

생활 폐기물 다중 객체 검출과 분류를 위한 i-YOLOX 구조에 관한 연구

왕웨이광*, 정 경 권**, 이 태 원***

요 약

생활 폐기물 쓰레기는 기후 변화, 자원 부족, 환경 오염을 불러오는 대표적인 문제로서, 이러한 문제를 해결하기 위해 지능적으로 쓰레기를 분류하는 방식을 연구하였고, 전통적인 분류 알고리즘부터 기계학습, 신경망에 이르기까지 많은 연구가 진행되고 있다. 그러나, 다양한 환경과 조건에서 쓰레기를 분류하기에는 여전히 데이터셋이 부족하고, 신경망 네트워크 구성 복잡도가 증가하며, 성능 측면에서도 실생활에 적용하기에 아직 미흡하다. 따라서 본 논문에서는 신속한 분류와 정확도 향상을 위해 i-YOLOX를 제안하고, 네트워크 매개변수, 검출속도, 정확도 등을 평가한다. 이를 위해 17개의 폐기물 범주를 포함하는 10,000개의 가정용 쓰레기 대상 샘플로 데이터 세트를 구성하고, YOLOX 구조에 Involution 채널 컨볼루션 연산자와 CBAM(Convolution Branch Attention Module)을 도입하여 i-YOLOX를 구성하고, 기존의 YOLO 구조와 성능을 비교한다. 실험 결과 복잡한 장면에서 쓰레기 객체 검출 속도와 정확도가 기존의 신경망에 비해 향상되어, 제안한 i-YOLOX 구조가 생활 폐기물 다중 객체 검출과 분류에 효과적임을 확인하였다.

A Study on the i-YOLOX Architecture for Multiple Object Detection and Classification of Household Waste

Weiguang Wang*, Kyung Kwon Jung**, Taewon Lee***

ABSTRACT

In addressing the prominent issues of climate change, resource scarcity, and environmental pollution associated with household waste, extensive research has been conducted on intelligent waste classification methods. These efforts range from traditional classification algorithms to machine learning and neural networks. However, challenges persist in effectively classifying waste in diverse environments and conditions due to insufficient datasets, increased complexity in neural network architectures, and performance limitations for real-world applications. Therefore, this paper proposes i-YOLOX as a solution for rapid classification and improved accuracy. The proposed model is evaluated based on network parameters, detection speed, and accuracy. To achieve this, a dataset comprising 10,000 samples of household waste, spanning 17 waste categories, is created. The i-YOLOX architecture is constructed by introducing the Involution channel convolution operator and the Convolution Branch Attention Module (CBAM) into the YOLOX structure. A comparative analysis is conducted with the performance of the existing YOLO architecture. Experimental results demonstrate that i-YOLOX enhances the detection speed and accuracy of waste objects in complex scenes compared to conventional neural networks. This confirms the effectiveness of the proposed i-YOLOX architecture in the detection and classification of multiple household waste objects.

Key words : i-YOLOX, Involution, CBAM, Household Waste

접수일(2023년 12월 11일), 수정일(2023년 12월 19일),
게재확정일(2023년 12월 20일)

* 동신대학교 컴퓨터학과 대학원(주저자)

** 동신대학교 전기공학과 교수(공동저자)

*** 동신대학교 컴퓨터학과 교수(교신저자)

1. 서 론

급격한 경제 성장과 함께 전 세계적으로 생활 폐기물, 특히 플라스틱 사용량이 급증하고 있으며, 재활용 규제에도 불구하고 폐기물 분류는 여전히 해결해야 할 과제로 남아 있고, 자원 낭비와 환경 오염을 초래하고 있는 실정이다. 이러한 문제를 해결하기 위해 자동 폐기물 분류 기술과 관련 연구가 활발히 진행되고 있다.

전통적인 접근 방식으로는 물리적 특성을 기반한 것으로서, 광학센서, 분광기, 유도 센서, 열화상 카메라, 카메라와 같은 기기를 이용하여 정보를 취득하고 컴퓨터비전 알고리즘을 적용하여 기계적으로 분류하는 방식을 들 수 있다.[1]

조금 더 진화한 방식으로 기계 학습 모델을 사용하여 쓰레기를 분류하려는 노력도 시도되었으며, 선형 판별 분석, 결정트리 방식, 베이지안 네트워크, SVM(Support Vector Machine) 알고리즘을 사용하여 자동으로 쓰레기를 분류하고자 하였다.[1] 그러나 이러한 방식은 제한적 조건과 환경에서 수행된 것으로서 실생활에 적용하기에는 미흡한 결과를 나타내고 있다.

