

다중 YOLO 모델을 활용한 방울토마토 성숙도 판별 성능 비교 연구

Performance Comparison Study of Cherry Tomato Ripeness Detection Using Multiple YOLO Models

양 다 연^{1,4}, 손 형 일^{1,2,3,*}, 주 찬 영^{1,4,*}
(Dayeon Yang^{1,4}, Hyoungh Il Son^{1,2,3,*}, and Chanyoung Ju^{4,*})

¹Department of Convergence Biosystems Engineering, Chonnam National University

²Interdisciplinary Program in IT-Bio Convergence System, Chonnam National University

³Research Center for Biological Cybernetics, Chonnam National University

⁴Purpose Built Mobility Group, Korea Institute of Industrial Technology

Abstract: Millions of tons of cherry tomatoes are produced annually, with the harvesting process being crucial. This paper presents a deep learning-based approach to distinguish the ripeness of cherry tomatoes during the harvesting process. It specifically analyzes the performance of the YOLO (You Only Look Once)v5 and YOLOv8 models, which are designed to rapidly detect the ripeness of cherry tomatoes in real time. The proposed system is expected to resolve rural issues caused by population decline and enhance productivity in cherry tomato harvesting environments. Future research will aim to apply this technology to harvesting robots for the automatic recognition and harvesting of cherry tomatoes.

Keywords: deep learning, harvesting robot, object detection, cherry tomato, YOLO (You Only Look Once)

I. 서론

방울토마토는 영양가가 높고 다양한 요리에 사용할 수 있어 많은 가정과 일상 생활에서 흔히 볼 수 있는 작물이다. 방울토마토 생산의 경우 수확 작업은 전체 인력의 40% 이상을 필요로 한다[1]. 방울토마토의 수동 수확 과정은 높은 인건비와 함께 수확기에 평소보다 많은 인력이 요구되는 어려움에 직면하고 있다. 농업은 기상 조건에 영향을 받을 뿐만 아니라 노동 시장에서도 많은 제약이 있는 작업이다. 실제 농업 환경에서는 충분한 인력을 확보하는 것이 어렵고, 기계적 수확 방법에 비해 시간이 더 많이 소요되는 지속 가능성 문제를 겪고 있다[2,3]. 이러한 문제를 해결하기 위해, 컴퓨터 비전과 딥러닝을 활용한 자동화 기술이 농업 분야에서 점점 더 중요한 역할을 하고 있다[4,5]. 특히, 농작물의 크기, 색상, 성숙도를 판단하여 작업의 효율성을 높일 수 있는 자동화된 시스템을 개발하기 위한 방울토마토 수확 로봇 연구는 계속해서 진행되어야 한다.

방울토마토 수확 로봇 작업에서 제일 중요한 핵심은 시각적 시스템을 통해서 정확한 방울토마토의 식별의 신뢰도를 높이는 것이다[6]. 실시간으로 작물의 이미지를 분석하여 로봇 스스로 작물을 분류하는 것은 농업 자동화 환경에서 필수적인 연구이다[7]. 수확로봇이 영상 또는 이미지로 방울토마토를 인식할 때는 시간에 따른 자연광, 잎이나 줄기에 의해

가려짐, 잎의 녹색과 비슷한 미성숙 토마토와의 구별과 같은 많은 어려움이 존재한다.

본 논문에서는 속도와 정확도 면에서 높은 성능을 제공하는 딥러닝 기반의 객체 검출 알고리즘인 YOLOv5 [8]와 YOLOv8 [9] 모델 및 YOLOv8에 ResNet50 백본을 적용한 모델을 비교 테스트하고 방울토마토의 성숙도를 실시간으로 판별하는 수확하는 시스템을 제안한다.

300장의 이미지를 라벨링한 후 데이터 증강 기법을 이용하여 데이터 양을 증가시켜 방울토마토의 성숙도를 학습함으로써 이미지 분류의 정확도를 평가하였다.

이와 관련된 자세한 내용은 2장에서는 딥러닝을 기본으로 농업분야에서 방울토마토를 성숙도 판별 자동화를 위해 컴퓨터 비전 및 머신러닝 기술을 이용한 관련 연구들을 소개하고 3장에서는 CNN기반의 방울토마토 성숙도 판별 알고리즘 구조 4장에서는 데이터셋으로 학습시킨 딥러닝 모델의 실험 과정 5장에서는 그 결과를 분석하는 내용을 다룰 것이다. 마지막 장에는 본 연구의 결론과 향후 연구 방향에 대해 논의한다.

