

ENGINEERING

Estimation of tomato maturity as a continuous index using deep neural networks

Taehyeong Kim¹, Dae-Hyun Lee^{2*}, Seung-Woo Kang², Soo-Hyun Cho², Kyoung-Chul Kim³

¹Artificial Intelligence Laboratory, Chief Technology Officer Division, LG Electronics, Seoul 06763, Korea

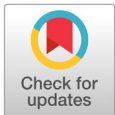
²Department of Biosystems Mechanical Engineering, Chungnam National University, Daejeon 34134, Korea

³Department of Agricultural Engineering, National Institute of Agricultural Sciences, Jeonju 54875, Korea

*Corresponding author: leedh7@cnu.ac.kr

Abstract

In this study, tomato maturity was estimated based on deep learning for a harvesting robot. Tomato images were obtained using a RGB camera installed on a monitoring robot, which was developed previously, and the samples were cropped to 128 × 128 size images to generate a dataset for training the classification model. The classification model was constructed based on convolutional neural networks, and the mean-variance loss was used to learn implicitly the distribution of the data features by class. In the test stage, the tomato maturity was estimated as a continuous index, which has a range of 0 to 1, by calculating the expected class value. The results show that the F1-score of the classification was approximately 0.94, and the performance was similar to that of a deep learning-based classification task in the agriculture field. In addition, it was possible to estimate the distribution in each maturity stage. From the results, it was found that our approach can not only classify the discrete maturation stages of the tomatoes but also can estimate the continuous maturity.



OPEN ACCESS

Citation: Kim T, Lee DH, Kang SW, Cho SH, Kim KC. Estimation of tomato maturity as a continuous index using deep neural networks. Korean Journal of Agricultural Science 49:785-793. <https://doi.org/10.7744/kjoas.20220072>

Received: September 07, 2022

Revised: November 02, 2022

Accepted: November 03, 2022

Copyright: © 2022 Korean Journal of Agricultural Science



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Key words: convolutional neural networks, deep learning, mean-variance loss, robot harvesting, tomato maturity

Introduction

최근 농업의 노동력 감소에 따라 생산량의 지속적인 확보를 위한 해결 방안이 시급히 요구되고 있으며(Baek et al., 2020), 기술적 대안으로 무인화 농업의 필요성이 대두되고 있다. 무인화 농업은 실시간 측정 기반으로 다양한 농작업을 자동 혹은 자율적으로 수행하는 기술집약적 농법이다. 특히 세계적인 인구 증가가 가속화되며 2050년 100억 명으로 예상됨에 따라 미래 지속적인 농업생산을 위해 무인화 농업에 대한 전세계적인 관심이 증대되고 있다(Kim et al., 2021).

농업용 로봇은 무인화 농업의 핵심기술 중 하나로써 다양한 작업 중 다른 산업의 로봇 기술의 활용도가 높은 수확작업을 중심으로 많은 관련 연구가 진행되고 있다. 수확로봇 관련 연구는 대부분 안정적인 주행 및 규격화된 재배환경을 가지는 시설온실을 대상으로 주로 수

행되고 있으며, 작업 대상체 인식, 접근 경로생성 및 추종, 장애물 회피, 대상물 분리 및 운반, 적재 등 복합적 기술로 구성된다. 인식은 로봇작업을 위해 가장 우선 수행되는 기술로 수확로봇의 경우 과일 등 작업 대상체의 위치, 자세 검출이 필요하며 주로 기계시각을 이용하여 기술개발이 수행되었다. 최근 스테레오 카메라, 딥스 카메라 등 3차원 영상측정 기술의 발전은 인식 성능의 향상을 가져다주고 있으며 특히, 딥러닝(deep learning) 기술의 등장으로 실질적 활용이 가능한 수준까지 기술 개발이 이루어지고 있다. 딥러닝과 같은 학습 기반 접근은 객체의 공간적 특징(spatial feature)의 자동추출이 가능하여 휴리스틱(heuristic) 기반 영상처리 등 기존 방식에 비해 높은 성능을 보여주고 있으며(Chun et al., 2018), 객체 분류의 경우 이미 인식율을 90% 이상까지 향상시켰다(Stefas et al., 2019).

