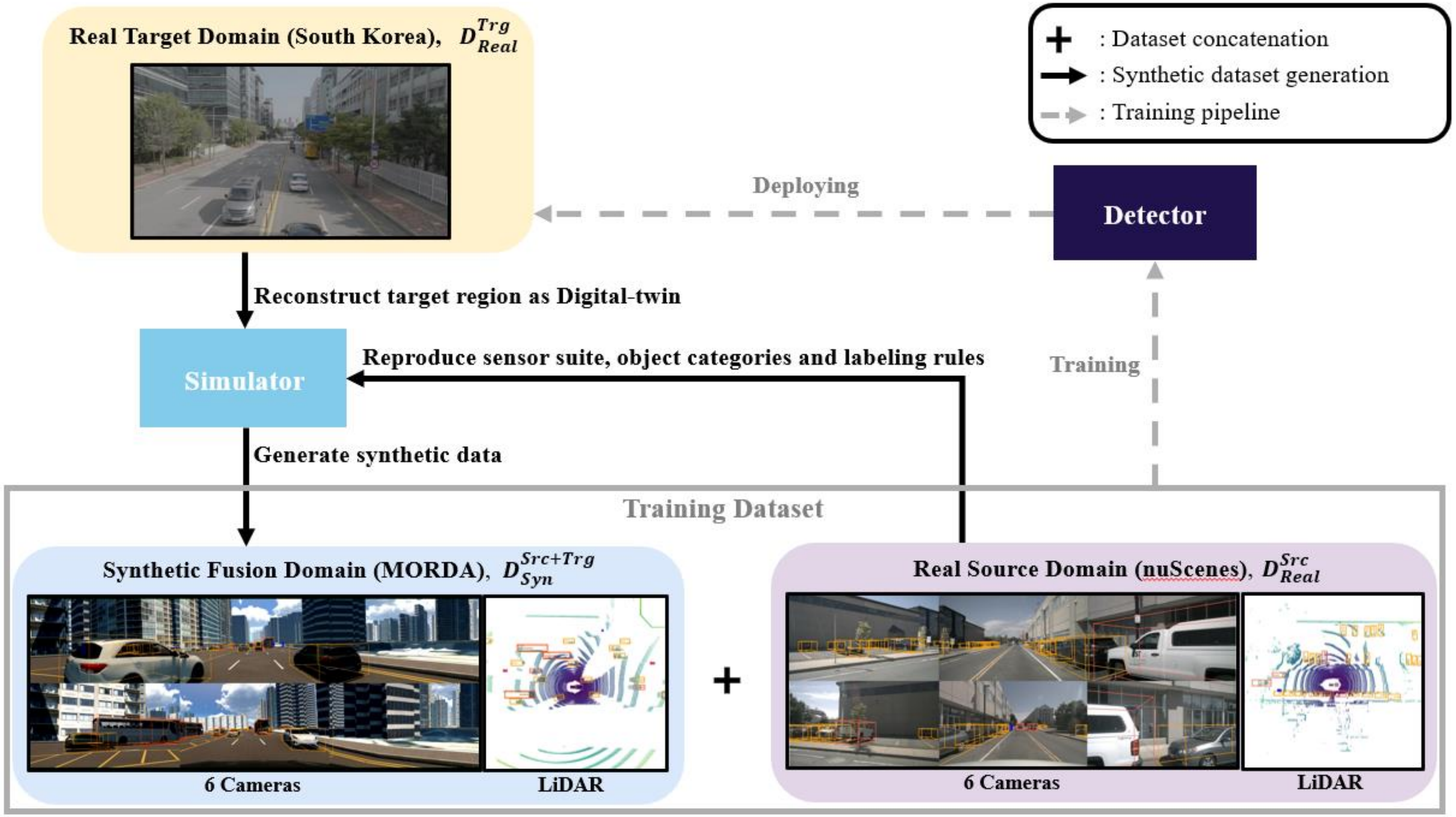
- Sim-to-Real Demo Using Cosmos-Transfer1
- 장점
  - 다양한 스타일의 이미지를 비교적 일관되게 생성 가능
    - 동일 영상내에서 그럴듯하게 합성
  - Sim-, Real- Data 모두 사용가능
  - Text Prompt를 이용한 날씨나 광조건 변경이 쉬움
  - Model & Weight 모두 공개
    - Post-training도 가능성

✓ 사실상 Inference-only
✓ 특정 날씨/광조건 극복을 위한 Augmentation 용으로 효율적

- 단점
  - 생각보다 번거로운 work-flow
    - Converting(Rendering) + Prompting -> Single-view -> Multi-view
  - Controllability 문제
    - 차량의 위치, 종류, 대수 등 변화
    - 신호등 모양 변화(미국 향 bias)
    - 시나리오는 잘 만들어져 있어야 함
  - 모델 크기 문제
    - 단일 GPU 기준 VRAM 80 GB 이상 필요
    - 40GB GPU 사용시, off-load 기능 이용하여 Inference만 가능
  - 속도 문제
    - 4초 분량 영상 합성을 위해 약 30분 소요

# MORDA (ICRA 2025) - Overview

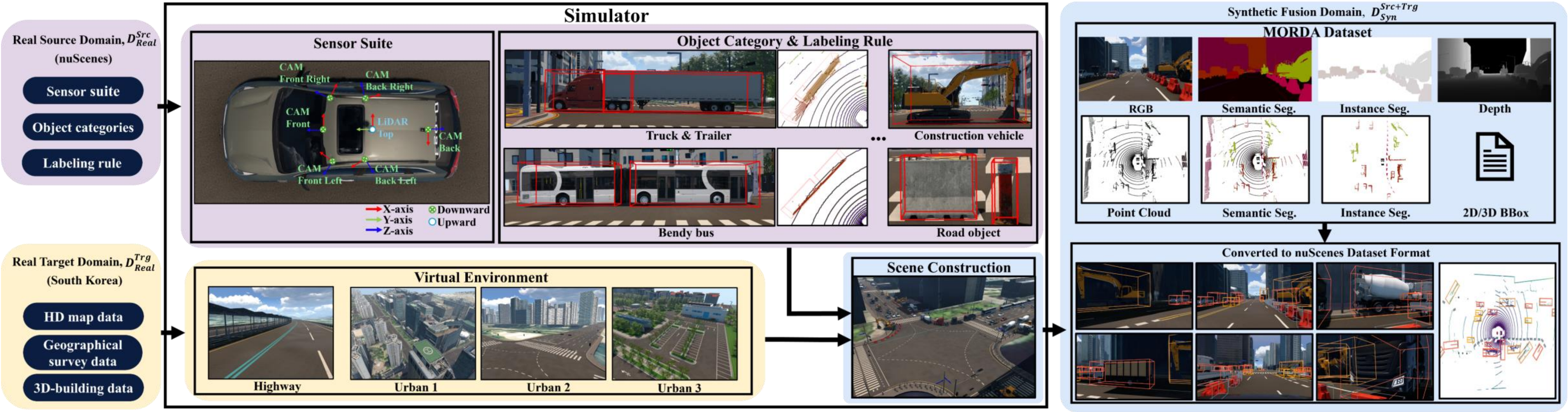
- Q. MORAI 데이터 학습에 도움되나요?
- In-distribution 및 Out-of-distribution 모두 성능 향상 확인




