

논문 2024-19-12

객체 영역에 특화된 뎁스 추정 기반의 충돌방지 기술개발 (Object-aware Depth Estimation for Developing Collision Avoidance System)

황규태, 송지민, 이상준*
(Gyutae Hwang, Jimin Song, Sang Jun Lee)

Abstract : Collision avoidance system is important to improve the robustness and functional safety of autonomous vehicles. This paper proposes an object-level distance estimation method to develop a collision avoidance system, and it is applied to golfcarts utilized in country club environments. To improve the detection accuracy, we continually trained an object detection model based on pseudo labels generated by a pre-trained detector. Moreover, we propose object-aware depth estimation (OADE) method which trains a depth model focusing on object regions. In the OADE algorithm, we generated dense depth information for object regions by utilizing detection results and sparse LiDAR points, and it is referred to as object-aware LiDAR projection (OALP). By using the OALP maps, a depth estimation model was trained by backpropagating more gradients of the loss on object regions. Experiments were conducted on our custom dataset, which was collected for the travel distance of 22 km on 54 holes in three country clubs under various weather conditions. The precision and recall rate were respectively improved from 70.5% and 49.1% to 95.3% and 92.1% after the continual learning with pseudo labels. Moreover, the OADE algorithm reduces the absolute relative error from 4.76% to 4.27% for estimating distances to obstacles.

Keywords : Collision avoidance, Computer vision, Deep learning, Continual learning, Object detection, Depth estimation, Object-aware depth estimation

1. 서론

한국은 일본이나 미국과 더불어 전체 인구 중 65세 이상의 고령자 인구가 21% 이상인 초고령사회에 도달했다 [1]. 이에 따라 다양한 산업에서 무인이동체에 대한 수요가 증가하고 자율주행의 안전성을 개선하기 위한 연구개발이 진행 중에 있다. 이와 같은 지능형 모빌리티의 자율주행 기술은 골프산업에서도 활용될 수 있다. 국내 골프장 이용자 수는 2019년 4,170만 명에서 2021년 5,073만 명으로 21.7%가 증가하여 골프산업은 지속적인 성장세를 보이고 있다 [2]. 고령자의 골프백과 함께 넓은 부지의 이동이 필요한 골프장 환경의 특성상 골프카트는 골프 산업에서 활용도가 높다. 기존 골프카트는 대부분 캐디가 직접 운전하는 방식이었으나, 최근에는 GPS나 자기장 기반의 유도센서를 기반으로 한 무인 운전 시스템이 적용되고 있다 [3, 4]. 무인으로 작동하는 골프카트는 장애물과의 충돌을 방지하기 위하여 일반적으로 초음파 센서 또는 마그넷 센서 기반의 충돌방지 시스템을 탑재한다. 하지만 이와 같은 기존 센서는 가격대가 높으며 유지보수가 어렵고 구현 가능한 기능과 서비스가 한정적이

라는 한계점이 있어서 기존 방식을 대체하기 위한 연구가 진행되고 있다 [5, 6].

카메라 영상 기반의 딥러닝 기술은 시각적인 정보를 분석함으로써 골프카트의 긴급 제어가 필요한 상황을 인지하는 데에 활용 가능하다. 또한, 단가가 낮고 소형이라는 카메라 센서의 특성상 유지보수가 용이하다는 장점을 가진다. 장애물 인식에 활용되는 대표적인 컴퓨터비전 및 딥러닝 기술에는 2차원 객체 검출 [7-9] 및 단안 깊이 추정 [10-13] 기술 등이 있다. 객체 검출 모델은 사전에 정의된 객체들에 대하여 시각적인 특징을 학습하고, 2차원 이미지 평면에서 바운딩박스로 픽셀 위치를 표현함으로써 해당 객체의 위치와 종류를 추론할 수 있다. 깊이 추정 모델은 입력 영상의 모든 픽셀들에 대한 물리적인 거리정보를 추론하는 딥러닝 모델을 의미하며, 일반적으로 LiDAR와 같은 외부의 거리센서 신호를 모델 훈련에 활용한다. 객체 검출 및 깊이 추정의 결과를 결합하여 주행환경에서 장애물과 관심객체에 대한 거리를 추정할 수 있으며, 이는 충돌방지 시스템의 요소기술로 활용이 가능하다. 본 연구는 골프카트의 충돌방지 시스템 구축을 위하여, 객체 검출 및 깊이 추정 모델을 활용한 객체 별 거리 추정 알고리즘의 개발을 목적으로 한다. 본 논문에서는 관심 객체의 영역에서 더 정확한 거리정보를 추론하기 위한 object-aware depth estimation (OADE) 기술을 제안한다.

지도학습 기반의 딥러닝 기술을 실제 산업 문제에 적용하는 과정에서 라벨 데이터를 생성하기 위한 레이블링 과정은

*Corresponding Author (sj.lee@jbnu.ac.kr)

Received: Dec. 28, 2023, Revised: Feb. 6, 2024, Accepted: Feb. 14, 2024.

G. Hwang: Jeonbuk National University. (M.S. Student)

J. Song: Jeonbuk National University. (M.S. Student)

S. J. Lee: Jeonbuk National University (Asst. Prof.)

※ 본 과제 (결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학 (LINC+) 육성사업의 연구결과입니다. 이 연구는 2023년도 산업통상자원부 및 산업기술평가관리원 (KETI) 연구비 지원에 의한 연구임 (20023305).

