

야지 자율 주행에서 통계적 접근 기반 앵커 박스를 이용한 객체 검출 연구

위 경 철*

A Study on Object Detection using Anchor Box based on Statistical Approach for Autonomous Off-Road Driving

Kyungchul Wee*

요 약

야지 자율 주행을 위해 주행요소에 대한 제어가 중요하므로 주변의 구조물 및 장애물 등 객체 검출은 필수적이다. 특히, 객체 검출에서 손실 함수 수렴과 정확성을 높이는 데 도움을 주는 앵커 박스(Anchor Box)는 매우 중요하다. 현재까지 연구된 앵커 박스는 중형비를 일정 비율을 활용하거나 클러스터링 기법을 사용하여 생성하는데, 출현 빈도가 적은 비율을 갖는 객체에 대해 정확성이 저하되는 단점이 있다. 따라서 본 논문에서는 앵커 박스의 크기와 비율에 대한 설정을 Pascal VOC07 데이터 셋의 GT(Ground-Truth) 박스의 통계를 활용하여 선정하는 기법을 제안하였다. 본 논문에서 정의한 ImageNet 기반 데이터 셋을 활용한 실험결과를 통해 제안 기법이 최신 알고리즘보다 객체 검출 성능이 우수함을 입증한다.

Abstract

It is essential to detect surrounding structures and obstacles because control of driving elements is important for autonomous driving in the field. In particular, the anchor box, which helps to converge the loss function and increase accuracy in object detection, is very important. The anchor box studied so far is created using a certain ratio of the aspect ratio or using a clustering technique, but there is a disadvantage in that accuracy is degraded for objects having a ratio with a small appearance frequency. Therefore, in this paper, we proposed a method of selecting the size and ratio of the anchor box using statistics of the GT box of the Pascal VOC07 dataset. Through the experimental results using the ImageNet-based dataset defined in this paper, the proposed technique proves that the object detection performance is superior to the latest algorithm.

Keywords

object detection, anchor box, statistics, labeling

* 국방과학연구소 연구원
- ORCID: <https://orcid.org/0000-0001-9823-5841>

· Received: Aug. 20, 2020, Revised: Sep. 15, 2020, Accepted: Sep. 18, 2020
· Corresponding Author: Kyungchul Wee
Dept. of Intelligence & Information Technology, Agency for Defence Development
Yuseong P.O.Box 35, Daejeon, 34186, KOREA
Tel.: +82-42-821-3442, Email: kcw@add.re.kr

I. 서 론

무인차량의 자율주행에는 객체 검출이 필수적이다. 자율주행 시 주변 환경의 객체를 검출함으로써 주행속도, 방향전환, 긴급제동 등의 주요 주행요소에 대한 제어가 필요하기 때문이다. 특히 야지 및 험지에서의 자율주행 시 객체 검출의 중요성은 더욱 높아진다. 왜냐하면 갑작스러운 객체의 출현 및 먼지, 안개 등의 열악한 환경이 대부분이기 때문이다. 이에 따라 야지환경의 경우, 출현할 수 있는 객체 등을 정의해야 하고, 이에 따른 객체 검출을 수행할 필요성이 있다.

객체 검출이란 영상 내에서 알고리즘을 학습시킬 때 사용된 클래스 정보에 속하는 객체를 분류(Classification)하고, 그 위치 정보(Localization)를 박스 형태로 알려주는 것이다. 객체 검출은 크게 2가지로 나뉜다. 영역 제안 네트워크(RPN, Region Proposal Network) 없이 한 번에 클래스 정보 및 위치정보를 찾는 1단계 검출기(One-stage detector) [1]-[4]와 영역 제안 네트워크를 이용하여 객체의 클래스 및 위치 정보를 찾는 2단계 검출기(Two-stage detector) [5]-[7]가 있으며, 모두 정확성 및 속도를 개선하기 위해 지금까지 활발한 연구가 진행되고 있고, 바운딩(Bounding) 박스 후보를 제안하는데 활발한 연구가 진행되었다.

