

# 열악한 환경에서의 자율주행을 위한 다중센서 데이터셋 구축

## Build a Multi-Sensor Dataset for Autonomous Driving in Adverse Weather Conditions

심성대<sup>†</sup> · 민지홍<sup>1</sup> · 안성용<sup>2</sup> · 이종우<sup>2</sup> · 이정석<sup>2</sup> · 배광탁<sup>3</sup> · 김병준<sup>3</sup> · 서준원<sup>3</sup> · 최덕선<sup>1</sup>  
Sungdae Sim<sup>†</sup>, Jihong Min<sup>1</sup>, Seongyong Ahn<sup>2</sup>, Jongwoo Lee<sup>2</sup>, Jung Suk Lee<sup>2</sup>,  
Gwangtak Bae<sup>3</sup>, Byungjun Kim<sup>3</sup>, Junwon Seo<sup>3</sup>, Tok Son Choe<sup>1</sup>

**Abstract:** Sensor dataset for autonomous driving is one of the essential components as the deep learning approaches are widely used. However, most driving datasets are focused on typical environments such as sunny or cloudy. In addition, most datasets deal with color images and lidar. In this paper, we propose a driving dataset with multi-spectral images and lidar in adverse weather conditions such as snowy, rainy, smoky, and dusty. The proposed data acquisition system has 4 types of cameras (color, near-infrared, shortwave, thermal), 1 lidar, 2 radars, and a navigation sensor. Our dataset is the first dataset that handles multi-spectral cameras in adverse weather conditions. The Proposed dataset is annotated as 2D semantic labels, 3D semantic labels, and 2D/3D bounding boxes. Many tasks are available on our dataset, for example, object detection and driveable region detection. We also present some experimental results on the adverse weather dataset.

**Keywords:** Multi-Sensor, Dataset, Autonomous Driving, Calibration, Adverse Weather

### 1. 서론

최근 무인차량 등의 연구가 활발히 진행되고 있다. 무인차량의 연구를 위해서는 자율주행을 위한 기술연구가 필수적이다. 과거에는 규칙기반의 알고리즘을 이용한 방법으로 가상환경 및 실제 환경에서 시험이 이루어지던 방식이었다. 하지만 영상을 활용한 인식이 딥러닝 기술의 발달에 따라 성능이 크게 좋아지면서, 딥러닝을 이용한 자율주행 기술 개발이 주로 이루어지고 있다. 또한, 딥러닝 기술은 현재 영상뿐만 아니라, 라이다 등의 데이터에도 적용되고 있다. 이러한 학습기반의



[Fig. 1] Data acquisition in adverse condition (examples of snowy and smoky conditions)

Received : May. 17. 2022; Revised : Jun. 2. 2022; Accepted : Jun. 20. 2022

※ This project was funded by Defense Acquisition Program Administration (DAPA)

※ The dataset can be served by an individual request and official procedures

1. Principal Researcher, ADD, Daejeon, Korea (happymin, tschoe@add.re.kr)

2. Senior Researcher, ADD, Daejeon, Korea (seongyong.ahn, jongwoo\_lee, jslee2011@add.re.kr)

3. Research Officer, ADD, Daejeon, Korea (gtbae, bjkim95, givenone@add.re.kr)

† Senior Researcher, Corresponding author: ADD, Daejeon, Korea (sdsim@add.re.kr)

알고리즘을 활용하기 위해서는 자율주행을 위한 센서 데이터가 확보되어야 한다. 자율주행을 위해 활용되는 대표적인 센서 데이터셋은 KITTI 데이터셋<sup>[1]</sup>이 있다. 최근에는 Cityscape<sup>[2]</sup>, Waymo dataset<sup>[3]</sup>, Argoverse dataset<sup>[4]</sup>, nuScenes dataset<sup>[5]</sup> 등이 대표적인 자율주행용 데이터 셋으로 적용되고 있다.

KITTI<sup>[1]</sup> 데이터셋은 자율주행을 위한 데이터셋의 시작이라고 볼 수 있다. 64채널의 라이다와 칼라 및 흑백 영상을 도심지에서 데이터를 획득하고, 물체 탐지, 위치추정 등에 주로 활용되며, 점차 활용분야를 넓혀가고 있다. Cityscape<sup>[2]</sup> 데이터는 차량시점에서 획득한 영상을 시맨틱분할의 형태의 라벨과 함께 제공한다. 자율주행 연구가 활발해지면서 Waymo, Argo AI, nuTonomy와 같은 자율주행을 연구하는 회사에서 Waymo<sup>[3]</sup>, Argoverse<sup>[4]</sup>, nuScenes<sup>[5]</sup> 데이터셋을 만들어 좀 더 다양한 환경에 대하여 제공하고 있다. 앞선 데이터셋은 일반적인 상황을 고려하여, 다중스펙트럼의 영상 등을 고려하고 있지 않다. KAIST<sup>[6]</sup> 데이터셋은 RGB 영상 및 열영상을 동시에 획득하여 다중스펙트럼의 영상을 활용하는 데이터셋을 제시하였으나, 라벨이 부족하여, 다양하게 활용되기 어렵다.

앞서 언급한 자율주행용 데이터셋은 대부분 상대적으로 좋은 환경에 대해서 데이터를 포함하고 있고, 눈이 오거나, 먼지나 연막이 생기는 등의 상황에 대한 고려는 전혀 없다. 하지만, 비포장길, 화재상황, 강설 등 특수목적이나 상황에서 자율주행이 필요한 경우, 이에 대한 고려가 필수적이다. 따라서 강설 등의 열악한 환경에서 자율주행 기술의 성능이 어느 정도인지 파악하고 관련 기술의 발달을 추구하기 위하여 열악한 환경의 데이터가 필요하다.

본 연구는 열악한 환경에서의 자율주행 관련 기술의 개발 및 평가를 위하여 [Fig. 1]과 같이 센서 시스템을 구성하여 강설, 강우, 연막, 먼지가 존재하는 환경에서의 멀티센서 데이터를 구성하였다. 또한, 다양한 센서의 특성을 확인하기 위하여 일반적인 자율주행용 데이터셋에서 제공하는 가시광영역의 영상 뿐만 아니라 근적외선 영역, 단파장 적외선, 열영상 등의 다중 스펙트럼의 카메라와 라이다 및 레이더를 함께 구성하여 데이터를 획득한다.

다중 스펙트럼의 영상 데이터와 라이다, 레이더 등의 데이터를 획득하기 위해서는 센서 데이터 간의 동기화, 센서좌표 간의 캘리브레이션 등의 기술이 필요하다. 섹션 2에서는 다중 스펙트럼 데이터 획득을 위한 센서 획득장치의 구성, 센서동기화 및 캘리브레이션에 대해 기술하고, 섹션 3에서 센서 데이터셋의 구성에 대해 기술한다. 마지막으로 섹션 4에서 본 데이터셋을 활용한 물체인식 및 주행가능영역 인식의 열악한 환경에서의 영향성에 대한 시험결과를 제시한다.