딥러닝에 의한 이미지 분류가 다양한 분야에 적용되고 있으며, 특히 CNN(Convolution Neural Network)은 학습에 의한 특징 추출에 장점을 가지고 있다. 이러한 이유로 쓰레기 분류 분야에서도 딥러닝을 활용한 연구가 보편화되고 있으며, Yang 등은 SIFT(Scale-Invariant Feature Transform)를 SVM과 결합하여 쓰레기 데이터셋인 Trash Net에서 63%의 정확도를 달성하였다.[1-3] 또한, Rabano et al., Sreelakshmi 등도 딥러닝 CNN을 적용하여 쓰레기 분류 정확도를 향상시켰다.[2] 최근에는 Bircano ğlu et al., Vo, Ye, Zhang et al. 등은 YOLO 를 활용하여 실생활에서 쓰레기를 분류하는 연구를 진행하였다. 그러나, 이러한 연구들은 제한적인 조건에서 연구가 진행되어 이미지 배경에서 정확하게 쓰레기를 분류하기에는 양질의 데이터셋이 부족하고, 여전히 검출 정확도와 검출 속도 향상은 해결해야 할 과제로 남아 있다.

본 논문에서는 데이터의 불균형, 샘플 희소성

을 보완하기 위해 데이터 증강 기술을 사용하여 데이터셋을 구성하고, CBAM 모듈과 채널 컨볼루션 기법을 통합한 i-YOLOX 객체 검출 알고리즘을 제안하고, 쓰레기 분류에 적합한 딥러닝 방식과 구조를 살펴보고, 정확도와 검출 속도를 비교 평가하고자 한다.

2. 생활 폐기물 데이터셋 구성

생활 폐기물 분류 및 재활용의 중요성이 높아짐에 따라 폐기물의 정확한 식별이 필요하게 되었으나, 폐기물 종류 및 특성의 다양성으로 인해 객체 검출에 어려움이 발생하였다. 이를 해결하기 위해 데이터셋을 풍부하게 하여 모델 일반화 및 정확도를 향상시키는 데이터 증강 기법이 활용되고 있다.

본 논문에서는 Mixup 및 Mosaic 데이터 증강 기법을 활용하여 사용 빈도가 높은 17개 일반 생활 폐기물 분야 10,000개의 이미지로 CDTD(Common Domestic Trash Dataset)를 생성하였다.

2.1 정크 이미지 데이터셋 구성

병, 캔과 같은 재활용품의 가치와 실제 사용 빈도를 고려하여 <표 1>과 같이 분류하고, 이미지 데이터를 jpg 포맷으로 확보하였고, 이미지에는 다수의 검출 대상이 포함된다. 또한, 효율적인 학습을 위해 획득한 이미지는 동일한 사이즈와 형태로 저장한다.

<표 1> 생활 쓰레기 분류 : 17분야, 10,000개

Category	수량	Category	수량
Battery	410	Plastic-hanger	300
Bottle	1730	Plastic-toy	400
Can	1900	Plush-toy	570
Cigarette	410	Pot	450
Clothes	350	Power-bank	380
Courier-bag	260	Shoes	420
Cutting-board	420	Tea-residue	360
Ointment	400	Wash-supply	900
Pillow	340		
Total		10,000	

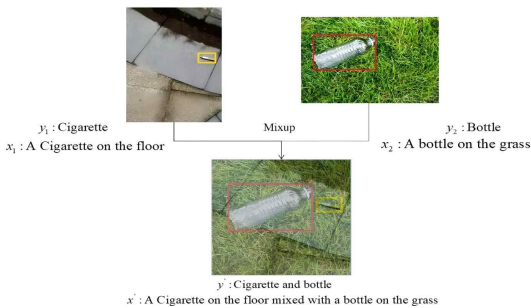
획득한 이미지를 식별하고 분류하기 위해 오픈

소스 도구 LabelImg를 이용하여 대상 정보를 XML 파일로 레이블링하여 저장하였다.