1단계 검출기인 YOLO(You Only Look Once)-v1은 입력영상을 소수의 바운딩 박스를 생성하여 예측을 진행하므로 작은 객체가 있을 경우 성능이 저하되는데, 이로 인해 기준이 되는 앵커 박스의 필요성이 대두되었다. 2단계 검출기인 R-CNN(Regions with CNN features), Fast R-CNN은 영역 제안 시 선택적 탐색(Selective search)을 통해 영상으로부터 객체가 존재할 위치에 바운딩 박스를 제안했는데 이는 외부 알고리즘으로 병목현상(Bottleneck)이 발생했다. 이로 인해 객체를 검출하는데 일정 기준을 가지고 있는 앵커 박스의 필요성이 증대되고 있었다.

객체 검출 기법에서 앵커 박스는 객체 검출을 용이하게 하고, 손실 함수(Loss function)를 빠르게 수렴하게 한다. 현재는 많은 논문에서 앵커 박스를 이용해 바운딩 박스를 효율적으로 생성하는 연구가

활발히 진행되고 있다. 그 방법에는 일정 비율로 만드는 방법과 유클리디안 거리 및 IoU(Intersection over Union)를 통해 클러스터링 기법을 사용하여 생성하는 경우도 있다. 그러나 이러한 기법들은 특정 범위의 객체만 검출하는 한계점이 있다.

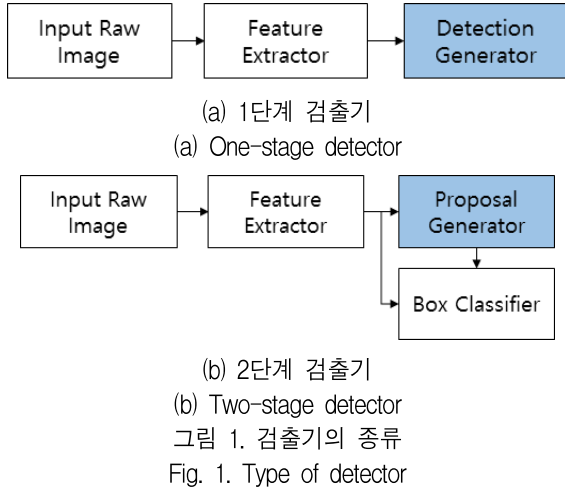
본 논문에서는 이러한 한계점을 극복하기 위해 통계적 기법을 활용한 앵커 박스의 높이, 너비, 그리고 종횡비(AR, Aspect Ratio)의 설정 기법을 제안하여 어떤 종횡비의 객체라도 검출이 용이하게 했다. 앵커 박스는 PASCAL VOC07 데이터 셋[8]의 GT 박스를 활용하여 생성하였고, ImageNet 데이터 셋[9]에 적용하여 실험한 결과 기존 알고리즘 대비 높은 성능을 보임으로써 제안된 방법의 우수성을 입증하였다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 연구에서 앵커 박스의 활용 방법 관련 연구를 기술하고, 3장에서는 본 논문에서 제안한 앵커 박스 생성 기법에 대해 설명한다. 4장에서는 야지 데이터 생성 방법 및 앵커 박스를 사용하는 최신 알고리즘인 YOLO-v3와 제안하는 기법의 성능을 정량적, 정성적으로 비교, 분석하였다. 끝으로, 5장에서는 결론 및 향후 과제에 대해 기술하였다.

II. 관련 연구

객체 검출 관련 연구 분야에서는 디폴트 박스(Default box)라고도 불리는 앵커 박스를 활용하는 방법이 매우 보편적이다. 그 이유는 앵커 박스는 영상 전반에 분포되어 있는 후보 바운딩 박스로써 초기 바운딩 박스를 예측하는데 도움을 주기 때문이다. 이는 학습 시 손실 함수를 비교적 빨리 수렴하게 하고, 추후 객체 검출의 정확성을 높이는 데도 영향을 준다.

앞 절에서 언급하였듯이 객체 검출기는 1단계 및 2단계로 분류될 수 있는데 이 두 가지 검출기 모두 그림 1과 같이 앵커박스가 활용된다. 1단계 검출기의 앵커 박스는 그림 1(a)와 같이 검출 발생기(Detection generator)에서 생성하여 사용되고, 이를 이용하여 학습 및 테스트를 하게 된다. 대표적인 것으로 SSD, YOLO-v2, YOLO-v3가 있다.