## 2. 센서 데이터 획득 장치

### 2.1 센서 구성

#### 2.1.1 다중 스펙트럼 영상 센서 구성

딥러닝의 발달과 함께 가장 많이 활용되는 센서는 영상센서다. 사람이 보고 판단하는 것과 유사하게 영상에서 정보를 추출하는 점 때문에 영상의 활용 가능성에 대해 오랜 기간 연구되고 있다. 또한, 카메라의 약점에도 불구하고 카메라의 가격이 다른 센서에 비하여 상대적으로 저렴하고, 색상 등 풍부한 정보를 제공할 수 있다는 점에서 많이 활용되고 있다.

일반적으로 다른 자율주행 데이터셋은 가시광 영역의 카메라를 활용하고 있다. 일반적인 영상으로, 다양한 영상처리 알고리즘 등을 활용할 수 있는 장점이 있으나, 본 논문에서 다루는 바와 같은 열악한 환경에서는 그 성능이 제한될 수 있다. 특히, 야간에 제공하는 정보가 부족하고, 강우, 강설 등에 따른 렌즈의 이물질에 의하여 정보가 왜곡될 수 있는 점이 있다. 따라서, 이를 극복할 가능성이 있는 다른 파장 대역의 영상을 활용할 수 있다면, 더 좋은 자율주행 알고리즘 개발에 활용할 수 있다. 본 논문에서는 이러한 특성을 확인하기 위하여 가시광 영역의 영상 이외에 근적외선영상(Near Infrared (NIR)), 단파장영상(Shortwave infrared (SWIR)), 열영상(Thermal (LWIR))을 동시에 포함시켰다. 각 영상의 특징으로, 근적외선 영상은 저조도 환경 등에서 유리한 특성을 가지고 있으며, 단파장 영상은 해무, 안개, 연막 등에 투과성이 상대적으로 높으며, 열영상 대비 선명도가 높기 때문에 활용성이 있다. 다만, 해당 파장대의 빛이 존재해야 하므로 일반적인 자연광 또는 특수 조명이 필요한 단점이 있다. 마지막으로 열영상은 앞서 다른 카메라 영상과 다르게, 물체에서 발산하는 원적외선 영상을 측정하여 온도정보를 파악함으로써, 가시광 영상 대비 야간에 유리하며, 조명환경에 영향이 적다는 장점이 있다. 반면, 해상력이 떨어지는 단점이 있다.

이러한 다중 스펙트럼의 영상을 하나하나 따로 관리하게 되면 다중 영상을 동시에 활용하여 정보를 융합하기에 어려움이 존재한다. 이를 위해서 센서 간의 캘리브레이션을 활용하게 되는데, 영상은 2차원의 정보로 표현되므로, 각 영상 간의 시차에 의한 오차를 최대한 줄이기 위해서 영상 센서의 배치에서 서로 간의 거리가 최소화되도록 할 필요가 있다.

#### 2.1.2 거리 센서 구성

자율주행을 위한 거리센서는 여러 데이터셋에서 라이다가 메인센서로 활용되고 있으며, Waymo, nuScenes 데이터셋 등은 레이더도 추가적으로 적용하고 있다. 본 논문에서 제시하

는 데이터셋에서의 거리센서는 차량 상부에 32채널의 회전형 라이더를 배치하고, 차량의 전방과 상부에 각각 레이더를 독립적으로 설치하는 구성으로 하였다.

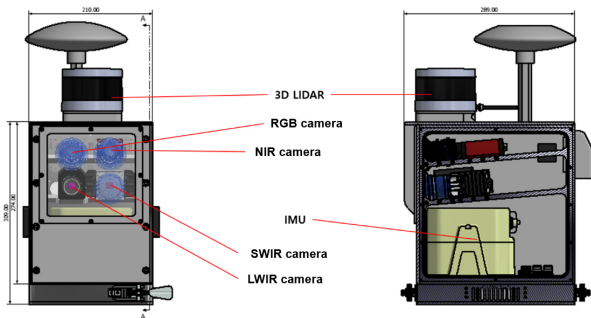
상부의 라이더는 상대적으로 부피가 작으면서 다중 채널을 제공하는 32채널의 증가형 3차원 라이더(VLP-32MR)를 활용하였으며, 레이더는 Smartmicro사의 UMR-96 FMCW 레이더를 적용하였다. 라이더와 레이더 모두 주변 광원에 영향을 받지 않으므로 낮밤에 상관없이 데이터 획득이 가능하다. 하지만, 라이더의 경우 먼지나 연막 등의 환경에서 투과성이 없으나 해상도가 높은 특징을 가진다. 반면, 레이더는 해상도는 낮으나, 악천후 등에서 투과성이 좋은 특성을 가지므로, 두 가지 모두 거리센서로 활용하는 것이 열악한 환경에서의 기술개발에 활용될 수 있다.

### 2.1.3 항법 센서 구성

항법센서는 GPS와 IMU를 통합하고, 차량의 휠인코더 정보를 활용하여 사후처리를 통하여 항법 데이터를 보정하는 방식으로 구성하였다. 이를 통하여 위치오차 2 cm 및 자세오차 0.02도 수준의 항법 데이터를 획득할 수 있다. 따라서 항법에서 발생하는 오차를 최소화하여 센서 데이터의 활용성을 높도록 설계하였다.

### 2.1.4 센서 배치

데이터 획득을 위한 센서 배치에서 고려해야 하는 점은 센서 간의 거리를 최소화하여 시차(parallax)로 인한 오차를 줄여야 한다는 것이다. KAIST 다중센서 데이터셋<sup>6)</sup>의 경우 빔스플리터를 이용하여 두 영상을 일치화하는 방법을 사용하였다. 하지만, 본 연구에서 활용되는 4대의 카메라를 빔스플리터를 사용하여 영상을 획득하는 것은 다음의 세 가지 문제가 존재한다. 첫째, 배치에서 센서의 부피를 크게 만들게 된다. 둘째, 받아들이는 광원의 에너지도 분산되게 된다. 셋째, 차량이 비포장 주행 등 외부 충격에 의한 영향을 쉽게 받을 수 있다.



