YOLOv7을 이용한 전자소자 인식 모델 개발

이희원[1], 안상태[1,2,*]
[1]경북대학교 전자공학부
[2]경북대학교 전자전기공학부
saebuk2000@knu.ac.kr. *stahn@knu.ac.kr

Development of Electronic Device Recognition Model using YOLOv7

Hee-Won Lee[1], Sangtae Ahn[1,2]
[1]School of Electronics Engineering, Kyungpook National University
[2]School of Electronic and Electrical Engineering, Kyungpook National University

요 약

회로설계 실험에서 사용되는 전자소자들은 실험이 종료된 후 다시 분류하여야 하는데, 소자가 작고 종류가 많아 분류에 많은 시간이 필요하게 된다. 본 논문에서는 이를 해결하기 위해 YOLOv7을 활용한 소자 인식 모델을 개발하여, 실험실 환경 유지를 위한 불필요한 인력 낭비를 줄이는데 도움을 주고자 한다.

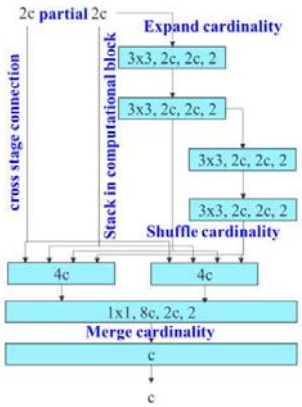
1. 서론

회로설계에 사용하는 전자소자들은 사용 후 분류하여야 하는데, 이는 일반적으로 관리자가 수동으로 분류하기 때문에 많은 시간이 소요된다. 이를 해결하기 위해 본 연구에서는 YOLOv7[1] 모델을 활용하여 자동으로 전자소자를 분류해주는 객체 인식 모델을 제안한다.

2. 본론

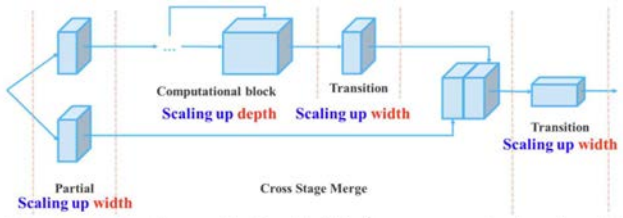
YOLO[2,3,4]는 You Only Look Once의 약자로 다른 모델들에 비해 빠른 처리 속도를 보여 실시간으로 객체 인식이 가능한 방법이다. YOLO는 단일 신경망 구조로 되어있어 구성이 단순하며 속도가 빠른 특성이 있다. 또한 이미지에서 주변 정보까지 학습하는 방법으로 이미지 전체를 처리하기 때문에, bounding box 예측을 할 때 예측 bounding box가 틀렸거나 IoU(Intersection over Union)가 사전에 설정한 기준에 못미치는 background FP(Flase Positive)[5]가 적다. 그렇기에 Training 데이터셋에 없었던 새로운 이미지에 대해서도 검출 정확도가 높다. 하지만 다른 객체 검출 모델과 비교하면 mAP(mean Average Precision)가 다소 떨어진다는 단점이 있다.

최근에 나온 YOLOv7[1]의 구조에서 매개변수와 계산을 효과적으로 활용할 수 있도록 “확장” 및 “복합 스케일링” 방법을 제안하였다. 이러한 방법을 위해 Extended Efficient Layer Aggregation Networks(E-ELAN) 구조를 제안하였다. E-ELAN은 확장, 셔플, 병합 카디널리티를 통해 계산 블록을 많이 쌓아도 학습 능력이 뛰어나며, 오직 계산 블록만 바뀌고 전이 계층은 바뀌지 않는 장점이 있다.



