

Effective Automatic Weed Detection With Improved YOLOv10

Hyeon-Jae Kwon*, Sangmin Suh**

*Student, Dept. of Information & Telecommunication Engineering, Gangneung-Wonju National University, Korea

**Professor, Dept. of Information & Telecommunication Engineering, Gangneung-Wonju National University, Korea

**CEO, Scalable-AI, Korea

[Abstract]

In this paper, we design an improved weed detection model using YOLOv10, a deep learning-based object detection algorithm. YOLOv10 improves its performance compared to previous versions by adding an attention module, the PSA module. PSA is strong at recognising complex patterns in large areas because it uses some features of its own attention to reduce computation and learn global information. However, it may be inefficient for certain problems, such as weeds, which are generally small objects. Therefore, in this paper, we propose an improved YOLOv10 by applying another attention module, SENet, instead of the PSA module. Since, SENet learns the importance between channels, it can learn the features of weeds in more detail than the PSA module. In addition, SENet is lighter, less computationally intensive, and faster than the PSA module, so we conducted experiments by replacing the PSA module with SENet, which is suitable for weed detection. The experiment consisted of 200 training runs with a total of 14 classes, and we compared the performance through various performance evaluations. The experimental results showed that the FPS increased from 476.19 to 526.32, which is about 9.52% processing speed improvement. The mAP50-95 value increased from 88.7% to 88.3%, which shows that the proposed model is lighter than the existing model and performs similarly to the existing model.

▶ **Key words:** Weed, YOLOv10, Detection, Deep learning, Agriculture

-
- First Author: Hyeon-Jae Kwon, Corresponding Author: Sangmin Suh
 - *Hyeon-Jae Kwon (hyeonjae0340@naver.com), Dept. of Information & Telecommunication Engineering, Gangneung-Wonju National University
 - **Sangmin Suh (sangminsuh@gwnu.ac.kr), Dept. of Information & Telecommunication Engineering, Gangneung-Wonju National University
 - **Sangmin Suh (sangminsuh@scalable-ai.io), Scalable-AI
 - Received: 2024. 10. 17, Revised: 2024. 11. 08, Accepted: 2024. 11. 08.

[요 약]

이 논문은 딥러닝 기반 객체 탐지 알고리즘인 YOLOv10을 활용하여 개선된 잡초 탐지 모델을 설계한다. 기존 YOLOv10에서는 Attention 모듈인 PSA 모듈을 추가하여 이전 버전들보다 성능을 개선하였다. PSA는 Self-Attention의 강력한 성능을 일부만 적용하여 연산량을 줄이고, 전역적 정보를 학습할 수 있어 큰 영역의 객체가 복잡한 패턴 인식에 강하다. 하지만 대체로 작은 크기의 객체인 잡초 같은 특정 문제에서는 비효율적일 수 있다. 따라서, 이 논문은 PSA 모듈 대신 다른 Attention 모듈인 SENet을 적용하여 개선된 YOLOv10을 제안한다. SENet은 채널 간 중요도를 학습하기 때문에 PSA 모듈보다 더 세밀하게 잡초의 특징을 학습할 수 있다. 또한, SENet은 PSA 모듈보다 더 가벼워 더 적은 연산을 수행하고, 더 빠른 속도로 탐지가 가능하여 잡초 탐지에 적합한 SENet으로 대체하여 실험을 진행했다. 실험은 총 14가지의 클래스로 200회 훈련을 수행했고, 다양한 성능평가를 통해 성능을 비교하였다. 실험 결과에 따르면, FPS는 476.19에서 526.32로 처리속도가 약 9.52%정도 향상되었다. mAP50-95값은 88.7%에서 88.3%로, 제안된 모델이 기존 모델보다 더 정량화된 모델임에도 불구하고 유사한 성능을 보인다.

▶ **주제어:** 잡초, 옴니v10, 검출, 딥러닝, 농업

I. Introduction

국내 인구 90% 이상이 도시지역에 거주하며, 농촌과의 거리는 점점 더 멀어지고 있다. 또한, 2022년도 초에 발생한 러시아-우크라이나 전쟁으로 발생한 곡물 파동, 기후재난 등은 세계적 식량 위기의 위험성을 경고하였다[1]. 때문에 농업 생산량과 농가 소득을 높이는 것이 오늘날 농업의 핵심 과제 중 하나이다. 농작물의 생산량을 방해하는 가장 큰 요소 중 하나는 바로 잡초이다. 잡초는 농작물의 영양분, 공간, 물 및 토양을 흡수하여 작물의 성장을 방해하고, 농작물에 기생하기 때문에 농작물 생산량과 품질을 크게 떨어뜨려 재정적 손실을 야기한다. 따라서 잡초를 방제하는 것이 매우 중요한 과제이다.