데이터 증강 기법으로는 랜덤 노이즈 삽입, 확대/축소, 회전, 밝기 변화, 랜덤 조각, 컬러 지터 등의 기법이 있으며, 본 논문에서는 위의 방식에 덧붙여 Mixup & Mosaic 방식을 이용하여 생활 폐기물의 데이터를 증강하였다.

2.2 Mixup 방식 생활 쓰레기 이미지 향상

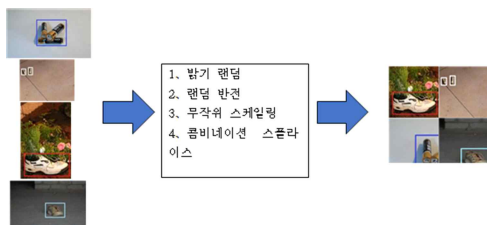
(그림 1)은 Mixup 과정을 보여주는 것으로, 두 개의 샘플을 선형적으로 결합하여 새로운 훈련 데이터를 생성하는 데이터 확대 기법으로서, 혼합 비율을 랜덤하게 선택하여 데이터 샘플을 다양하게 생성할 수 있어 데이터 일반화에 도움을 준다.



(그림 1) Mixup 융합 과정

2.3 Mosaic 방식 생활 쓰레기 이미지 향상

(그림 2)는 Mosaic 과정을 보여주는 것으로, 여러 개의 이미지를 하나의 큰 이미지로 조합하여 데이터의 다양성을 향상시키는 데이터 증강 방법이다.[10] 랜덤 커팅, 스케일링 및 회전을 통해 학습 데이터를 풍부하게 하여 딥러닝 모델 일반화 및 정확도를 향상시키므로, 다중 객체 및 소규모 생활 폐기물 탐지에 효과적인 방법이다.



(그림 2) Mosaic 융합 과정

2.4 데이터 증강에 따른 데이터셋 구성

본 논문에서는 데이터 증강 기법에 의해 생성된 데이터셋의 품질을 확인하기 위해 실험을 진행하였다. 객관적이고 효율적인 비교를 위해 Fast R-CNN과 같은 기존의 2단계 검출기를 능가하는 속도와 정확성으로 알려진 SSD(Single Shot MultiBox Detector)를 테스트 대상 검출 알고리즘으로 진행하였고, PyTorch를 사용했다.

학습은 속도와 모델 일반화를 위해 기율기 0.93과 바이어스 0.0005를 가진 확률적 경사 하강법을 사용했다. 최대 및 초기 학습률은 각각 0.002와 0.00002인 동적 학습률을 사용하였다. 모델 정확성을 위해 배치 크기 36, 작업 스레드 30개를 사용하였다. 학습은 100개의 에포크로 구성되었으며, 각각의 데이터 세트는 전체 데이터 세트를 포함하였다.

비교 실험에서는 랜덤 밝기(RB), 랜덤 컷(RC), 랜덤 스케일링 및 회전(RR), 컬러 디터링(CD), 랜덤 노이즈(RN)의 데이터 증강 기법으로 얻은 데이터셋과 Mixup(BE+Mixup)과 Mosaic(BE+Mixup+Mosaic) 방식의 데이터 증강 기법으로 얻은 데이터에 대해 학습 지속시간, 정확도, 리콜, 평균 정밀도(mAP)의 네 가지 성능 지표를 사용하여 비교하였으며 <표 2>와 같다.

<표 2> 데이터 증강 기법 데이터 품질 실험

Scheme	Training duration/h	Precision/%	Recall/%	mAP/%
No enhancement	1.57	77.48	76.18	81.02
RB	1.57	79.87	75.29	81.59
RB+RC	1.58	76.36	77.80	82.04
RB+RC+RR	1.59	80.18	75.75	82.15
RB+RC+RR+CD	1.59	83.16	79.17	83.06
RB+RC+RR+CD+RN	1.59	81.02	75.54	81.40
BE+Mixup	1.65	82.40	78.45	84.23
BE+Mixup+Mosaic	1.67	84.61	80.16	85.11

<표 2>는 데이터 증강 기법으로 생성된 데이터 품질 실험 결과이다. 모든 방법이 결합된 경우에도 학습 시간의 변화는 미미하여 데이터 증강 기법 확대에 의한 영향은 크지 않은 것으로 확인되었다.