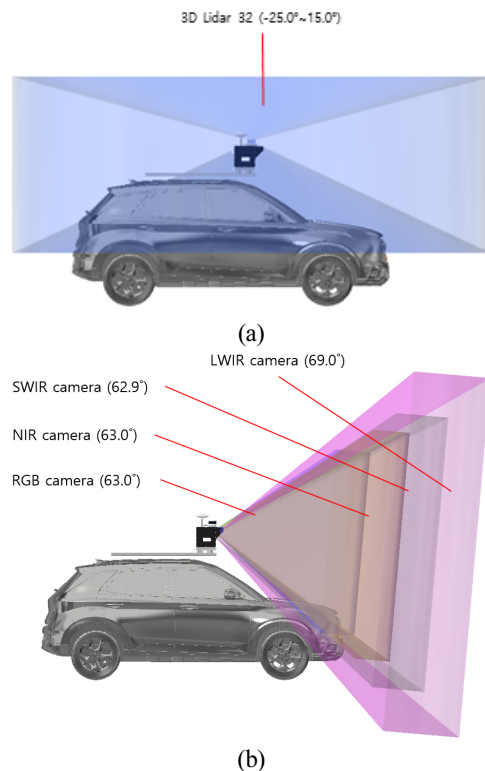
[Fig. 2] Sensor configuration with minimum gap of multiple sensors: 4 imaging sensors, 1 lidar and navigation sensors. 2 radars are installed on the vehicle

본 연구에서는 4대의 카메라의 배치를 물리적 거리를 최대한 가까이하는 방식으로 진행하여 불필요한 빔 스플리터 등을 사용하지 않도록 하였다. 이로 인하여 작은 크기의 센서모듈을 구성할 수 있었다. 또한 라이더도 가까운 위치로 배치하여 카메라와 라이더 사이의 시차를 최소화 하도록 설계하였다. 또한, 항법 센서를 영상과 라이더 센서모듈과 같이 통합하여 항법과 센서 사이의 캘리브레이션을 적용하였다. [Fig. 2]는 본 연구에서 설계된 센서 배치이다.

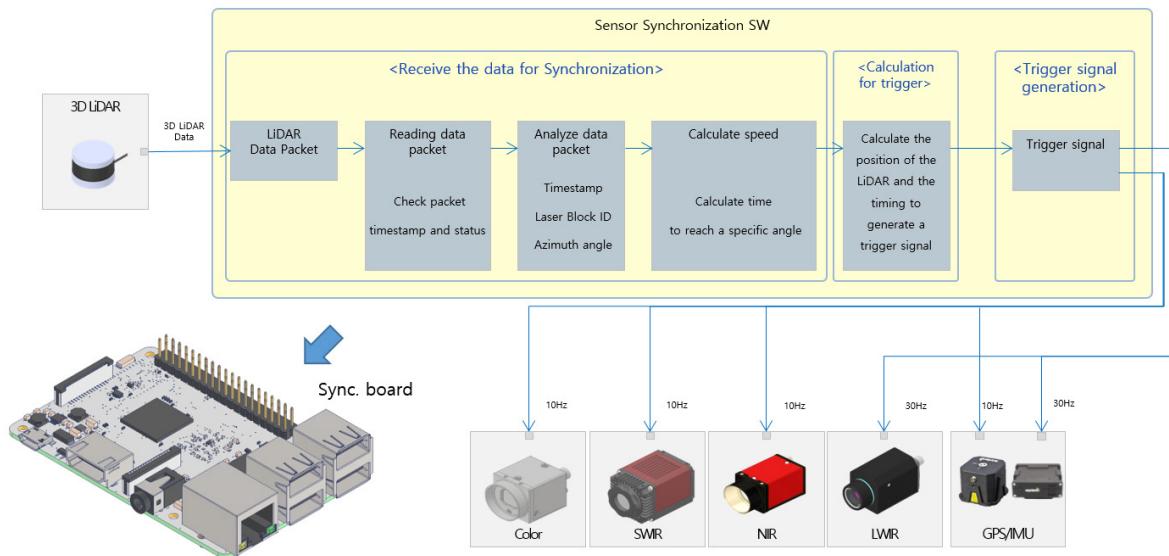
대부분의 자율주행 데이터셋에서 정면을 바라보는 영상을 획득하는데 본 연구의 센서 배치시 영상은 하향 5도를 지향하도록 하였다. 이를 통하여 영상에서 실제 주행에 활용되는 지면 및 물체 등에 좀더 많은 부분을 획득할 수 있다. 또한, 라이더의 수직시야각의 중심과 일치할 수 있다. 센서 배치를 통하여 분석된 센서 데이터의 획득 범위는 [Fig. 3]과 같이 표현될 수 있다.

### 2.2 센서 동기화

자율주행용 센서 데이터셋을 획득하는데 가장 중요한 부분 중의 하나는 센서 간의 동기화이다. 센서 간의 동기화를 하는 방법은 크게 두 가지로 구분할 수 있다. 첫째, 센서데이터 획득 시간을 기준으로 서로 가장 가까운 시간대의 데이터를 연관시



[Fig. 3] Vertical field of view of sensors: (a) lidar FOV is vertically 40° and horizontally 360°, (b) vertical FOV of four cameras are similar with about 65°



[Fig. 4] Synchronize multiple sensors using lidar rotating angle by analyze packet analysis. Synchronization algorithm is implemented on a tiny circuit board (Rasberry Pi 3)



[Fig. 5] Near IR image according to trigger timing: (a) 0° trigger timing, (b) 45° trigger timing, When lidar shots forward, the return of laser affects on near IR image

키는 것이다. 둘째, 하드웨어적으로 트리거 신호를 생성하여 센서 획득 시간의 동기화를 수행하는 것이다. 본 논문에서는 두 번째의 방법을 적용하되, 하드웨어적인 동기화를 완전히 지원하지 않는 LWIR 영상은 초기 동기화 트리거 이후 첫 번째의 방법을 이용하여 동기화 된 영상으로 구성하였다. 하드웨어적인 방법을 사용하는 방법은 주행 중에 획득하는 데이터 셋이므로, 상호 간의 데이터 획득 시간 오차가 적을수록 보다 정확한 데이터셋을 얻을 수 있고, 센서 간의 외부 파라미터만 고려하면 되기 때문이다.

센서를 동기화하는 방법은 [Fig. 4]와 같이 하드웨어 및 소프트웨어를 구성하고, 자율주행의 메인센서인 3차원 라이다를 기준으로 하여, 회전식 라이다의 특정 회전각도를 읽고, 이를 바탕으로 라이다의 스캐닝 영역을 계산하고, 그 타이밍을 맞춰 영상 센서의 트리거를 보내는 방식으로 구성된다. 이때, 항법 신호를 같이 수신하여, 해당하는 타임스탬프 및 항법 정보를 포함하도록 한다. LWIR 카메라(FLIR A65)는 외부 트리

거에 의한 임의적인 프레임 조정은 불가하나, 트리거를 기점으로 동작할 수 있어, 동기화 신호를 입력으로 하여 동일한 주기로 획득하도록 하였다.