II. 관련 연구

컴퓨터비전 및 머신러닝을 활용하여 농작물의 상태를 관찰하고 농작물의 성장 과정에서 방해 요소나 병충해를 자동

* Co-Corresponding Author

Manuscript received August 9, 2024; revised September 11, 2024; accepted October 16, 2024

양다연: 전남대학교 융합바이오시스템기계공학과 대학원생(ydy4890@kitech.re.kr, ORCID[®] 0009-0001-4372-0679)

손형일: 전남대학교 융합바이오시스템기계공학과 교수(hison@jnu.ac.kr, ORCID[®] 0000-0002-7249-907X)

주찬영: 한국생산기술연구원 목적기반모빌리티그룹 선임연구원(cyju@kitech.re.kr, ORCID[®] 0000-0001-9848-6819)

※ 본 연구는 농림축산식품부 및 과학기술정보통신부, 농촌진흥청의 재원으로 농림식품기술기획평가원의 스마트팜다부처패키지혁신기술개발사업의 지원을 받아 연구되었음(No. 421032-04).

으로 인식하려는 시스템을 농업분야에서 많은 연구가 진행되고 있다. 수확로봇에서 방울토마토의 성숙도를 구별하기 위해 CNN(Convolutional Neural Network)기반의 YOLOv5 모델을 사용하여 농작물을 탐지하고 선별하는 연구를 진행해 왔다.

본 섹션에서는 기존 연구에서 YOLO 모델을 활용한 농작물 탐지의 성공 사례를 검토하고, 이를 최신 YOLOv8 모델과의 비교 분석을 통해 그 향상된 활용도를 탐색한다.

알고리즘 모델로 실시간으로 대상을 인식하고 탐지하는 것을 객체 탐지라고 하며 농업분야에서는 이러한 딥러닝 방법이 많이 사용되어 왔다.

[10]연구에서는 채소 질병을 탐지하기 위하여 YOLO 시리즈의 백본 네트워크 중 하나인 CSPDarknet을 개선하여 CSP-TR (Cross Stage Partial Transformer)구조로 채소 질병을 탐지하는 알고리즘을 제안하였고 이 모델은 기존 YOLOv5s 모델과 비교한 결과 객체 탐지 정확도는 93.1%였다.

[11]은 복잡하고 막혀있는 토마토 환경에서 딥러닝에서 많이 사용되는 RNN (Recurrent Neural Network), CNN 기반의 아키텍처를 사용하지 않았다. 오직 Transformer Attention 메커니즘만을 사용하여 토마토 성숙의 클래스를 정확하게 분류하였고 토마토 수확, 등급 지정 및 품질 관리에서 높은 효율성과 정확성을 향상시켰다.

[12]연구에서는 최신 실시간 객체 탐지 기술 알고리즘인 YOLOv9로 토마토의 성숙도를 분류하고 개수를 세는 방법을 제안하였다. 본 알고리즘은 정보의 손실을 방지하는 PGI, 경량 네트워크 구조인 GELAN을 두 가지 기술을 새로 도입하여 기존의 YOLO 알고리즘을 개선하였다. 이는 0.882의 성능을 보여주었다.

[13]은 MHSA 어텐션 메커니즘을 사용하여 YOLOv8의 백본을 개선하였다. 이는 실제 생산 환경의 복잡성을 고려하여 새로운 객체 탐지 방법이며 토마토 성숙도의 등급과 계수 모델을 구축하여 토마토의 성숙도를 구별하였다. 이 모델은 0.864의 mAP를 보여주었다.

[14]은 기존의 Conv 레이어를 대체하여 DSConv 레이어를 사용하고 DPAG (Dynamic Pyramid Attention Gate), FAM (Feature Attention Module) 모듈을 추가하여 객체의 특징을 강조 하고 정밀도를 향상시키며 계산 복잡도를 줄였다. 실험결과 mAP가 93.4%를 보여주며 농업 환경에서 실시간으로 토마토 감지 요구 사항을 충족하였다.

[15]은 YOLOv8과 CenterNet 객체 감지를 위한 딥러닝 모델을 사용하여 토마토의 성숙도를 분류하였다. 각 객체의 중심을 키포인트로 간주하고 해당 중심점을 기준으로 크기와 위치를 계산하는 CenterNet 모델은 YOLOv8을 사용한 모델이 99.5%로 더 높은 정확도를 나타내었다.

[16]은 ResNet 네트워크 아키텍처를 사용하여 Mask R-CNN 딥러닝 알고리즘으로 객체 탐지 및 세그먼트화 작업을 수행하였다. Mask R-CNN이 생성한 분할 결과를 후 처리하여 배경에서 탐지된 과일을 제거하고 각 객체의 마스크에 해당하는 픽셀의 중간 깊이 값을 매트랩으로 계산하여 특정 임계값을 초과하면 해당 객체를 전경으로 간주하였다. 그리하여 온실에서 간단하고 저렴한 카메라 세트를 사용하여 토마토를 감지하는데 효과적이었다.