반면, 랜덤 스케일링 및 회전, 컬러 디터링은 정밀도와 Recall 지표에서 상당한 개선 효과를 가져왔고, 이는 다양한 모양과 색상 특징 정보 추출에 도움이 되는 기법임을 알 수 있다.

랜덤 노이즈는 성능이 악화된 것으로 나타나 데이터셋 생성에서 제외되었다. 기본 영상이미지 자체 관련 없는 정보가 다량으로 포함되어 있어 학습에 의한 성능 향상이 없을 것으로 판단되었다.

Mixup의 추가는 증강을 사용하지 않은 훈련에 비해 mAP를 3.21% 향상시켰고, Mixup과 Mosaic을 결합하면 증강을 사용하지 않은 것에 비해 mAP가 4.09% 증가해 가장 높은 정밀도, 리콜 및 평균 정밀도(mAP)를 달성했다. 위 방식을 통해 데이터셋 신뢰도가 향상될 수 있음을 확인하였다.

3. i-YOLOX 구조 객체 검출

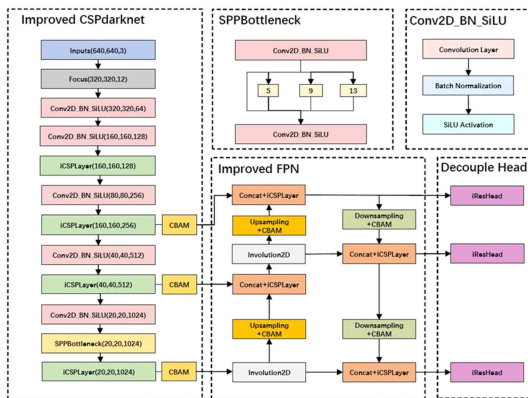
3.1 I-YOLOX 개요

i-YOLOX는 생활 폐기물의 영상검출 및 분류 성능을 향상시키기 위해 YOLOX를 개선한 구조이다. (그림 3)은 i-YOLOX의 기본 구조를 나타낸다.

(그림 3)에서 YOLOX의 백본 네트워크에서 CSP Layer 부분에 특징 추출을 향상시키는 Involution 연산자를 통합하여 구조를 최적화하였다.

또한, 백본 네트워크 3개의 출력 각각에 CBAM 모듈을 추가하여 특징 추출 효율을 향상시키고자 하였고, PANet 구조에 Involution과 CBAM 블록을 도입하여 연산속도를 개선하고, 파라미터를 줄이면서 특징 추출을 강화하였다.

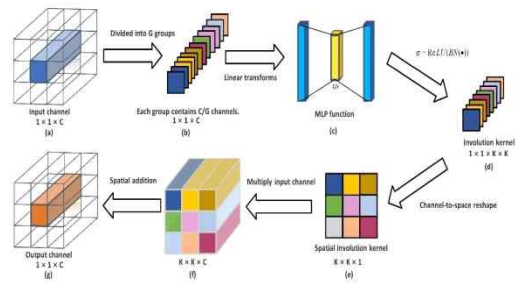
마지막으로, i-YOLOX의 검출 및 분류 성능을 더욱 향상시키기 위해 involution 연산자를 활용하여 잔류 디커플링 헤드를 구성하였다.



(그림 3) i-YOLOX 구조도

3.2 Involution 채널 컨볼루션

Involution은 Li 등[4]에 의해 소개된 채널별 컨볼루션 연산의 형태이다. 전통적인 컨볼루션 방식이 공간과 채널별로 포괄적으로 정보를 처리하면서 유사한 특징 추출과 파라미터 중복 등의 문제가 발생하고 있지만, Involution에서는 채널 내에 국한된 정보를 처리하는데 중점을 두고 있어서 유사 특징 추출로 인한 비효율성을 개선할 수 있다.



(그림 4) Involution 데이터 처리

(그림 4)는 Involuton 채널 연산 방식을 나타내고 있다. (그림 4)의 (a)는 입력 특징 채널을 나타내고, (b)는 모든 입력 채널을 G 그룹으로 나누고, (c)에서는 선형 활성화 계층으로 분류하고, (d)에서는 ReLU 활성화 후 채널을 정렬하고, (e)에서는 공간상 채널로 재정렬하고, (f)에서는 공간상 정렬된 채널에 입력 채널과 곱셈연산을 수행하고, (g)에서는 공간상 범위에 걸쳐 (f)를 합산하여 특징 추출 채널을 얻는다. 이러한 과정을 거쳐 계산 효율성을 보장하면서 보다 효과적인 특징을 추출할 수 있다.

