

논문 2023-18-32

Refinement Module 기반 Three-Scale 보행자 검출 기법 (A Three-scale Pedestrian Detection Method based on Refinement Module)

정 경 민, 박 수 용, 이 현*
(Kyungmin Jung, Sooyong Park, Hyun Lee)

Abstract : Pedestrian detection is used to effectively detect pedestrians in various situations based on deep learning. Pedestrian detection has difficulty detecting pedestrians due to problems such as camera performance, pedestrian description, height, and occlusion. Even in the same pedestrian, performance in detecting them can differ according to the height of the pedestrian. The height of general pedestrians encompasses various scales, such as those of infants, adolescents, and adults, so when the model is applied to one group, the extraction of data becomes inaccurate. Therefore, this study proposed a pedestrian detection method that fine-tunes the pedestrian area by Refining Layer and Feature Concatenation to consider various heights of pedestrians. Through this, the score and location value for the pedestrian area were finely adjusted. Experiments on four types of test data demonstrate that the proposed model achieves 2-5% higher average precision (AP) compared to Faster R-CNN and DRPN.

Keywords : Deep Learning, Pedestrian Detection, Refinement Module

I. 서 론

보행자 검출 기법은 Object Detection [1, 2]의 특정 객체 검출 방법으로 카메라를 통해 여러 보행자를 검출하는 방법이다. 보행자 검출은 카메라 성능에 따라 보행자의 인상착의, 높이, 폐색현상 등으로 다양한 보행자 특징을 추출하는 것이 중요하다. 특히 사진 속 같은 보행자라도 다른 크기로 보여 보행자를 검출하지 못할 수도 있다. 이를 해결하기 위해 Scale-aware Fast R-CNN (SAF R-CNN) [3], Double Region Proposal Network (DRPN) [4]에서는 보행자의 높이 차이 문제를 초점을 맞춘 모델이다. SAF R-CNN, DRPN은 데이터 셋의 보행자 평균 높이 값을 활용하여 보행자를 Large-scale, Small-Scale 두 가지 Scale로 구분하였다. 하지만 보행자는 두 가지 크기로 구분하기 어렵다. 예를 들어 보행자를 나이별로 나눈다면 어린이, 청소년, 성인 나눌 수 있으며, 높이 차이는 크게 3가지 나뉘질 수 있다. 만약 SAF R-CNN, DRPN에 적용한다면 청소년에 해당하는 보행자의 검출은 부정확할 수 있다. 따라서 본 논문에서는 보행자의 세 가지 그룹의 보행자를 구분하기 위해 Refining Layer [5]와 Feature Concatenation [6, 7]을 결합한 보행자 검출 방법을 제안한다. Refining Layer는 보행자의 영역이 부정확한 문제를 해결하기 위해 고안된 방법으로 보행자 영역에 대한 점수와 위치값을 미세조정한다. Feature

Concatenation은 Refining Layer의 데이터 손실을 방지하기 위해 Refining Layer 이전 Feature와 Refining Layer 이후 Feature를 결합하는 방법이다. 이 두 가지 방법을 통해 Refinement Module을 구성하였다. 이를 통해 세가지 그룹에 대한 보행자 영역을 미세 조정하여 정확한 보행자를 검출하고자 한다. 이후 실험에서는 Faster R-CNN (SRPN) [8], DRPN 과의 Average Precision (AP) [9], 비교를 통해 우위를 가려보려고 한다.

2장에서는 관련 연구에서는 두가지 크기의 보행자 검출 기법, 미세 조정 모듈인 Refining Layer와 Feature Concatenation에 대해 설명한다. 3장에서는 본 논문에서 제안하는 기법에 대한 Refinement Module에 대해 설명한다. 4장에서는 Medium-scale의 최적의 Threshold에 대한 실험, Dataset, Experiment Setting, 평가 기준을 통해 SRPN, DRPN, Three-scale Region Proposal Network (TRPN)의 성능을 비교하였다. 최종적으로 5장에서는 결론과 향후 연구를 언급하고 마치고자 한다.