센서 동기화를 통하여 영상을 획득하는 시간과 3차원 라이다 등의 데이터를 획득하는 시간을 맞추어 동일한 시간대의 데이터 획득이 가능해지는데, 3차원 라이다의 파장이 NIR 카메라의 감지파장과 겹치는 문제로 [Fig. 5]의 (a)와 같이 NIR 영상에 레이저빔이 표기되어 영상이 왜곡되는 현상이 있을 수 있다는 것이다. 이를 해결하기 위하여 영상의 트리거 타이밍을 NIR에 레이저가 나타나지 않도록 하면서 서로 가장 가까운 시간대에 트리거 신호를 생성하도록 계산하여 적용하였다. 이에 대한 알고리즘은 [Fig. 4]의 왼쪽 아래와 같이 Raspberry Pi 3에 구현되어 별도의 동기화 기능을 수행하도록 하였다.

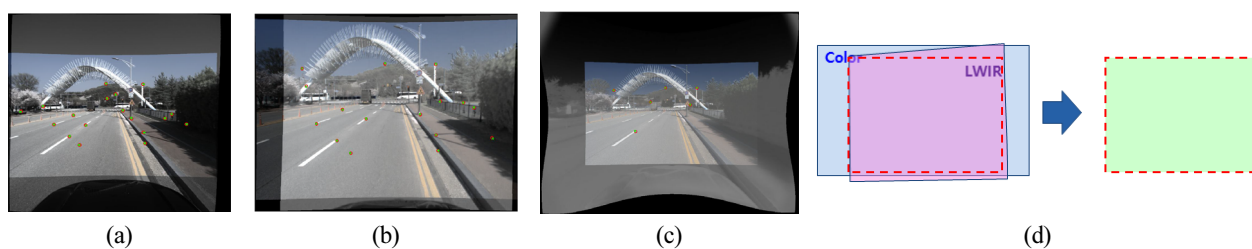
### 2.3 센서 간 캘리브레이션

#### 2.3.1 다중 카메라 간 캘리브레이션

본 데이터셋에서 적용되는 다중 스펙트럼 영상을 활용하기 위해서 다중 카메라 간의 캘리브레이션이 필요하다. 먼저 개별 카메라의 카메라 내부 파라미터는 일반적인 카메라 캘리브레이션 방법<sup>[7,8]</sup>을 이용하여 구한다. 여기에서 추정된 렌즈 왜곡 파라미터를 바탕으로 렌즈왜곡이 보정된 영상으로 변환한다.

다중 카메라를 이용할 때, 서로 간의 정보를 융합하기 위해서는 데이터를 하나의 공통된 축으로 융합을 하여야 한다. 이때, 빔스플리터 등을 이용하여 동일한 광축을 가지도록 설계를 하는 방법이 있을 수 있으나, 앞서 배치 관련 설명에서 공간





[Fig. 6] Fusion of two images using homography matrix (a) NIR to RGB, (b) SWIR to RGB, (c) LWIR to RGB, (d) Crop the common region after fusing images



[Fig. 7] Projection of point clouds onto RGB image (example in snowy condition): colored points represents the distance and red star points are gathered from radar

의 문제 등이 존재하게 된다. 따라서 본 연구에서는 다중 카메라 사이의 캘리브레이션은 영상들 간의 호모그래피를 이용하여 영상을 융합하는 방향으로 하였다. 호모그래피는 평면에서 다른 평면으로의 이동을 표현하는 것이다. 따라서 2차원의 영상을 다른 영상으로 변환할 수 있다. 이처럼 호모그래피를 이용하여 2차원 영상을 서로 변환할 수 있으나, 문제는 2차원 영상에서 획득하는 정보는 3차원의 공간으로부터 생성되므로, 서로 다른 위치의 두 영상은 시차에 따른 오차가 반드시 존재하게 된다. 하지만 주행데이터 영상의 대부분은, [Fig. 6]과 같이 지면에서 먼 곳의 소실점까지의 거대한 가상의 평면에 데이터가 존재하는 것으로 가정을 하고, 이 평면에 해당할 수 있는 대응점을 이용하여 호모그래피를 구하게 되면 호모그래피와 데이터의 특성을 고려한 영상의 융합이 이루어질 수 있게 된다. 4대의 카메라 중 메인이 되는 센서는 텍스처 등이 가장 풍부한 RGB 이미지를 기준으로 하여 호모그래피로 영상을 변형(warping)하고, 이로부터 [Fig. 6]의 (d)와 같이 공통된 영역을 추출하여 사용하도록 하였다. 이렇게 함으로 인해서 시차가 존재하는 문제를 완벽히 제거하지는 못하지만, 최소화하되, 영상을 융합하고, 하나의 라벨을 공유하는 형태의 데이터를 생성할 수 있다. 본 연구에서 다중 스펙트럼의 영상은 1666×1213의 해상도를 가지도록 설정되었다. 또한, 다중스펙트럼 카메라 각각의 특성을 그대로 활용하기 위하여 각 카메라가 받아들이는 raw 데이터를 uint16의 형태로 저장하는 방식을 이용하였다.

### 2.3.2 카메라-라이다, 레이다 간 캘리브레이션

카메라와 라이다 데이터 간의 캘리브레이션은<sup>[9]</sup>의 방법을 적용하여 V자 형태의 타겟을 이용하여 카메라와 3차원 라이다 사이의 회전 및 이동 파라미터를 구하여 적용하였다. 다중 영상은 RGB 영상에 통합되므로, RGB 영상화 3차원 라이다 사이의 외부 파라미터를 추정하여 적용하였다. 그리고 레이다 데이터는 라이다 좌표를 기준으로 설계치의 파라미터를 활용하여 후보정하여 적용하였다. 3차원 라이다와 레이다의 캘리브레이션 결과를 확인하기 위하여, RGB 영상에 해당 포인트 클라우드를 투영하여 [Fig. 7]과 같다.

## 3. 센서 데이터 구성

### 3.1 센서 데이터 획득 환경 구분

열악한 환경에서의 센서 데이터 획득은 센서 데이터를 획득하기 위한 환경을 분석하는 것으로부터 시작한다. 자율주행이 이루어지는 환경은 맑은 날과 같은 좋은 환경인 환경도 있을 수 있으나, 강우, 강설, 먼지, 연막과 같은 환경에서도 자율주행이 이루어질 수 있어야 한다. 강우, 강설은 날씨에 따른 변화이고, 먼지나 연막은 비포장도로나 주변 인위적인 환경에 의한 것이다.