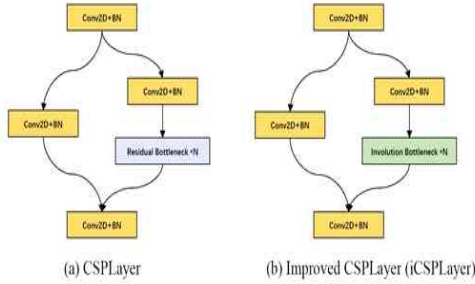
본 논문에서는 Involution 방식을 생활 쓰레기 검출 및 분류 분야에 응용하여 관련분야에서 활용할 수 있는 데이터셋과 파라미터 등을 제시한다.

3.3 Improved CSPdarknet

백본 네트워크에서 CSPLayer는 방대한 영상 데이터를 처리하기 위해 필수적이며, 기본 컨볼루션 레이어와 Subresidual 블록을 갖는 백본 블록, CSPLayer의 출력에 직접 연결된 residual 블록으로 구성된다.

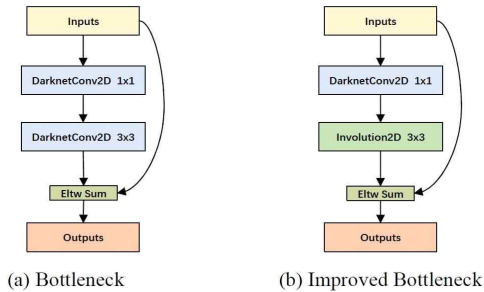
특징 모델링을 강화하고 생활 폐기물 영상 처리에

서 중복성을 줄이기 위해 i-YOLOX는 (그림 5)와 같이 CSPLayer의 잔여 병목 모듈을 involution 병목 모듈로 대체한다. 또한 CBAM 모듈은 백본 네트워크의 세 개의 iCSPLayer에 추가되어 출력 계층에서 특징 추출 레이어 성능을 개선한다.



(그림 5) CSPLayer 구조도

Sub-residual Bottleneck은 백본 네트워크에서 특징 추출에 핵심적인 역할을 하며, (그림 6)의 (a)와 같이 주로 적층된 1x1 컨볼루션 레이어와 3x3 컨볼루션 레이어로 구성되며, 이는 잔차 연결을 통해 컨볼루션 출력을 입력과 합산하여 기울기 분산 문제를 극복하는데 도움이 된다. 그러나 3x3 컨볼루션 레이어에는 많은 파라미터가 존재하여 파라미터가 중복될 수 있다. 본 논문에서는 파라미터 수를 줄이기 위해 3x3 Darknet Conv2D를 Involution 2D 레이어로 대체하여 네트워크 효율성을 높이고 성능을 개선한다. Involution Residual Bottleneck 구조는 (그림 6)의 (b)와 같다.



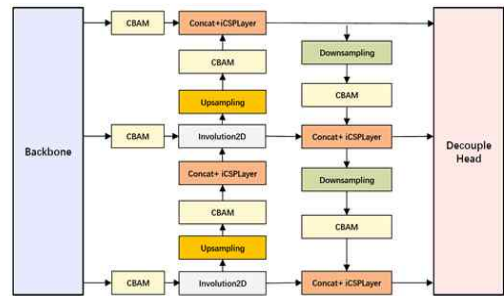
(그림 6) Sub-residual block 구조도

3.3 CBAM을 활용한 iPANet

FPN과 PANet은 YOLOX[5]에서 백본 특징 추출 신경망의 출력 레이어로서, 특징 추출을 향상시키기 위한 모듈이다.

iCSPLayer 블록 처리 후 채널 컨볼루션 주변 정보가 국지적인 간섭을 초래할 수 있으므로, 이를 해결하기 위해 업샘플링 모듈과 다운 샘플링 모듈 다음에 CBAM 모듈 4개를 추가하였다. CBAM 모듈을 결합한 PANet은 서로 다른 크기에서 발생된 특성을 더 잘 통합하고 추출하여 검출 정확도를 향상시킨다.