[17]에서는 기존의 YOLO 모델의 성능을 개선하기 위하여 YOLOv2 모델을 수정하여 M3 모델로 나뉘는 내에 있는 과일을 탐지한다. 먼저 M1 모델에서는 입력 이미지를 13×13 그리드 대신 26×26 그리드로 분할하여 작은 객체를 더 잘 탐지할 수 있도록 하였지만 훈련, 탐지 속도가 늦어진 것을 개선하기 위해 M2 모델에서 일부 컨볼루션 레이어와 풀링 레이어를 제거하여 모델을 얇게 만들어 속도를 높였다. 하지만 이 과정에서는 약간의 정확도 손실이 발생하여 M2 모델을 기반으로 Splitter와 Joiner 블록을 추가한 M3 모델을 사용하였다. 이는 더 높은 해상도에서 이미지를 처리할 수 있게 하여 작은 객체를 더 잘 탐지할 수 있게 하고 정확도를 향상시키는 동시에 적절한 속도를 유지할 수 있도록 하였다.

성숙도 구별에 있어서 작물의 색상은 중요한 역할을 하며 우리가 흔히 볼 수 있는 사과, 오렌지, 바나나와 같은 작물들은 고유한 색상을 가지고 있다. 하지만 작물 중에 일부 과일은 완전히 익어도 초록색 잎과 유사한 색상을 가져서 색상만으로 작물과 잎을 구별하기는 어렵다[18]. 오이와 같이 잎과 비슷한 색상을 가진 문제를 극복하기 위해 [19]에서는 I-RELIEF 알고리즘과 MPCNN (Multi-Path Convolutional Neural Network)을 사용하였다. 오이와 배경을 구분하는데 유리한 상위 3개의 색상 구성 요소를 선택하여 가중치를 계산하고 배경 간의 차이를 감소시켜 색상 구성 요소로부터 각각 별도의 경로로 입력 받아 개별적인 컨볼루션 레이어를 통해 특징을 추출하였다. 이러한 방법을 기반으로 오이 인식을 한 결과 90% 이상의 높은 정확도를 달성하였다.

Faster R-CNN 알고리즘은 높은 정확도를 제공하지만 처리 속도가 느려 실시간 탐지가 필요한 방울토마토 성숙도 판별을 위해 YOLO 알고리즘을 사용할 것이다. 본 연구에서는 YOLOv8의 백본을 ResNet50으로 변경하고 다양한 모델 크기에 대한 비교 실험을 통해 실시간 수확 로봇 환경에서 높은 정확도와 효율성을 동시에 달성하고자 한다.

III. 성숙도 판별 알고리즘

본 장에서는 방울토마토의 성숙도를 판별하기 위한 딥러닝 기반의 객체 검출 알고리즘인 YOLO 시리즈 모델들과 YOLOv8 변형 모델을 직접 실험하여 시스템의 성능을 비교하고 평가한다. YOLO 알고리즘은 모델은 구조가 단순하고 신경망을 통해 직접 바운딩 박스와 위치의 카테고리를 출력할 수 있어 모델의 크기가 작고 계산 속도가 빠르다. YOLO는 최종 탐지 결과를 얻기 위해 한 번의 처리 과정만 있기 때문에 실시간으로 탐지가 가능하며, 높은 일반화 능력을 가지고 있어 다른 분야로의 전이 학습도 가능하다[20].

1. YOLO

농작물의 성숙도를 분류하는 문제는 성숙도 클래스별로 차이가 커서 분류 난이도가 높고 한 장면에 많은 개체들이 동시에 나타나 객체탐지가 필요하다. 그리하여 본 연구는 시스템의 구현을 위하여 CNN의 기반인 YOLOv5와 YOLOv8 모델을 비교 테스트하였다. 기존의 YOLOv5 모델은 표적 탐지의 정확성을 극대화하기 위해 다양한 데이터 증강기법인 이미지의 회전, 채도 조정, 노출 조절 등이 포함되며 더욱