자율주행 등에 적용할 수 있는 알고리즘의 성능을 분석하기 위해서는 열악한 환경의 정도에 따라 어떠한 성능 저하가 있는지에 대한 파악을 위하여 일정한 기준이 존재하여야 한다. 이를 위하여 강우 및 강설의 환경은 강수량을 기준으로 하게 되면 정량적으로 환경을 구분할 수 있다. 강수량은 비가 오거나 눈이 오는 환경 모두에 적용될 수 있는 측정치이고, 본 연구에서는 이러한 정도를 정량적 수치를 기준으로 하여 Low, Medium, High의 세 가지 레벨로 구분하여 데이터를 획득하고, 구분하였다. 강우 및 강설의 기준치는 [Table 1]과 같다. 강수량은 기상청의 강수량 정보<sup>[10]</sup>를 활용하여 매분 또는 10/30/60분 단위의 데이터를 획득할 수 있고, 기준치는 시간당 강수량이 된다. 이때, 기존에 내린 비 또는 쌓인 눈은 정략적인 수치를 측정할 수 없으므로 고려하지 않는다.

[Table 1] Quantitative classification of adverse weather

	Low	Medium	High
Rainy	~5 mm/hr	5~20 mm/hr	20 m/hr~
Snowy	~0.1 mm/hr	0.1~0.5 mm/hr	0.5 mm/hr~

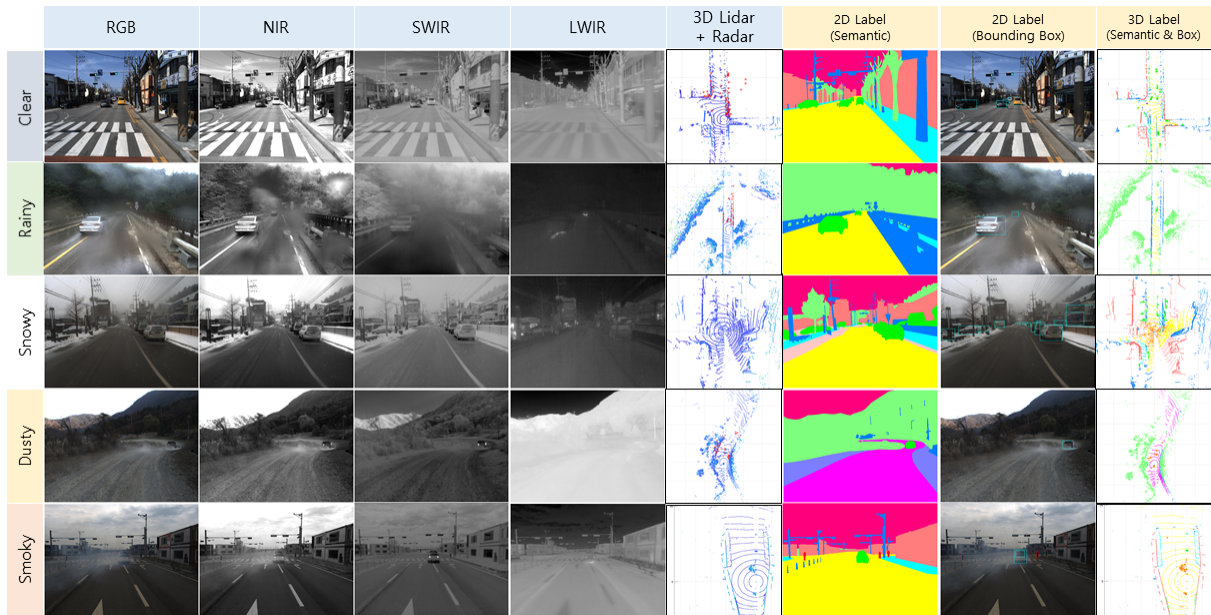
먼지나 연막의 경우, 명확한 정량적 지표를 이용하여 열악한 환경의 정도를 결정하기 어려운 점이 있다. 이는 먼지나 연막이 바람 등의 영향으로 일정하게 주어지지 않기 때문이다. 본 연구에서는 이러한 문제를 RGB 영상에서 사람에 의해서 물체의 인식 정도를 개별적으로 판단하여 Low, Medium, High의 세 가지로 구분하였다.

3.2 센서 데이터 구축

데이터를 학습 및 평가 등에 활용하기 위해서는 데이터에

참값을 적용하여야 한다. 데이터는 연속된 데이터를 얻게 되므로, 데이터는 1 Hz 단위로 샘플링하여 사용하고, 정렬된 영상에서 시맨틱 분할을 위한 분할정답 값을 생성한다. 이때, 보행자, 차량 등의 물체는 아이디를 부여하여, 바운딩 박스를 자동으로 생성하도록 한다. 또한, 3차원 데이터를 활용하기 위하여 3차원 데이터의 시맨틱 분할을 적용하고, 물체에 해당하는 클래스는 3차원 바운딩 박스를 구한다. 이때, 3차원의 포인트 클라우드를 그 정보가 들쭉날쭉하므로, 작업자가 판단할 수 있도록 항법 정보를 기반으로 포인트 데이터를 누적하여 라벨을 할당하고 원래의 프레임으로 변환하는 방식으로 적용한다.

본 연구에서 획득되는 데이터셋과 기존의 데이터셋은 [Table 2]와 같이 비교할 수 있다. 본 데이터셋은 열악한 환경에 대한 데이터를 포함하고 있으며, 이를 활용하기 위한 다중 스펙트럼의 영상을 포함하고 있다. 또한, 2차원과 3차원 데이



[Fig. 8] Examples from Our Driving Dataset in various adverse weather conditions. From left to right, we show the RGB, NIR, SWIR, LWIR, 3D point cloud, and annotations (semantic segmentation and bounding boxes in 2D and 3D). Red star points in point clouds view represent information from radars

[Table 2] Driving dataset comparison. Aforementioned datasets are compared with our dataset. Annotation frame numbers represent image frames have 2D bounding boxes or semantic segmentation

Dataset	imgs	PC lidar	radar	Multi-spectrum	Ann. frames (2D)	3D semantic	3D boxes	GPS	Conditions Night/Rain/Snow/Dust/Smoke	Map	Classes
KITTI <sup>[1]</sup>	15 k	15 k	0	0	15 k	0	200 k	Y	N / N / N / N / N	N	8
Cityscape <sup>[2]</sup>	25 k	0	0	0	25 k	0	0	Y	N / N / N / N / N	N	30
Waymo <sup>[3]</sup>	1 M	200 k	0	0	<b>200 k</b>	<b>200 k</b>	12 M	Y	Y / Y / N / N / N	N	23
Argoverse <sup>[4]</sup>	490 k	44 k	0	0	22 k	0	993 k	Y	Y / Y / N / N / N	Y	15
nuScenes <sup>[5]</sup>	<b>1.4 M</b>	<b>400 k</b>	1.3 M	0	40 k	40 k	1.4 M	Y	Y / Y / N / N / N	Y	23
KAIST <sup>[6]</sup>	8.9 k	8.9 k	0	8.9 k	8.9 k	0	0	Y	Y / N / N / N / N	N	3
<b>Ours</b>	70 k	70 k	70 k	<b>70 k</b>	70 k	4 k	120 k	Y	Y / Y / Y / Y / Y	N	12

[Table 3] Annotation classes for image and point clouds

	Classes (12)	
	Common (11)	Specific (1)
Image	Pedestrian, Car, Cycle, Paved Road, Unpaved Road, Sidewalk, Other Traversable Ground, Vegetation, Building, Other Object, Snowed Ground	Sky
Point Cloud		Obscurant

터에 대한 라벨을 추가하여, 2차원, 3차원 물체 탐지 및 주행가능영역 분석 등 열악한 환경에서의 다양한 형태의 과업을 도출할 수 있다. [Fig. 8]은 본 연구에서 획득한 다중스펙트럼 영상 및 라이다, 레이더 정보와 라벨을 환경별로 나타낸 데이터셋의 예시이다.