II. 관련 연구

1. 두가지 크기의 보행자 검출 기법

SAF R-CNN은 Fast R-CNN [10]의 작은 보행자의 검출률을 높이기 위해 보행자를 Large-scale과 Small-scale로 구분하여 검출을 진행하여 문제를 해결하였다. 보행자의 크기는 학습 데이터의 보행자 평균 높이 값을 기준으로 구분하였다. 평균 높이보다 높은 경우 Large-scale, 낮은 경우는 Small-scale로 구분하였다. Large-Scale은 수식 (1), Small-Scale은 수식 (2)를 통해 가중치를 계산하였다. 수식

*Corresponding Author (mahyun91@sunmoon.ac.kr)

Received: Apr. 6, 2023, Revised: May 17, 2023, Accepted: Jul. 16, 2023.

K. M. Jung: Sunmoon University (M.S. Student)

S. Y. Park: University of Rochester (Undergraduate Student)

H. Lee: Sunmoon University (Assoc. Prof.)

※ 본 논문은 문화체육관광부 및 한국콘텐츠진흥원의 2022년도 문화체육관광 연구개발사업으로 수행되었음 (과제명: 인지·신체 복합중재 재활운동 증강 디바이스 기술 개발, 과제번호: SR202106002, 기여율: 100%).

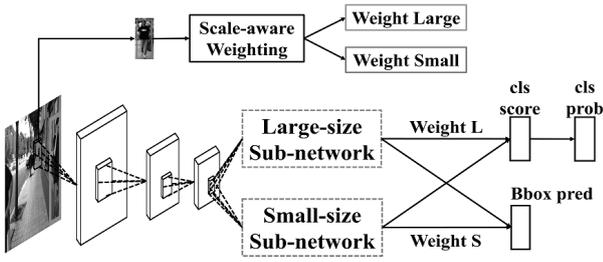


그림 1. SAF R-CNN의 흐름도
Fig. 1. Flowchart of SAF R-CNN

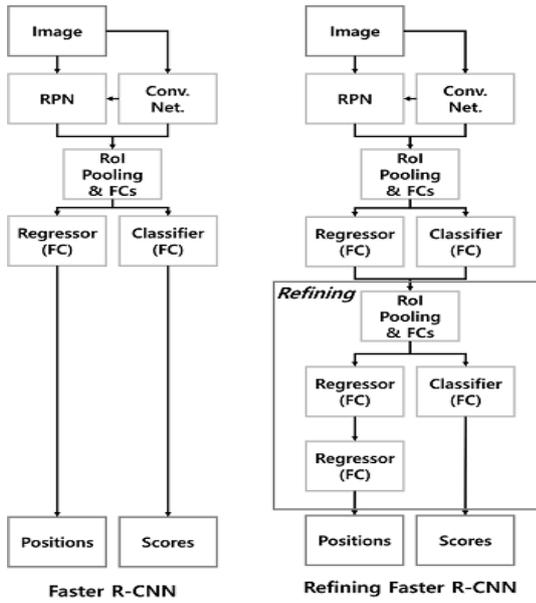


그림 2. RF R-CNN과 Faster R-CNN의 동작 구조
Fig. 2. RF R-CNN and Faster R-CNN Operational Structure

(1)의 $\alpha=1$, $\beta=10$ 으로 초기 설정값을 설정하였고, h 는 구분할 보행자 Proposal의 높이, \bar{h} 는 Dataset의 평균 높이 값을 사용한다.

$$w_l = \frac{1}{1 + \alpha \exp\left(-\frac{h - \bar{h}}{\beta}\right)}, \quad (1)$$

$$w_s = 1 - w_l. \quad (2)$$

수식 (1-2)를 통해 Large-scale, Small-scale은 1에 가까운 가중치를 가지게 된다. 계산된 가중치는 그림 1과 같이 Large-size Sub-network, Small-size Sub-network의 Fully Connected Layer 과정을 거친 Feature vector에 곱해서 보행자를 검출한다. 보행자 높이에 맞는 계산을 진행하여 각 크기에 대한 검출률을 높였다. 하지만 SAF R-CNN 방식은 여전히 Fast R-CNN의 기반 방법으로 CPU 연산으로 학습하여 병목 현상을 일으켰다.

DRPN은 Faster R-CNN의 작은 객체의 검출률을 높이기 위해 SAF R-CNN의 Scale 구분 방식을 사용하였다. 이전

의 Faster R-CNN보다 작은 객체에 대한 인식률을 개선하여 전체적인 인식률이 높아졌다.