시간과 비용을 초래한다. 따라서, 본 연구에서는 적은 비용으로 고성능의 객체 검출 모델을 개발하기 위하여 연속학습(continual learning)을 기반으로 객체 검출 모델을 개발하였다. 기 학습된 객체 검출 모델을 활용하여 라벨이 없는 영상 데이터에서 객체의 대략적인 위치에 대한 pseudo 레이블을 생성하여 학습에 적용하였다. 기존의 깊이 추정 모델 훈련에 사용되는 LiDAR 센서 신호는 밀도가 낮다(sparse)는 한계가 있다. 이러한 한계 때문에 대부분의 LiDAR 센서는 카메라보다 FoV(field of view)가 작으며, 영상의 상단과 하단에는 깊이 추정 모델의 학습에 활용 가능한 거리정보가 없는 영역이 발생한다. 따라서, 언덕과 내리막길이 다수 존재하는 골프장 환경의 특성상 관심객체에 대한 거리정보가 없는 경우가 발생하며, 이에 대한 거리 추정 성능이 급격히 하락한다. 본 논문에서는 LiDAR의 FoV에 걸쳐있는 객체들을 코너케이스로 간주하고, 이를 해결하기 위한 기법인 object-aware LiDAR projection(OALP)을 제안한다. OALP는 검출된 관심객체 영역 내에서 LiDAR의 센서 정보를 보정함으로써, 코너케이스에서의 성능을 개선하기 위한 방법이다. 제안하는 방법이 객체 별 거리추정 정확도 개선에 효과가 있음을 보이기 위하여 실제 골프장 환경에서 데이터셋을 구축하고 country club(CC)의 약자를 사용하여 CC-데이터셋이라고 명명하였다. 제안하는 기술의 단계별 성능을 평가함으로써 딥러닝 모델 및 제안 기술의 영향을 체계적으로 분석하였다.

II. 본 론

1. CC-데이터셋 구성

그림 1은 실제 골프장 환경에서 CC-데이터셋의 수집을 위하여 골프카트에 구축한 임베디드 보드 및 다중 센서들의 구성을 보여준다. 영상과 거리정보의 획득을 위하여 Intel사의 Real sense D435i 카메라와 오이스터사의 OS1 LiDAR를 활용하였다. 카메라로 획득한 영상은 객체 별 거리 추정 알고리즘의 입력 데이터로 사용되며, LiDAR 센서의 거리정보는 깊이 추정을 위한 딥러닝 모델의 학습에 활용된다. 카메라 센서의 가로와 세로 방향의 FoV는 69°와 42°이며, LiDAR 센서의 세로 방향 FoV는 22.5°이다. 따라서 획득 영상의 상하단에 해당하는 영역에는 거리정보 획득이 불가능하다. 데이터 획득 및 딥러닝 모듈의 구동을 위한 임베디드 보드는 Nvidia사의 Jetson AGX orin을 사용하였으며 ROS(Robot Operating System)를 기반으로 다중 센서들의 동기화된 데이터를 취득하였다. 고가의 LiDAR 센서는 데이터 수집용 골프카트에서만 필요로 하며, 상용화 예정인 카트에서는 카메라 센서만을 이용하여 거리정보의 추정이 가능하다. 따라서, 제안하는 기술은 비용 대비 효과적인 충돌 방지 시스템을 구축하는 데에 활용할 수 있다.

실제 데이터 획득 과정에서는 골프카트와 사람이 관측되는 다양한 시나리오 및 날씨의 다양성을 고려하였다. CC-데이터셋은 다양한 시공간적 환경을 고려하여 수집되었으며, 3

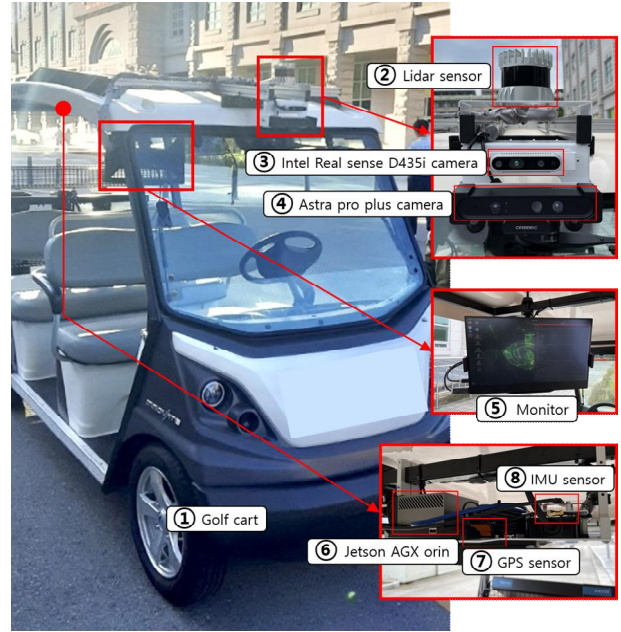


그림 1. 골프카트에 설치된 임베디드 보드 및 다중 센서 구성
Fig. 1. Configuration of embedded board and sensors mounted on golfcart

군데의 골프장과 32개의 주행 시퀀스로 구성된다. 야외인 골프장 환경의 공간적인 다양성과 계절 및 날씨를 고려한 데이터 수집 시나리오를 계획하였으며, 현재까지 여름과 가을, 맑은 날과 우천 시 데이터를 획득하였다. 그림 2는 실제 골프장 환경에서 데이터를 수집하는 과정을 보여준다. 예를 들어, 김제 스프링스CC의 데이터는 우천 상황의 데이터 획득과정을 보여준다. 데이터 수집 과정에서는 두 대 이상의 골프카트를 활용하였고 객체 검출 모델 훈련을 위한 레이블링은 골프장 환경에서 충돌 시 위험요소가 큰 골프카트와 사람 객체를 대상으로 하였다. 영상에 ‘Golfcart’ 객체가 나타나도록 데이터를 수집하기 위하여 데이터 수집 골프카트 전방에 선행 골프카트를 배치하였고, 20 km/h 이하의 속도로 주행하며 LiDAR 범위 내에서 두 카트 사이의 거리를 조절하였다. 또한, 다수의 ‘Person’ 객체에 대한 영상데이터를 수집하기 위하여 골프장마다 하나 이상의 주행 시퀀스에 대하여 선행 카트의 인원들이 하차하여 자유롭게 걸어 다니는 시나리오를 구성하고 다수의 관심 객체들이 나타나는 영상 데이터를 취득하였다.

