Delving into Localization Errors for Monocular 3D Object Detection

설재민 연구원

Software Engineer

자율주행팀

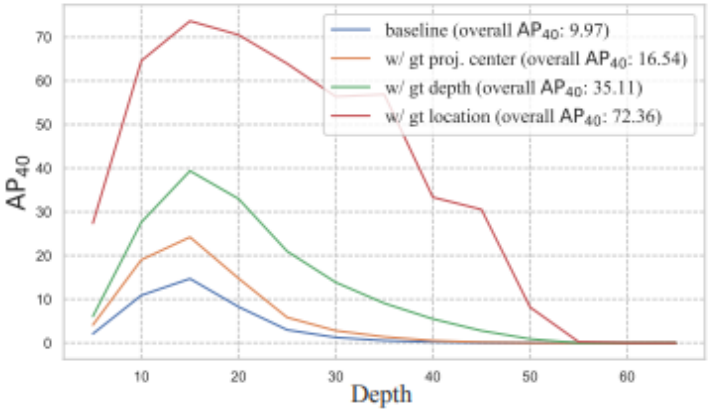
MORAI

Observation 1

- Localiazation을 사용하니 성능이 향상하여 Lidar 기반과 비슷한 성능을 보임
  - Localiozation Error가 Key factor
- 3D 물체의 중심점으로 projection하는것 또한 성능 향상에 영향을 끼침
  - 2D bbox의 중심점과 3D 물체의 projection된 점 사이의 좌표가 같지 않으면 재정렬
- 2D object Dectection related branch들을 유지하는것이 중요
  - 3D detection의 feature를 학습하는데 사용 가능한 auxuiliary task

Observation 2

- 거리가 멀어짐에 따라 정확도가 급격히 감소
  - 거리가 멀수록 localization error는 증가하며 이러한 문제는 피할 수 없음
- 먼 물체와 가까운 물체는 네트워크의 입장에서 학습 시 차이가 굉장히 큼
  - 먼 물체에 대한 데이터셋을 trainset에서 제거하여 trainloss를 감소시킴



Copyright © MORAI Inc. 2021 All rights reserved.

Observation 3

- 기존의 방법들은 각각의 요소(component)들을 독립적으로 따로 optimize
- 각각의 요소들의 contribution을 반영하지 못하면 적절한 optimize가 불가능할 수 있음
- 3D size estimatio을 위한 IOU Loss라는 것을 제안
- 각각의 요소들에 대한 contribution에 따라 loss를 적용

Copyright ⓒ MORAI Inc. 2021 All rights reserved.

3

Problem Definition

- 각 물체에 대한 카테고리를 나타내는 2D 바운딩 박스와 3D 바운딩 박스가 존재
- 2D Bbox의 정보를 사용하여 3D object detection
- RGB 이미지와 이에 대응하는 camera parameter가 input으로 주어질 때, 3D공간에서 물체에 대한 interest를 classify하고 localize

Copyright © MORAI Inc. 2021 All rights reserved.

CenterNet

- **Backbone : DLA-34(speed-accuracy trade off를 위해 사용)**
- **7개의 light weight head가 그 위에 구현 (3 X 3 conv 와 1 X 1 conv)**
  - 2D detectio과 3D detection을 위해 사용
- **2D detection 과 3D detection 부분이 나누어있음**
- **2D detection**
  - 모델1은 detect한 모델의 classification score와 중심 좌표 C = (u,v)로 표현하는 heatmap을 출력
  - C는 bbox의 센터인 GT데이터로, supervised learning으로 학습
  - 모델 2는 2D bbox와 대략적인 중심점의 차이를 예측하는 모델
  - 모델3은 2d bbox의 크기 [w`,h`]를 예측하는 모델
  - 즉 모델1과 모델2를 이용하여 중심좌표에서 얼만큼 떨어져있는 오브젝트인지 예측하며 모델3을 이용하여 해당 물체의 크기를 예측

Copyright ⓒ MORAI Inc. 2021 All rights reserved.