2단계 검출기는 그림 1(b)와 같이 영역 발생기 (Proposal generator)에서 생성되어 사용된다. 대표적인 것으로 Faster R-CNN이 있다.

앵커 박스를 정하는 가장 중요한 요인은 박스의 종횡비인 높이와 너비를 설정하는 데 있다[3]. Faster R-CNN은 영역 제안 단계에서 선택적 탐색이 아닌 매우 작은 CNN인 영역 제안 네트워크로 해결하였다. 이 때 앵커 박스의 개념이 처음으로 도입되었는데, 박스를 3가지 스케일과, 1:1, 1:2, 2:1 등의 3가지 종횡비를 설정하였다. 그러나 앵커 박스를 종횡비를 일정한 비율로 설정을 하게 되면 영상 내 전혀 다른 비율의 객체가 출현할 경우 우수하게 검출하지 못하는 단점이 있다[3].

YOLO-v1은 전체 영상에서 98개의 소수의 바운딩 후보 박스와 CNN을 통과한 후 마지막 층 특징맵만 활용하기 때문에 한정된 정보만 존재하여 정확도가 저하되는 단점이 있다. SSD(Single Shot multiplex Detector)는 이를 극복하기 위해 컨볼루션 과정을 거치는 스케일에 따른 중간 특징맵마다 앵커 박스를 적용함으로써 세부적인 정보들이 사라지는 문제점을 해결하였고, 작은 객체들도 보다 우수하게 검출할 수 있었다. 그러나 여전히 앵커 박스는 종횡비를 일정 비율로 지정하여 출현 빈도가 적은 비율을 갖는 객체에 대해서는 학습 속도 및 검출의 정확성이 저하되었다.

YOLO-v2 및 YOLO-v3는 앵커 박스를 단순히 일정한 종횡비를 사용하지 않고, 데이터 셋 GT 박스에 따른 IoU를 거리의 개념을 활용한 K-평균 클러

스터링(K-means clustering) 기법을 통해 앵커 박스의 종횡비를 정했다. 이는 Faster R-CNN, SSD의 앵커 박스에 비해 영상 내 객체들의 차원을 축소시켜 클러스터링을 진행하기 때문에 일반적이고, 효율적인 박스를 생성했다. 이로써 YOLO-v2 및 YOLO-v3는 YOLO-v1에 비해 재현율이 크게 개선되었다. 하지만 K-평균 클러스터링 기법은 매번 특정한 값을 지정해줘야 하고, 주변 잡음까지 같은 그룹으로 묶어버리는 단점이 있다. 따라서 앵커 박스를 활용한 객체 검출에서의 성능개선을 위해서는 일반적으로 활용하는 데이터 셋 GT 박스의 높이와 너비의 상대적 분포에 따른 앵커 박스의 비율을 고려할 필요가 있다.

III. 제안 방법

3.1 GT 박스 기반 개선된 앵커 박스 생성

본 논문에서는 그림 2에서 보는 바와 같이 GT 박스 추출, 정규화, 분류 작업을 통해, 분류별 평균, 표준편차를 구하여 앵커 박스를 생성하였다. 우선 GT 박스 추출을 위해 Pascal VOC07 데이터 셋 9,963장을 활용하여 일반적으로 영상 내 객체 GT 박스를 추출하였다. 앵커 박스 생성 시 과적합(Over-fitting)을 피하기 위해 실험에 이용한 ImageNet 기반 야지 데이터 셋은 활용하지 않았다.

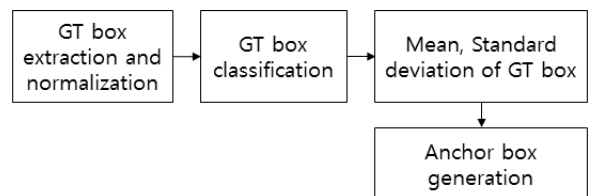


그림 2. 제안 방법 블록도

Fig. 2. Block diagram of proposed method

추출된 GT 박스를 기준으로 그림 3처럼 영상 너비 및 높이가 1일 때 GT 박스를 크기 조정(Resize)을 하는 GT 박스 정규화 과정을 거친다. 이 때 GT 박스의 너비는 w , 높이는 h 이고, 0~1 사이의 값을 갖는다. 정규화 후 GT 박스를 분류하는데 대부분은 너비, 중분류는 식 (1)과 같이 종횡비로 설정한다.