2. RF R-CNN

RF R-CNN는 R-CNN [11] 기반의 모델들의 문제 중 하나인 작은 객체의 Proposal이 사라지는 문제를 Refining Layer를 통해 해결하였다. Refining Layer는 Classification refinement와 Regression refinement 두 단계로 진행하였다. 그림 2의 classification refinement는 이전 Convolution layer의 Feature map과 이전 Regressor의 RoI를 입력 데이터로 사용한다. 이후 FC로 구성된 classifier를 통해 보행자에 대한 Score를 한 번 더 계산한다. Regression refinement는 FC Layer로 구성되어 있는 Regressor 과정을 통해 보행자 영역에 대해 한 번 더 계산한다. RF-R-CNN은 RoI Pooling 이후의 한 번 더 검출하는 Refining Layer를 결합함으로써 classification score vector, bounding box vector에 대한 미세조정이 이루어졌고, 보행자의 영역이 비교적 정확해졌다.

III. 본 론

1. TRPN의 구분법

DRPN은 그림 3의 DRPN과 같이 보행자를 Large-scale, Small-scale로 구분한 후 보행자를 검출하였다. 하지만 그림 3의 TRPN과 같이 보행자가 세 가지 Scale로 존재하는 경우 보행자가 애매하게 구분되어 부정확한 보행자를 검출한다. 따라서 사진 속 보행자에 따른 정확한 높이 구분을 위해 새로운 Scale 구분법인 TRPN을 도출하였다.

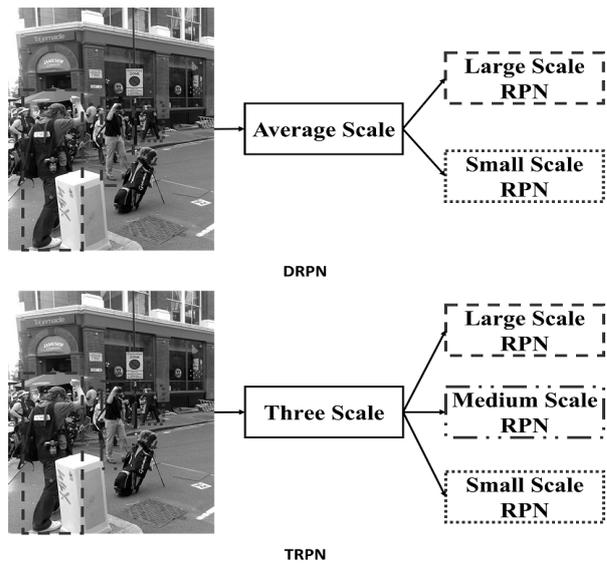


그림 3. DRPN, TRPN 각 모델에 대한 Scale 구분법에 따른 보행자 크기 구분

Fig. 3. Pedestrian size classification by scale classification for DRPN and TRPN models

$$op_m = \frac{op_h}{length(op)}, \quad (3)$$

$$img_h = \frac{input_h}{feat_{stride}}, \quad (4)$$

$$l_h = op_m + img_h, \quad (5)$$

$$s_h = op_m - img_h. \quad (6)$$

새로운 구분법은 DRPN에서 사용하던 수식 (1-2)를 수식 (3-6)으로 변경하여 세 가지 Scale을 구분하는 수식으로 변경하였다.

수식 (3)은 RPN을 통해 추출되는 RoI를 사용하여 추출되는 Object Proposal을 사용하여 계산하였다. $length(op)$ 는 Object Proposal의 최종 개수이고, op_h 는 Object Proposal 높이의 총합이다. 두 개의 값을 나누어 Object Proposal의 평균 높이를 추출하였다. 이후 수식 (4-6)에 활용하였다. 수식 (4)에서 $feat_{stride}$ 의 경우 VGG16 [12]의 Layer의 깊이 값인 16의 값을 사용하였고, $input_h$ 는 학습 데이터의 높이인 624의 값을 사용하였다. 두 개 값을 나누어 최종 Convolution 과정이 끝날 때 산출되는 높이 값과 동일한 값을 산출하였다. 이후 수식 (5)에서 수식 (3), 수식 (4)를 더한 값을 Large-Scale의 기준으로 사용하였고, 수식 (6)에서 수식 (3), 수식 (4)를 뺀 값을 Small-Scale의 기준으로. 수식 (5)보다 큰 경우 Large-scale, 수식 (5) 보다 작고, 수식 (6)보다 큰 경우 Medium-Scale, 수식 (6)보다 작은 경우는 Small-Scale로 구분하였다.