그림 3은 GPS 위치정보를 기반으로 데이터 취득 경로를 시각화한 것을 보여주며, 3곳의 골프장에서 54개의 홀을 주행하며 총 22 km의 주행거리에 해당하는 데이터를 획득하였다. 데이터 수집 골프카트에서 10 Hz로 동기화된 영상 시퀀스와 다중 센서 데이터를 취득하였으며, 총 89,850 쌍의 1280x720 해상도 영상과 LiDAR point cloud를 수집하였다. 학습데이터는 골프장 마다 4,800 장씩 샘플링 하여 총 14,400장의 영상으로 구성된다. 개발 기술의 엄밀한 성능검증을 위하여 독립된 시퀀스의 테스트 데이터를 구성하였다. 테스트 데이터는 일반적인 상황 및 코너케이스에서 각각



익산 쌍테힐CC (2023.09.04.)

김제 스파힐스CC (2023.09.20.)

청주 세레니티CC (2023.10.16.)

그림 2. CC-데이터셋 수집 환경
Fig. 2. Data acquisition environments



익산 쌍테힐CC (2023.09.04.)

김제 스파힐스CC (2023.09.20.)

청주 세레니티CC (2023.10.16.)

그림 3. GPS 기반 데이터 취득 경로 시각화
Fig. 3. Visualization of GPS-based data acquisition path

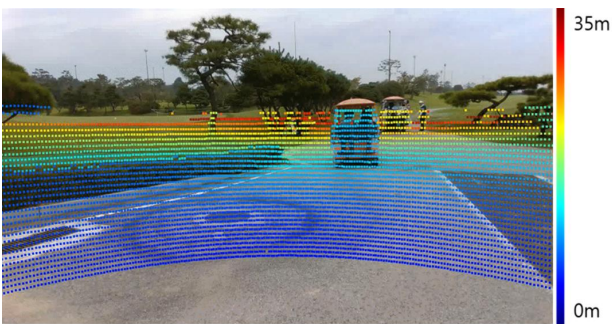


그림 4. LiDAR projection 시각화 예시
Fig. 4. Visualization of LiDAR projection

300 프레임과 50 프레임의 길이의 연속된 시퀀스 2개씩으로 구성하였으며, 따라서 테스트 데이터는 총 700장의 영상으로 구성된다.

골프카트에 설치된 센서들의 융합을 위하여 실험시마다 다중 센서들 사이의 캘리브레이션을 진행하여 수집 데이터

의 정밀도를 높였다. 본 논문에서는 OpenCV를 활용하여 카메라 내부 파라미터인 K 를 계산하였다. 또한 Tsai et al. [14] 이 제안한 알고리즘을 활용하여 카메라와 LiDAR 사이의 상대적인 유클리디안 변환 관계를 표현하는 변환행렬인 $[R | t]$ 를 계산했다. 카메라 내부 파라미터 K 와 외부 파라미터인 변환행렬 $[R | t]$ 는 수식 (1)과 같이 LiDAR의 3차원 관측 좌표 P 를 카메라 영상에 대응되는 2차원 좌표 p 로 투영하는 과정에서 사용된다.

$$p = K[R|t]P. \quad (1)$$

깊이 추정 모델의 학습을 위하여 카메라와 LiDAR 사이의 캘리브레이션 파라미터를 기반으로 LiDAR 센서 신호의 투영 영상을 생성하였다. 이와 같은 투영 영상을 LiDAR projection이라고 하며, 그림 4는 2차원 좌표 p 에 대응되는 RGB 이미지를 LiDAR projection과 함께 시각화한 예시를 보여준다. LiDAR projection은 RGB 이미지와 크기가 동일하고 깊이 추정 모델 훈련과정에서 레이블로 사용된다.