지금까지의 잡초 방제 방법은 수동 잡초 방제와 화학적 잡초 방제가 있다. 수동 잡초 방제는 많은 노동력을 필요로 하여 효율적이지 않으며, 화학적 잡초 방제는 제초제를 살포하여 잡초의 성장 속도를 줄여 수동 잡초 방제보다는 효율적이지만 넓은 지역에 제초제를 뿌리는 것은 제초제와 노동의 낭비뿐만 아니라 환경 오염 및 건강에 해로운 토양 환경을 야기한다. 또한, 잡초 방제를 위해서 정확한 잡초 식별을 하여 해당 잡초의 식물학적 특성에 맞는 제초제를 사용해야 한다[2]. 따라서 잡초만을 정확하고 효율적으로 식별하고 잡초의 종류와 밀도에 따라 정량적으로 살포하는 것이 중요하다.

최근 이미지 분류, 객체 검출 등 다양한 딥러닝 기반의 인공지능 기술이 크게 발전하여[3-4], 다양한 인공지능 기

술이 농업 생산 기여에 사용되었다[5-7]. 또한 객체 검출 기술을 활용하여 이미지나 비디오에서 객체를 식별하고 위치를 찾을 수 있다. 객체 검출모델은 대표적으로 크게 SSD(Single Shot Detection)와 YOLO(You Only Look Once)가 있다[8-9]. SSD는 속도는 YOLO보다 느리지만, 성능이 우수하고, YOLO는 속도가 매우 빠르며 SSD보다 성능이 낮아[10], 잡초를 실시간으로 감지하기 위해서는 속도가 빠른 YOLO를 사용하는 것이 적합하다.

이 논문에서는 기존 YOLOv10에서 사용한 PSA(Partial Self-Attention) 모듈보다 좀 더 가벼운 SENet(Squeeze and Excitation Network)을 사용하여 연산량과 파라미터 수를 줄여 속도를 높임으로써 잡초 검출에 효율적으로 사용할 수 있는 개선된 YOLOv10을 제안한다.

논문의 구성은 다음과 같다. 제2장에서는 YOLOv10에 사용되는 모듈들을 소개한다. 3장에서는 잡초 데이터 세트와 개선된 YOLOv10 모델을 설계한다. 제 4장에서는 훈련 결과에 대한 성능을 평가하고, 마지막 장에서 결론을 내린다.

II. Preliminaries

이 장에서는 YOLOv10과 그 내부에서 사용되는 모듈들을 설명한다.

1. Related works

1.1 YOLO

잡초를 식별하기 위해 실시간 객체 검출 모델로 유명한 YOLO를 사용하였다. YOLO는 이미지를 동일한 그리드 영역으로 나누어 각 그리드 영역에 대해 객체가 존재하는지의 여부를 바운딩 박스와 신뢰도 점수로 예측을 한다. 또한, YOLO는 Classification과 Localization을 동시에 하는 단일 단계 검출 방식을 사용하여 속도가 매우 빨라 실시간 객체 검출이 가능하다.

YOLO는 2016년에 처음 소개되었으며, 이후로 YOLO는 여러 버전으로 발전해왔다[11]. 가장 최신 버전은 YOLOv10이며, YOLOv10은 실시간 객체 검출 기술에 새로운 접근 방식을 도입하여 이전 YOLO 버전에서 발견된 모델의 부족한 부분들을 해결하였다. YOLOv10은 이전 버전까지 사용했던 비최대 억제 기법(NMS)을 제거하여, 많은 연산량을 줄여 추론 속도를 크게 개선하였다[12]. 또한 Attention 메커니즘을 추가하여 큰 계산 비용 없이 성능을 개선하였다.