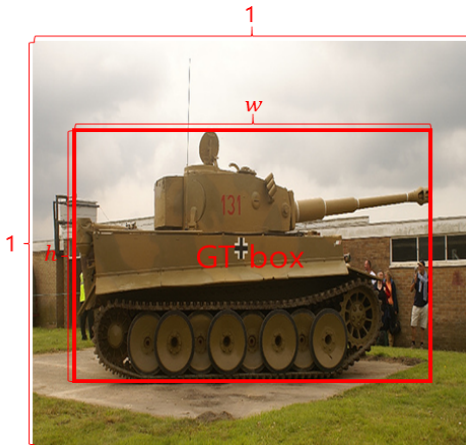


그림 3. 정규화된 GT 박스
Fig. 3. Normalized GT box

$$\text{Aspect Ratio (AR)} = h/w. \quad (1)$$

야지 주행을 할 때 물체는 주로 좌우로 이동하므로 너비는 중요한 요인이고, 일정 범위의 너비에서 높이 비교를 위해 종횡비가 중요하기 때문이다. 대분류, 중분류가 정해지면 Pascal VOC07의 GT 박스를 너비, 종횡비에 따라 균등하게 나누기 위해 그림 4처럼 누적분포함수(CDF, Cumulative Distribution Function)를 활용한다[10].

우선 x 축을 너비로 하는 CDF의 1/3, 2/3에 해당되는 지점에 대한 t, s 값을 구하여 너비를 나누고, 각 너비 범위에 대해 x 축을 종횡비로 하는 CDF의 t, s 값을 구하여 중분류를 나누어 표 1처럼 총 9개의 범위를 설정하였다. 다음 과정은 분류된 9개의 GT 박스 그룹에 대해 GT 박스들의 너비 및 높이의 평균과 표준편차를 구한다. 너비의 평균 μ_w 는 식

(2), 높이의 평균 μ_h 는 식 (3)과 같다. n 은 해당 앵커 박스에 해당하는 GT 개수이다.

$$\mu_w = \sum_{i=1}^n w/n, \quad (2)$$

$$\mu_h = \sum_{j=1}^n h/n, \quad (3)$$

너비의 표준편차 σ_w 는 식 (4), 높이의 표준편차 σ_h 는 식 (5)와 같다.