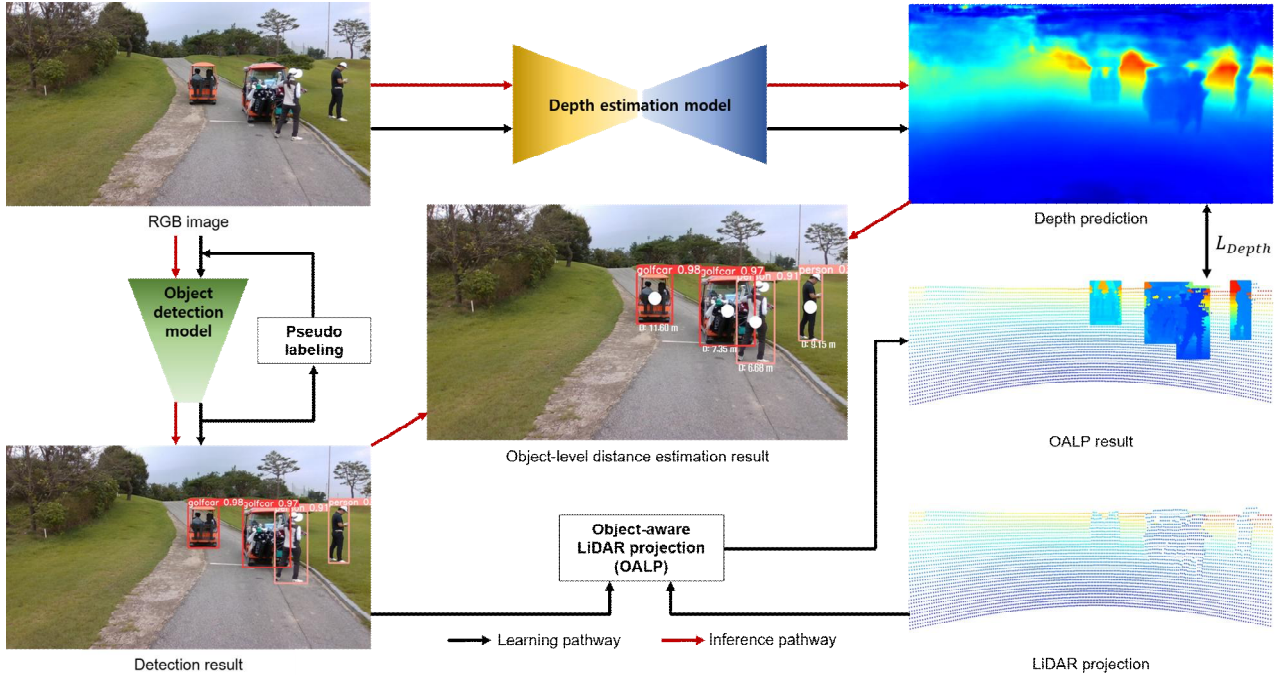


그림 5. 제안방법인 OADE의 학습 파이프라인
Fig. 5. Training pipeline of the proposed method

2. OADE 파이프라인과 연속학습 및 OALP과정

본 논문에서 제안하는 객체 별 거리 추정을 위한 OADE 는 객체 검출과 깊이 추정을 위한 두 가지 딥러닝 모델을 기반으로 설계되었다. 제안하는 알고리즘의 핵심은 객체 검출 모델의 연속학습 및 OALP 기반의 객체 별 거리추정 모델의 성능 고도화 방법에 있다. 그림 5는 본 논문에서 제안하는 방법인 OADE의 학습 및 추론 과정을 보여준다. 학습과정에는 RGB 이미지 및 LiDAR projection이 사용되었고, 흑색 경로를 따라 두 가지 모델을 훈련시킨다. 객체 검출 모델은 학습 데이터셋의 pseudo 레이블링을 통한 3회의 재학습 과정으로 직접 레이블링을 하는 시간을 최소화하는 동시에 성능을 최대화하였다. 깊이 추정 모델은 객체 검출 모델의 추론 결과를 기반으로 바운딩박스 내부의 깊이 정보가 유사한 것을 가정하여 LiDAR의 FoV가 벗어난 영역을 맵핑하는 기법인 OALP를 통해 객체 별 거리의 손실을 최소화하였다. 결과적으로 추론 과정에서는 한 장의 RGB 이미지가 적색 경로를 따라 병렬적으로 객체 검출 및 깊이 추정 모델을 통과하고, 두 가지의 추정 결과를 이미지상에 표현한다.

객체 검출 모델의 연속학습 과정에는 일부 영상데이터에 대한 바운딩박스 레이블을 기반으로 학습된 객체검출 모델을 활용한 pseudo 레이블을 활용하였다. 그림 6에서 그 과정을 볼 수 있으며, n , X , X_n , \hat{y}_n , y_n 과 D_n 은 각각 반복횟수, 전체 RGB 이미지 데이터셋, n 번째 하위 입력 데이터셋, n 번째 모델 예측 결과, n 번째 pseudo 레이블과 n 번째 학습된 모델을 의미한다. Pseudo 레이블링 과정은 반복되며, 첫 번째 과정은 구축된 X 의 하위 입력 데이터셋과 pseudo 레

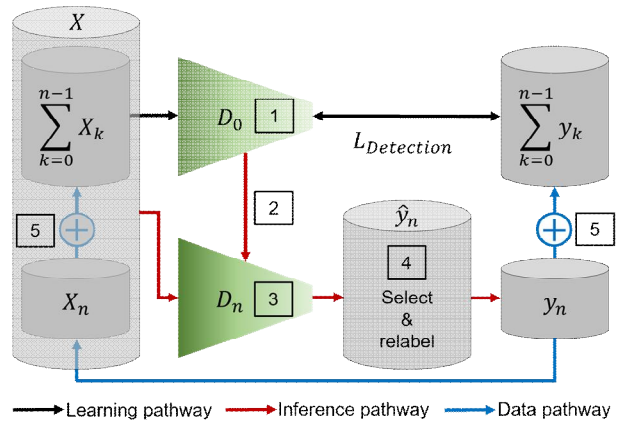


그림 6. Pseudo 레이블링 과정
Fig. 6. Pseudo labeling for object detection model

이블을 통한 초기 검출모델 D_0 의 학습이다. 두 번째는 n 번째로 학습된 초기 모델의 파라미터를 D_n 에 전이하는 과정이며, 세 번째는 D_n 을 통해 X 에 대한 모든 \hat{y}_n 을 추론하는 과정이다. 네 번째 과정에서 \hat{y}_n 에서 잘 검출된 데이터의 선별 및 바운딩박스의 수정이 이루어지며, 그 결과 y_n 을 얻는다. 마지막으로 다섯 번째 과정에서 y_n 과 대응되는 X_n 이 D_{n+1} 을 학습하기 위해 각각 이전까지 구축된 하위 데이터셋에 병합된다. 그림 6에서 흑색, 적색 및 청색 화살표는 각각 학습, 추론 및 데이터 이동 경로를 의미하며, 실험에서 반복횟수 n 은 3으로 설정되었다. 초기 하위 데이터셋 X_0 와