Source: https://arxiv.org/abs/2501.04950

# Method: Synthetic Fusion Domain

- Architecture of The Proposed Method

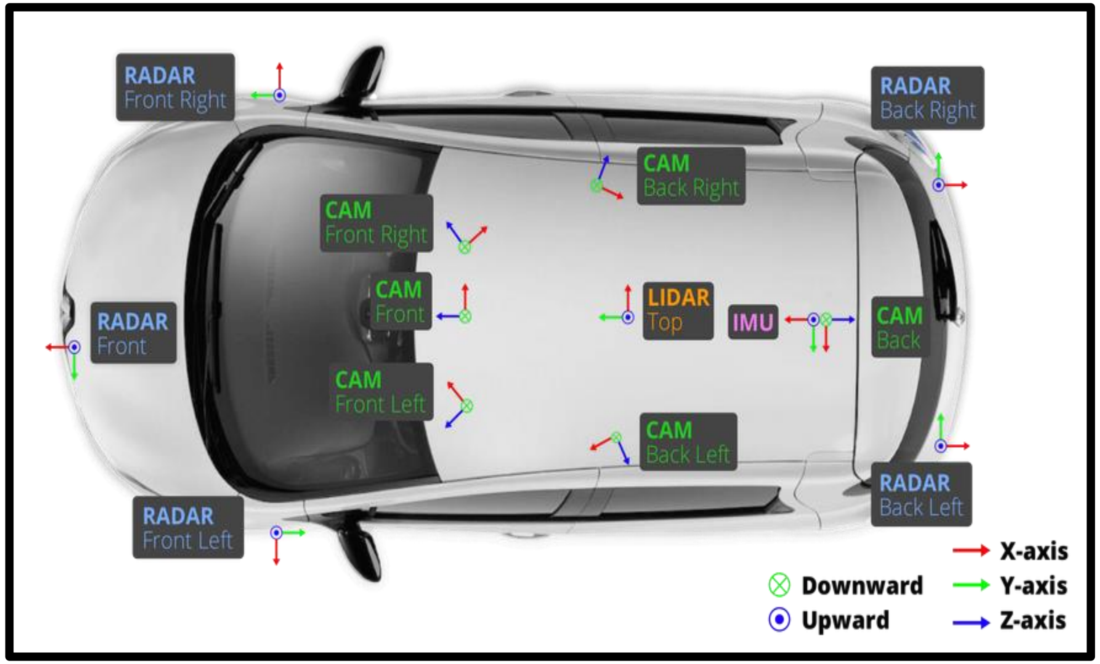



- **Reproducing characteristics (data-acquisition framework and labeling policy) of Real Source Domain, nuScenes ($D_{Real}^{Src}$)**
  - ✓ To suppress discrepancies in data distribution caused by different sensor spec, placement, and annotation rules.
- **Reproducing characteristics (map) of Real Target Domain, South Korea ($D_{Real}^{Trg}$)**
  - ✓ To reflect the geographical features of $D_{Real}^{Trg}$ in the virtual environment.
- **Blend the mentioned characteristics in a simulator to construct the Synthetic Fusion Domain ($D_{Syn}^{Src+Trg}$)**
  - ✓ To mimic real-world data that could be obtained by deploying sensors of $D_{Real}^{Src}$ to target region of $D_{Real}^{Trg}$.

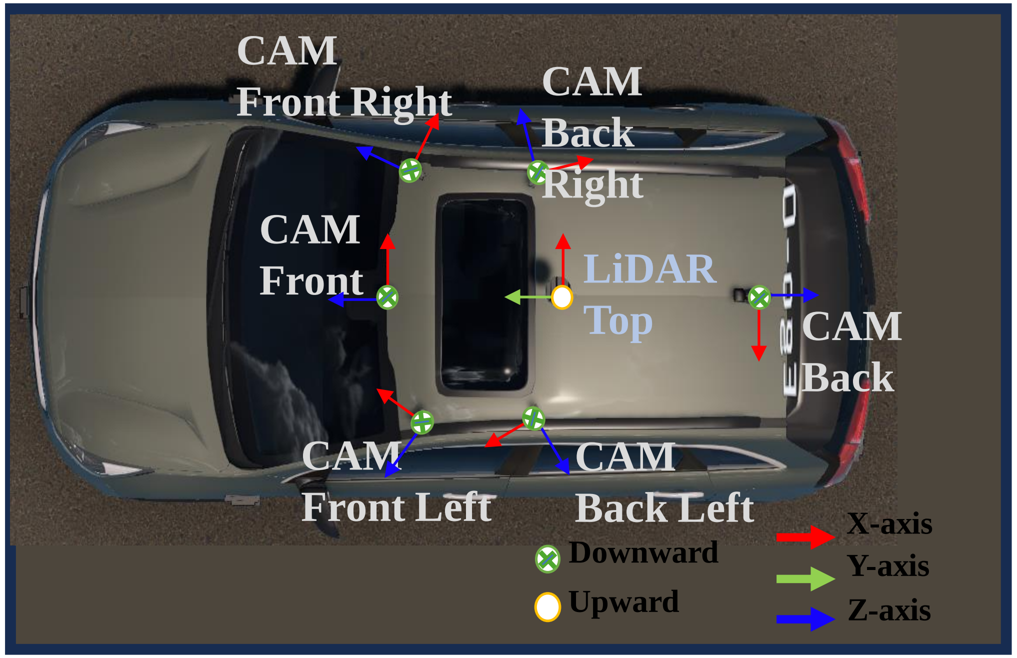
# Method: Synthetic Fusion Domain

- Reproducing Characteristics of Real Source Domain: Sensor Suite

**nuScenes**



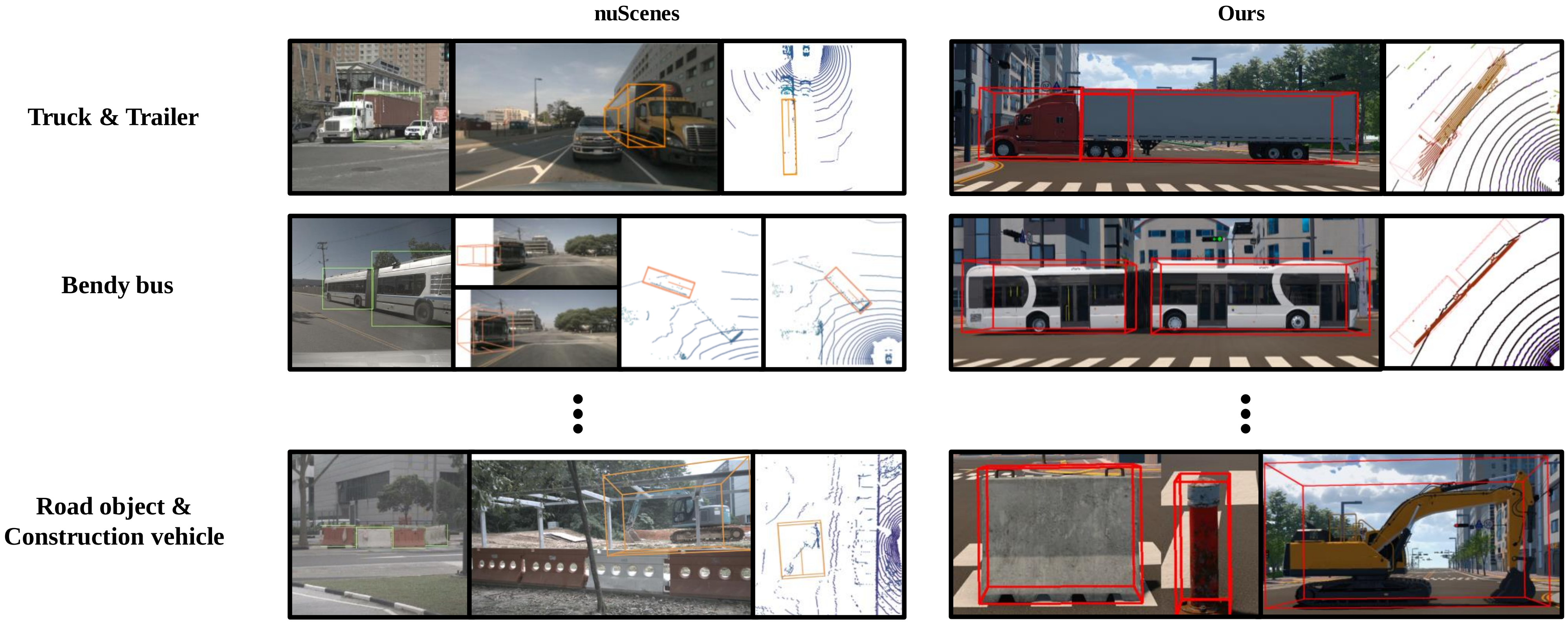

**Ours**




- **Sensor**
  - ✓ 6x cameras (1600 x 900 resolution)
  - ✓ 1x LiDAR (32-channel spinning type)
- **Placement (Position & Orientation)**
  - ✓ Following the configuration of nuScenes