5

CenterNet

- **3D detection**
  - 모델1은 Projection된 3D detection의 중심을 예측
  - 카메라 내부 파라미터 K를 이용하여 3차원의 좌표 계산 가능
  - 예측한 3D detection의 중심과 z값과 인트린직 파라미터의 인버스를 곱하여 3차원 좌표 구할 수 있음
  - 모델 2는 3D 공간의 Heading Angle 예측
  - 모델 3은 3D bbox의 size 예측

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3D} = \mathbf{K}^{-1} \begin{bmatrix} \mathbf{c^w} \cdot z \\ z \end{bmatrix} = \mathbf{K}^{-1} \cdot \begin{bmatrix} x^w \cdot z \\ y^w \cdot z \\ z \end{bmatrix}_{2D},$$

Copyright © MORAI Inc. 2021 All rights reserved.
6

Method	Extra data	3D			BEV			AOS			Runtime
		Easy	Mod.	Hard	Easy	Mod.	Hard	Easy	Mod.	Hard	
Decoupled-3D [4]	Yes	11.08	7.02	5.63	23.16	14.82	11.25	87.34	67.23	53.84	-
AM3D [26]	Yes	16.50	10.74	9.52	25.03	17.32	14.91	-	-	-	~400 ms
PatchNet [25]	Yes	15.68	11.12	10.17	22.97	16.86	14.97	-	-	-	~400 ms
D4LCN [12]	Yes	16.65	11.72	9.51	22.51	16.02	12.55	90.01	82.08	63.98	-
Kinematic3D [3]	Yes	19.07	12.72	9.17	26.69	17.52	13.10	58.33	45.50	34.81	120ms
GS3D [22]	No	4.47	2.90	2.47	8.41	6.08	4.94	85.79	75.63	61.85	~2000 ms
MonoGRNet [30]	No	9.61	5.74	4.25	18.19	11.17	8.73	-	-	-	60 ms
MonoDIS [36]	No	10.37	7.94	6.40	17.23	13.19	11.12	-	-	-	-
M3D-RPN [2]	No	14.76	9.71	7.42	21.02	13.67	10.23	88.38	82.81	67.08	161 ms
SMOKE [24]	No	14.03	9.76	7.84	20.83	14.49	12.75	92.94	87.02	77.12	**30 ms**
MonoPair [10]	No	13.04	9.99	8.65	19.28	14.83	12.89	91.65	86.11	76.45	57 ms
Ours	No	**17.23**	**12.26**	**10.29**	**24.79**	**18.89**	**16.00**	**93.46**	**90.23**	**80.11**	40 ms
Improvement	-	+2.47	+2.27	+1.64	+3.77	+4.06	+3.11	+0.52	+3.21	+2.99	-

Table 3: **Performance of the Car category on the KITTI *test* set.** Methods are ranked by moderate setting (same as KITTI leaderboard). We highlight the best results in **bold** and the second place in underlined.

Method	3D@IOU=0.7			BEV@IOU=0.7			3D@IOU=0.5			BEV@IOU=0.5		
	Easy	Mod.	Hard	Easy	Mod.	Hard	Easy	Mod.	Hard	Easy	Mod.	Hard
CenterNet [44]	0.60	0.66	0.77	3.46	3.31	3.21	20.00	17.50	15.57	34.36	27.91	24.65
MonoGRNet [30]	11.90	7.56	5.76	19.72	12.81	10.15	47.59	32.28	25.50	48.53	35.94	28.59
MonoDIS [36]	11.06	7.60	6.37	18.45	12.58	10.66	-	-		-	-	
M3D-RPN [2]	14.53	11.07	8.65	20.85	15.62	11.88	48.53	35.94	28.59	53.35	39.60	31.76
MonoPair [10]	16.28	12.30	10.42	24.12	18.17	15.76	55.38	42.39	**37.99**	**61.06**	**47.63**	**41.92**
Ours	**17.45**	**13.66**	**11.68**	**24.97**	**19.33**	**17.01**	**55.41**	**43.42**	37.81	60.73	46.87	41.89
Improvement	+1.17	+1.36	+1.26	+0.85	+1.16	+1.25	+0.03	+1.03	-0.18	-0.33	-0.80	-0.03

Table 4: **Performance of the Car category on the KITTI *validation* set.** Methods are ranked by moderate setting (same as KITTI leaderboard). We highlight the best results in **bold** and the second place in underlined.

Copyright © MORAI Inc. 2021 All rights reserved.