또한, Conv2D 모듈을 Involution2D 모듈로 교체하여 특징 추출 네트워크에서 파라미터의 수를 줄이고자 하였다. 개선된 PANet(iPANet)의 구조는 (그림 7)과 같다.

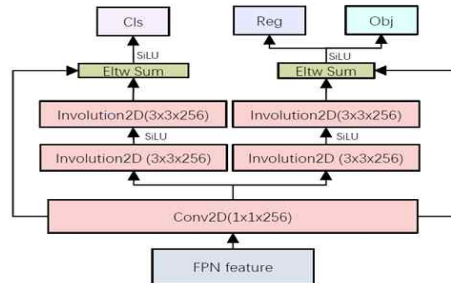


(그림 7) iPANet 구조

3.3 개선된 iResHead

i-YOLOX의 검출 헤드는 분류와 회귀 연산을 동시에 수행하는 채널 컨볼루션 iResHead(잔류 디커플링) 구조를 설계하였다.

iResHead의 설계는 DSSD(Deconvolutional Single Shot Detector) 목표 검출 알고리즘[6]을 활용하여 변수의 수를 효과적으로 줄이고, 경사 손실을 방지하며, involution을 통해 모델 정확도를 유지할 수 있도록 설계되었으며, 여기에는 Cls(분류), Reg(회귀), Obj(목표 신뢰도) 출력 속성 계층이 포함된다.



(그림 8) iResHead 구조

이러한 설계를 통해 시험모델의 성능, 신뢰성, 계산 효율을 향상시키는 동시에 매개변수의 수를 감소할 수 있다. (그림 8)은 iResHead의 구조를 나타낸 것으로, Cls는 (H×W×C)의 출력 레이어 형태로 객체의 클래스를 예측하고, Reg는 객체의 위치와 크기를 예측하며, 출력 레이어 형태는 (H×W×4)이다. Obj는 객체가 그 위치에 존재하는지 여부를 예측하며, 출력 레이어 형태는 (H×W×1)이다.

4. 실험 및 결과

본 논문에서는 모델 설계 및 구현을 위해 PyTorch 프레임워크를 사용하여, 모델 학습은 운동량 매개 변수를 0.93으로 설정하고 가중치 감쇠를 0.0005[7]로 설정하는 확률적 구배 하강 알고리즘을 사용했다. 초기 속도 0.01에서 시작하고 cosine learning rate reduction strategy[8]를 사용하여 동적 학습률 전략을 적용하였다. 훈련 배치 크기는 36개이며, 30개의 작업 스레드가 효율성과 정확성을 향상시켰다. 훈련은 150개의 에포크에 걸쳐 진행되었으며, 유효성 검사 세트에 대해 매 5개의 에포크마다 Precision(P), Recall(R), mAP, GFLOP(Giga Floating Point Operations Per Second), FPS(Frames per second) 의 항목을 평가하였다.

4.1 i-YOLOX 효율성 실험 결과

본 논문에서는 YOLOX 기준 알고리즘에 iCSPLayer 만 적용했을 경우, iCSPLayer+iPANet 의 경우, iCSPLayer+iPANet+ResHead 경우, iCSPLayer+iPANet+iResHead 의 경우로 구분하여 실험을 진행하였고, <표 3>과 같다.

<표 3> YOLOX 기반의 실험

Scheme	iCSPLayer	iPANet	Residual Head	Involution Head
iCSPLayer	√			
iCSPLayer-iPANet	√	√		
iCSPLayer-iPANet-ResHead	√	√	√	
iCSPLayer-iPANet-iResHead	√	√	√	√

<표 4>는 P, R, mAP, 파라미터 수, GFLOP, FPS 지표에 대한 <표 3>의 4가지 경우에 대한 결과이다. 실험 결과 i-YOLOX는 P, R, mAP에서 기존의 YO

LOX 보다 우수한 성능을 보였으며, GFLOPs도 감소하여 성능이 개선되었음을 확인할 수 있었다.