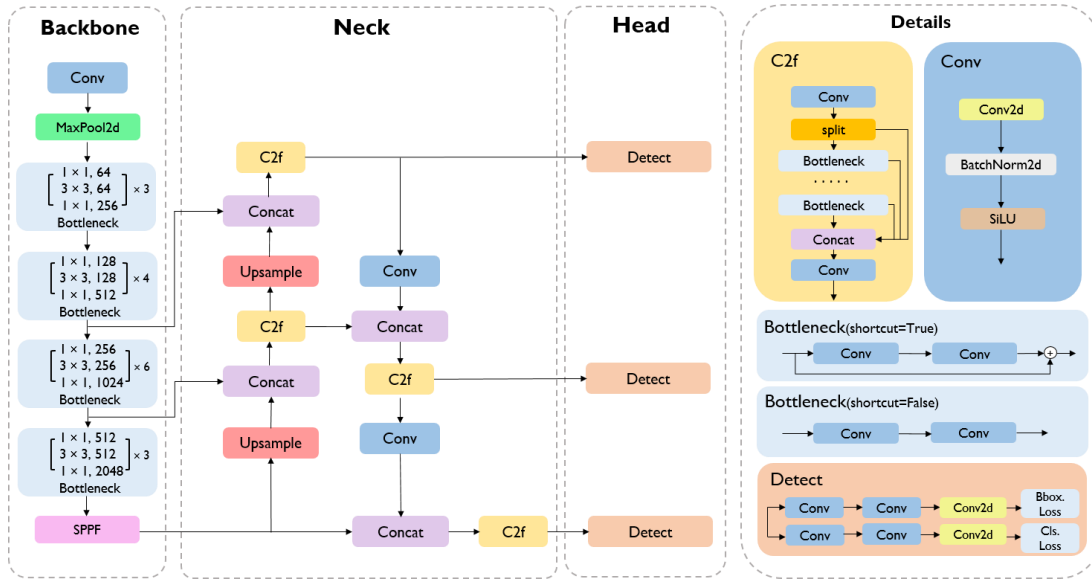


그림 1. 방울토마토 탐지를 위해 ResNet50의 Bottleneck 구조의 Residual Block을 사용하는 YOLOv8 아키텍처.

Fig. 1. YOLOv8 architecture using ResNet50's bottleneck residual block for cherry tomato detection.

많은 배경에서 객체를 정확하게 식별할 수 있도록 한다. YOLOv5는 훈련 과정 중에 훈련 데이터셋에 가장 잘 맞는 anchor boxes를 동적으로 계산하는 기능을 가지고 있어 네트워크가 자동으로 최적의 anchor를 찾아내 성능을 개선할 수 있다. 하지만 배경이 복잡한 이미지에 대해서는 농작물의 생육, 질병, 해충을 선별하는 능력이 다소 부족하여 인식 성능이 낮아 이를 보완한 것이 YOLOv8이다.

YOLOv5는 Pytorch로 구현하여 설계한 모델로 YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x로 나뉘며 n이 가장 가벼운 모델이다. 이 알고리즘은 주로 Backbone, Neck, Head 부분으로 구성된다.

최신 모델인 YOLOv8에서는 YOLOv5에서 사용되었던 CSP Layer를 더 효율적이고 간결한 C2f 모듈로 대체하여 모델의 구조가 단순화되었으며 연산 효율성을 향상시켰다. 또한 SPPF (Spatial Pyramid Pooling Fast) 레이어를 도입하여 다양한 크기의 이미지 특징들을 고정된 크기의 맵으로 풀링함으로써 연산 속도를 더욱 가속화했다. 첫 번째 컨볼루션 레이어의 커널 크기를 6x6에서 3x3으로 줄여 총 파라미터 수를 절반 이하로 감소시켰고, 이에 따라 모델의 크기와 메모리 사용량을 크게 줄였다. 이러한 개선 사항들은 모델의 추론 속도와 성능을 크게 향상시켜, 실시간 객체 탐지 작업에서의 효율성을 증대시켰다. 따라서, 방울토마토 성숙도 분류 작업에 있어 모델들의 성능 차이를 실험을 통해 구체적으로 분석한다.

2. Improved YOLO

이미지 특징 추출 또는 인식 시스템에서는 신경망의 깊이가 모델의 성능에 중요한 영향을 미치지만 네트워크가 깊어질수록 발생할 수 있는 기울기 손실 또는 폭발 문제로 성능 저하가 우려된다. 이를 해결하기 위해 ResNet [21]이 도입되었으며 이 구조는 Residual block을 활용하여 기울기 소실 문제 현상을 보완하고 연산 비용을 효율적으로 계산한다. ResNet은 입력 데이터의 정보를 보존하면서 데이터의 크기를

줄이고 잔차 함수를 학습함으로써 신경망의 깊이가 증가하더라도 뛰어난 성능을 유지할 수 있도록 설계되었다. ResNet은 첫 번째 1x1 컨볼루션에서 입력차원을 64차원으로 줄여 차원축소를 하고 3x3 컨볼루션에서 차원을 유지하며 중간 레이어에서 실제 특징을 추출한다. 마지막 두 번째 1x1 컨볼루션에서 다시 차원을 확장시켜 출력을 생성하여 처음 입력 차원과 동일하게 맞춘다. 특히 Shortcut Connection을 통해 레이어를 건너뛰어 각 레이어가 입력데이터에 대한 잔차 $F(x)$ 만을 학습 하여 최종 출력은 $y=F(x)+x$ 가 되어 기울기가 소실되는 문제 없이 안정적인 성능을 유지한다.



그림 2. 방울토마토의 클래스 분류. (a) 성숙한 방울토마토. (b) 미성숙한 방울토마토.

Fig. 2. Class classification of cherry tomatoes. (a) Ripe cherry tomato. (b) Unripe cherry tomato.

본 연구에서는 이미지 탐지 정확성을 높이기 위해 기존의 YOLOv8 모델의 백본을 ResNet50으로 교체하여 적용하였다. ResNet50은 총 50개의 레이어로 구성되어 각 레이어는 깊어지는 신경망의 복잡성과 학습시간을 효율적으로 관리하기 위해 bottleneck 구조를 도입하였다. 이 구조는 1x1 컨볼루션 레이어를 사용하여 시작과 끝에 채널 차원을 조절함으로써 연산 비용을 최소화하면서도 정보의 손실을 줄인다. 그림 1은 ResNet50으로 백본을 변형한 YOLOv8의 구조를 나타낸다.