# Method: Synthetic Fusion Domain

- Reproducing Characteristics of Real Source Domain: Object Category & Labeling Rule




- **10 detection classes of nuScenes are supported.**
  - ✓ Including rare classes like construction vehicles.
- **3D-BBox annotation policy is reimplemented.**
  - ✓ Including those for objects with two rigid sections (e.g., truck&trailer and bendy bus).

# Method: Synthetic Fusion Domain

- Reproducing Characteristics of Real Target Domain: Virtual Environment

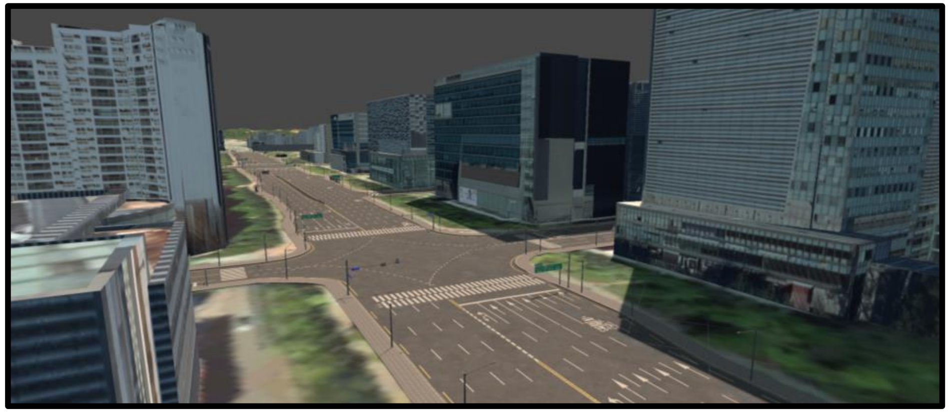
**Real Target Domain,** $D_{Real}^{Trg}$

**(South Korea)**

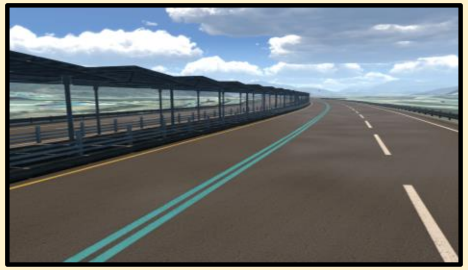
**HD map data**

**Geographical survey data**

**3D-building data**

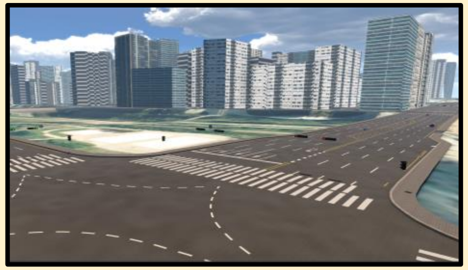


**Virtual Environment**



**Highway**

**Urban 1**



**Urban 2**

**Urban 3**

- In total, 4 digital-twin cities are leveraged.
- They are enriched with geographical features including lane markings, road surface, traffic infrastructure, and terrain.

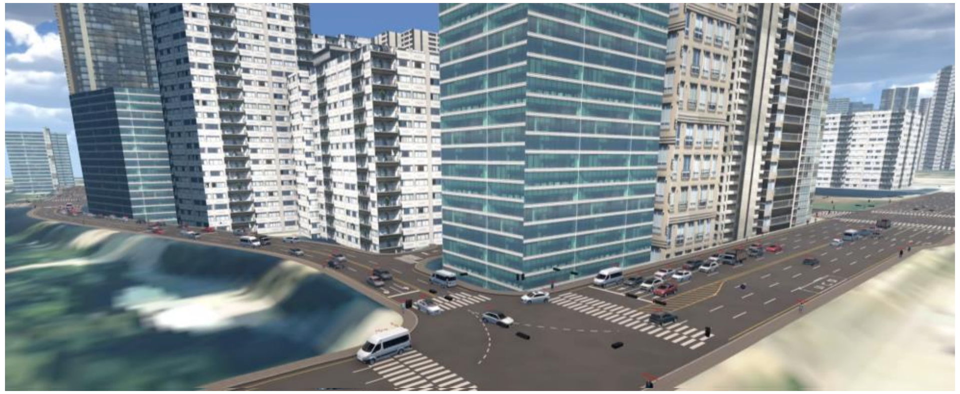
# Method: Synthetic Fusion Domain

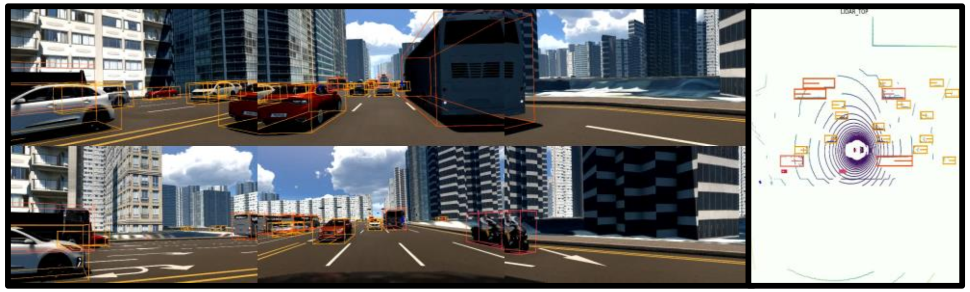
- Scene Construction & Sensor Data Generation

**Static Scene**

- Only ego vehicle is moving.
- Larger number of barrier, traffic cone, truck&trailer, construction vehicles.