<표 4> 성능 지표별파라미터별 실험 결과

Method	P/%	R/%	mAP/%	Param/M	GFL OPs	FPS
YOLOX(base)	85.78	80.06	85.68	8.94	26.79	26.37
iCSPLayer	86.43	81.92	86.15	7.12	21.86	43.50
iCSPLayer-iPANet	87.47	79.34	86.29	8.42	25.97	35.46
iCSPLayer-iPANet-ResHead	86.43	81.92	86.99	8.42	25.97	42.33
iCSPLayer-iPANet-iResHead	87.72	81.96	87.15	6.86	17.28	37.01

iCSPLayer에서 CBAM 모듈과 Involution을 도입하여 P, mAP를 약간 우수한 성능을 확보했으며, FPS를 17.13% 향상하였다.

iPANET은 6개의 CBAM 모듈을 추가하여 매개변수 수는 증가했으나, mAP 성능을 0.14% 개선하였다.

스킵 연결 ResHead 방식의 성능개선은 미미하였고, iResHead를 추가한 구조에서는 P, mAP가 각각 1.29%, 0.16% 향상되었고, 파라미터 수가 급격하게 감소하였으며, GFLOPs에서 성능 개선을 확인할 수 있었다.

4.2 검출 구조별 비교 실험

본 논문에서는 테스트 데이터셋에 대하여 기존의 객체 검출 방식과 제안된 방식의 성능을 비교하기 위해 <표 5>와 같이 P, R, mAP, GFLOPs, FPS 등의 지표로 실험을 수행하였다.

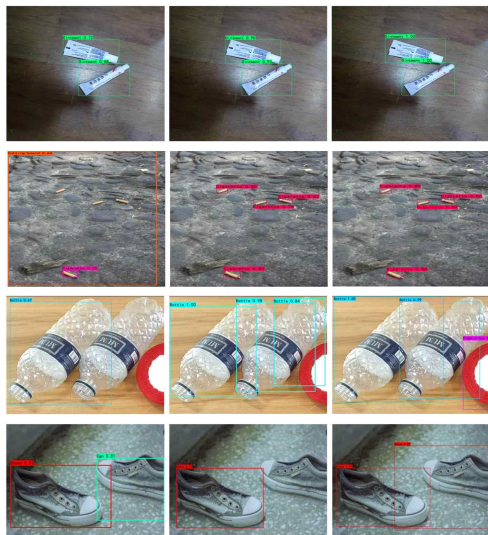
<표 5> 검출 구조 비교 실험 결과

Algorithm	P/%	R/%	mAP/%	Param/M	GFLOPs	FPS
Faster-RCNN	54.67	86.14	81.70	137.04	370.14	11.79
YOLOv4	84.34	58.68	75.13	64.02	60.07	16.98
SSD	84.15	77.15	82.08	25.88	62.45	40.13
YOLOv5	83.30	54.49	75.63	47.06	115.91	24.55
YOLOX	85.78	80.06	85.68	8.94	26.79	26.37
i-YOLOX (Proposed)	87.72	81.96	87.15	6.86	17.28	37.01

<표 5>의 결과에서 보듯이, YOLOv4와 YOLOv5는 Faster-RCNN보다 빠르지만 정확도가 떨어지며, YO

LOX는 속도와 정확도의 우수한 성능을 보였다. SSD의 파라미터 수와 계산 복잡도는 YOLOX, i-YOLOX보다 높지만 FPN 블록을 포함하지 않아 검출 속도가 더 빠름을 알 수 있었다. i-YOLOX는 원래의 YOLOX에 비해 매개 변수의 수, 계산 복잡도, 속도 및 정확성 측면에서 우수한 성능을 보임을 확인할 수 있었다.

실제 응용분야에서 I-YOLOX의 성능을 테스트하기 위해 (그림 9)처럼 4개의 이미지 세트에 YOLOv4, YOLOX, I-YOLOX를 적용하여 성능을 테스트하였다.



(a)YOLOv4 (b)YOLOX (c) i-YOLOX
(그림 9) 생활쓰레기 검출 테스트

두 번째 이미지에서는 YOLOv4가 한 개의 담배꽂초만 검출하였고, YOLOX와 i-YOLOX는 모든 담배꽂초를 정확하게 식별했으며 i-YOLOX의 정확도가 더 우수하였다. 작은 규모의 객체에 대해 i-YOLOX가 다른 구조보다 더 정확하게 검출함을 확인할 수 있었다.

세 번째 이미지에서, YOLOX는 플라스틱 장난감을 병으로 검출하였고, YOLOv4는 한 개의 병만 검출했지만, I-YOLOX에서는 두 개의 병이 정확히 검출되었다.