IV. 성숙도 판별 실험

1. 데이터셋

본 논문에서 사용하는 방울토마토 이미지는 머신러닝 데이터셋 플랫폼 Kaggle에서 제공하는 토마토 성숙도 데이터셋인 Tomato Fruits Dataset를 사용하였다. 딥러닝 모델을 학습시키기 위해 중요한 작업인 이미지 애노테이션을 수행하였다. 방울토마토의 색이 빨간색으로 명확한 것은 'ripe' 비교적 주황색이나 초록색을 띄는 것은 'unripe'로 두 가지 클래스로 나누어 라벨링을 하였고 이를 그림 2에 나타내었다.

Roboflow를 사용하여 데이터를 애노테이션하였으며 학습 비율은 학습용 데이터셋 70%, 검증용 데이터셋 20%, 테스트용 데이터셋 10%로 설정하였다. 방울토마토 열매를 각각 bounding box로 애노테이션하여 이 영역은 객체가 위치 영역과 경계를 정확하게 나타낸다. 익은 방울토마토를 수확하기 위한 시스템을 학습시키기 때문에 방울토마토가 빨간색이 아니라면 모두 익지 않은 방울토마토로 인식이 될 수 있다. 데이터 증강 기법은 방울토마토의 성숙도를 정확하게 판별하는 모델의 성능을 높이고 과적합을 방지하기 위하여 데이터의 다양성을 높이는 것이 중요하다. 날씨 변화로 인해 열매의 밝기가 달라지고, 다양한 환경 조건이 존재하기 때문에

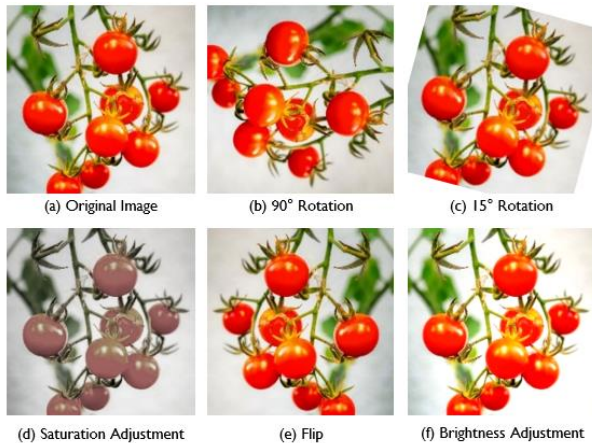


그림 3. 방울토마토 이미지에 적용된 다양한 데이터 증강 기법으로 (a) 원본 이미지, (b) 90° 회전, (c) 15° 회전, (d) 채도 조정, (e) 확대, (f) 밝기 조정.

Fig. 3. Various data augmentation techniques applied to cherry tomato images, showing. (a) Original Image. (b) 90° Rotation. (c) 15° Rotation. (d) Saturation Adjustment. (e) Zoom. (f) Brightness Adjustment.

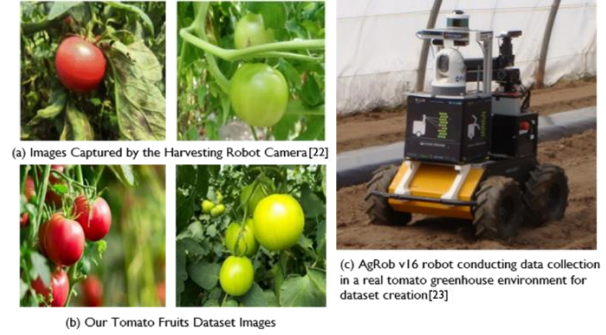


그림 4. (a) 실제 로봇에 장착된 카메라로 촬영한 토마토 이미지[22]. (b) Tomato Fruits Dataset의 토마토 이미지. (c) 실제 토마토 온실 환경에서 데이터셋 수집을 위해 촬영을 수행한 AgRob v16 로봇[23].

Fig. 4. (a) Tomato images captured by cameras mounted on the actual harvesting robot [22]. (b) Tomato images from the Tomato Fruits Dataset. (c) AgRob v16 robot conducting data collection in a real tomato greenhouse environment for dataset creation [23].

본 실험에서는 그림 3과 같이 원본 이미지에 대하여 회전(rotation), 채도(saturation), 밝기(brightness), 노출(exposure), 뒤집기(flip)를 포함한 데이터 증강 기법을 사용하여 다양한 데이터셋을 구축하였다. 특히, 이러한 데이터 증강을 통해 수확 로봇이 실제로 작동하는 다양한 환경 조건과 유사한 데이터를 확보하여 모델이 실제 로봇에 적용될 때에도 높은 성능을 발휘할 수 있도록 하였다. 그림 4는 실제 수확 로봇에 장착된 카메라로 수집된 방울토마토 이미지와 본 연구에서 구축한 데이터셋 간의 유사성을 보여주며, 데이터셋 생성을 위해 실제 방울토마토 온실 환경에서 데이터 수집을 수행하는 데 사용된 AgRob v16 로봇도 함께 제시한다[22,23].