영상에 대한 라벨의 종류와 3차원 포인트에 대한 라벨은 각각 12개의 클래스로 구성되고, 두 데이터에서 11개의 클래스는 공유되, 1개의 클래스는 각각의 센서 특성에 맞도록 개별적으로 할당하였다. 영상의 경우, 주행영상의 상단부의 많은 부분이 하늘에 해당되는 반면, 거리센서에서는 하늘에 해당하는 라벨을 할당할 수 없으며, 대신 강우, 강설, 먼지 및 연막 등에 의해서 거리 센서의 데이터에서 공중에 부양한 데이터는 특정 물체나 지형 등에 해당되지 않으므로 애매함(obscurant)으로 할당한다. [Table 3]은 본 데이터셋에서 구성되는 클래스이다.

### 3.3 센서 데이터셋 태스크

본 연구에서 획득한 데이터셋은 열악한 환경에서 자율주행 등에서 활용 가능한 환경인식 알고리즘의 성능을 확인하고, 개선하기 위한 것이다. 자율주행의 대표적인 환경인식으로 물체 인식 및 주행가능영역이 있다.

#### 3.3.1 물체인식

물체 인식은 자율주행에서 주행 간에 차량의 안전을 보장하고 효과적인 자율주행을 수행하기 위해서 필수적인 분야 중 하나이다. 과거에는 주로 RGB 영상 기반의 물체 인식이 주로 이루어졌지만, 최근에는 다양한 대역대의 영상이나 포인트 클라우드에서 물체를 인식하는 기술들이 활발하게 연구되고 있다. 본 데이터셋을 이용하면 이와 같은 최근의 추세를 반영하여 다양한 형태의 센서 데이터에 대한 물체 탐지를 수행하는 것이 가능하다.

본 연구에서는 구축된 데이터셋에서 다양한 환경 변화가 물체탐지 알고리즘에 미치는 영향을 확인하기 위해서 잘 알려진 물체탐지 알고리즘들인 YOLOv3<sup>[11]</sup>과 CenterPoint<sup>[12]</sup> 알고리즘을 선택하여 성능을 시험하고 환경 영향성을 평가하였다.

YOLOv3 알고리즘은 빠른 인식속도와 높은 정확도로 최신 알고리즘은 아니지만 널리 활용되므로 영상 기반 물체인식에 적합하다고 판단된다. CenterPoint 알고리즘은 3차원의 포인트 클라우드에서 물체를 빠르고 정확하게 찾아낼 수 있는 알고리즘이다. 해당 알고리즘들을 이용한 성능 평가 결과는 시험 결과 세션에서 제시한다.

#### 3.3.2 주행영역인식

주행영역을 인식하는 것은 자율주행을 위해서 주변의 환경을 인식하는데 중요한 분야 중 하나이다. KITTI 데이터셋<sup>[1]</sup>에서는 주행가능한 영역을 바이너리 클래스로 정의하여 시험을 진행하였다. 본 데이터셋을 이용하여 동일한 과업을 수행할 수 있다. 12개의 클래스 중에서 주행이 가능한 paved road, unpaved road, other traversable ground, snowed ground에 대해서 주행이 가능한 영역으로 정의하고, 주행가능영역을 영상과 라이다 정보를 이용하여 인식하는 것이다.

본 연구의 주된 목적이 열악한 환경에서의 인식결과이므로, 몇 가지 기본 알고리즘을 선택하여 다른 데이터셋에서 미리 학습된 모델과 열악한 환경 등에서 얻은 데이터를 활용한 모델의 결과 분석 등을 통하여 환경의 영향성을 평가하도록 하였다.

본 연구에서 성능 평가를 위해서 사용된 알고리즘은 KITTI 데이터셋<sup>[1]</sup> 벤치마크에서 상위권에 위치한 PLARD<sup>[13]</sup>와 LoDNN<sup>[14]</sup>을 선택하였다. 두 알고리즘 모두 벤치마크 상에서 실행속도가 빠른 장점이 있으므로, 실제 적용성이 높을 수 있기 때문에 해당 알고리즘을 기본 성능확인을 위한 기준으로 선택하였다. 각 알고리즘의 환경에 따른 인식 성능 시험 결과는 시험세션에서 제시한다.

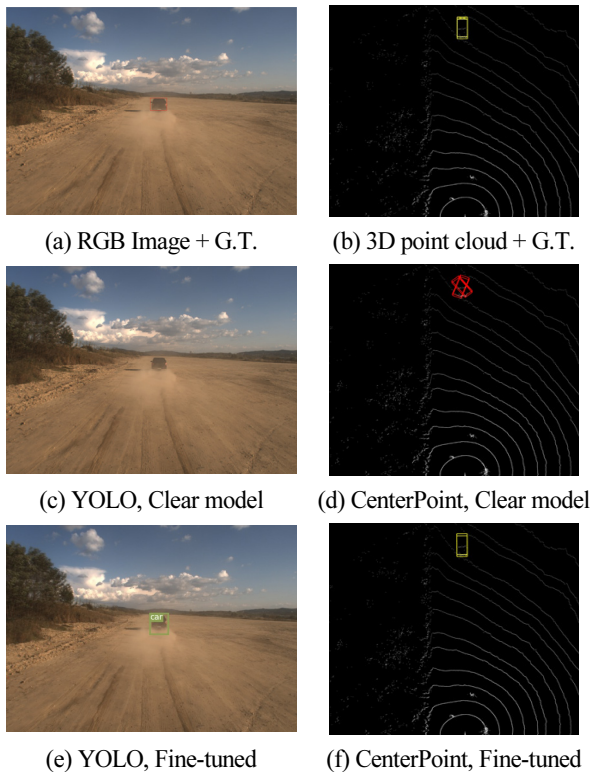
## 4. 시험결과

### 4.1 물체인식 성능 시험

이번 장에서는 앞 절에서 소개된 바와 같이 YOLOv3<sup>[11]</sup>과 CenterPoint<sup>[12]</sup>의 두 가지 알고리즘을 선택하여 열악한 환경에서의 물체 인식 성능을 확인해보도록 한다. YOLOv3 알고리즘을 통해서 RGB 영상에 대한 물체 탐지 성능을 평가하였고, CenterPoint 알고리즘을 통해서 포인트 클라우드에 대한 물체탐지 성능을 평가하였다. 성능의 평가를 위한 인식 대상의 클래스는 pedestrian과 car의 두 가지 종류이다. 본 데이터셋에는 pedestrian과 car 외에도 cycle의 클래스가 존재하지만, cycle에 대한 객체 수가 다른 클래스들에 비해서 많이 적고, 특히 열악한 환경에서는 더더욱 찾아보기 어려운 관계로 cycle을 인식 대상 클래스에서 제외하였다.