1.2 YOLOv10 algorithm

YOLOv10 모델은 Backbone, Neck, Head로 구성된다. Backbone에서는 특징을 추출하여, Neck과 Head로 전달하는 역할을 한다. Neck에서는 추출된 특징들을 융합하여 Head로 전달한다. Head는 최종 객체 위치와 클래스를 예측한다. Backbone은 Conv, C2f, SCDn, PSA 모듈로 구성되어있다. Conv, C2f를 사용하여 다양한 특징들을 추출해내고, SCDn를 사용하여 특징 맵의 크기를 줄이면서 중요한 특징 정보를 압축합니다. PSA는 입력 특징 맵을 두 부분으로 분리(Split)하여 MHSA(Multi-Head self-Attention)을 통해 전역적인 피쳐들 간의 관계를 학습한다. PSA(Partial Self-Attention) 구조는 다음과 같다.

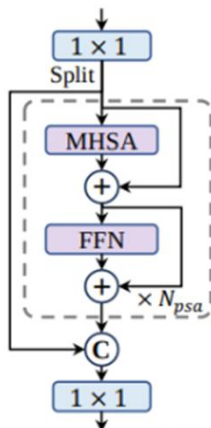


Fig. 1. The Partial Self-Attention module

PSA는 Self-Attention의 강력한 성능을 일부만 적용하여 연산량을 줄이고, 전역적 정보를 학습할 수 있어 큰 영역의 객체나 복잡한 패턴 인식에 유리하다. 하지만 작은 크기의 객체인 잡초와 같은 특정 문제에서는 비효율적일 수 있다. 또한, 실시간으로 처리하기에 있어서 여전히 많은 연산량, 연산 비용, 메모리 사용량을 보인다.

III. Proposed Model Design

이 장에서는 잡초 데이터 세트와 개선된 yolov10 모델을 제안한다.

1. Dataset

이 논문에서 사용하는 데이터 세트는 aihub 사이트에서 구하였다[13]. 데이터 세트는 총 15가지 종류의 잡초로 구성하였고, 다음과 같다.

Table 1. weed class

class	contents
0	gangppi
1	gaebireum
2	gaeyeuggui
3	ggaepul
4	gangpul
5	muldalgaebi
6	gamagsali
7	balaeng-i
8	soebileum
9	olbanga
10	golaeng-i
11	jomyeong-aju
12	hanlyeoncho
13	hunimyong-aju

각각 클래스마다 360장씩 모아 총 5040장의 데이터 세트를 구성하였다. 또한, 각 클래스마다 8:2의 비율로 나누어 5040장 중 3920장은 학습을 하기 위한 학습 데이터 세트(train data), 1120장은 검증을 하기 위한 검증 데이터 세트(validation data)로 나누었다. 또한, YOLO 형식에 맞게 라벨링 파일을 수정해주었다. YOLO의 라벨링 파일은 txt 형식으로 되었으며, txt 파일에는 객체에 대한 클래스와 바운딩 박스 좌표 정보(x, y, width, height)를 담고 있다.

2. Proposed model

1.1 SENet

잡초의 종류는 매우 다양하며, 모양새 또한 매우 비슷한 잡초들이 많다. 따라서, 이 논문에서는 잡초를 보다 더 세밀하고, 빠른 속도로 탐지하기 위해 SENet을 Backbone 네트워크에 도입한다. SENet은 현재 YOLOv10에 있는 PSA보다 모델 구조가 단순하여 더 빠른 훈련과 추론이 가능하며, 각 채널 간의 세밀한 특징을 잡는 데에 더 효율적이다. SENet은 각 채널이 포함한 정보의 중요도를 학습하여, 중요한 채널에 가중치를 더 부여하는 방식으로 모델의 성능을 향상시킨다. SENet의 구조는 다음과 같다.

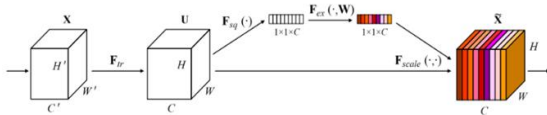


Fig. 2. A Squeeze-and-Excitation block

SENet은 크게 Squeeze, Excitation, Scale 세 가지 단계로 나뉘며, 이를 통해 채널 간 상관관계를 학습한다.

1. Squeeze

이 과정에선 각 채널이 이미지에서 얼마나 중요한지를 측정할 수 있다. GAP(Global Average Pooling)을 사용하여 입력 특징 맵의 각 채널 모든 위치에 있는 값들의 평균을 계산한다. Squeeze의 수식은 다음과 같다.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

$u_c(i, j)$ 는 입력 특징 맵의 c 번째 채널이고, i, j 는 해당 채널의 위치를 나타낸다. z_c 는 c 번째 채널의 모든 위치에 있는 값들의 평균이다.