총 742개의 이미지가 생성되었으며 이 중 663개의 이미지는 딥러닝 모델을 훈련하는 데 사용되었고, 나머지 63개의 이미지는 모델 성능을 평가한다.

2. 방울토마토 성숙도 판별 실험

본 연구에서는 방울토마토의 성숙도를 판별하는 실험을 하기 위해 커스텀 데이터셋을 활용하여 다양한 YOLO 모델을 학습시켰다. 학습 과정에서 각 모델의 성능을 최적화하기 위해 손실 함수를 적용하였고 손실 함수는 바운딩 박스의 위치, 클래스 확률, 객체 존재 여부에 대한 오차를 최소화하기 위해 설계되었다. 이를 통해 모델이 방울토마토의 정확한 위치와 성숙도 상태를 효과적으로 예측할 수 있도록 학습된다.

모델의 성능을 평가하기 위하여 precision, recall, mAP (mean average precision)을 식 (1)~(3)에 나타내었다.

$$\text{precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

$$\text{recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

표 1. YOLO 모델의 실험 성능 비교.

Table 1. Experimental performance comparison of YOLO models.

Model	precision	recall	mAP	Layers	Parameters	GFLOPs
YOLO-v5s	0.657	0.751	0.701	157	7015519	15.8
YOLO-v9s	0.679	0.755	0.738	917	7318368	27.6
YOLO-v8s	0.639	0.739	0.654	168	11126358	28.4
YOLO-v8m	0.604	0.774	0.67	218	25840918	78.7
YOLO-v8l	0.626	0.729	0.671	268	43608150	164.8
YOLO-v8x	0.616	0.717	0.657	268	68125484	257.4
YOLO-v8s(ResNet50)	0.668	0.784	0.721	350	30990102	86.7
YOLO-v8m(ResNet50)	0.677	0.76	0.733	378	39229974	111.8
YOLO-v8l(ResNet50)	0.707	0.775	0.757	406	49231382	151.7
YOLO-v8x(ResNet50)	0.718	0.768	0.745	406	62828886	196.6

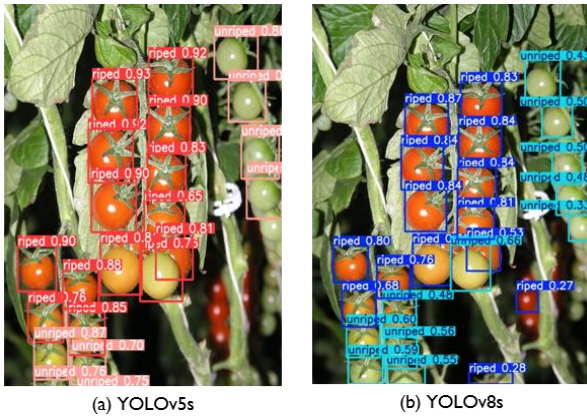


그림 5. 방울토마토의 숙성 감지 및 분류 결과.

Fig. 5. Detection and classification of cherry tomato ripeness.

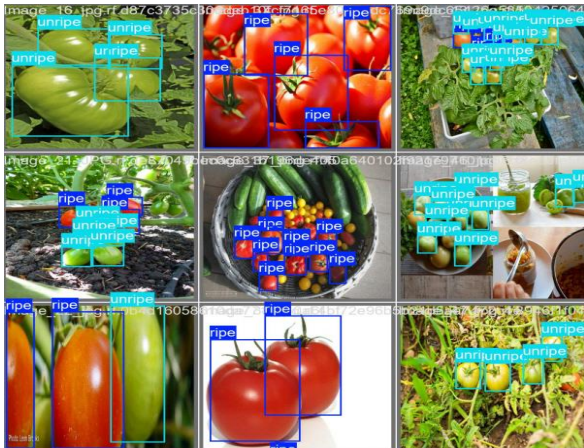


그림 6. 검증 데이터셋의 예측 결과.

Fig. 6. Prediction results of the validation dataset.

식 (3)의 N 은 이미지 데이터셋 내 객체 클래스의 총 수를 나타내고 AP_i 는 Class i 의 평균 정밀도를 의미한다. 여기서 True Positive는 모델이 익은 토마토를 정확히 익은 토마토로 정확히 예측하였을 때 수치를 의미한다. False Positive는 모델이 안 익은 토마토를 잘못 예측하여 익은 토마토로 예측할

경우의 수치이고 False Negative은 모델이 익은 토마토를 잘못 예측하여 안 익은 토마토로 예측하였을 때 수치를 의미한다.