$$\sigma_w = \sum_{i=1}^n (\mu_w - w)^2/n, \quad (4)$$

$$\sigma_h = \sum_{j=1}^n (\mu_h - h)^2/n. \quad (5)$$

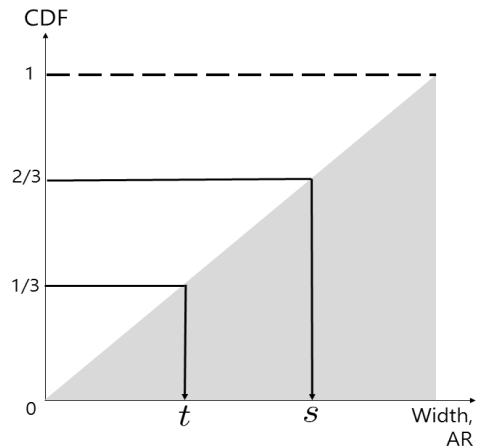


그림 4. 너비 및 종횡비에 대한 누적분포함수
Fig. 4. CDF for width and aspect ratio

표 1. 앵커 박스의 너비, 높이의 평균 및 표준편차

Table 1. Mean and standard deviation of the width and height of anchor box

Width	AR (Aspect Ratio)	μ_w	σ_w	μ_h	σ_h
$0 \leq w \leq 0.3020$	$0 \leq AR \leq 1.2830$	0.1840	0.0703	0.1638	0.0784
	$1.2830 < AR \leq 2.1684$	0.1809	0.0722	0.3061	0.1326
	$2.1684 < AR$	0.1429	0.0747	0.4492	0.2194
$0.3020 < w \leq 0.6458$	$0 \leq AR \leq 0.9423$	0.4804	0.0988	0.3276	0.1117
	$0.9423 < AR \leq 1.4341$	0.4777	0.1003	0.5630	0.1399
	$1.4341 < AR$	0.4335	0.0916	0.7803	0.1442
$0.6458 < w \leq 1$	$0 \leq AR \leq 0.7403$	0.8597	0.1074	0.4730	0.1319
	$0.7403 < AR \leq 0.9813$	0.8629	0.1085	0.7461	0.1147
	$0.9813 < AR$	0.8092	0.1088	0.9027	0.0909

무수히 많은 데이터 셋을 모집단, Pascal VOC07 데이터 셋을 표본집단이라 할 때 모집단의 평균은 $(\mu \pm 1.96\sigma / \sqrt{n})$ 의 신뢰구간(Confidence interval)에 95%의 확률로 존재한다. 따라서 도출된 GT 박스 평균은 앵커 박스를 생성하는데 이용된다. 앵커 박스의 너비는 \hat{w} , 높이는 \hat{h} 라 정의하고, 이는 제안기법의 베이스라인(Baseline)인 YOLO-v3의 입력영상 416x416 해상도에서 설정되어야 한다. 따라서 9가지 범위에 해당하는 μ_w, μ_h 에 416을 각각 곱한 후 반올림하면 표 2와 같이 9개의 앵커 박스가 도출된다.

표 2. 앵커 박스의 416x416 해상도에서 너비, 높이 및 종횡비

Table 2. Width, height, and aspect ratio of anchor box at 416x416 resolution

Anchor box	(\hat{w}, \hat{h}) at 416x416 resolution	AR (Aspect ratio)
1	(77, 68)	0.8902
2	(75, 127)	1.6921
3	(59, 187)	3.1435
4	(200, 136)	0.6819
5	(199, 234)	1.1786
6	(180, 325)	1.8000
7	(358, 197)	0.5502
8	(359, 310)	0.8646
9	(337, 376)	1.1155

IV. 실험결과 및 분석

제안한 기법에 대한 우수성을 입증하기 위해 3.1절에서 언급하였듯이 제안기법의 베이스라인은 본 논문에서 소개된 알고리즘 중 가장 속도가 빠르고, FPN(Feature Pyramid Network)를 사용하여 정확성이 높은 YOLO-v3로 설정하였고, 제안기법과 YOLO-v3를 비교 실험하였다[4]. YOLO-v3 구현의 객관성을 유지하기 위하여 관련 저자가 공개한 DarkNet-53 네트워크를 활용하였고, 제안기법은 해당 코드를 기반으로 앵커 박스 부분의 크기 및 비율을 수정 구현하였다. 동등 비교를 위해 배치 사이즈는 16으로 설정하였다. YOLO-v3 알고리즘은 252개 층의 CNN에서 앞의 249개 층을 고정시키고 뒤의 3개를 학습시키는 Freezing 과정과, 전체 252개 층을 학습시키는 Unfreezing 과정이 있는데, Freezing 50회, Unfreezing

50회를 설정하였다. Nvidia Titan X기준으로 Freezing 과정은 회당 평균 1,143초, Unfreezing 과정은 회당 평균 4,947초가 소요되었다. 4.1절에서는 레이블링을 통한 야지 데이터 셋 구성, 4.2절에서는 실험결과와 정량적 분석, 4.3절에서는 실험결과와 정성적 분석으로 나누어 분석한다.

4.1 야지 데이터 셋 구성

본 제안 기법은 야지 및 험지 환경에서의 객체 검출을 목표로 한다. 따라서 인공지능 학습 및 테스트 데이터 셋은 ImageNet의 야지에서 출현 빈도가 높은 객체들로 구성하였다. 클래스(Class)는 개(Bird-dog), 고양이(Jungle-cat), 군인(Soldier), 말(Wild-horse), 멧돼지(Wild-boar), 사슴(Deer), 소(Cow), 염소(Wild-goat), 전차(Tank), 토끼(Rabbit), 트랙터(Tractor) 등 11개이고, 총 5,600장이다.