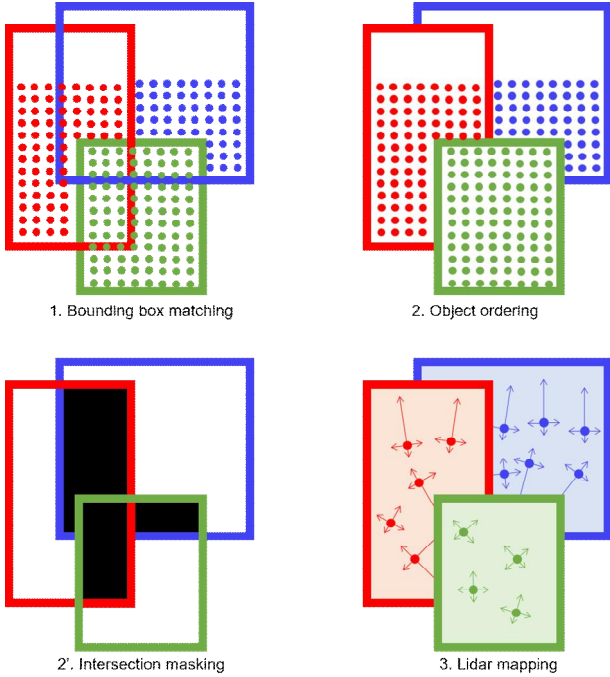


그림 7. Object-aware LiDAR projection 과정 및 occlusion 처리
 Fig. 7. Process of object-aware LiDAR projection and occlusion

레이블 y_0 는 177쌍으로 직접 레이블링을 수행하였다. 최종 객체검출 모델 D_3 는 이후 OALP와 OADE 알고리즘에서 객체 추론을 위하여 활용된다.

깊이 추정의 관점에서, 입력 영상과 LiDAR의 FoV의 불일치로 인하여 발생하는 코너케이스를 해결하기 위하여 OALP를 도입하였으며, 그림 7에서 OALP의 생성 과정을 볼 수 있다. OALP는 LiDAR의 FoV에 걸쳐있는 객체에 대한 거리추정 성능을 향상시키기 위하여 해당 바운딩 박스에 없는 깊이 레이블을 가장 가까운 LiDAR point 값으로 매핑하는 방식이다. OALP의 과정은 bounding box matching, object ordering, LiDAR mapping으로 이루어진다. Bounding box matching은 객체 검출 모델 D_3 를 이용하여 학습 데이터셋 X 의 객체를 모두 추론한 이후, 최소한 LiDAR projection 이미지에 바운딩박스의 좌표를 표시하는 과정이다. 이후 object ordering에서는 이미지 내의 객체가 지면과 닿아있다는 가정하에, 지면과 닿는 바운딩박스의 좌표에 따라 카메라와 가까운 거리에 따른 객체의 순서를 정의한다. Occlusion이 발생한 영역의 경우에는 가까운 객체의 LiDAR point를 맵핑하기 위하여 멀리 있는 객체의 바운딩박스의 겹친 영역을 0 값으로 마스킹한다. 마지막으로 LiDAR mapping과정에서 객체 별로 비어있는 픽셀에 nearest neighbor search기법인 KDtree [15]를 활용하여 LiDAR point 맵핑을 수행하였다. OALP 과정에서 바운딩박스에 LiDAR point가 없는 경우는 제외하였으며, 그림 5의 OALP result에서 객체 별로 밀도가 높은 (dense) 깊이가 생성된 레이블 이미지의 예시를 볼 수 있다.

3. 객체 검출 및 깊이 추정 알고리즘

객체 검출 모델은 대중적이며, 추론 정확도가 높은 You only look once (YOLO) 계열의 모델을 채택하였고, 이 중에서 YOLOv5 모델 [16]을 사용하였다. YOLOv5 모델의 구조는 크게 특징을 추출하는 백본 네트워크, 각 레이어의 특징을 융합하는 neck 네트워크, 바운딩박스 좌표 및 클래스를 추정하는 헤드로 구성된다. 백본 네트워크는 cross stage partial (CSP)-DarkNet [17]을 사용하며, 특징맵의 일부를 분산시켜 컨볼루션을 진행함으로써 연산량을 줄이고 정확도를 유지하였다. Neck 네트워크에서는 다중 스케일의 특징맵을 융합하는 feature pyramid network (FPN) [8] 중에 낮은 레벨의 특징에 집중한 path aggregation network (PAN) [18]를 적용하였다. 헤드는 Neck 네트워크에서 출력된 특징맵으로부터 모델의 최종 출력인 바운딩박스 파라미터 및 클래스 확률을 추정하는 컨볼루션 레이어로 구성된다. 본 논문에서 다루는 객체 클래스가 2가지인 것을 고려하여, YOLOv5 모델 중 크기가 작고 빠른 s모델을 사용하였으며, D_0 모델의 학습에 Microsoft COCO 데이터셋 [19]에서 학습된 전이학습 파라미터를 사용하였다. 객체 검출 모델의 학습을 위한 그림 6의 손실함수는 다음과 같은 수식으로 정의된다.

$$L_{detection} = L_{reg} + L_{conf} + L_{da}. \quad (2)$$

수식 (2)에서 L_{reg} 은 라벨과 모든 grid cell에서 예측된 바운딩박스의 중심 좌표 및 너비와 높이의 오차 제곱이다. L_{conf} , L_{da} 는 각각 모든 grid cell에서 예측한 클래스에 속하거나 클래스가 존재할 확률에 대한 손실함수이며, 전체적인 손실함수는 L_{reg} , L_{conf} 와 L_{da} 의 합으로 이루어진다.