3. 실험 환경

본 연구에서 실험을 위해 구축한 환경은 우분투(ubuntu) 22.04 운영 체제에서 CUDA 12.2 버전으로 수행되었으며, 컴퓨팅 자원으로는 NVIDIA GeForce RTX 4070 GPU를 활용하였다. 모델은 총 100회의 학습 에포크(epoch)와 초기 학습은 0.01으로 학습되었다. 다양한 YOLO 모델들을 비교하기 위해 YOLO 시리즈 모델과 YOLOv8의 다양한 변형 모델들을 학습시켰으며, YOLOv8 모델에 ResNet50 백본을 결합한 모델도 학습시켰다. 이를 통해 각 모델의 성능을 비교하고 분석하였다.

V. 실험 결과

본 연구에서는 YOLOv5와 YOLOv8 모델들의 성능을 평가하였고 그림 5에서 방울토마토의 성숙도에 따라 익은 토마토와 덜 익은 토마토를 분류하는 모델의 능력을 그림 6은 검증 데이터셋에서의 예측 결과를 시각적으로 보여준다. 객체 알고리즘의 테스트 결과는 표 1에서와 같이 탐지 성능, 모델의 크기와 속도를 보여준다. YOLOv5s 모델의 경우 precision은 0.657, recall은 0.751, mAP는 0.701로 측정되었다.

특히 기존 YOLOv8 모델들의 백본을 ResNet50으로 교체하여 성능을 정밀도(precision), 재현율(recall), 평균 정밀도(mAP), 레이어수, 파라미터수 그리고 GFLOPs (Giga Floating Point Operations Per Second)을 기준으로 평가한다. 백본을 ResNet50으로 교체한 YOLOv8 모델들은 일반적으로 기존 모델들에 비해 정밀도가 향상되었다. YOLOv8s 모델은 기존 정밀도 0.639에서 ResNet50 백본 사용 시 0.668로 개선되었으며, mAP 역시 0.654에서 0.721로 증가하였다.

실험 과정에서 모델의 구조 및 파라미터 설정이 성능에 미치는 영향을 확인하였으며, YOLOv8l (ResNet50) 모델은 가장 높은 0.757 mAP를 기록하였다. 이는 방울토마토 수확을 위한 높은 정밀도로 실시간 작업 수행에 최적화되어 있음을 나타낸다. 농업 환경에서의 실시간 방울토마토 감지 및 수확 작업에 이 모델을 적용함으로써 높은 수확 효율과 빠른 처리를 기대할 수 있다.

VI. 결론

본 논문에서는 방울토마토 수확 로봇을 위한 실시간 방울 토마토 판별 객체 검출 알고리즘을 제안하였다. 이 연구는 YOLO 딥러닝 모델을 활용하여 객체 인식의 속도와 정확도를 향상시키는 것을 목표로 하였고 기존의 300장의 데이터셋을 데이터 증강 기법을 통해 총 742장으로 확장하여 새로운 데이터셋을 구축하였으며, 이러한 데이터셋은 기존 데이터보다 성능 면에서 우수함을 보였다.

이후 연구에서는 방울토마토 판별 알고리즘을 더욱 세분화하여 방울토마토 열매만을 타깃으로 한 세그멘테이션 작업을 진행할 계획이다. 이를 통해 로봇이 각 방울토마토의 위치를 정확히 탐지하고, 매니퓰레이터를 사용하여 방울토마토의 성숙도에 따라 자동으로 수확할 수 있도록 한다. 현재 데이터셋의 양이 다소 적어 실험 결과가 높지 않았으나 향후 데이터셋을 확장하여 추가 실험을 진행할 예정이다. 또한 해당 모델을 실제 환경에서 하여 알고리즘의 효율성과 실용성을 검증하고 YOLO 알고리즘보다 Faster R-CNN을 통해 높은 정확도와 빠른 속도로 실시간 탐지를 필요로 하는 실제 농업 환경에서의 적용 가능성을 평가하고자 한다.