**Dynamic Scene**





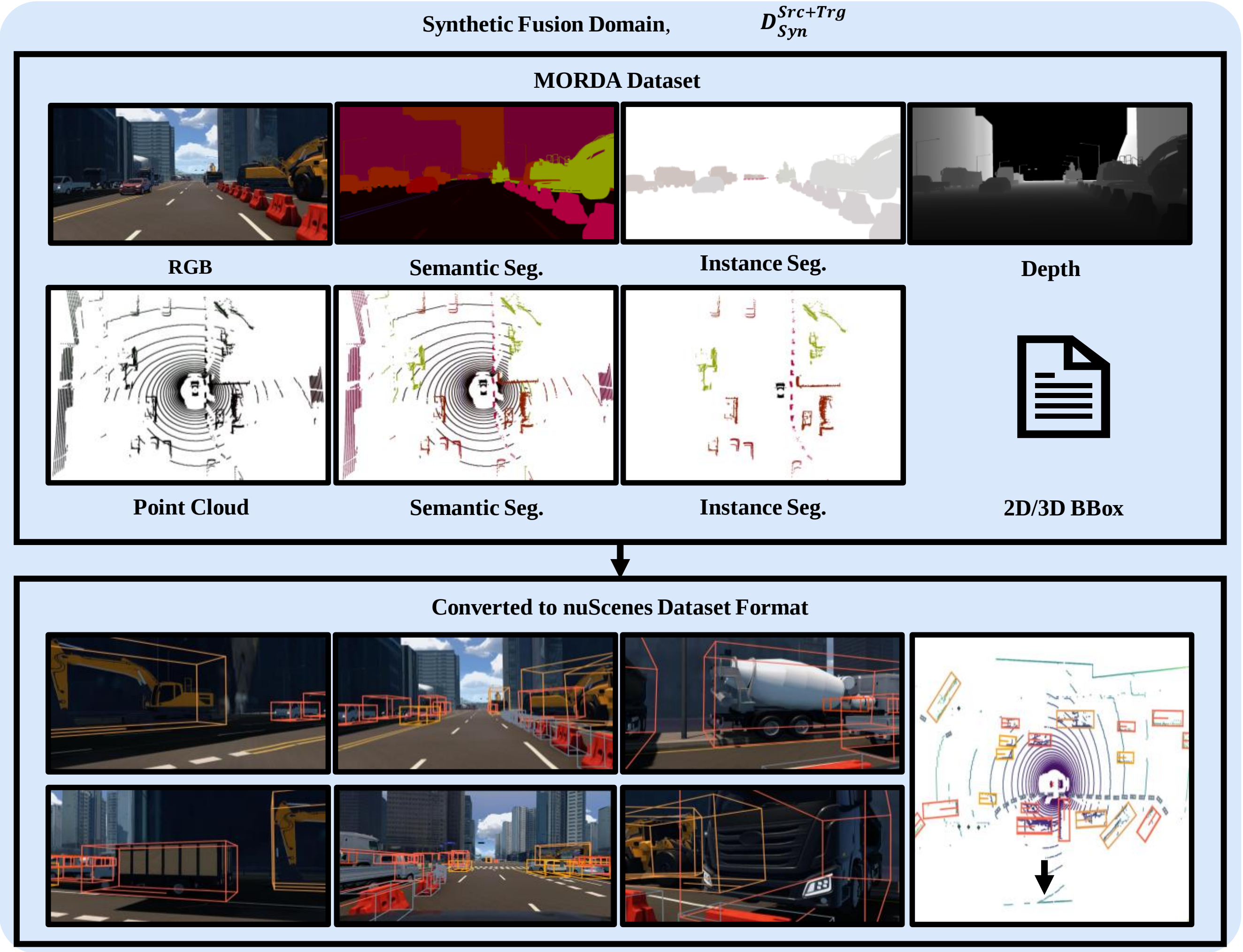
- Surrounding vehicles are also moving.
- Density around ego vehicle is maintained to mimic dense traffic scene.

# MORDA Dataset

- MORDA Dataset from Synthetic Fusion Domain

**Content**

- **Total Frames :** ~37K frames
- **Sensor data at 20Hz**
  - ✓ 6x camera image (1600 x 900 resolution)
  - ✓ 1x 32-channel spinning LiDAR point cloud
- **Corresponding GT labels**
  - ✓ Semantic segmentation
  - ✓ Instance segmentation
  - ✓ 2D / 3D BBox annotation
  - ✓ Pixel-wise depth
- **Vehicle Info**
  - ✓ Ego-vehicle pose data




**Conversion to nuScenes Format**

- **Frames :** ~3.7K keyframes and ~33K sweeps
- **Frequency:** 20Hz → 2Hz

# Experiment

- Datasets and Pre-processing

- **AI-Hub Dataset***
  - ✓ Employed to evaluate the efficacy of our method (MORDA).
  - ✓ Collected in real-target domain ($D_{Real}^{Trg}$) of this research, South Korea.
  - ✓ Real-world multimodal driving dataset (5x cameras, 1x LiDAR).
  - ✓ All data are used solely for validation, not training.
- **Pre-processing details**
  - ✓ Only images from the front camera, which co-exist in all datasets, are used.
  - ✓ For AI-Hub, images are cropped to 1600x900.
  - ✓ For AI-Hub, only annotations whose class labels are one of car, truck, bus, or pedestrian are kept.
  - ✓ Annotated frames in original format are divided into Keyframes (w/ annotation) and sweeps (w/o annotation), following nuScenes's frequency of 2Hz.

Summary of properties by dataset. The arrow indicates the applied pre-processing. T and V denote train and validation, respectively.

| | #Annot. Frame | #CAM. | Resolution | #Beam | #Det. Class | Freq. (Hz) |
|---|---|---|---|---|---|---|
| nuScenes (T) — $D_{Real}^{Src}$ | 28K | 6→1 | 1600×900 | 32 | 10 | 2 |
| **MORDA (T)** $D_{Syn}^{Src+Trg}$ | 37K→3.7K | 6 →1 | 1600×900 | 32 | 10 | 20 →2 |
| AI-Hub (V)* — $D_{Real}^{Trg}$ | 80K→8.8K | 5→ 1 | 1920×1200 →1600×900 | 128 | 8→4 | 10→2 |

# Results

Impact of MORDA in 2D/3D object-detection performance (trained and evaluated on nuScenes).

C, L, CV, and MC denote Camera, LiDAR, Motorcycle and Construction Vehicle, respectively. mAP is computed by averaging Aps for all 10 detection classes.

| Task | Modality | Model | Remark | Training Dataset | Evaluation Dataset nuScenes ($D^{Src}_{Real}$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Truck | Bus | Trailer | CV | MC | Bicycle | mAP† | NDS |
| 2D | C | Faster-RCNN | ✗ | nuScenes | 31.4 | 48.0 | 16.1 | 4.2 | 20.9 | 19.8 | 27.3 | ✗ |
| 2D | C | Faster-RCNN | ✗ | nuScenes + MORDA | 31.0 | 50.0 | 16.5 | 3.7 | 22.0 | 20.7 | **27.8** | ✗ |
| | | | | | –0.4 | +2.0 | +0.4 | –0.5 | +1.1 | +0.9 | +0.5 | ✗ |
| 3D | L | PointPillars | SECFPN | nuScenes | 39.2 | 51.7 | 28.6 | 5.4 | 21.5 | 0.9 | 34.94 | 50.02 |
| 3D | L | PointPillars | SECFPN | nuScenes + MORDA | 40.5 | 53.1 | 30.8 | 7.3 | 22.9 | 0.8 | **35.45** | **50.44** |
| | | | | | +1.3 | +1.4 | +2.2 | +1.9 | +1.4 | –0.1 | +0.51 | +0.42 |
| 3D | L | SSN | ✗ | nuScenes | 49.5 | 65.8 | 33.7 | 17.1 | 52.6 | 23.3 | 48.29 | 59.42 |
| 3D | L | SSN | ✗ | nuScenes + MORDA | 52.2 | 66.4 | 33.1 | 16.8 | 52.5 | 24.6 | **48.92** | **60.41** |
| | | | | | +2.7 | +0.6 | –0.6 | –0.3 | –0.1 | +1.3 | +0.63 | +0.99 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes | 49.0 | 63.3 | 31.4 | 10.9 | 41.4 | 18.6 | 48.96 | 59.42 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes + MORDA | 49.5 | 64.3 | 32.8 | 14.3 | 44.4 | 13.9 | **49.43** | **59.76** |
| | | | | | +0.5 | +1.0 | +1.4 | +3.4 | +3.0 | –4.7 | +0.47 | +0.34 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes | 53.2 | 66.5 | 36.0 | 15.0 | 55.2 | 36.8 | 56.23 | 64.51 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes + MORDA | 55.1 | 69.2 | 37.3 | 18.5 | 57.0 | 38.1 | **57.80** | **65.47** |
| | | | | | +1.9 | +2.7 | +1.3 | +3.5 | +1.8 | +1.3 | +1.57 | +0.96 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes | 55.1 | 67.9 | 34.9 | 15.2 | 55.8 | 36.1 | 56.95 | 65.40 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes + MORDA | 55.9 | 67.1 | 35.7 | 16.1 | 58.4 | 40.1 | **58.02** | **66.02** |
| | | | | | +0.8 | –0.8 | +0.8 | +0.9 | +2.6 | +4.0 | +1.07 | +0.62 |