이 데이터셋에 대해 레이블링(Labeling)을 진행하였는데, 소에 대해 예를 들어보면 그림 5와 같이 좌측 상단을 시작점으로 하여 가로를 x축, 세로를 y축으로 정의하여 물체를 감싸는 바운딩 박스의 좌상단 좌표(x_{min}, y_{min}) 및 우하단 좌표(x_{max}, y_{max})를 기록하였다. 이를 표 3과 같이 TXT 형식으로 저장하였고, 순서대로 해당 영상 경로, 바운딩 박스 최소값 $x(x_{min})$, 최소값 $y(y_{min})$, 최대값 $x(x_{max})$, 최대값 $y(y_{max})$, 클래스 값(class value)의 값으로 표현하였다.

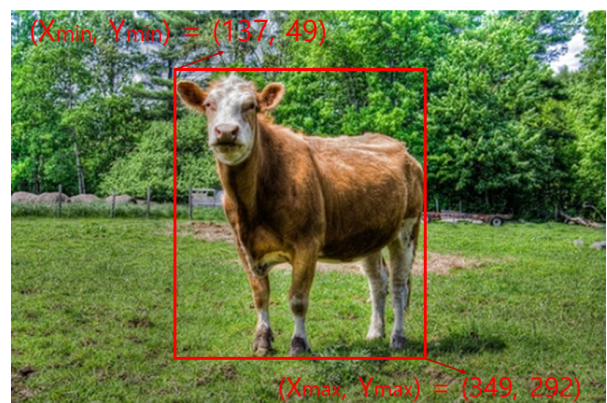


그림 5. 영상의 바운딩 박스
Fig. 5. Bounding box of image

표 3. TXT 형식

Table 3. TXT format

(Image path)	(x_{min}),(y_{min}),(x_{max}),(y_{max}),($class\ value$)
1.jpg	137,49,349,292,1

클래스 값은 표 4와 같이 설정하였다. 레이블링 된 11개의 클래스 데이터(총 5,600장) 중 학습 및 테스트 데이터의 비율을 1:1로 나누었고, 각 클래스 별 학습 및 테스트 비율을 1:1에 가깝게 설정하였다. 학습 및 테스트 데이터 셋 기준은 무작위로 선별하여 활용하였다.

표 4. 클래스 값

Table 4. Class value

Class	Class value
bird-dog	0
cow	1
deer	2
jungle-cat	3
rabbit	4
soldier	5
tractor	6
wild-boar	7
wild-goat	8
wild-horse	9
tank	10

4.2 정량적 분석

객체 검출의 성능을 측정하는 지표는 AP(Average Precision)이다. AP는 특정 클래스에서 재현율 0에서 1까지 0.1단위로 증가할 때 각각에 대응하는 정밀도(Precision) 값들의 평균이다. mAP(mean Average Precision)는 각 클래스의 AP값의 평균값이다. 즉, 최종적으로 모든 클래스를 고려한 정밀도의 평균을 제공하기 위해서 mAP를 사용한다. 제안기법 및 YOLO-v3 모두 GT 및 예측(Predicted) 박스 상호간의 IoU가 0.5이상이면 참 양성(True positive)로 설정하였다. 이와 관련한 mAP를 그림 6에서 그래프로 나타내었다. 각 클래스별 비교를 하면 트랙터, 전차는 두 기법 모두 수치가 높았다. 한편, 사슴, 염소는 테스트 분석 결과 서로의 특징이 비슷해 서로 거짓 양성(False positive)이 있어 수치가 상대적으로 저하되었다.