깊이 추정 모델로 사용한 BTS [10] 모델은 엔코더-디코더 구조로 이루어지며, 엔코더는 특징 추출기 및 atrous spatial pyramid pooling (ASPP) [20]으로, 디코더는 업샘플링 및 local planar guidance (LPG) 레이어로 구성된다. 엔코더의 특징 추출기에서 출력된 중간 특징맵들은 디코더의 동일한 레이어의 특징에 스킵 커넥션으로 연결되어 낮은 레벨의 특징을 전달한다. 특징 추출기의 최종 출력은 ASPP를 통하여 수용영역의 크기를 키우고 다중 스케일의 커널을 통한 컨볼루션으로 세밀한 부분의 특징을 학습하고자 하였다. 디코더의 각 레이어에서는 특징맵의 스케일 별로 업샘플링, 스킵 커넥션과 LPG 레이어가 융합되며, 특히 LPG 레이어는 특징맵이 fine-grained 할수록 더 세밀한 부분을 학습하도록 설계되었다. 본 연구에서 사용한 특징 추출기는 DenseNet161 [21]이며, CC-데이터셋이 야외환경임을 고려하여 KITTI 데이터셋 [22]에서 학습된 전이학습 파라미터를 사용하였다. 깊이 추정 모델의 학습을 위한 그림 5의 손실함수는 다음과 같이 정의된다.

$$L_{depth} = \alpha \sqrt{\frac{1}{T} \sum_i g_i^2 - \left(\frac{1}{T} \sum_i g_i\right)^2} + (1-\lambda) \left(\frac{1}{T} \sum_i g_i\right)^2. \quad (3)$$

수식 (3)에서 g_i 는 $\log \hat{d}_i - \log d_i$ 를 의미하며, T 는

LiDAR point의 개수이다. 실험에서 상수 λ 와 α 는 각각 0.85 및 10으로 설정되었다.

4. 실험 결과

본 논문에서의 실험을 위한 하드웨어 환경은 Intel Core i9-10940X CPU, 64 GB DDR4 RAM 그리고 NVIDIA Geforce RTX 3090 Ti로 이루어져있다. 그리고 제안하는 알고리즘의 구현 및 성능평가를 위한 소프트웨어 환경은 파이썬 및 파이토치를 사용하였다. 실험 결과는 OADE 알고리즘을 통한 객체 별 거리추정 성능, pseudo 레이블링을 통한 객체 검출 모델의 성능과 OALP를 통한 깊이 추정 모델의 성능을 순차적으로 보이고 결과를 분석하였다. 객체 별 거리추정 성능은 전체 테스트셋 및 코너 케이스에 대한 성능을 구분하였으며, 바운딩박스 중앙의 깊이 값에 대한 절대상대오차 (ARE, absolute relative error)로 평가하였다. 객체 검출 모델은 분류 결과를 정밀도 (Precision) 및 재현율 (Recall)로, 예측 바운딩박스의 결과를 mAP (metric average precision)50 및 mAP50-95로 평가하였다.

제안하는 알고리즘의 객체 별 거리추정 성능을 평가하기 위하여 수집된 CC-데이터셋에서 실험하였다. 표 1은 제안하는 알고리즘 OADE를 적용하기 전과 후의 객체 별 거리추정 비교실험 결과를 보여주며, 전체 테스트셋 및 코너케이스에서의 성능으로 구분하였다. 그림 8의 코너케이스는 객체의 바운딩박스에 LiDAR point가 절반 이상 존재하지 않는 시퀀스가 50프레임 이상 유지된 경우로 선정하였다. 객체의 거리 기준은 검출된 바운딩박스과 깊이 이미지를 매칭한 이후 바운딩박스의 중앙 픽셀의 깊이 값을 사용하였다. 실험 결과에서 볼 수 있듯이, OADE 알고리즘을 적용하면 그렇지 않았을 때보다 대부분 향상된 거리 추정 성능을 보인다. 특히 코너케이스에서 0.0262의 절대상대오차 결과를 보이며, 이는 OALP를 통한 LiDAR 맵핑이 효과적임을 의미한다. 객체 별 거리 추정은 예측된 바운딩박스를 기반으로 수행되었기 때문에 객체 검출 모델의 성능을 함께 고려할 필요가 있다. 이러한 측면에서 클래스 'Person'에서의 성능 하락은 객체 검출 모델의 평가 결과와 함께 분석할 필요가 있다.

표 1. 객체 별 거리 추정 성능 비교 실험

Table 1. Performance comparison of object-aware depth estimation

Case	Class	Method	ARE
Total test sequence	All	Vanilla	0.0476
	Golfcart		0.0467
	Person		0.0675
	All	OADE	0.0427
	Golfcart		0.0410
	Person		0.0707
Corner case	All	Vanilla	0.0362
		OADE	0.0262



그림 8. 코너케이스 예시
Fig. 8. Example of corner case

표 2는 pseudo 레이블링을 적용하기 전과 후의 객체 비교실험 결과를 보여주며, 모든 클래스에 대한 D_1 과 D_3 모델의 성능을 평가하였다. 모든 평가지표에서 pseudo 레이블링을 사용하였을 때 검출 성능이 개선된 것을 볼 수 있고, 특히 'Person' 클래스에서 mAP50-95 측면의 결과가 두드러진다. 이를 통하여 'Person' 클래스에서 D_1 의 낮은 검출 성능으로 'Person' 클래스에 해당하는 객체가 많이 검출되지 않았음을 알 수 있다. 또한 적은 예측 결과가 객체 별 거리추정 성능에도 영향을 주었을 것으로 해석할 수 있다. 그림 9의 정성적 객체 검출 결과에서는 첫 번째 및 두 번째 열이 각각 D_1 과 D_3 모델의 검출 결과를, 각 행은 제한적인 상황을 의미한다. 정성적 결과를 통하여 D_1 모델이 검출하지 못했던 작은 객체, 블러한 이미지나 객체의 일부가 가려진 경우 등을 D_3 모델은 정확하게 검출한 것을 볼 수 있다. 정량적, 정성적인 객체 검출 결과를 통하여 pseudo 레이블 기반의 연속학습이 시간 효율과 알고리즘의 성능 개선에 효과적임을 알 수 있다.