딥러닝 기반 수확 대상체 인식은 다양한 작물을 대상으로 진행되고 있으며, 토마토(Benavides et al., 2020; Rong et al., 2021; Zhang et al., 2022), 파프리카(Li et al., 2018), 구아바(Lin et al., 2019) 등 시설원에 작물뿐만 아니라 사과(Fu et al., 2020; Kang et al., 2020), 리치(Li et al., 2020) 등 노지환경에서도 최근 연구들이 추진되고 있다. 특히, Benavides 등(2020)은 토마토를 대상으로 과일뿐만 아니라 수확을 위한 절단부인 꽃자루(peduncle)까지 검출하였으며, 여러 개의 과일의 중심위치와 꽃자루를 2차원 좌표로 87.5%의 정확도로 인식이 가능하였다. 또한, Zhang 등(2022)의 연구에서는 토마토 과일과 꽃자루를 동시에 3차원 좌표로 검출하였으며 최대 94%의 높은 정확도를 보여주었다.

수확로봇의 경우 대상 수확물(과일 등)의 위치, 자세뿐만 아니라 완숙 여부의 판단이 필요하다. 이를 위한 연구가 토마토의 성숙도 평가를 중심으로 수행되었으며(Seo et al., 2021; Begum and Hazarika, 2022), Seo 등(2021)의 연구에서는 후숙과인 토마토의 다양한 수확시기 결정을 위해 색상공간(color space) 분석을 수행하여 6단계(green, breaker, turning, pink, light red, red)의 성숙도 분류를 90% 이상의 정확도를 보여주었다. 하지만 성숙도의 경우 중간 단계에 대한 정의가 모호하며 지도학습의 경우 데이터 라벨 작업에 성능이 의존적이기 때문에 다양한 환경에 대해 보편적인 특징 표현을 얻기 어렵다. 따라서, 분류학습을 통해 정해진 부류로 한정하기보다 연속적인 값으로 평가할 수 있는 방법이 다양한 환경에 활용하기 효과적이다.

본 연구에서는 토마토의 실시간 성숙도 평가를 위해 수행되었으며, 합성신경망(convolutional neural networks, CNN)과 평균-분산 손실(mean-variance loss)을 이용하여 성숙도를 연속적인 값으로 평가하였다. 실제 재배환경에서 영상수집 및 학습데이터를 구축하였으며, 합성신경망의 분류학습을 수행한 후 추론 시 분류된 각 결과의 확률값의 기댓값을 이용하여 성숙도를 평가하였다.

Materials and Methods

토마토 영상 수집 및 학습데이터 구성

토마토 성숙도 평가를 위한 영상데이터는 국립농업과학원 농업공학부에 위치한 시설온실 테스트베드에서 2021년 6월경 수집하였으며 개발된 모니터링 로봇(Seo et al., 2021)을 이용하여 원격제어 및 영상을 저장하였다. 토마토 영상의 촬영을 위해 모니터링 로봇의 측면에 카메라(RealSense D435, Intel, CA, USA)가 설치되어 있으며, 영상은 30 fps 속도로 800 × 600 픽셀의 정밀도로 촬영 및 저장되었다. 학습데이터 구성을 위해 수집된 영상 내 토마토 과일 영역의 주석작업을 수행하였으며, 토마토 영역만을 가지는 샘플들을 확보하기 위해 Fig. 1과 같이 객체 중심의 bounding box 및 의미분할(semantic segmentation)을 위한 polygon 경계정보 주석작업을 수행하고 이를 이용하여 배경을 제거하였다. 이때, bounding box는 토마토 과일 영역을 포함하는 최소크기의 사각형이다. 배경이 제거된 샘플들은 학습데이터로 사용하기 위해 모델의 입력크기인 128 × 128 픽셀 크기로 해상도를 변경하였으며, 성숙도 4단계(green, turning, pink, red)에 대한 부류 라벨링을 수행하였다. 이때, 성숙도 일반적으로 6단계로 분류되나

(Seo et al., 2021), 육안으로 붉은색의 점유율을 판단하기 쉽지 않으며, 본 연구에서는 분류 정확성이 아닌 분류학습을 통해 각각의 부류 내 분포를 고려하여 연속적인 지표로 최종 평가한다. 따라서, 영상의 성숙도는 육안으로 쉽게 분류 가능한 4단계로 축소하여 진행하였다.