3. Refinement Module

TRPN에서는 수식 (1-6)을 통해 세분화된 높이 계산을 통해 가중치를 계산하였다. 추가로 Refining Layer 통해 보행자 영역을 미세조정하고 가중치를 통하여 조금 더 정확한 Bounding Box vector, class score vector를 계산을 하였다. Refining Layer는 그림 4과 같다. Refine Bounding Box Layer는 Bounding Box Position를 입력 데이터로 사용하였고, Refine Score Layer는 Score Feature vector를 입력 데이터로 사용하였다. 각 Refine Layer는 그림 4과 같이 2개의 Hidden Layer를 가진다. Refine Layer들은 입력 데이터를 20588로 변환해 준 후 4096으로 변환해 주는 작업을 진행한다. 마지막 Layer는 입력 노드와 같은 크기로 변환해 준다. 이후 Refining Layer에서 나온 Feature vector와 각 Scale 별로 계산된 가중치를 곱해주는 작업을 진행한다. 이러한 과정을 통해 Bounding Box Position, Score Feature vector는 한 번 더 검출하는 Refining Layer 보행자의 영역을 미세조정하였다.

Feature Concatenation은 UNet에서 CNN구조의 깊어지는 Layer로 인해 Feature Map이 성능이 떨어지는 문제를 해결한 방법이다. 본 논문의 Feature Concatenation 모듈은 기존 RoI Fooling 과정을 거친 데이터와 Refining Layer의 결과를 결합하여 사용하였다. Refining Layer에서 나온 Feature vector와 가중치를 곱해주어 Weighted Feature vector를 생성한다. Refining Layer는 미세조정을 진행하였지만, Layer

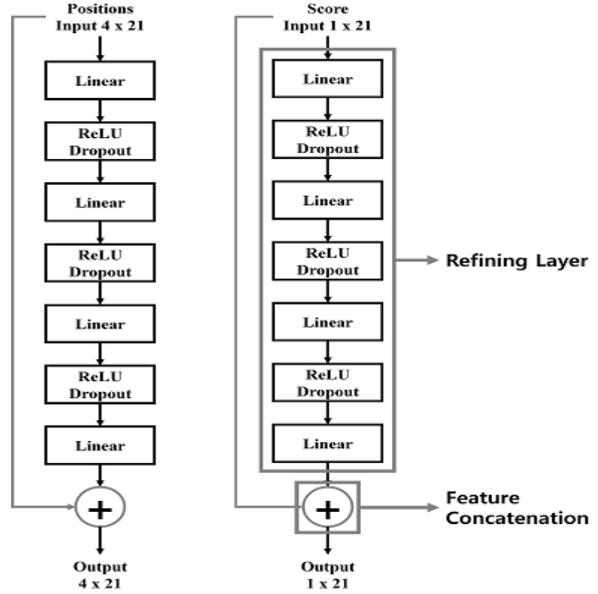


그림 4. Feature Concatenation과 Refining Layer의 결합을 통한 미세조정 모듈의 전체 구조

Fig. 4. Overall structure of fine-tuning module by combining Feature Concatenation and Refining Layer