제안하는 OALP 기법과 기존의 깊이 추정 모델의 정성적인 비교 실험 결과는 그림 10에서 볼 수 있으며, 각 열은 코너케이스의 예시를 의미하고 각 행은 순서대로 RGB 이미지, 기존 모델 및 OALP를 적용한 모델의 예측 결과를 의미한다. 입력 및 예측 영상에서의 관심 객체는 황색 상자로 표시하였다. 기존 모델의 예측 결과에서는 LiDAR의 FoV에서 벗어난 영역부터 거리를 급격하게 가깝게 추정하는 경향을 보이며, 객체의 윗부분 또한 객체보다 가깝게 추정하였다. 이와 다르게 제안하는 알고리즘의 예측 결과에서는 FoV

표 2. 객체 검출 모델의 성능 비교 실험

Table 2. Performance comparison of object detection models

Class	Method	Precision	Recall	mAP50	mAP-95
All	Without pseudo labeling (D_1)	0.84	0.706	0.75	0.574
Golfcart		0.975	0.921	0.957	0.867
Person		0.705	0.491	0.543	0.281
All	With pseudo labeling (D_3)	0.975	0.934	0.983	0.904
Golfcart		0.998	0.948	0.991	0.964
Person		0.953	0.921	0.975	0.844

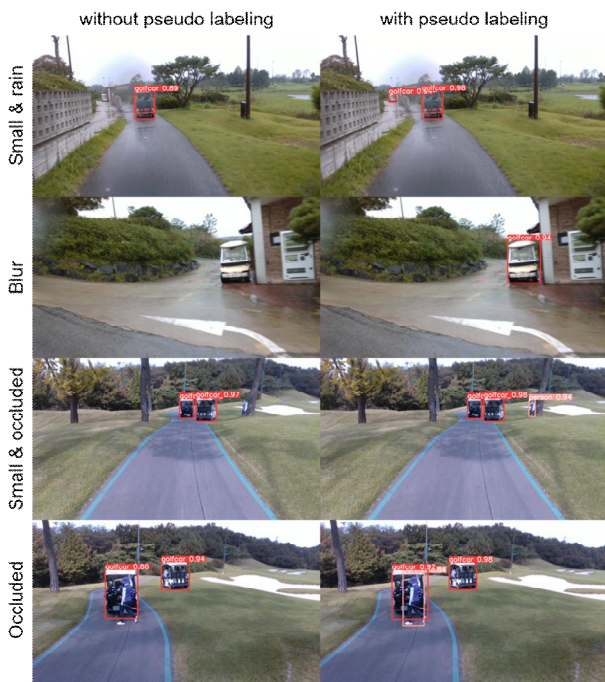


그림 9. 정성적 객체 검출 결과 비교

Fig. 9. Comparison of qualitative object detection results

의 경계에서의 급격한 변화가 사라진 것을 볼 수 있다. 이러한 결과에서 OALP를 통한 LiDAR 맵핑으로 코너케이스에서의 성능 개선에 효과가 있음을 알 수 있으며, 관심 객체 영역에서 LiDAR point가 없는 구역과 존재하는 구역의 추정된 거리가 일정한 것을 볼 수 있다.

본 논문에서 제안하는 객체 별 거리추정 기술의 처리 과정은 영상의 입력력 과정과 객체 검출 모델의 처리과정, 그리고 깊이 추정 모델의 추론과정으로 구성된다. 골프카트에 설치된 Jetson AGX orin 기반의 임베디드 시스템에서 제안 알고리즘의 단계별 처리시간은 표 3과 같다. 제안 알고리즘은 입력 영상에 있는 모든 객체들에 대하여 거리를 추정하는 데에 약 2.38초가 소요된다. 과도한 처리시간으로 인하여 제안기술이 골프카트 시제품에 즉시 적용되기는 어려우나, 딥러닝 모델의 경량화 기술을 추가적으로 연구하여 실시간 처리가 가능한 거리추정 기술을 개발할 계획이다.

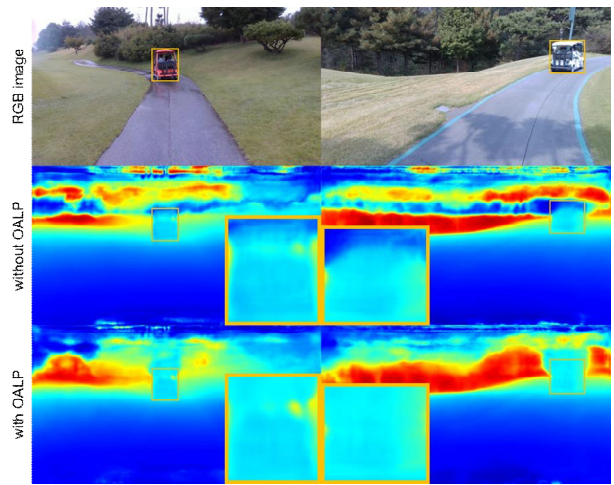


그림 10. 정성적 깊이 추정 결과 비교

Fig. 10. Comparison of qualitative depth estimation results

표 3. 알고리즘 단계별 처리시간

Table 3. Processing time of the proposed method

Processing steps	Duration (sec)
Preprocessing & post-processing	0.0226
Object detection model	0.0546
Depth estimation model	2.3111
Total	2.3883

III. 결론

본 연구에서는 카메라 기반의 충돌방지 기술 개발을 위한 CC-데이터셋과 OADE 알고리즘을 제안한다. OADE를 구성하는 객체 검출 모델의 학습에는 pseudo 레이블링 기반의 연속학습을, 깊이 추정 모델의 학습에는 OALP를 설계하였다. CC-데이터셋에서 OADE 알고리즘의 객체 별 거리추정 결과는 테스트셋 뿐만 아니라 특히 코너케이스에서 더 좋은 성능을 보인다. 객체 검출 클래스를 사람 및 골프카트로 설정하였을 때 높은 검출 및 추정 성능을 보였으며, 이에 따라 골프장 주행 환경에서 장애물로 인식되는 다른 객체들을 검출 대상에 확장하여 높은 성능 기대할 수 있다. 향후 연

