

새로운 깊이 예측 백본 네트워크 모델을 사용한 새로운 딥 러닝 기반 객체 검출 방법

경희대학교 컴퓨터공학과
KYUNGHEE UNIV.

Department of Computer Engineering
kairos9603@khu.ac.kr 신은섭
ocn54321@hanmail.net 표승우
shbae@khu.ac.kr 배성호

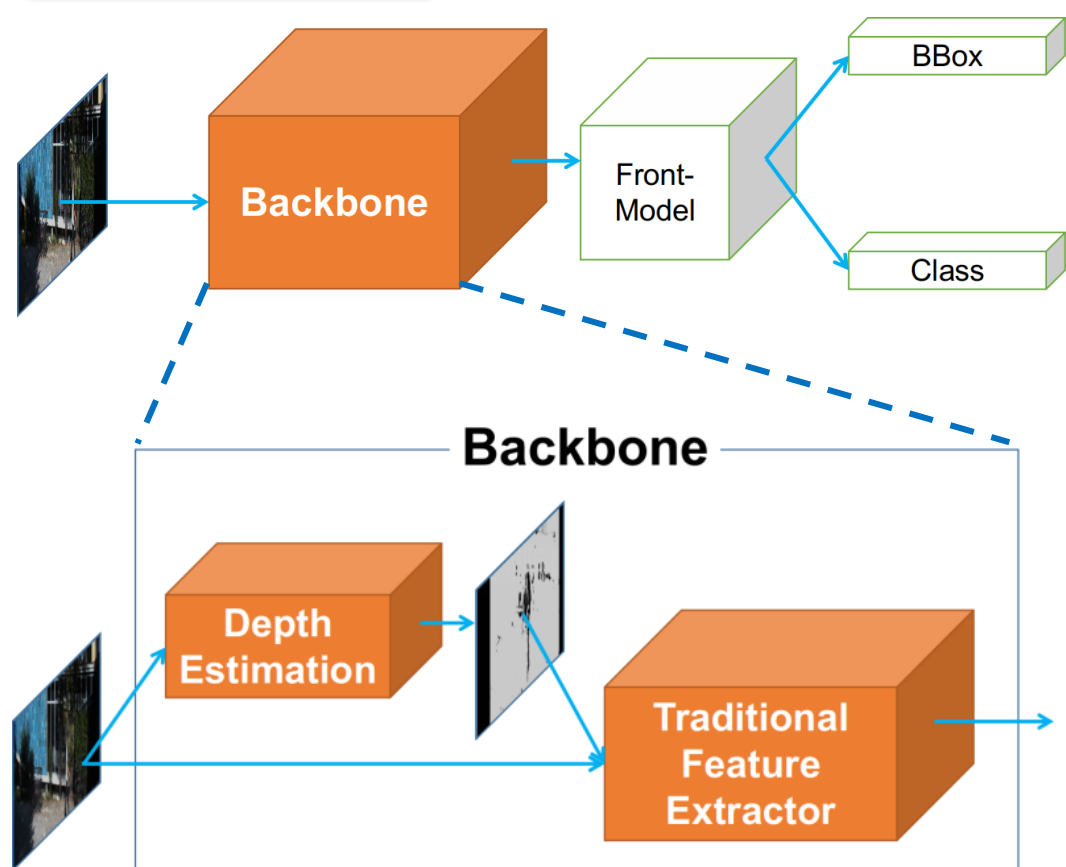
연구배경

최근 딥 러닝을 이용한 객체 검출 방법은 괄목할 만한 성능을 보임. 그러나 물체가 **배경 텍스처**와 유사하거나, 물체의 **크기가 작을 경우** 검출율이 떨어지는 문제가 존재함.
이러한 문제를 해결 하고자 RGB이미지로 부터 추출해 내는 특징뿐만 아니라, **자체 예측한 깊이 맵**에서의 특징도 추출하는 **새로운 백본 모델**을 제안함.