(그림 1) YOLOv7 모델 구조

YOLOv7에서 제안하는 모델 스케일링 방법은 compound model scaling 방법이다. YOLOv7이 사용하는 스케일링 모델인 concatenation-based model에서 scaling-up 또는 scaling-down을 할 때 전이 계층의 채널이 변경된다. 이는 전이 계층의 입력과 출력 채널이 모두 변경되고, 출력 채널이 변경된다는 것은 그 다음 전이 계층의 입력 채널 또한 변경

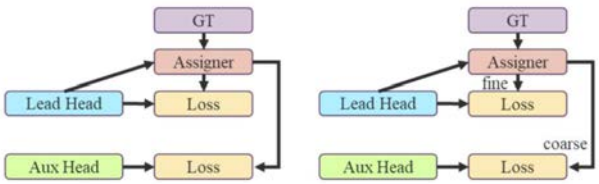


(그림 2) compound model scaling 방법

되고 반복적으로 발생한다. 이로 인해 scaling factor에 대해 분석이 불가능하기에 YOLOv7에서는 compound model scaling 방법을 제안하였다.

compound model scaling 방법은 계산 블록에만 깊이 scaling-up을 적용하고 이에 맞춰 출력 채널을 변경한다. 계산 블록의 출력 채널이 증가하였으므로, 전이 계층의 너비 scaling-up을 적용한다. partial 부분의 입력 feature는 이전 전이 계층의 출력 feature이므로 똑같이 partial 부분도 너비 scaling-up이 된다. 이를 통해 모델이 초기 설계 시 가졌던 특성을 유지할 수 있고, 최적의 구조를 유지할 수 있다.

YOLOv7의 모델 학습 구조는 RepVGG[6]를 약간 변경하여 사용한다. 기존의 RepVGG는 재매개화 방법을 사용하여, 학습할 때 병렬로 Conv+BN들을 학습하다가 inference 시에 해당 Conv+BN들을 하나의 Conv로 재매개화한다. 하지만 RepVGG의 RepConv의 identity connection은 다른 모델을 파괴한다. 그래서 YOLOv7에서는 RepConv에서 identity connection이 없는 RepConvNet을 제안한다. 결과적으로 residual이나 concatenation connection은 재매개화되는 구조이다.



(그림 3) YOLOv7이 제안하는 2가지 Deep supervision label assignment 구조

3. 실험

3-1. 데이터 셋

소자 인식 모델 개발을 위해 python에서 selenium을 이용해 google에서 이미지를 내려받았다. 내려받은 이미지를 전처리를 통해 분류하였고, data augmentation을 사용하여, 결과적으로 약 3,500개의 이미지 데이터 셋을 구축하였다. labelImg 모듈을 사용하여 이미지에서 전자소자에 해당하는 위치에 bounding box를 레이블링하면, 이미지에서의 bounding box의 좌푯값과 class 정보가 텍스트 파일로 저장된다. 생성된 3,500개의 데이터 셋을 train 70%, validation 20%, test 10%로 분배하였다. 이때 train과 validation, test에 사용할 이미지 파일이름을 목록으로 저장하여 각각 학습과 평가, 검증에 사용할 수 있도록 설정하였다. 데이터 셋의 class 개수는 5개로 설정하였다. class 명은 resistors, Electrolytic capacitors, LED, IC_chip, bipolar junction Transistor(BJT)이다.

3-2. YOLOv7 학습

Geforce RTX 3090Ti 3개를 가지고 YOLOv7 학습하였다. 학습에 필요한 데이터셋과, YOLOv7 모델을 github에서 git clone하여 사용하였다.

데이터셋에 최적화된 학습을 하기 위해, 하이퍼파라미터 값을 직접 설정하였다. 이는 표 1에 기재되어있다. 멀티 GPU로 학습하기 위해서는 total batch size는 GPU 3개를 사용하기 때문에, total batch size의 값을 3의 배수로 설정해야 한다.

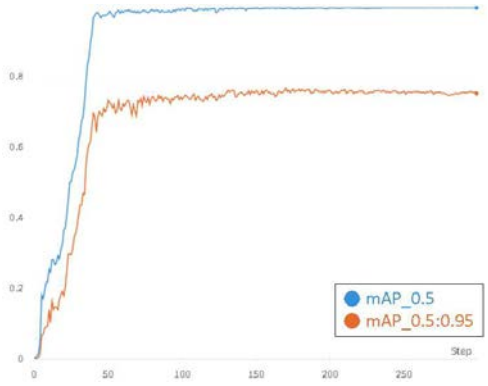
<표 1> 모델 학습 파라미터 정의

Parameter	Value
batch-size	16
epochs	300
(total batch size)	48
img-size	640 640
cfg	YOLOv7-e6e.yaml
data	custom_dataset/data.yaml