네 번째 이미지에서, YOLOv4는 한 개의 신발만 정확하게 검출했고, 다른 한 개의 신발은 캔으로 오인식하였다. YOLOX는 한 개의 신발만 검출

하였으나, i-YOLOX는 모든 신발을 성공적으로 검출함을 알 수 있었다.

실제 이미지 실험에서도 제안한 i-YOLOX가 기존의 YOLOX 객체 검출 알고리즘에 비해 생활 폐기물에 대한 보다 정확한 검출 성능을 보임을 확인할 수 있었다.

5. 결론

본 논문에서는 딥러닝 기술을 활용하여 폐기물 분류의 효율성을 개선하고자 하였다.

먼저, 쓰레기 객체 검출에 널리 사용될 수 있는 생활 폐기물 데이터 세트를 데이터 증강기법을 사용하여 생성하였다.

또한, 특성 추출의 효율성과 정확성을 향상시키고 매개변수 수를 줄이기 위해 Involution 채널 컨볼루션 메커니즘을 도입하고, CBAM 메커니즘과 채널 컨볼루션 메커니즘을 결합한 i-YOLOX 객체 검출 모델을 제안하였다.

i-YOLOX 모델은 실험을 통해 널리 사용되고 있는 객체 검출 구조보다 더 우수한 성능을 보였으며, 서로 다른 데이터셋과 생활폐기물 대상을 보다 정확하게 검출하여 실제 폐기물 분류 적용 시 정확한 결과를 제공할 수 있을 것으로 기대된다.

i-YOLOX의 구조를 개선하고 훈련 데이터셋을 보강하여 공개함으로써, 쓰레기 객체 검출 분야 연구가 지속적으로 이루어지고, 실제 생활 현장에 적용해서 결과를 다시 보강하는 일련의 과정을 통해 관련 분야 연구가 활성화 되기를 기대한다.

참고문헌

- [1] Chen Z, Yang J, Chen L, et al. Garbage classification system based on improved ShuffleNet v2 [J]. Resources, Conservation and Recycling, 178: pp, 106090, 2022.
- [2] Bochkovskiy A, Wang C-Y, Liao H-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection[C]. proceedings of the Computer Vision and Pattern Recognition (CVPR). 2020.

- [3] Ge Z, Liu S, Wang F, et al. YOLOX: Exceeding YOLO Series in 2021 [J]. arXiv preprint arXiv:210708430, arXiv: 2107.08430, 2021.
- [4] Li D, Hu J, Wang C, et al. Involution: Inverting the inherence of convolution for visual recognition[C]. proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp, 12321-12330, 2021.
- [5] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]. proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp, 936-944, 2017.
- [6] Fu C-Y, Liu W, Ranga A, et al. Dssd: Deconvolutional single shot detector [J]. arXiv preprint arXiv:170106659, 2017.
- [7] Mamalis T, Stipanovic D, Voulgaris P. Stochastic Learning Rate Optimization in the Stochastic Approximation and Online Learning Settings [J]. arXiv e-prints, arXiv:2110.10710, 2021.
- [8] Loshchilov I, Hutter F. SGDR: Stochastic Gradient Descent with Warm Restarts [J]. arXiv preprint arXiv:160803983, 2016.
- [9] Zhang H, Cisse M, Dauphin Y N, et al., "mixup: Beyond Empirical Risk Minimization." proceedings of the International Conference on Learning Representations, 2018.
- [10] Bochkovskiy A, Wang C-Y, Liao H-Y. "YOLOv4: Optimal Speed and Accuracy of Object Detection." proceedings of the Computer Vision and Pattern Recognition (CVPR), 2020.

————— [저자소개] —————



왕웨이광 (Weiguang Wang)
2023년 12월 동신대 컴퓨터학 박사
재학중
email : wangweiguang@gmail.com



정 경 권(Kyung Kwon Jung)
2000년 2월 동국대학교 전자공학과
공학석사
2003년 2월 동국대학교 전자공학과
공학박사
2023년 12월 동신대학교 전기공학과
교수 재직중
email : kkjung@dsu.ac.kr



이 태 원 (Taewon Lee)
1998년 2월 전남대학교 전자공학과
공학석사
2003년 2월 전남대학교 전자공학과
공학박사
2023년 12월 동신대학교 컴퓨터학과
교수 재직중
email : taewon@dsu.ac.kr