전체 학습데이터는 총 1,888장이며 성숙도별 샘플은 Table 1과 같이 green, turning, pink, red 단계 각각 374, 124, 402, 988장으로 구성되었다. 이때, 전체데이터는 학습, 검증, 시험데이터로 분류하였으며, 각각 472, 472, 944장이다.

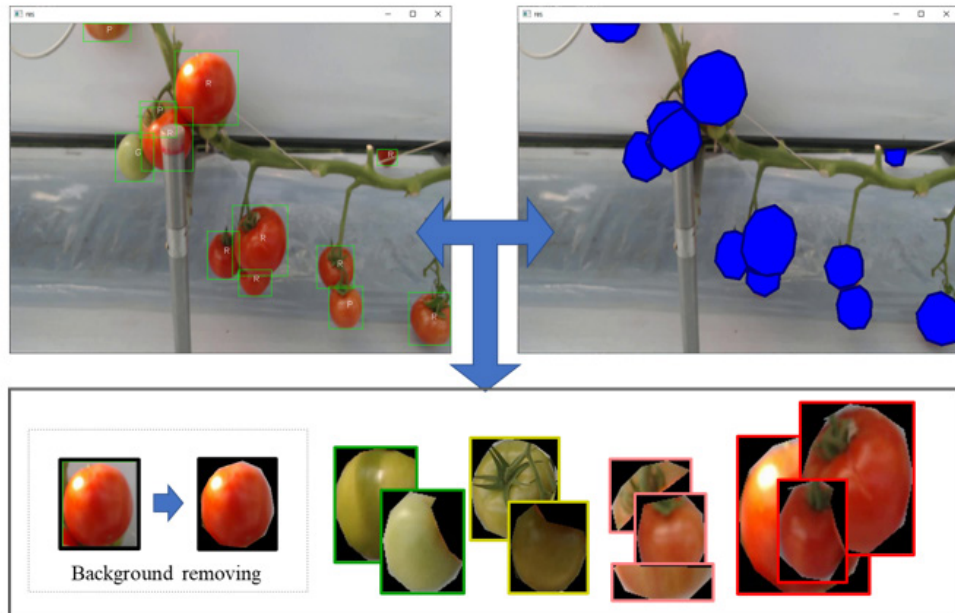


Fig. 1. Tomato annotation and background removal processes for dataset construction.

Table 1. Dataset composition with number of samples by maturity stage.

Type	Green	Turning	Pink	Red	Total
Train	94	31	101	247	472
Validation	93	31	100	247	472
Test	187	62	201	494	944
Total	374	124	402	988	1,888

심층신경망을 이용한 토마토 성숙도 평가 모델

본 연구에서는 입력된 토마토 영상의 성숙도 분류를 기반으로 연속적인 성숙도 값을 평가하였다. 성숙도 분류 학습은 심층신경망(deep neural networks) 모델을 이용하였으며 성숙도는 미숙단계(green)에서 완숙단계(red)의 성숙도를 0 - 1 사이의 연속적인 값으로 표현하였다.

Fig. 2는 성숙도 학습을 위한 모델 구성을 보여주고 있으며, 입력된 영상은 심층신경망을 통해 특징표현이 된 후 이를 이용한 성숙도 부류 선정 후 주어진 라벨과 비교 및 손실을 계산하여 모델 가중치를 갱신하게 된다. 심층신경망의 특징 추출은 합성곱신경망(convolutional neural networks, CNN)을 통해 수행되며, 본 연구에서 사용된 합성곱신경망 구조는 Fig. 3과 같이 4개의 합성곱신경망 층(layer), 1개의 완전연결망(fully connected network) 층과 softmax 분류기로 구성되었다. 입력영상의 크기는 $128 \times 128 \times 3$ (width \times height \times channel) 픽셀이며, 합성곱신경망의 필터 크기는 3×3 으로 전체 층에서 동일하고, 활성화 함수로는 ReLU (rectified linear unit)을 이용하였다. 입력된 영상의 분류 결과는 softmax 분류기에서 4개의 부류에 대한 확률값으로 출력되며, 이 중 가장 높은 확률값을 가진 부류로 분류된다.