2. Excitation

Squeeze 연산에서 나온 값들을 두 개의 FC(Fully Connected) Layer를 사용하여 채널 간 상관관계를 학습한다. Excitation의 수식은 다음과 같다.

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

s 는 각 채널의 최종 가중치이고, z 는 Squeeze 연산을 통해 나온 평균값을 나타낸다. W_1 은 첫 번째 FC Layer의

가중치 행렬이며, W_2 는 두 번째 FC Layer의 가중치 행렬이다. δ 는 활성화 함수인 ReLU 함수이다. σ 는 활성화 함수인 Sigmoid이다. Sigmoid를 통해 출력 값들을 0~1 사이로 정규화하여, 각 채널에 대한 가중치로 사용한다.

3. Scale

마지막으로, 이전의 연산을 통해 얻은 가중치를 원래의 입력 채널에 곱하여 채널 값의 중요도를 재구성한다. Scale의 수식은 다음과 같다.

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (3)$$

\tilde{x}_c 는 최종적으로 출력된 c 번째 채널 값이고, s_c 는 Sigmoid 함수를 통해 계산된 가중치이며, u_c 는 c 번째 채널의 입력값이다.

IV. Performance evaluation

이 장에서는 잡초 데이터 세트를 훈련한 모델에 대해 다양한 성능평가 지표를 사용하여 기존의 YOLOv10과 제안하는 YOLOv10의 성능을 비교한다.

1. Evaluation method

YOLOv10에서 성능평가를 하기 위해 성능평가 지표로 정밀도(Precision), 재현율(Recall), mAP@0.5(Mean Average Precision at IoU 0.5), mAP@0.5:0.95(Mean Average Precision at IoU 0.5 to 0.95)를 사용하였다. <표 2>는 Precision, Recall, F1 score에 사용되는 성능평가 지표를 나타낸다.

Table 2. TP, FP, FN, TN

		Predict	
		True	False
Actual	True	TP	FN
	False	FP	TN

여기서 TP는 실제 잡초를 정확히 잡초로 검출한 경우이고, FP는 실제 잡초가 아닌 것을 잘못 검출한 경우, TN은 잡초가 없는 것을 정확히 잡초가 아니라고 한 경우, FN은 실제 잡초가 있지만 모델이 검출하지 못한 경우를 뜻한다. <표 2>를 이용한 정밀도, 재현율 수식은 다음과 같다.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

mAP은 모델이 객체를 얼마나 정확하게 검출하는지를 평가하는 주요 지표이다. YOLOv10에서 mAP@0.5와 mAP@0.5:0.95를 사용하였다. mAP@0.5는 IoU 임계값을 0.5로 설정하여 계산한 평균 정확도를 뜻하며, mAP@0.5:0.95는 IoU 임계값을 0.5에서 0.95까지 0.05 간격으로 변화시키며 계산한 평균 정밀도를 뜻한다. mAP와 AP의 수식은 다음과 같다.

$$AP = \sum_n (R_n - R_{n-1}) \times P_n \quad (6)$$

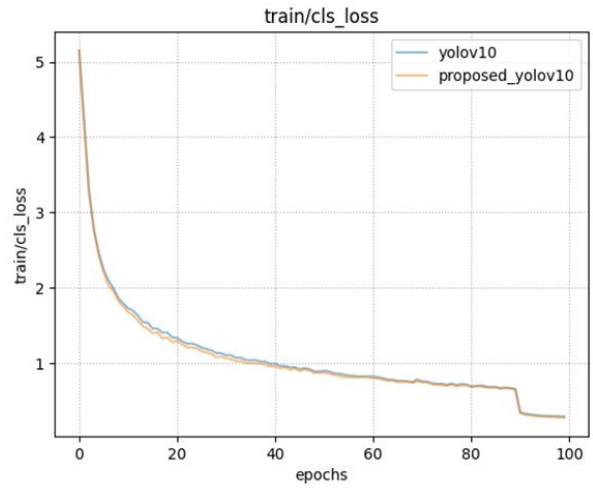
R_n 은 n 번째 지점에서의 Recall 값이고, P_n 은 n 번째 지점에서의 Precision 값을 나타낸다.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (7)$$