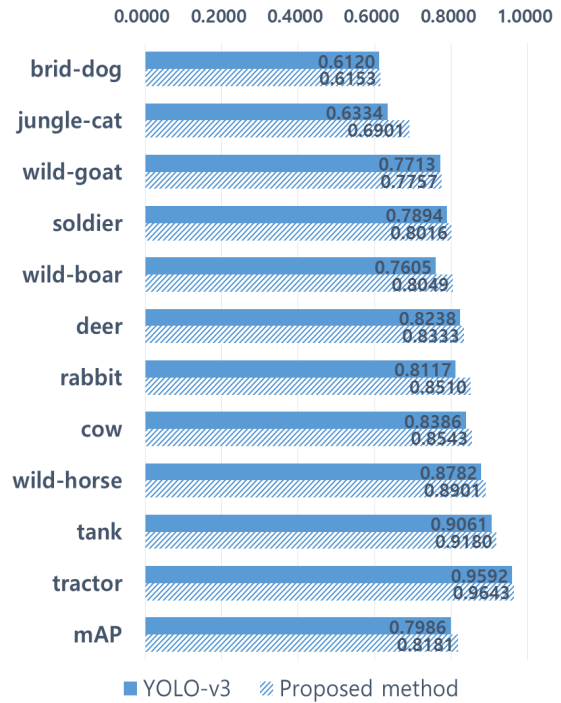


그림 6. 각 클래스 별 AP 및 mAP (IoU = 0.5)
Fig. 6. AP for each class and mAP (IoU = 0.5)

각 기법에 대한 비교를 하면 mAP는 YOLO-v3는 79.86%, 제안 기법은 81.81%의 정확도를 보였다.

4.3 정성적 분석

그림 7은 테스트 결과를 정성적으로 표현한 그림으로써, 영상 내 파랑 실선은 GT, 한 가지 모양의 초록 점선은 참 양성, 두 가지 모양의 빨강 점선은 거짓 양성이다. 개, 전차의 경우 두 기법 모두 참 양성으로 객체 검출을 성공하였으나, YOLO-v3는 개를 군인으로, 전차를 트랙터로 클래스를 잘못 분류하는 거짓 양성도 검출하였다. 이는 앵커박스의 효율적인 생성이 되지 않아 학습을 저하시키고, 결국 테스트 시 각 클래스에서 정확성이 낮아지는 요인을 보였다.

또한, 개, 전차의 경우 참 양성일 때 제안기법이 YOLO-v3에 비하여 IoU 수치가 높다. 이는 제안기법의 앵커 박스가 학습 데이터 셋의 GT 박스와 유사한 너비 및 높이를 갖게 되어 손실 함수가 보다 빠르게 수렴하여, 테스트 시 예측 박스가 GT 박스에 정확하게 위치화가 된 것이다.