구에서는 봄, 겨울과 안개, 눈 등 다른 계절과 날씨의 데이터를 추가적으로 수집하여 제안하는 CC-데이터셋에 추가할 예정이다. 또한 OADE 알고리즘을 저전력 임베디드 시스템에서 실시간으로 구동하기 위한 최적화 및 경량화를 수행할 예정이다. 본 논문에서 제안하는 OADE 알고리즘은 객체를 인식하고 거리를 추정하는 점에서 자율주행 및 매니플레이션 분야 등 다양한 어플리케이션에서 활용될 수 있을 것이라 기대된다.

References

- [1] H. Kang, "고령화 사회와 환대산업의 전망," Journal of Korea Entertainment Industry Association, Vol. 10, No. 1, pp. 10-16, 2018.
- [2] M. Kim, S. Kang, "Demand Prediction Analysis of the Domestic Golf Courses Applying Exponential Smoothing Model," Korean Journal of Convergence Science (KJCS), Vol. 11, No. 11, pp. 113-131, 2022.
- [3] J. Kim, S. Ye, Y. Kang, S. Han, "Design and Implementation of Android Based Golf Cart Autonomous Driving Simulation System," Proceedings of the Korea Information Processing Society Conference, Vol. 19, No. 1, pp. 843-845, 2012.
- [4] J. Choi, I. Yeo, J. Jo, "Autonomous Traveling of Unmanned Golf-Car using GPS and Vision System," Journal of the Korean Society for Precision Engineering, Vol. 26, No. 6, pp. 74-80, 2009.
- [5] H. Park, J. Lee, K. Lee, I. Choi, "Development of Cost-effective Self-Driving Module for Golf Cart," Proceedings of Institute of Control, Robotics and Systems, pp. 208-209, 2020.
- [6] J. Kim, G. Son, Y. Jung, "Yolo v3-based Golf Course Environmental Object Detection Method," The Journal of Korean Institute of Communications and Information Sciences, pp. 1164-1165, 2020.
- [7] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," Proceedings of Advances in Neural Information Processing Systems, Vol. 28, 2015.
- [8] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, "Feature Pyramid Networks for Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2117-2125, 2017.
- [9] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko, "End-to-End Object Detection with Transformers," European Conference on Computer Vision (ECCV), pp. 216-229, 2020.
- [10] J. Lee, M. Han, D. Ko, I. Suh, "From Big to Small: Multi-Scale Local Planar Guidance for Monocular Depth Estimation," ArXiv:1907.10326, 2019.
- [11] S. Bhat, I. Alhashim, P. Wonka, "AdaBins: Depth Estimation Using Adaptive Bins," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4009-4018, 2021.
- [12] Z. Li, X. Wang, X. Liu, J. Jiang, "BinsFormer: Revisiting Adaptive Bins for Monocular Depth Estimation," ArXiv:2204.00987, 2022.
- [13] 송지민, 이상준, "비지도학습 기반의 딥스 추정을 위한 지식 증류 기법," 대한임베디드공학회논문지, 제17권, 제4호, pp. 209-215, 2022.
- [14] D. Tsai, S. Worrall, M. Shan, A. Lohr, E. Nebot, "Optimising the Selection of Samples for Robust Lidar Camera Calibration," Proceedings of 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 2631-2638, 2021.
- [15] J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching," Communications of the ACM, Vol. 18, pp. 509-517, 1975.
- [16] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
- [17] C. Wang, H. Liao, Y. Wu, P. Chen, J. Hsieh, I. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 390-391, 2020.
- [18] S. Liu, L. Qi, H. Qin, J. Shi, J. Jia, "Path, "Aggregation Network for Instance Segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8759-8768, 2018.
- [19] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. Zitnick, "Microsoft COCO: Common Objects in Context," Proceedings of European Conference on Computer Vision (ECCV), pp. 740-755, 2014.
- [20] M. Yang, K. Yu, C. Zhang, Z. Li, K. Yang, "DenseASPP for Semantic Segmentation in Street Scenes," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3684-3692, 2018.
- [21] G. Huang, Z. Liu, L. Maaten, K. Weinberger, "Densely Connected Convolutional Networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700-4708, 2017.
- [22] A. Geiger, P. Lenz, R. Urtasun, "Are we Ready for Autonomous Driving? The KITTI Vision Benchmark Suite," Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3354-3361, 2012.

Gyutae Hwang (황규태)



2022 Electronics Engineering from Jeonbuk National University (B.S.)
2022~Division of Electronics and Information Engineering from Jeonbuk National University (M.S.)

Career:

2022 Electronics and Telecommunications Research Institute (Researcher)

Field of Interests: Deep learning, Computer vision, Biometric estimation, Model compression

Email: gyutae741@jbnu.ac.kr

Jimin Song (송 지 민)



2022 Electronics Engineering from Jeonbuk National University (B.S.)
2022~Division of Electronics and Information Engineering from Jeonbuk National University (M.S.)

Field of Interests: Artificial intelligence, Computer vision, Deep learning, Robotics

Email: jimn_song@jbnu.ac.kr

Sang Jun Lee (이 상 준)



2011 Electrical Engineering from POSTECH (B.S.)
2018 Electrical Engineering from POSTECH (Ph.D. & M.S.)

Career:

2018~2020 Samsung Advanced Institute of Technology (Senior Researcher)

2020~Jeonbuk National University (Assistant Professor)

Field of Interests: Deep learning, Computer vision, Image processing, Robotics

Email: sj.lee@jbnu.ac.kr