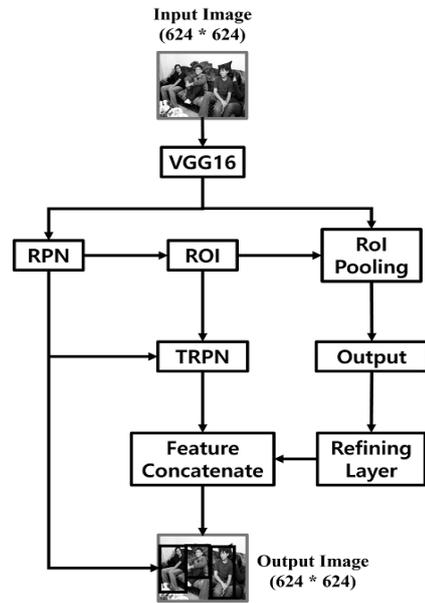


그림 5. 본 논문에서 제안하는 구조

Fig. 5. the structure proposed in this paper

가 깊어져 복잡성이 커져 작은 Proposal에 대한 데이터가 손실될 가능성이 클 수 있다. Feature Concatenation 개념을 활용하여 Refining Layer 이전의 데이터와 결합하였다. Feature vector와 Weighted Feature vector를 결합함으로써 작은 Proposal 데이터의 손실 문제를 해결하고자 하였다. 이를 통해 그림 5와 같은 본 논문에서 제안하는 구조를 완성하였다.

IV. 실험 및 결과

1. Medium Scale 구분을 위한 최적의 Threshold 값 탐색

Medium-Scale Proposal 높이에 따라 Large-scale에 해당하거나, Small-Scale에 해당하는 경우가 존재하였다. 검출을 진행하는 경우 표 1과 같이 낮은 가중치로 인해 부정확한 검출이 진행된다. 더 정확한 구분을 위해 TRPN의 각 Scale에 해당하는 값을 분석을 진행하였다.

표 2와 같이 Medium-Scale의 경우 0.2743으로 낮은 값의 가중치를 갖게 된다. 기존 DRPN보다 정확한 구분을 하였지만, 오히려 잘못된 계산을 통해 낮은 검출률을 보일 수 있다. 따라서 Medium-Scale에 해당하는 Proposal의 분포도를 통해 추가 데이터 분석을 진행하였다.

그림 6는 Medium-Scale이 많이 해당하는 부분인 170부터 190 사이의 분포도이다. 대부분 180부터 185에 분포되어 있고, 182에서 183 사이의 값이 제일 많은 값을 갖는 것으로 나타났다. 분포도를 통해 180~185 사이의 값을 Medium-Scale의 Threshold로 활용하였고 표 3과 같은 결과를 나타냈다.

표 3은 180 ~ 185까지의 값을 Threshold 값으로 사용하여 PASCAL VOC 2007/2012를 학습, 검증 데이터로 사용하여 학습한 결과이다. 정확도는 Validation Loss가 가장 낮았을 때 Test Dataset의 Average Precision이다. 각 Large-weight는 모두 0.95의 값을 갖게 학습하였다. 표 3의 결과에서 볼 수 있듯이 180의 Threshold 값이 제일 좋은 결과를 보였다. 본 논문에서는 Threshold 값을 180을 설정하고 추가 가중치 값은 0.5로 설정하였다.

표 1. 학습데이터 PASCAL VOC 2007/2012 대한 평균 가중치, 높이 값의 결과

Table 1. Average Weight, Height Results for Learned Data PASCAL VOC 2007/2012

Method	Scale	Weight	Height
TRPN	Large	0.9847	351.3248
	Medium	0.45	209.8727
	Large	0.08	141.4313
	Medium		
	Small	0.0038	102.5570
DRPN	Large	0.7736	290.6966
	Small	0.0486	117.4244

표 2. TRPN의 각 Scale에 해당하는 높이의 평균, 최소 최대값

Table 2. Average, minimum and maximum height for each scale of TRPN

Method	Scale	Avg	Wt	Min	Max
TRPN	Large	351.3248	0.9847	624.0000	176.1880
	Medium	170.0642	0.2743	270.1848	100.5497
	Small	102.5570	0.0038	196.3789	14.9760

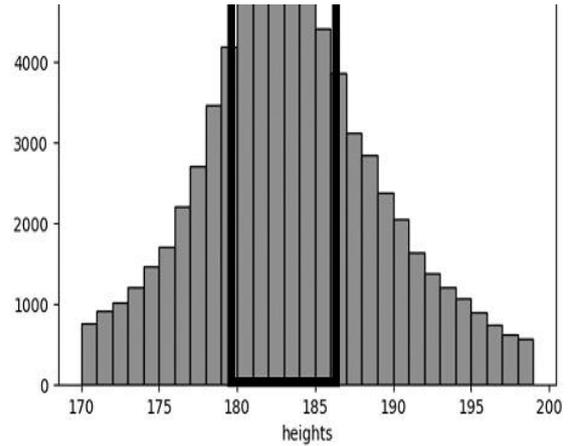


그림 6. Medium-Scale의 170 ~ 190 사이의 값에 대한 분포도
Fig. 6. A distribution diagram of a medium scale for values 170-190