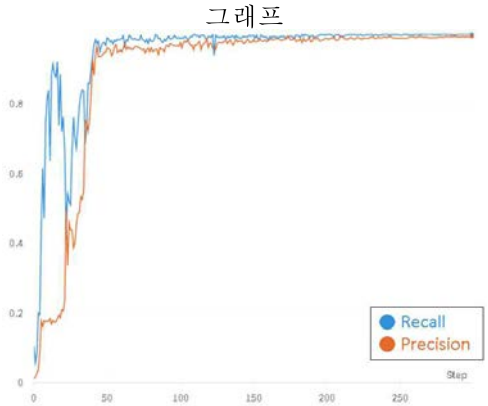
학습하는데 총 6시간 25분 26초가 소요되었고, 학습 결과 train dataset의 object loss는 0.001679, class loss는 0.0003172, box loss는 0.01754이다. mAP_0.5는 0.9935, mAP_0.5:0.95는 0.7515, precision은 0.9938, recall은 0.9989로, 높은 정확도로 전자소자를 인식한다는 것을 확인할 수 있었다. 학습 결과로 그려진 Loss와 mAP의 그래프는 그림 4에서 볼 수 있다.

4. 결론

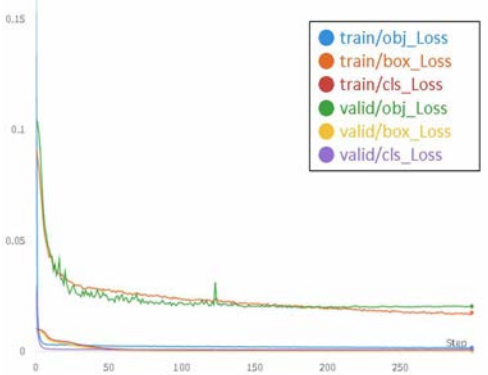
학습 결과로 나온 가중치 파일을 이용해 레이블링이 없는 이미지에서 전자소자 객체 검출을 진행했을 때, 그림 5와 같이 전자소자를 높은 정확도로 인식하는 것을 볼 수 있다. 본 연구결과를 이용하여 추후에는 하드웨어와 결합하여 실제 전자소자 분류기를 개발할 예정이다.



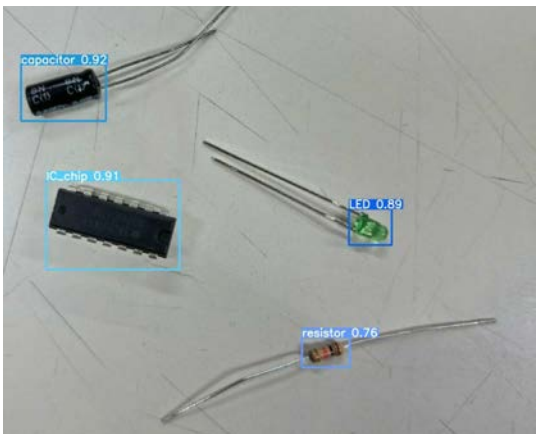
(그림 4-1) mAP_0.5, mAP_0.5:0.95 그래프



(그림 4-2) Recall, Precision 그래프



(그림 4-3) train과 validation의 object, box, class Loss



(그림 5) 전자소자 인식 결과

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2022R1A4A1023248).

참고문헌

[1] Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors". arXiv, 2022

[2] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection". Computer Vision and Pattern Recognition, pp. 779 - 788, 2016.

[3] Joseph Redmon and Ali Farhadi. "YOLO9000: better, faster, stronger". Computer Vision and Pattern Recognition, pages 7263 - 7271, 2017.

[4] Joseph Redmon and Ali Farhadi. "YOLOv3: An incremental improvement". arXiv. 2022

[5] Thanh Binh Nguyen, Van Tuan Nguyen, Sun-Tae Chung, Seongwon Cho. "Real-time Human Detection under Omni-irectional Camera based on CNN with Unified Detection and AGMM for Visual Surveillance". Journal of Korea Multimedia Society, 2016.

[6] Xiaohan Ding, Xiangyu Zhang, Ningning Ma, Jungong Han, Guiguang Ding, Jian Sun. "RepVGG: Making VGG-style ConvNets Great Again. arXiv. 2022