각각의 알고리즘에 의한 실험 결과는 [Fig. 9]와 같이 나타난다. 이 그림에서는 맑은 날의 데이터를 이용하여 학습된 모델과 열악한 환경 데이터를 추가하여 학습된 모델 사이에서 차이가 발생하고 있는 것을 확인할 수 있다. 정량적인 결과로 비교하면 그 차이를 더 명확하게 확인할 수 있다. [Table 4]와 [Table 5]에서는 YOLOv3 알고리즘에 의한 영상 물체 인식과 CenterPoint 알고리즘에 의한 3차원 포인트 클라우드 물체 인식 성능을 각각의 열악한 환경별로 비교하여 나타내었다. 이 결과에서 확인할 수 있듯이 맑은 날의 데이터로 학습된 모델은 다양한 열악한 환경에서 물체 탐지를 수행 시 성능이 저하되는 것을 확인할 수 있다. 이와 같은 상황을 극복하기 위해서 추가로 다양한 종류의 열악한 환경에서 얻어진 데이터를 학습 데이터에 추가해 주는 것만으로도 열악한 환경에서의 물체 인식 성능을 크게 높일 수 있는 것을 확인할 수 있다. 위 실험을 통해서 본 논문에서 수행하는 바와 같이 다양한 환경에 대한 데이터셋을 구축하는 것이 다양한 열악한 환경에서 물체탐지를 수행하기 위해서 매우 중요한 일임을 확인할 수 있다.



[Fig. 9] Sample of detection results with YOLOv3 and CenterPoint algorithms: RGB Image with (a) Ground Truth (red box), (c) YOLOv3 Clear model result (brown box) and 3D point clouds with (b) Ground Truth (yellow region), (d) CenterPoint clear model result (red regions), (f) Center Point fine-tuned result (yellow region)

#### 4.2 주행영역인식 성능 시험

앞 장에서 언급한 바와 같이 열악한 환경에서 주행가능영역을 찾는 것은 중요한 과업 중 하나이고, 이를 위한 기준이 되는 알고리즘은 KITTI 데이터셋<sup>[11]</sup> 벤치마크에서 상위권에 위치한 PLARD<sup>[13]</sup>과 LoDNN<sup>[14]</sup>를 선택하였다. [Fig. 10]은 PLARD 알고리즘의 실험 결과이다. 일반적인 데이터셋과 같이 맑은 날의 데이터를 이용하여 학습한 모델과 열악한 환경 데이터를 추가하여 학습한 모델 사이의 결과에 아주 큰 차이가 있음을 보인다. [Table 6]의 결과에서와같이, 맑은 날의 데이터를 학습한 모델은 비가 오거나, 연막이 있거나, 눈이 오는 경우에 성능이 저하되는 것을 확인할 수 있다. 이를 극복하기 위한 가장 쉬운 방법인 데이터 제공을 통하여 일정부분 극복할 수 있는 것으로 해석할 수 있다. 비나 눈이 오는 경우에 대한 학습을 통해 정확도를 높일 수 있고, 연막의 경우 제한적이긴 하지만 기본 모델에 비해 정확성을 높은 것을 확인할 수 있다.

추가로 라이다 데이터만 사용한 방법 중 하나인 LoDNN의 결과를 [Table 4]에서 확인할 수 있다. 강우나 강수의 경우 라이다 데이터에서 눈 또는 비에 의해 공중에 맺히는 포인트 데이터가 발생하게 되고, 이에 따라 그러한 데이터가 없이 학습된 모델은 열악한 환경에서 그 성능이 더 저하되는 것을 알 수 있다.

이와 같은 결과로부터 열악한 환경에서는 단일 센서를 활용한 방식을 적용하는 것보다, 다중센서를 활용하는 것이 더 성능이 높은 것을 확인할 수 있으며, 라이다보다 영상을 활용

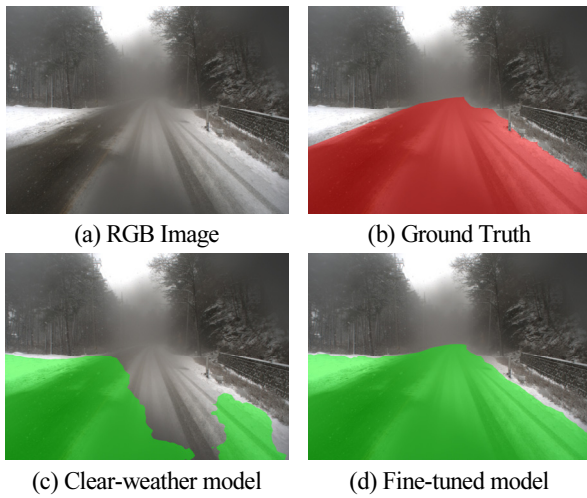
[Table 4] Quantitative comparison for YOLOv3 algorithm object detection result (Average Precision, AP) between clear model (A) and fine-tuned model (B) (Input: RGB images)

	All	Sunny	Rainy	Smoky	Snowy	Dust
Model (A)	11.95	<b>43.65</b>	5.17	6.27	4.73	14.25
Model (B)	<b>27.83</b>	34.14	<b>26.48</b>	<b>16.49</b>	<b>25.42</b>	<b>55.57</b>

[Table 5] Quantitative comparison for CenterPoint algorithm object detection result (Average Precision, AP) between clear model (A) and fine-tuned model (B) (Input: 3D point clouds). We used IOU threshold 0.5 for both bird's eye view matching (bev) and 3D bounding matching (3D)

	Sunny	Rainy	Snowy	Dust	Smoky	All
Model (A) (bev, IOU.0.5)	25.79	16.79	24.17	44.62	40.10	29.92
Model (B) (bev, IOU.0.5)	<b>55.08</b>	<b>45.24</b>	<b>34.89</b>	<b>59.51</b>	<b>59.42</b>	<b>53.75</b>
Model (A) (3D, IOU.0.5)	21.24	11.38	18.81	36.86	27.30	22.36
Model (B) (3D, IOU.0.5)	<b>50.78</b>	<b>39.93</b>	<b>29.53</b>	<b>56.63</b>	<b>59.00</b>	<b>49.40</b>