표 3. PASCAL VOC 2007/2012 데이터셋을 통하여 Threshold를 구분 한 후 학습한 결과

Table 3. Results learned after distinguishing Threshold through the PASCAL VOC 2007/2012 dataset

	Threshold					
	180	181	182	183	184	185
Acc	70.9703	69.7215	70.3642	70.5902	70.2511	70.1705
add Wt	0.50	0.48	0.46	0.44	0.41	0.38

표 4. 세가지 모델에서 사용한 데이터셋에 대한 세부 정보

Table 4. Details on the datasets used by the three models

Dataset	Size	Number of images	Type
ETH	624x624	Test : 999	Pedestrian
INRIA	624x624	Test : 288	Person
PASCAL	624x624	Train : 3,019	Person
		Val : 3,076	
KITTI	374x1224	Test : 1,814	Pedestrian
		Test : 1,779	

2. 데이터 셋

데이터 셋은 사람의 동작 여부에 따라 두 가지로 구분하였다. Person의 경우 사람의 움직임이 없었을 때 촬영된 사진이고 Pedestrian의 경우는 사람의 움직임 있었을 때 촬영된 사진이다. PASCAL VOC Dataset [13]은 대표적인 Object Detection Dataset 중 하나이다. 총 20개의 클래스(사람, 동물, 사물)를 가지고 있는 데이터 셋이다. 기존 SRPN, DRPN 과의 최대한 동일한 환경에서의 비교를 위해 person 클래스만 사용하였다. 2007/2012의 Dataset만 사용하였다. 학습에서는 2007/2012를 사용하여 총 3,019장과 평가에서는 2012를 사용하여 1814장을 사용하였다. INRIA Person Dataset [14]은 PASCAL VOC Dataset 과같이 인물 중심의

Dataset이다. 평가에만 사용하였고, 288장을 사용하였다. ETH Dataset [15]는 보행자로 구성되어 있는 Dataset이다. SRPN, DRPN, TRPN을 비교 분석하기에는 충분한 가치가 있는 Dataset이다. 평가에만 사용하였고, 999장을 사용하였다. KITTI Dataset [16]은 자율 주행에서 많이 사용하는 데이터 셋이며 보행자, 나무, 자전거 등 11개의 클래스로 구성되어 있다. 보행자 클래스가 있는 데이터만 사용하였으며 1779장을 평가에 사용하였다.

표 4는 본 논문에서 사용한 데이터 셋의 세부 정보를 나타낸다. 평가 데이터는 PASCAL VOC 2007, INRIA Person Dataset, ETH Dataset, KITTI Dataset을 이용한다. 평가에 사용되는 학습할 때 사용하던 이미지 크기와 동일한 624 * 624로 사용하였다. KITTI Dataset의 경우 이미지 크기를 624 * 624로 진행하는 경우 이미지가 과도하게 축소되어 특징 추출에 문제가 발생하였다. 따라서 KITTI Dataset의 경우는 기존 크기인 374*1224로 사용하여 평가를 진행하였다.

3. 실험 환경

본 논문에서 제안한 TRPN은 SAF R-CNN, DRPN을 활용하여 개발하였다. 학습, 검증, 평가 환경은 NVIDIA GeForce RTX 2080TI 12GB에서 진행하였다. Backbone Network는 SRPN, DRPN, TRPN 모두 VGG16으로 사용하였다. Optimizer function, Loss function 모두 DRPN과 동일하게 SGD [17, 18]와 Smooth L1, CrossEntropyLos를 사용하였다. learning rate의 초깃값은 0.001이며, Epoch 9부터는 0.0001로 변경하여 학습을 진행하였다. 학습 데이터인 PASCAL VOC Dataset의 경우는 624*624의 고정된 크기를 사용하였다. 학습에서 NMS의 Threshold는 0.7로 설정하였고 평가에서는 0.5의 값을 사용하였다. 모델에서 사용하는 이미지는 Flip을 적용하지 않고 진행하였다.