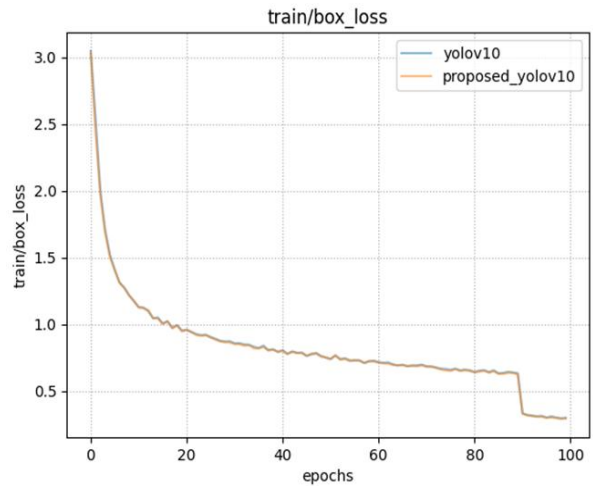
N 은 총 클래스의 개수이며, 각 클래스에 대한 AP의 평균값을 의미한다.

2. Performance comparison

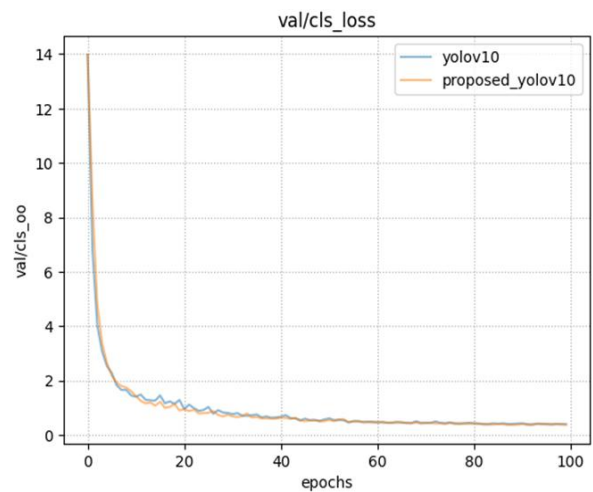
사전에 학습된 가중치를 사용하지 않은 YOLOv10 모델과 이 논문에서 제안하는 YOLOv10을 훈련하여 각각의 성능을 비교하였다. Optimizer로 AdamW를 사용하였고 Learning rate는 0.01을 사용하였으며, 훈련은 총 100회 진행하였다. Fig. 3은 훈련 100회를 진행하여 나온 train과 validation의 손실함수이다.