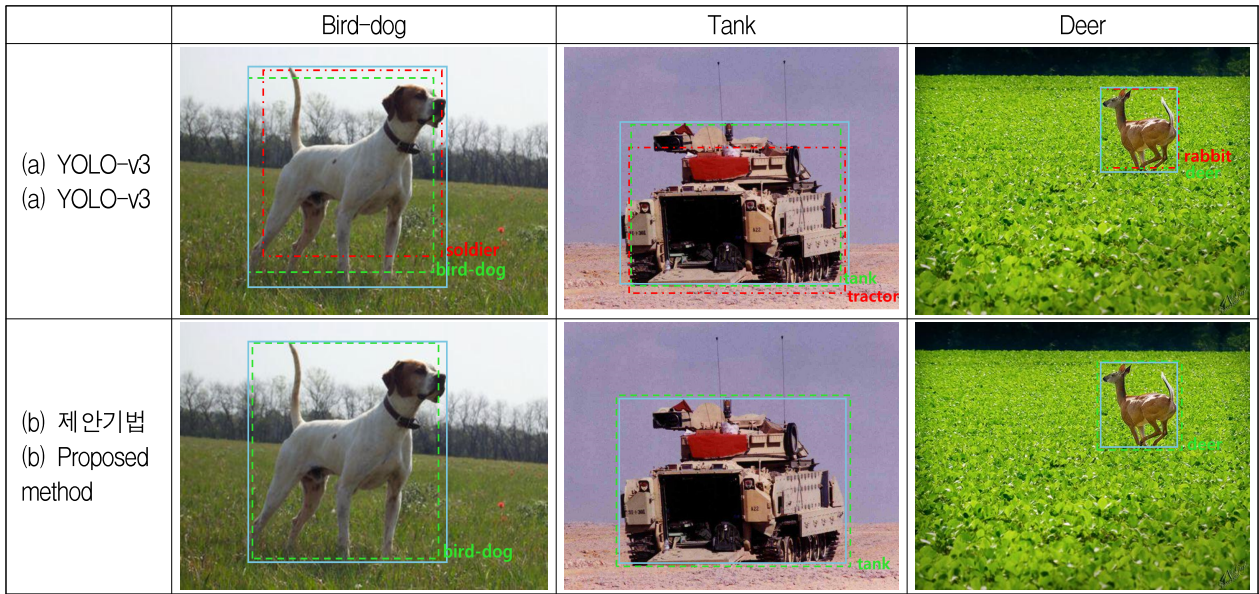


그림 7. YOLO-v3 및 제안기법에 대한 테스트 영상 예
Fig. 7. Examples of test images about YOLO-v3 and proposed method

사슴의 경우 제안기법은 영상 내 물체 크기가 작은 경우에도 학습이 잘 되어 참 양성을 검출했지만, YOLO-v3는 참 양성 이외 사슴을 토끼로 오인하는 거짓 양성을 검출했다. 이는 작은 물체일수록 데이터 셋의 종횡비 변화가 큰데 제안기법은 균등한 범위로 나누어 앵커 박스를 효율적으로 생성하였기 때문에 작은 물체도 정확하게 검출했다.

V. 결론 및 향후 과제

본 논문에서는 야지에서 출현할 수 있는 객체를 효율적으로 검출하기 위해 앵커 박스를 통계적으로 접근하여 우수한 결과를 도출하였다. 그리고 적절한 앵커 박스를 설정하기 위해, Pascal VOC07 데이터 셋 9,963장을 사용하였고, 제안기법의 우수성을 입증하기 위해 5,600장의 ImageNet 기반 야지 데이터 셋을 구성하여 레이블링을 진행하고, YOLO-v3와의 비교 실험을 통해 성능을 정량적, 정성적으로 분석하였다.

본 논문에서는 Pascal VOC07 데이터 셋의 GT 박스에 기반하여 앵커 박스를 설정하였다. 추후 MS COCO[11], Pascal VOC12 데이터 셋[12]을 활용하거나 새로운 데이터 셋을 활용한다면 더욱 일반적인 앵커 박스를 생성하여 성능을 개선할 수 있다. 더

나아가 제안 기법은 RGB 영상만을 사용하였는데, 추후 야지에서 RGB 및 라이다 영상을 정합하여 클래스 및 위치정보를 측정하면 객체를 검출하는데 수월할 것이다.

References

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", in Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, pp. 779-788, Jun. 2016.
- [2] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A.C. Berg, "Ssd: Single shot multibox detector", in Proceedings of the Springer conference on european conference on computer vision, Amsterdam, Netherlands, pp. 21-37, Oct. 2016.
- [3] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger", in Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, pp. 6517-6525, Jul. 2017.
- [4] J. Redmon and A. Farhadi, "YOLOv3: An

incremental improvement", arXiv:1804.02767, Apr. 2018.

- [5] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", in Proceedings of the IEEE conference on computer vision and pattern recognition, Columbus, OH, pp. 580-587, Jun. 2014.
- [6] R. Girshick, "Fast R-CNN", in Proceedings of the IEEE international conference on computer vision, Santiago, Chile, pp. 1440-1448, Dec. 2015.
- [7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", Advances in neural information processing systems, 2015.
- [8] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results", 2007.
- [9] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database", in Proceedings of the IEEE conference on computer vision and pattern recognition, Miami, FL, USA, pp. 248-255, Jun. 2009.
- [10] W. K. Pratt, "Digital Image Processing", New York: Wiley, 1978.
- [11] COCO:Common Objects in Context (2016). <http://mscoco.org/dataset/#detections-leaderboard>. [accessed: 25 Jul. 2016]
- [12] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge", International journal of computer vision, Vol. 88, No. 2, pp. 303-338, Sep. 2009.

저자소개

위 경 철 (Kyungchul Wee)



영상처리

2014년 2월 : KAIST 전기 및
전자공학과(공학사)
2017년 2월 : KAIST 전기 및
전자공학부(공학석사)
2017년 2월 ~ 현재 : 국방과학
연구소 연구원
관심분야 : 인공지능, 자율주행,