선행연구

선행연구	해당연구
깊이 맵 정보를 외부에서 제공[1]	자체적인 깊이 맵 예측
Classification 과업에서 좋은 성능을 보인 모델을 백본으로 사용[1,2]	기존 특징 추출 및 자체 생성한 깊이 정보로부터 추가적인 특징을 추출하는 새로운 백본 사용
이미지와 깊이 맵을 이용하여 각각 학습[2]	이미지와 깊이 맵을 동일한 네트워크로 한번에 학습

제안하는 모델



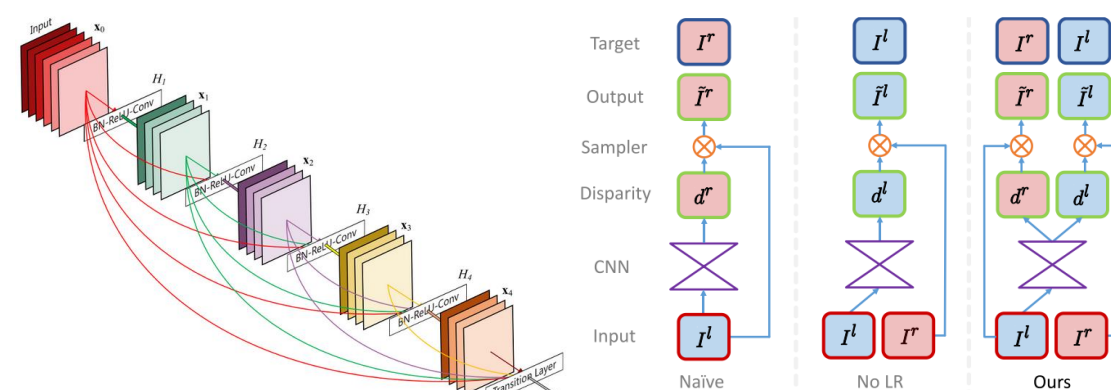
[그림 1] 제안하는 모델의 구조

기존의 객체 검출 모델 구조에서 Backbone은 Feature Extractor로서 사용되고 여기서 추출된 Feature를 이용하여 Front Model에서 실질적인 Object Detection을 수행함.

기존의 모델들은 Backbone을 Classification 과업에서 좋은 성능을 보인 모델을 사용함. 제안하는 모델은 기존의 **Backbone 내부에서 깊이 맵을 예측**, Feature를 추출, 이로서 RGB영상에는 없는 **Geometric정보**를 추가적으로 확보 가능.

실험 설계

▶ 모델 설계



[그림2] DenseNet

[그림3] C.Gordard 모델

실험에 사용된 모델의 전체적인 구조는 [그림1]과 같음.
Depth Estimator 로 C.Gordard[그림 3]의 모델을, Feature Extractor로 DenseNet[그림 2]을 사용함. Front Model은 RetianNet을 이용함.

▶ 데이터 셋

KITTI의 Object Detection Dataset에서 **Left RGB 이미지**와 Bounding Box와 Class가 기록된 **Annotaion** 데이터를 사용하였고, 추가적인 깊이 맵을 만들기 위해 Right RGB 이미지를 사용함. KITTI 데이터셋에서는 **깊이 맵**을 제공하지 않아 Left, Right 이미지를 입력으로 **openCV를 이용하여 생성**.

실험결과

	Total Loss	Regression Loss	Classification Loss	mAP
Baseline	0.3071	0.2720	0.0351	0.4036
Ours	0.1002	0.0973	0.0029	0.4006

본 실험에서 사용된 Baseline은 Front model로 사용한 2018년 가장 좋은 성능을 보인 모델인 RetinaNet을 사용.

실험 결과 Baseline 보다 **loss가 많이 줄었음**을 볼 수 있음, Object Detection 모델의 주요 성능 지표로 사용되는 **mAP** 또한 Baseline과 **비슷함**을 보여줌.

결론 및 향후연구

본 연구에서는 Feature Extractor로서 사용되는 **Backbone 네트워크의 새로운 구조**를 제안함. Backbone 네트워크 내부에서 **깊이 맵을 예측**하여 깊이 맵과 관련 깊은 새로운 특징을 추출 하도록 설계함.

학습결과 기존의 모델보다 좋은 성능을 보이진 못하였는데 OpenCV로 생성한 깊이 맵이 부정확 한 것이 문제로 보임. 이 점을 보완하여 더 좋은 성능을 낼 수 있도록 할 것임.