- When MORDA is incorporated during training,
  - ✓ Classwise AP scores fluctuate but meaningful AP gain can be observed in general.
  - ✓ mAP and NDS scores on nuScenes consistently increase across modalities and networks.
- Overall, **MORDA does not degrade the detection ability on $D^{Src}_{Real}$ (nuScenes).**

# Results

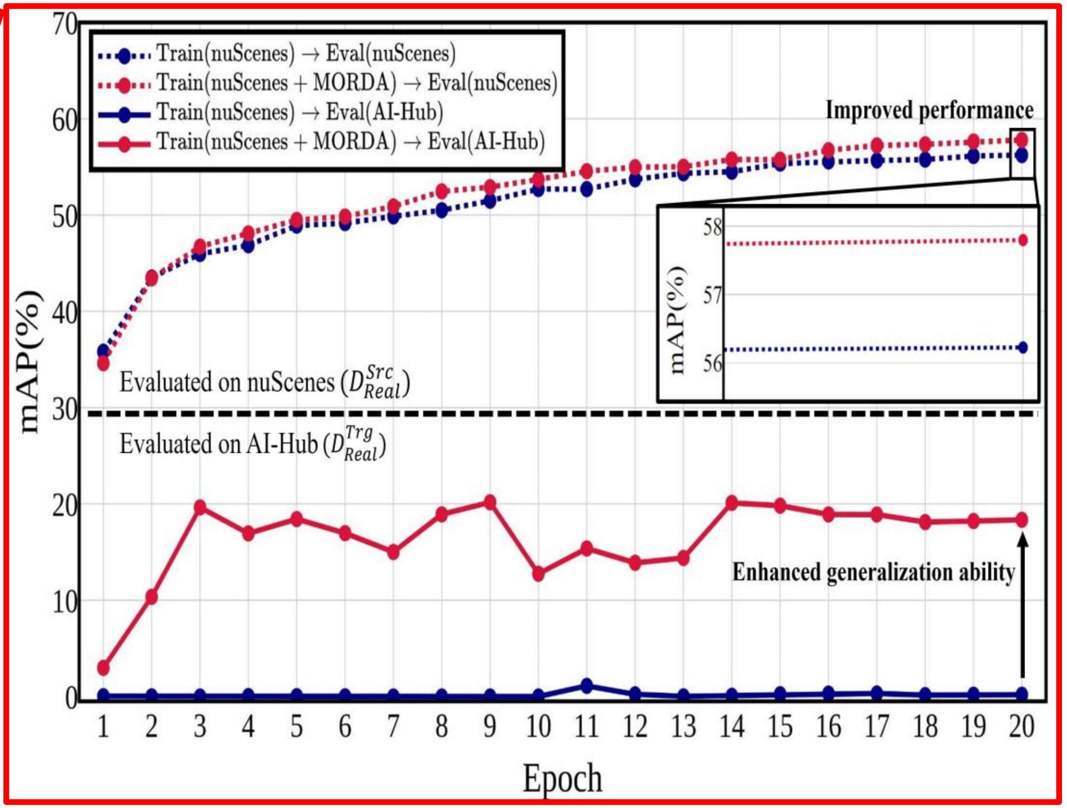
Impact of MORDA in 2D/3D object-detection performance (trained on nuScenes and evaluated on unforeseen AI-Hub).

| Task | Modality | Model | Remark | Training Dataset | Evaluation Dataset: AI-Hub ($D^{Trg}_{Real}$) Car | Truck | Bus | Ped | mAP‡ | NDS |
|---|---|---|---|---|---|---|---|---|---|---|
| 2D | C | Faster-RCNN | ✗ | nuScenes | 19.5 | 13.8 | 16.8 | 3.3 | 13.35 | ✗ |
| 2D | C | Faster-RCNN | ✗ | nuScenes + MORDA | 27.2 | 18.6 | 24.5 | 8.3 | **19.7** | ✗ |
| | | | | | +7.7 | +4.8 | +7.7 | +5.0 | +6.35 | ✗ |
| 3D | L | PointPillars | SECFPN | nuScenes | 24.93 | 3.425 | 11.70 | 10.93 | 12.74 | 21.95 |
| 3D | L | PointPillars | SECFPN | nuScenes + MORDA | 29.67 | 5.59 | 12.28 | 11.99 | **14.88** | **22.63** |
| | | | | | +4.74 | +2.165 | +0.58 | +1.06 | +2.14 | +0.68 |
| 3D | L | SSN | ✗ | nuScenes | 27.30 | 9.93 | 19.94 | 15.48 | 18.16 | 26.92 |
| 3D | L | SSN | ✗ | nuScenes + MORDA | 36.76 | 13.79 | 24.16 | 20.76 | **23.87** | **30.21** |
| | | | | | +9.46 | +3.86 | +4.22 | +5.28 | +5.71 | +3.29 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes | 0.0* | 0.0* | 0.065 | 0.0* | 0.016 | 2.64 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes + MORDA | 14.73 | 0.012 | 2.01 | 0.0* | **4.19** | **11.38** |
| | | | | | +14.73 | +0.012 | +1.945 | 0.0 | +4.17 | +8.74 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes | 0.717 | 0* | 0* | 0* | 0.179 | 7.203 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes + MORDA | 63.44 | 5.682 | 4.264 | 0* | **18.34** | **22.93** |
| | | | | | +62.72 | +5.682 | +4.264 | 0.0 | +18.16 | +15.72 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes | 0.197 | 0.0* | 0.0* | 0.0* | 0.049 | 4.19 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes + MORDA | 63.28 | 1.0 | 24.6 | 0.0* | **22.22** | **23.71** |
| | | | | | +63.08 | +1.0 | +24.6 | 0.0 | +22.17 | +19.52 |

- When MORDA is incorporated during training, mAP and NDS on AI-Hub increase across modalities and networks.
- In other words, **MORDA provides useful synthetic features for previewing and indirectly learning about unforeseen $D^{Trg}_{Real}$ (AI-Hub).**