REFERENCES

- [1] Z. Miao, X. Yu, N. Li, Z. Zhang, C. He, Z. Li, C. Deng, and T. Sun, "Efficient tomato harvesting robot based on image processing and deep learning," *Precision Agriculture*, vol. 24, no. 1, pp. 254–287, 2023.
- [2] S. A. Magalhães, L. Castro, G. Moreira, F. N. Dos Santos, M. Cunha, J. Dias, and A. P. Moreira, "Evaluating the single-shot multibox detector and yolo deep learning models for the detection of tomatoes in a greenhouse," *Sensors*, vol. 21, no. 10, p. 3569, 2021.
- [3] C. Mitaritonna and L. Ragot, "After covid-19, will seasonal migrant agricultural workers in europe be replaced by robots," *CEPII Policy Brief*, vol. 33, pp. 1–10, 2020.
- [4] M. K. Kim, S. M. Lim, and M. C. Won, "Seedling area recognition algorithm for autonomous rice planter based on deep learning and image processing," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 29, no. 3, pp. 245–251, 2023.
- [5] S. M. Lim, M. K. Kim, D. H. Lee, and M. C. Won, "Vision-based farmland boundary detection algorithm for automation of an agricultural tractor," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 29, no. 3, pp. 208–216, 2023.
- [6] O. M. Lawal, "Development of tomato detection model for robotic platform using deep learning," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26 751–26 772, 2021.
- [7] J. E. Kim, J. H. Seol, and H. I. Son, "Preliminary experimental results of a deep learning-based intelligent spraying system for pear orchard," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 26, no. 1, pp. 23–28, 2020.
- [8] G. Jocher, "Ultralytics yolov5," 2020. [Online]. Available: <https://github.com/ultralytics/yolov5>
- [9] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics yolov8," 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [10] J. Li, Y. Qiao, S. Liu, J. Zhang, Z. Yang, and M. Wang, "An improved yolov5-based vegetable disease detection method," *Computers and Electronics in Agriculture*, vol. 202, p. 107345, 2022.
- [11] A. Khan, T. Hassan, M. Shafay, I. Fahmy, N. Werghi, S. Mudigansalage, and I. Hussain, "Tomato maturity recognition with convolutional trans-formers," *Scientific Reports*, vol. 13, no. 1, p. 22885, 2023.
- [12] H.-T. Vo, K. C. Mui, N. N. Thien, and P. P. Tien, "Automating tomato ripeness classification and counting with YOLOv9," *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 4, 2024.
- [13] P. Li, J. Zheng, P. Li, H. Long, M. Li, and L. Gao, "Tomato maturity detection and counting model based on MHSA-YOLOv8," *Sensors*, vol. 23, no. 15, p. 6701, 2023.
- [14] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, "A lightweight yolov8 tomato detection algorithm combining feature enhancement and attention," *Agronomy*, vol. 13, no. 7, p. 1824, 2023.
- [15] B. Xiao, M. Nguyen, and W. Q. Yan, "Fruit ripeness identification using YOLOv8 model," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 28039–28056, 2024.
- [16] M. Afonso, H. Fonteijn, F. S. Fiorentin, D. Lensink, M. Mooij, N. Faber, G. Polder, and R. Wehrens, "Tomato fruit detection and counting in greenhouses using deep learning," *Frontiers in Plant Science*, vol. 11, p. 571299, 2020.
- [17] K. Bresilla, G. D. Perulli, A. Boini, B. Morandi, L. Corelli Grappadelli, and L. Manfrini, "Single-shot convolution neural networks for real-time fruit detection within the tree," *Frontiers in plant science*, vol. 10, p. 421226, 2019.
- [18] Y. Edan, S. Han, and N. Kondo, "Automation in agriculture," *Springer handbook of automation*, pp. 1095–1128, 2009.
- [19] S. Mao, Y. Li, Y. Ma, B. Zhang, J. Zhou, and K. Wang, "Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion," *Computers and Electronics in Agriculture*, vol. 170, p. 105254, 2020.
- [20] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of yolo algorithm developments," *Procedia Computer Science*, vol. 199, pp. 1066–1073, 2022.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [22] O. M. Lawal, "Development of tomato detection model for robotic platform using deep learning," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26751–26772, 2021.
- [23] S. A. Magalhães, L. Castro, G. Moreira, F. N. Dos Santos, M. Cunha, J. Dias, and A. P. Moreira, "Evaluating the single-shot multibox detector and YOLO deep learning models for the detection of tomatoes in a greenhouse," *Sensors*, vol. 21, no. 10 p. 3569, 2021.



양 다 연

2021년 공주대학교 정보통신공학과 학사 졸업. 2023년~현재 전남대학교 융합 바이오시스템기계공학과 석사 과정. 관심분야는 딥러닝, 모바일로봇 SLAM.



손 형 일

1998년 부산대학교 생산기계공학과(공학사). 2000년 부산대학교 지능기계공학과(공학석사). 2010년 KAIST 기계공학과(공학박사). 2003~2005년 LG전자 생산기술원 선임연구원. 2005~2009년 삼성 전자 ITD센터 책임연구원. 2010년 일본 동경대학교 생산기술연구소 박사후 연구원. 2010~2012년 독일 막스플랑크연구소 박사후 연구원. 2012~2015년 삼성중공업 중앙연구소 수석연구원. 2015~현재 전남대학교 융합바이오시스템기계공학과 교수. 관심분야는 field robotics, hybrid systems, systems and synthetic biology.



주 찬 영

2017년 전남대학교 지역·바이오시스템 공학과(공학사). 2019년 전남대학교 지역·바이오시스템공학과과(공학석사). 2022년 전남대학교 지역·바이오시스템공학과(공학박사). 2022~2022년 UC Berkeley 박사 후 연구원. 2022~현재 한국생산기술연구원 목적기반모빌리티그룹 선임연구원. 관심분야는 supervisory control, field robotics, hybrid systems, artificial intelligence.