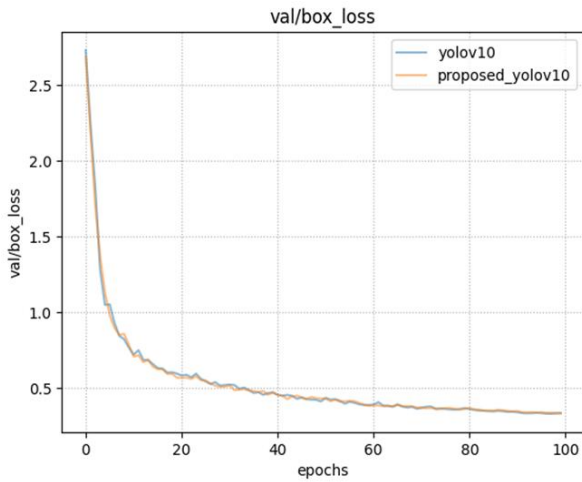
(a) train/cls loss



(b) train/box loss



(c) train/box loss



(d) train/box loss

Fig. 3. Loss function

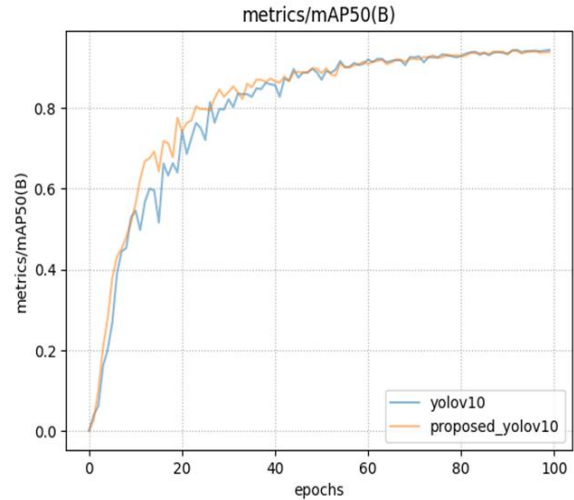
Fig. 3을 통해 학습이 반복될수록 class loss와 box loss가 점점 감소하는 것을 볼 수 있다. 이를 통해 두 모델이 잡초 데이터 세트에 대해서 학습을 잘하고 있다는 것을 확인할 수 있으며, 두 모델의 loss가 매우 유사한 것을 볼 수 있다. 이 결과는 더 경량화된 모델을 사용했음에도 불구하고, 여전히 기존의 YOLOv10과 비슷하게 loss가 감소하는 모습을 보이고 있다.

YOLOv10과 개선된 YOLOv10을 각각의 성능을 비교하기 위해 GFLOPS (Giga Floating Point Operations Per Second), Parameter, FPS (Frames Per Second), Inference Time을 사용하였다.

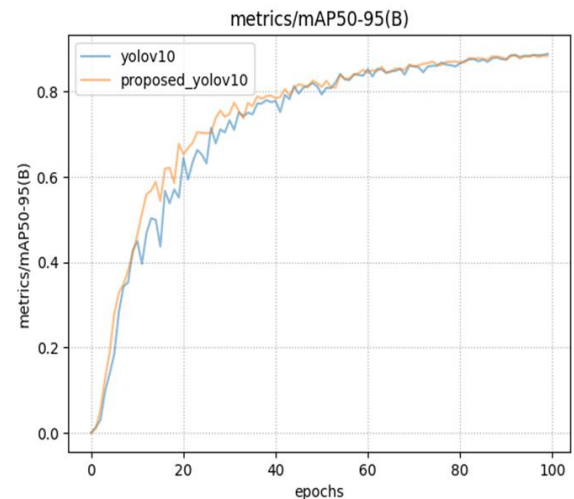
Table 3. Comparison of the performance of YOLOv10 and proposed YOLOv10

	YOLOv10	Proposed YOLOv10	Delta
GFLOPS	12.4	11.8	4.84%
Params(M)	8.1	7.0	13.58%
FPS	476.19	526.32	10.52%
Inference Time	0.0021	0.0019	9.52%
mAP50	94.3%	93.8%	-0.53%
mAP50-95	88.7%	88.3%	0.45%

표에 따르면, 전체적으로 제안된 YOLOv10 모델의 크기가 15.7% 작으며 추론 속도는 빠르다는 것을 알 수 있다.



(a) mAP50



(b) mAP50-95

Fig. 4. Comparison of mAP50 and mAP50-95

Fig. 4를 통해 기존 YOLOv10 모델과 제안된 YOLOv10 모델이 거의 유사한 성능을 보이는 것을 알 수 있다. 이는 제안된 YOLOv10 모델이 더 가벼운 모델임에도 불구하고, 정확도 측면에서 기존 모델과 유사한 성능을 보인다는 것을 의미한다.

이 논문에서 제안하는 모델을 노지에 적용하기 위해선 몇 가지 방법을 고려해볼 수 있다. 첫 번째로, 고해상도 카메라로 확인하고자 하는 전체 영역을 촬영한 후, 그 이미지를 분할하여 제안하는 모델에 차례로 적용해서 잡초를 구분하는 방법이고, 또 다른 하나의 방법은 카트나 작은 로봇에 카메라를 장착하여 실시간으로 노지를 스캔하여 잡초를 찾는 방법이다. 첫 번째 방법은 초고해상도의 고가의 카메라가 필요하고, 두 번째 방법은 로봇이 움직일 수 있는 최소한의 이동 공간이 필요하다는 단점이 존재한다.

그러므로 실사용자는 노지의 상황에 맞는 시스템 구성을 고려해야 한다.