# Results

| Task | Modality | Model | Remark | Training Dataset | Evaluation Dataset: AI-Hub ($D^{Trg}_{Real}$) Car | Truck | Bus | Ped | mAP‡ | NDS |
|---|---|---|---|---|---|---|---|---|---|---|
| 2D | C | Faster-RCNN | ✗ | nuScenes | 19.5 | 13.8 | 16.8 | 3.3 | 13.35 | ✗ |
| 2D | C | Faster-RCNN | ✗ | nuScenes + MORDA | 27.2 | 18.6 | 24.5 | 8.3 | **19.7** | ✗ |
| | | | | | +7.7 | +4.8 | +7.7 | +5.0 | +6.35 | ✗ |
| 3D | L | PointPillars | SECFPN | nuScenes | 24.93 | 3.425 | 11.70 | 10.93 | 12.74 | 21.95 |
| 3D | L | PointPillars | SECFPN | nuScenes + MORDA | 29.67 | 5.59 | 12.28 | 11.99 | **14.88** | **22.63** |
| | | | | | +4.74 | +2.165 | +0.58 | +1.06 | +2.14 | +0.68 |
| 3D | L | SSN | ✗ | nuScenes | 27.30 | 9.93 | 19.94 | 15.48 | 18.16 | 26.92 |
| 3D | L | SSN | ✗ | nuScenes + MORDA | 36.76 | 13.79 | 24.16 | 20.76 | **23.87** | **30.21** |
| | | | | | +9.46 | +3.86 | +4.22 | +5.28 | +5.71 | +3.29 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes | 0.0* | 0.0* | 0.065 | 0.0* | 0.016 | 2.64 |
| 3D | L | CenterPoint | Pillar (0.2m) | nuScenes + MORDA | 14.73 | 0.012 | 2.01 | 0.0* | **4.19** | **11.38** |
| | | | | | +14.73 | +0.012 | +1.945 | 0.0 | +4.17 | +8.74 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes | 0.717 | 0* | 0* | 0* | 0.179 | 7.203 |
| 3D | L | CenterPoint | Voxel (0.1m) | nuScenes + MORDA | 63.44 | 5.682 | 4.264 | 0* | **18.34** | **22.93** |
| | | | | | +62.72 | +5.682 | +4.264 | 0.0 | +18.16 | +15.72 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes | 0.197 | 0.0* | 0.0* | 0.0* | 0.049 | 4.19 |
| 3D | L | CenterPoint | Voxel (0.075m) | nuScenes + MORDA | 63.28 | 1.0 | 24.6 | 0.0* | **22.22** | **23.71** |
| | | | | | +63.08 | +1.0 | +24.6 | 0.0 | +22.17 | +19.52 |




- We observe that incorporating MORDA during training could mitigate this generalization failure.
- In short, MORDA contributes to **mitigating the impact of domain shift without degrading the detection performance on the trained domain (nuScenes)**.

# Results

Trained on nuScenes $(D^{Src}_{Real})$

nuScenes $(D^{Src}_{Real})$

Reliable detection performance

Deploying

AI-Hub $(D^{Trg}_{Real})$

Generalization failure

Car
Truck
Bus
Pedestrian
Others

Trained on nuScenes $(D^{Src}_{Real})$ + MORDA $(D^{Src+Trg}_{Syn})$

nuScenes $(D^{Src}_{Real})$

Reliable detection performance

Deploying

AI-Hub $(D^{Trg}_{Real})$

Better generalization and adaptation

# Results

Comparison on mAP gains for Faster-RCNN with respect to synthetic datasets.

mAP is calculated averaging AP scores for 10 detection classes of nuScenes.

| **Training Dataset** | **#Frame** | **mAP (%, ↑)** |
|---|---|---|
| nuScenes | 28K | 27.3 |
| nuScenes + VKITTI2 | 28K + 2.5K | 27.0 (–0.3) |
| nuScenes + SYNTHIA-AL | 28K + 7.3K | 27.4 (+0.1) |
| nuScenes + SHIFT | 28K + 30K | 27.7 (+0.4) |
| **nuScenes + MORDA (Ours)** | 28K + 3.7K | **27.8** (+0.5) |

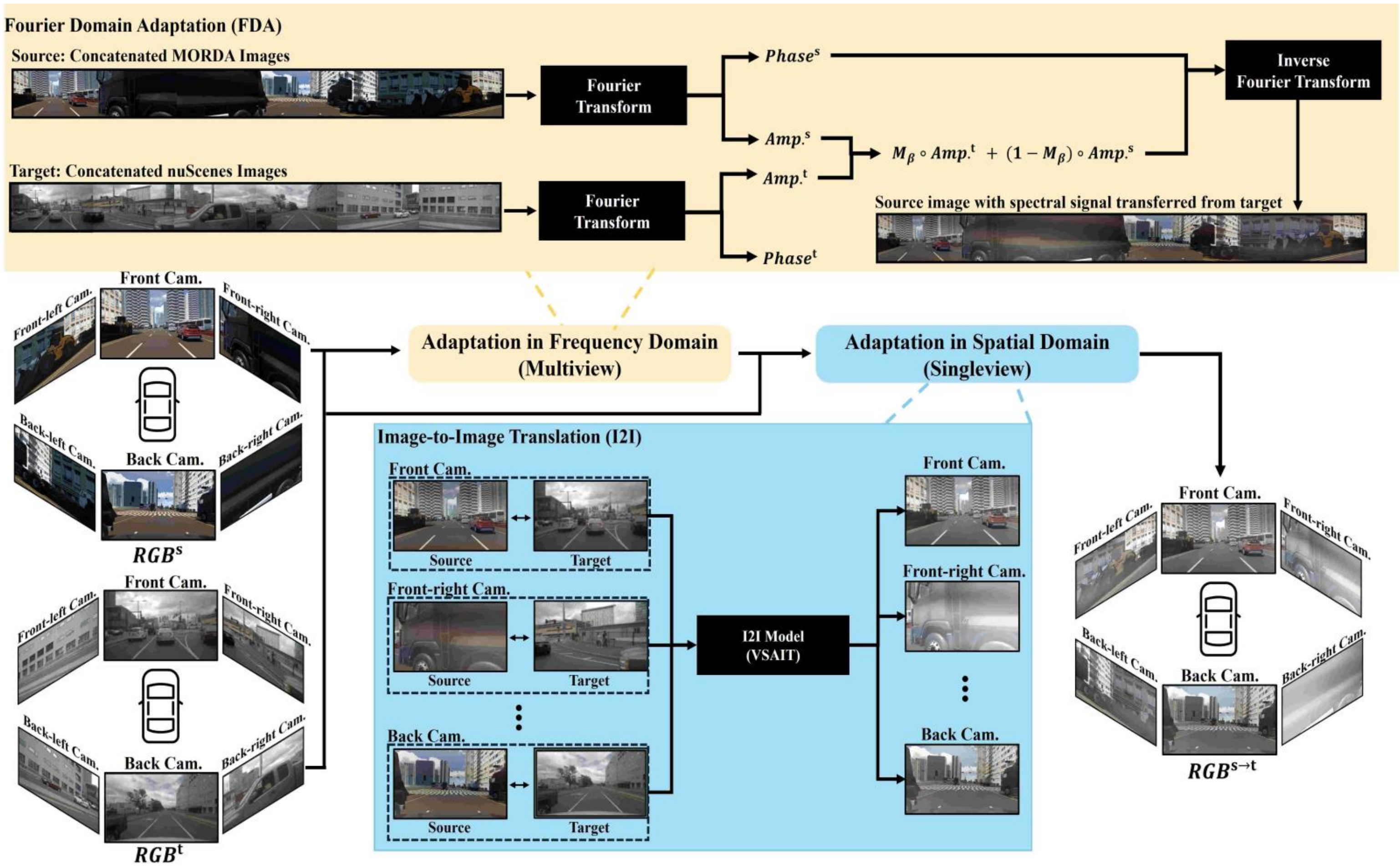
- **Compared with Virtual KITTI2 (VKITTI2), SYNTHIA-AL, and SHIFT Datasets.**
- **Pre-processing**
  - ✓ Only images from the front camera are used.
  - ✓ VKITTI2 (Assumed to be 10 Hz as KITTI), SYNTHIA-AL, and MORDA are downsampled to 2 Hz.
  - ✓ SHIFT is downsampled by a factor of 5 due to its exceptional size.
  - ✓ Class mapping
    - ✓ 2D BBoxes for Cyclist and Bicycle of SYNTHIA-AL are merged to bicycle.
    - ✓ 2D BBoxes for car and van of VKITTI2 are merged to car.