[Fig. 10] Sample of results with PLARD and snowy weather data: (a) RGB Image, (b) Ground truth image, red region is traversable region labeled by human, (c) Result of the model trained on clear data, (d) Result of the model trained on adverse weather data. Green color represents an inferred traversable region which the probability is over 0.5

[Table 6] Quantitative result (Max F score) comparison of traversable region detection using PLARD<sup>[13]</sup> and LoDNN<sup>[14]</sup> with between clear model and fine-tuned model. The clear models are trained on sunny data and the fine-tuned models are trained on all adverse weather data

	Sunny	Rainy	Snowy	Dust	Smoky	All
PLARD (Sunny)	<b>98.30</b>	96.71	96.18	96.23	92.44	96.06
PLARD (All)	97.94	<b>97.81</b>	<b>97.12</b>	<b>98.34</b>	<b>96.26</b>	<b>97.96</b>
LoDNN (Sunny)	92.01	74.76	78.37	87.57	80.82	84.27
LoDNN (All)	<b>92.76</b>	<b>81.09</b>	<b>83.50</b>	<b>97.22</b>	<b>83.51</b>	<b>89.73</b>

하는 것이 더 성능이 높은 경향이 있다. 이는 라이다 데이터에서 발생하는 노이즈 데이터를 영상으로 보완할 수 있기 때문이다.

## 5. 결론

본 연구를 통하여 열악한 환경에서 자율주행에 필요한 기술을 개발하기 위한 데이터셋을 확보하였다. 이를 통하여 주행가능영역과 물체를 인식함에 있어 환경조건에 따라 기존 데이터셋에 학습된 모델의 한계를 확인하였다. 또한 다중스펙트럼의 영상을 포함하는 데이터셋으로, 향후 본 연구의 데이터셋을 활용하여 열악한 환경에서 해당 환경을 극복하기 위한 방법 등 다양한 연구가 이루어질 수 있을 것이다.

## References

- [1] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 2012, DOI: 10.1109/CVPR.2012.6248074.
- [2] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes dataset for semantic urban scene understanding," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 2016, DOI: 10.1109/CVPR.2016.350.
- [3] Waymo, *Waymo Open Dataset: An autonomous driving dataset*, [Online], <https://waymo.com/open>, Accessed: August 29, 2019.
- [4] M-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays, "Argoverse: 3d tracking and forecasting with rich maps," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long beach, USA, 2019, DOI: 10.1109/CVPR.2019.00895.
- [5] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, and O. Beijbom, "nusenes: A multimodal dataset for autonomous driving," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, DOI: 10.1109/CVPR42600.2020.01164.
- [6] Y. Choi, N. Kim, S. Hwang, K. Park, J. Yoon, K. An, and I. S. Kweon, "KAIST multi-spectral day/night data set for autonomous and assisted driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, Mar., 2018, DOI: 10.1109/TITS.2018.2791533.
- [7] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, Nov., 2000, DOI: 10.1109/34.888718.
- [8] MATLAB camera calibration toolbox, [Online] <https://www.mathworks.com/help/vision/camera-calibration.html>, Accessed: April 10, 2018.
- [9] S. Sim, J. Sock, and K. Kwak, "Indirect correspondence-based robust extrinsic calibration of LiDAR and camera," *Sensors*, vol. 16, no. 6, Jun., 2016, DOI: 10.3390/s16060933.
- [10] Automatic Weather System, Korea Meteorological Agency, [Online], [http://www.weather.go.kr/weather/observation/aws\\_table\\_popup.jsp](http://www.weather.go.kr/weather/observation/aws_table_popup.jsp), Accessed: Jul. 28, 2020.
- [11] R. Joseph and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018, DOI: 10.48550/arXiv.1804.02767.
- [12] Y. Tianwei, X. Zhou, and P. Krahenbuhl, "Center-based 3d object detection and tracking," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, USA, 2021, DOI: 10.1109/CVPR46437.2021.01161.
- [13] C. Zhe, J. Zhang, and D. Tao, "Progressive lidar adaptation for road detection," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 3, May, 2019, DOI: 10.1109/JAS.2019.1911459.
- [14] L. Caltagirone, S. Scheidegger, L. Svensson, and M. Wahde, "Fast LIDAR-based road detection using fully convolutional neural networks," *2017 IEEE Intelligent Vehicles Symposium (IV)*, Redondo Beach, USA, 2017, DOI: 10.1109/IVS.2017.7995848.



### 심 성 대

2004 경북대학교 전자전기공학부(학사)  
2006 포항공과대학교 전자전기공학과(석사)  
2006~현재 국방과학연구소 선임연구원

관심분야: 센서 캘리브레이션, 환경인식, 로봇비전



### 배 광 탁

2019 KAIST 전기및전자공학부(학사)  
2019~현재 국방과학연구소 현역연구원

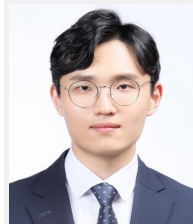
관심분야: 컴퓨터비전, 3차원 인식, 로봇비전



### 민 지 흥

2000 중앙대학교 전기전자공학부(학사)  
2002 중앙대학교 첨단영상대학원(석사)  
2012 KAIST 전기전자공학(박사)  
2002~현재 국방과학연구소 책임연구원

관심분야: 컴퓨터비전, 환경인식



### 김 병 준

2019 KAIST 전기및전자공학부(학사)  
2019~현재 국방과학연구소 현역연구원

관심분야: 컴퓨터비전, 3차원 인식, 로봇비전



### 안 성 용

2008 연세대학교 전기전자공학과(학사)  
2010 연세대학교 전기전자공학과(석사)  
2010~현재 국방과학연구소 선임연구원

관심분야: 로보틱스



### 서 준 원

2021 서울대학교 컴퓨터공학부(학사)  
2021~현재 국방과학연구소 현역연구원

관심분야: 컴퓨터비전, 환경인식, 로봇비전



### 이 종 우

2008 포항공과대학교 전자전기공학과(공학사)  
2019 포항공과대학교 전자전기공학과  
(공학박사)  
2019~현재 국방과학연구소 선임연구원

관심분야: 컴퓨터비전, 영상처리, 환경인식



### 최 덕 선

2003 연세대 전기전자공학과(공학석사)  
2004 연세대 전기전자공학과(공학석사)  
2016 연세대 전기전자공학과(공학박사)  
2004~현재 국방과학연구소 책임연구원

관심분야: 장애물 탐지/추적/예측, 장애물 인식



### 이 정 석

2001 포항공과대학교 기계공학과(학사)  
2003 포항공과대학교 기계공학과(석사)  
2010 포항공과대학교 기계공학과(박사)  
2011~현재 국방과학연구소 선임연구원

관심분야: 컴퓨터비전, 영상처리, 환경인식