V. Conclusions

이 논문에서는 YOLOv10을 활용하여 개선된 YOLOv10을 설계했다. 기존의 YOLOv10 모델에서 사용하는 PSA 모듈에서 SNet 모듈로 대체하여, Parameter 수와 연산량을 줄여 YOLOv10 모델을 더 가볍게 만들었다. 결과에 따르면 YOLOv10은 12.4 GFLOPS, 개선된 YOLOv10은 11.8로, 개선된 YOLOv10 모델이 더 적은 계산량으로 동작했다, Parameter 수는 YOLOv10이 8.1M이고 개선된 YOLOv10이 7.0M으로 연산량이 더 줄어든 것을 확인했다. YOLOv10 모델은 1초에 약 476장의 이미지를 처리했으며, 개선된 YOLOv10 모델은 1초에 약 526장의 이미지를 처리하였다. 이는 처리 속도가 약 9.52%정도 향상되었다는 결론이다. mAP 또한, 기존의 YOLOv10과 거의 유사한 성능을 보였다. 이러한 성능평가를 통해, 제안된 YOLOv10 모델이 성능 손실 없이 연산량을 줄여 효율성을 높인 모델임을 검증하였다.

ACKNOWLEDGEMENT

This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2024-RS-2023-00260267)

REFERENCES

- [1] Campaigns surveys <https://campaigns.do/surveys/329>
- [2] Yoo Kwang Hyun, Vo-Hoang Trong, Kim Jin Young, "Deep Learning Research Trends For Weed Identification", Journal of The Korean Institute of Communication Sciences, No.8, 2021.
- [3] S. Suh, "Improved Motion Recognition Based On Convolution Neural Network For High Accuracy", journal of Knowledge Information Technology and Systems(JKITS), No. 3, 2021.
- [4] S. Suh, "Human Activity Recognition Systems Based On Maximum Color Difference And Deep Learning", Journal of Knowledge Information Technology and Systems(JKITS), No.4, 2021.
- [5] Gwanghyun Yu, Jaewon Lee, Vo Hoang Trong, Dang Thanh Vu, Huy Toan Nguyen, HooHwan Lee, Dosung Shin, Jinyoung Kim, "Alien Weed Classification: A Hierarchical Structure Based On Convolutional Neural Networks", Journal of Korean Institute of Information Technology, No. 12, Dec, 2019.
- [6] Harshita Shri Panati, Gopika P, Diana Andrushia A and Mary Neebha T, "Weeds And Crop Image Classification Using Deep Learning Technique", International Conference on Advanced Computation and Communication Systems(ICACCS), pp 117-122, DOI:10.1109/ICACCS57279.2023.1011958.
- [7] Y. Dandekar, K. Shinde, J. Gangan, S. Firdausi and S. Bharne, "Weed Plant Detection From Agricultural Filed Images Using YOLOv3 Algorithm", International Conference On Computing, Communication, Control And Automation(ICCUBEA), pp 1-4, DOI: 10.1109/ICCUBEA54992.2022.10011010.
- [8] Liu, W. et al. "SSD: Single Shot MultiBox Detector", ECCV Lecture notes in Computer Science, vol 99095. DOI: https://doi.org/10.1007/978-3-319-46448-0_2.
- [9] J. Redmon, S.Divvala, R. Girshick and A. Farhadi, "YOU Only Look Once: Unified, Real-Time Object Detection", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 779-788, DOI: 10.1109/CVPR.2016.91.
- [10] S. Tariyal, R. Chauhan, Y. Bijalwan, R. Rawat and R. Gupta, "A comparative study of MTCNN, Viola-Jones, SSD and YOLO face detection algorithms," International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), pp. 1-7, doi: 10.1109/IITCEE59897.2024.10467445.
- [11] YOLOv10 object detection Better, Faster and Smaller, visionplatform, <https://visionplatform.ai>
- [12] YOLOv10-docs <https://docs.ultralytics.com/ko/models/yolov10/>
- [13] Dataset, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=527>

Authors



Hyeon-Jae Kwon received a B.S degree in information & Telecommunication Engineering at Gangneung-Wonju National University, Korea in 2023. He is currently pursuing an M.S degree at the Graduate School of General

Studies, Gangneung-Wonju National University, Korea.



Sangmin Suh received B.s., M.s., and Ph.D degree in the electronics engineering in the Hanyang university of Seoul Korea in 1991, 1994, and 2003 respectively. From 1994 to 1999, he has been with Daewoo

telecommunication. From 2003 to 2019, he was in Samsung electronics. He currently holds assistant professor in the department of information and telecommunication engineering at Gangneung-Wonju national university. His current research interests include artificial intelligence and robotics.