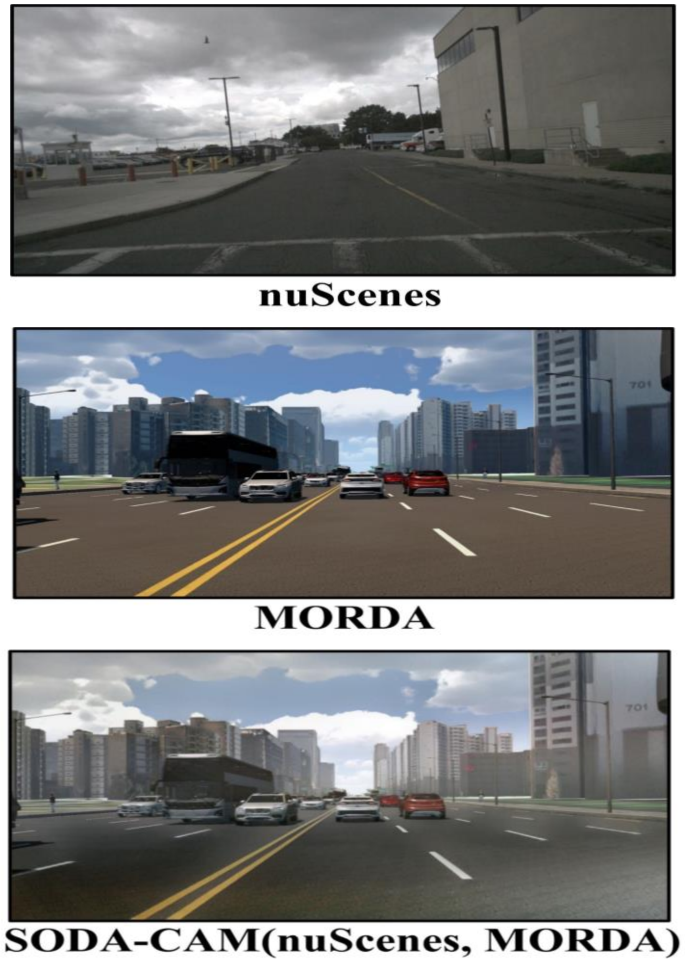
- **MORDA brings the biggest gain in mAP amongst compared synthetic datasets.**
  - ✓ SHIFT presents comparable gain but with ~8x more images.

# SODA (IV2025)

- Image-to-Image Translation (Camera Sim-to-Real)

SODA (IV2025)

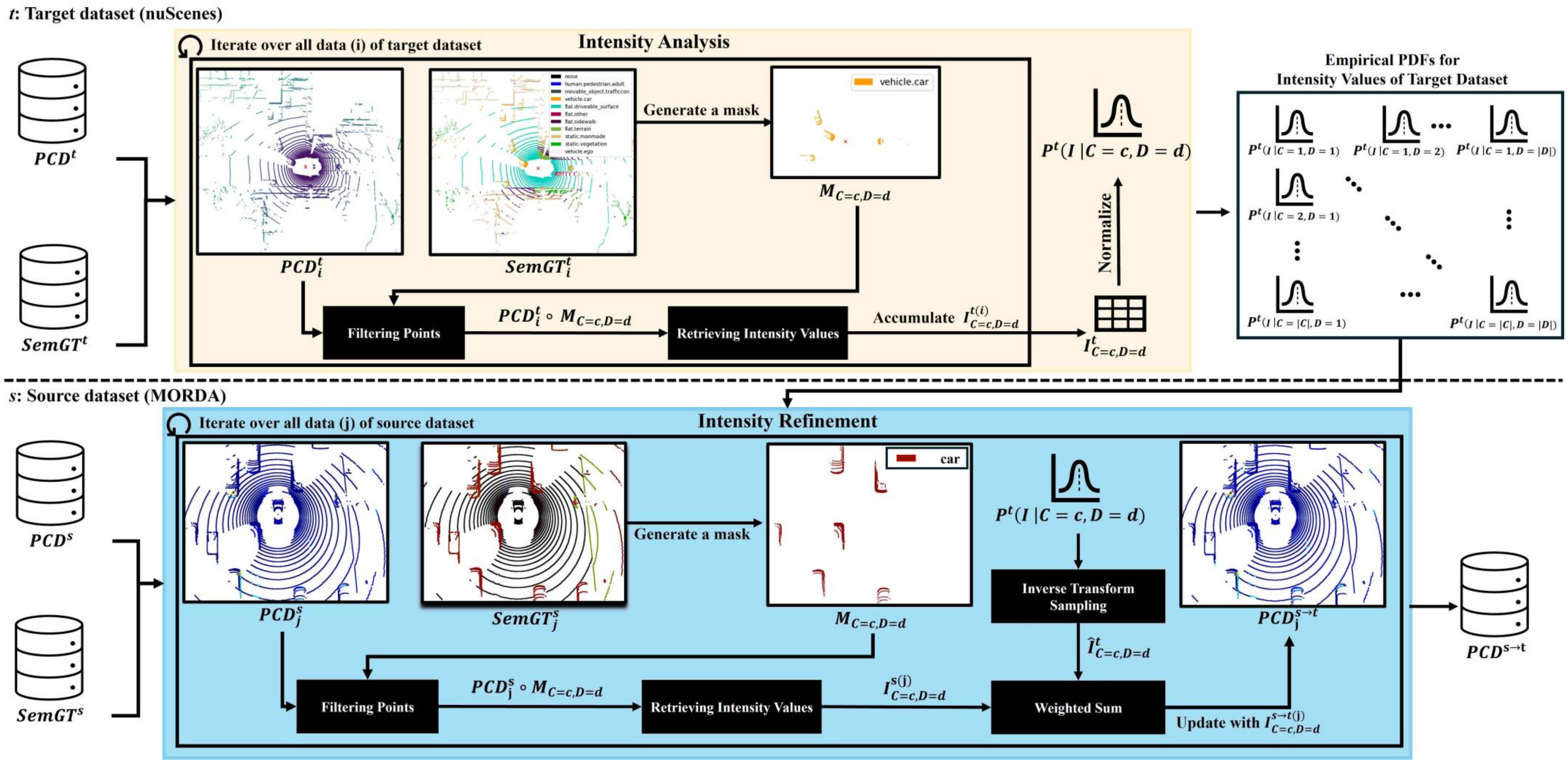



Sensor-aware Offline Synthetic-to-Real Adaptation for Multimodal Driving Datasets: Shape Preserved, Signal Improved, IV2025, Hojun Lim et al.

# SODA (IV2025)

- LiDAR Intensity Refinement (LiDAR Sim-to-Real)



# Results

| Training Dataset | Similarity | | | Perception | |
|---|---|---|---|---|---|
| | Image | Point Cloud | | BEVFusion (Camera + LiDAR) | |
| | FID (↓) | $W_1$ (↓) | $L_1$ (↓) | NDS (%, ↑) | mAP (%, ↑) |
| nuScenes | N/A | N/A | N/A | 70.4 | 66.8 |
| + MORDA | 73.1 | 0.3478 | 11.41 | **71.1** | 67.9 |
| **+ MORDA (SODA)** | **62.9** | **0.1606** | **6.68** | **71.1** | **68.3** |

***Wasserstein-1 Distance ($W_1$) and L1 distance compute the distance between two intensity distributions from MORDA and nuScenes. For better readability, $W_1$ scores are scaled by factor of 100.**