4. 평가 기준

Average Precision (AP)을 통해 SRPN, DRPN, TRPN의 평가를 진행하였다. PASCAL VOC Dataset 학습 데이터로 사용한 세 개의 모델에 대한 평가를 진행하였다. AP는 Precision-recall curve (PR curve)에서 그래프 선 아래의 면적을 계산한 값으로 객체 검출 성능 평가에서 많이 사용된다. 식 (7)에서 P_n 은 예측한 영역, G_n 은 정답 영역 (ground truth)를 의미한다. P_n 과 G_n 의 두 영역의 겹쳐있는 정도를 의미하는 Intersection over Union (IOU)을 계산한 후 일정 값 보다 경우에만 정답이라고 판단한다. 본 논문에서 0.5보다 높은 경우 정답이라고 판단하였다. 0.5보다 높은 값을 가지는 영역은 수식 (8, 9)을 활용하여 Precision과 Recall을 구하였고, PR Curve를 통해 AP를 계산하였다. 최종 계산된 AP는 모델의 평가 기준으로 사용하였다.

5. 평가 결과

PASCAL VOC 2007/2012로 학습한 세 가지 모델에 대해 표 4에 표기되어 있는 데이터 셋에 대한 평가를 진행하였다.

표 5. 세가지 모델에 대해 네 가지 데이터셋의 AP 평가를 진행한 결과 Table 5. Based on AP evaluations of four datasets on three models

Dataset	SRPN	DRPN	TRPN
ETH	42.3783	39.1220	43.7061
INRIA	90.7618	87.8314	91.7802
PASCAL	69.4101	68.0247	70.9703
KITTI	38.1386	34.1173	42.2679
Total	60.1722	57.2739	62.1811

평가는 위에서 설명한 AP를 통해서 수치상의 성능을 평가하였다. 표 5은 ETH, INRIA, PASCAL VOC 2007, KITTI 총 4가지 데이터 셋을 통해 세 가지 모델을 평가한 결과이다.

$$t_0 = \frac{area(P_n \cap G_n)}{area(P_n \cup G_n)} \Bigg\} 0.5, \quad (7)$$

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{all\ detections}, \quad (8)$$

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{all\ ground\ truths}. \quad (9)$$

표 5의 결과에서 ETH Dataset은 1.4 ~ 4.6%, INRIA Dataset은 1.0 ~ 2.9%, PASCAL VOC Dataset은 1.5 ~ 2.9%, KITTI Dataset은 4.1 ~ 8.1%로 SRPN, DRPN보다 좋은 결과를 나타냈다. 최종적으로 4가지 데이터 셋의 평균 결과인 Total에서는 2% - 4.9%로 좋은 결과를 나타냈다.

V. 결론

본 논문에서는 Faster R-CNN의 Object Proposal을 3개의 Scale로 구분하는 모델을 제안하였다. 기존 방법인 SAF R-CNN, DRPN은 데이터 셋의 평균 높이를 통하여 Scale 구분법을 제안하였다. 하지만 SAF R-CNN, DRPN에서 사용하던 Scale 구분법을 통해 Medium-scale을 계산할 경우 모호한 값이 추출되었다. 문제 해결을 위해 Medium-scale에 대한 새로운 기준을 제시하여 3가지 scale로 구분하였다. 또한 Medium-scale에 해당하는 보행자의 높이를 계산하였을 때, Large-scale, Small-scale에 인접한 값들이 존재하였다. 이후 실험을 통해 Medium-Scale의 높이를 Threshold 기준으로 나누어 Medium-Large-scale, Medium-Small-scale로 세분화하였다. 구분된 보행자의 영역은 Refinement Module을 통해 미세조정하여 최종적으로 모델의 구조를 완성하였다. 표1, 2의 결과처럼 2개의 scale보다 3개의 scale로 구분하는 것이 1에 가까운 가중치를 얻게 되었고, 표 5과같이 비교적 높은 검출률을 얻게 되었다.

이러한 방법을 통해 보행자를 단일 크기로 검출하는 것보다 다중 검출을 하는 것이 더 좋은 결과를 보여준다는 것을 확인하였고, 이러한 연구는 보행자 검출 방법이 활용되는 자율주행 자동차나, CCTV 속 보행자를 검출하는 분야에서 활용될 수 있다. 추후 더 세분화된 보행자 구분법을 연구함으로써 보다 정확한 보행자를 구분하고자 한다.