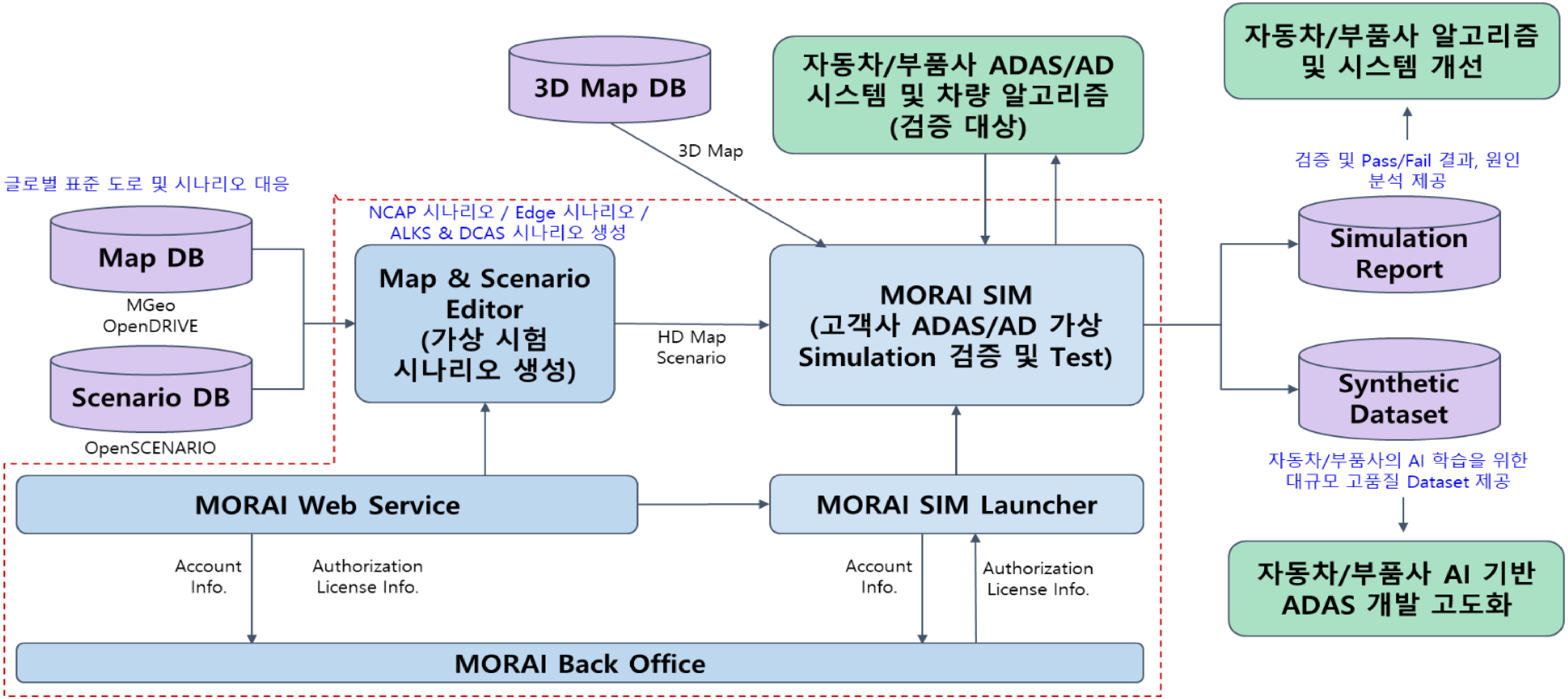
- ► No training strategy to explicitly bridge the domain gap other than simple dataset concatenation was used.
- ► Consistent Improvement of fidelity in image (FID), intensity distribution ($W_1$ and $L_1$), and perception performance (**mAP**)

# Results

| Training Dataset | Rainy | | Night | |
|---|---|---|---|---|
| | NDS (%, ↑) | mAP (%, ↑) | NDS (%, ↑) | mAP (%, ↑) |
| nuScenes | 72.0 | 67.5 | 45.0 | 41.2 |
| + MORDA | 71.8 (-0.2) | 68.5 (+1.0) | 45.3 (+0.3) | 41.4 (+0.2) |
| **+ MORDA (SODA)** | **72.1 (+0.1)** | **69.6 (+2.1)** | **46.1 (+1.1)** | **43.1 (+1.9)** |

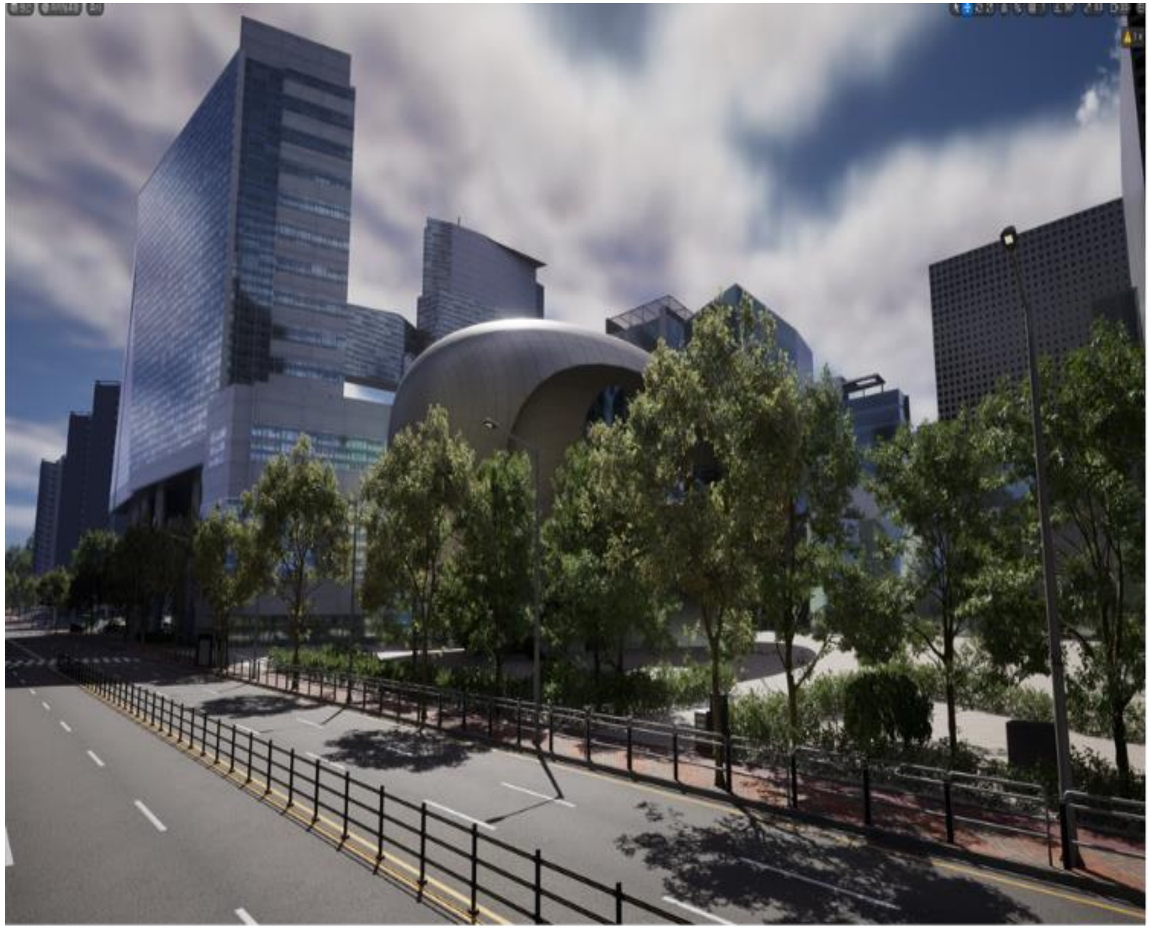
**Perception performance of BEVFusion evaluated on subsets (rainy and night) of nuScenes's validation set.**

# MORAI Virtual simulation 활용 구성



# Digital twin

상암 자율주행 지구 디지털 트윈



# MORAI SIM