References

- [1] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770-778, 2016.
- [2] Z. Zou, K. Chen, Z. Shi, Y. Guo, J. Ye, "Object Detection in 20 years: A Survey," arXiv preprint arXiv:1905.05055, 2019.
- [3] J. Li, X. Liang, S. M. Shen, T. Xu, J. Feng, S. Yan, "Scale-aware Fast R-CNN for Pedestrian Detection," IEEE Transactions on Multimedia, Vol. 20. No. 4, pp. 985-996, 2017.
- [4] S. H. Song, H. B. Hyeon, H. Lee, "A Pedestrian Detection Method using Deep Neural Network," Journal of KIISE, Vol. 44, No. 1, pp. 44-50, 2017.
- [5] M. C. Roh, J. Lee, "Refining Faster-RCNN for Accurate Object Detection," 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). IEEE, pp. 514-517, 2017.
- [6] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015: 18th International Conference, Proceedings, Part III 18, pp. 234-241, Springer International Publishing, 2015.
- [7] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
- [8] S. Ren, K. He, R. Girshick, J. Sun. "Faster R-cnn: Towards Real-time Object Detection with Region Proposal Networks," Advances in Neural Information Processing Systems 28, pp. 91 - 99, 2015.
- [9] P. Dollár, C. Wojek, B. Schiele, P. Perona, "Pedestrian Detection: A Benchmark," 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 304-311, 2009.
- [10] R. Girshick. "Fast r-cnn," arXiv Preprint, 2015.
- [11] R. Girshick, J. Donahue, T. Darrell, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, 2014.
- [12] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint 1409.1556. 2014.
- [13] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, A. Zisserman, "The Pascal Visual Object Classes (voc) Challenge," International Journal of Computer Vision 88.2, pp. 303-338, 2010.
- [14] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, pp. 886-893, 2005.
- [15] A. Ess, B. Leibe, L. Van Gool, "Depth and Appearance for Mobile Scene Analysis," 2007 IEEE 11th International Conference on Computer Vision. IEEE, pp. 1-8, 2007.
- [16] A. Ess, B. Leibe, K. Schindler, L. Van Gool, "A Mobile Vision System for Robust Multi-person Tracking," 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 1-8, 2008.
- [17] A. Geiger, P. Lenz, R. Urtasun, "Are we Ready for Autonomous Driving? the Kitti Vision Benchmark Suite," 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 3354-3361, 2012.
- [18] S. Ruder, "An Overview of Gradient Descent Optimization Algorithms," arXiv preprint arXiv:1609.04747 (2016).

Kyung-Min Jung (정 경 민)



2021 Computer Science and Engineering from Sunmoon University, Asan, Republic of Korea (B.E.)

2021~Computer and Electronics Convergence Engineering from Sunmoon University, Asan, Republic of Korea (M.S.)

Field of Interests: Image Processing, Pedestrian Detection, Object Detection, Bio BigData

Email: rudals2366@sunmoon.ac.kr

Soo-Yong Park (박 수 용)



(Expected 2024) Computer Science from University of Rochester, Rochester, New York (B.E.)

Field of Interests: Human-Computer Interaction, Image Processing, Object Detection

Email: spark180@u.rochester.edu

Hyun Lee (이 현)



1998 Computer Science and Engineering from Sunmoon University, Asan, Republic of Korea (B.E.)

2002 Computer Science and Engineering from Sunmoon University, Asan, Republic of Korea (M.S.)

2010 Computer Science and Engineering from the Univ. of Texas at Arlington, Arlington, U.S. (Ph.D.)

2012~Computer Science and Engineering from Sunmoon University, Asan, Republic of Korea (Prof.)

Career:

2016, 2018~2020 Highly Cited Researcher (Clarivate Analytics)

2019~Int. J. Systems Science, Int. J. Control, Autom. & Systems, Associate Editor

Field of Interests: Real-time decision-making systems, Autonomic computing, Human-Care System, Bio BigData, Hyperscale AI

Email: mahyun91@sunmoon.ac.kr