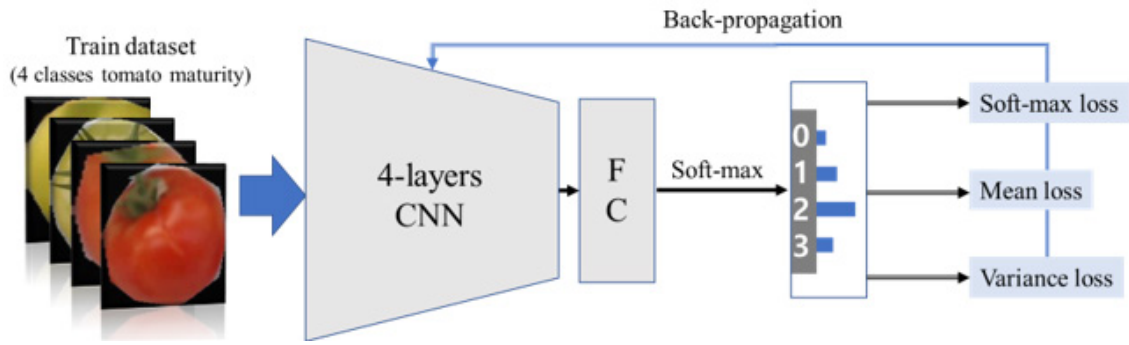


Fig. 2. Configuration of the model training with mean-variance loss for estimating the continuous tomato maturity. CNN, convolutional neural network; FC, fully connected.

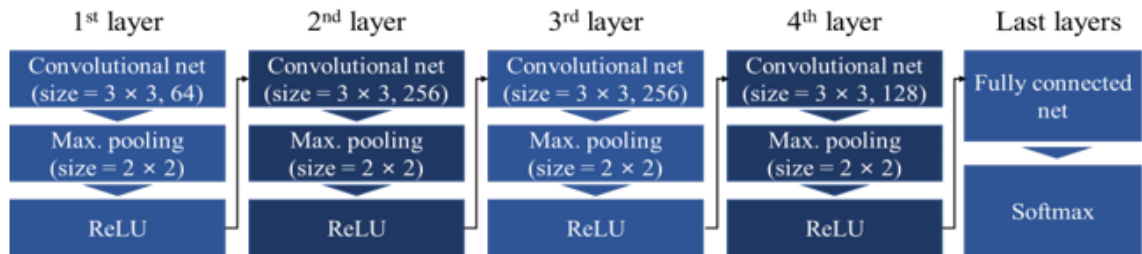


Fig. 3. Convolutional neural networks for classification of tomato maturity. ReLU, rectified linear unit.

분류 학습의 결과만을 이용하여 토마토 성숙도를 평가하면 4단계 중 하나만 선택되게 되고, 이는 중간 단계의 다양한 성숙 단계에 대한 반영이 어렵다. 따라서, 분류학습이라도 각 단계 내 차이에 대한 반영이 가능하고 이를 통해 연속적인 성숙도로 표현하는 것이 필요하다. 본 연구에서는 분류학습 시 기본적인 손실인 소프트맥스 손실(softmax loss) 외 식(1), (2)와 같은 평균-분산 손실(mean-variance loss)을 이용하였으며(Pan et al., 2018), 평균-분산 손실은 토마토 성숙도와 같이 부류의 선정이 모호하여 부정확한 라벨을 가지는 경우 부류의 확률적 분포(distribution) 학습 및 내재적인 모호성(ambiguity)의 반영이 가능한 장점이 있다. 학습 후 토마토 영상에 대한 성숙도 최종 추론은 식(3)과 같이 소프트맥스 분류기를 통해 출력되는 확률분포와 클래스 라벨과의 내적을 이용하여 계산된 값을 정규화하여 0에서 1 사이의 값을 가지게 한다. 이때, 정규화는 클래스가 0 - 3번의 정수값을 가지므로 최댓값인 3을 나눠준다.

$$L_m = \frac{1}{2N} \sum_{i=1}^N (m_i - y_i)^2 \quad (1)$$

$$L_v = \frac{1}{N} \sum_{i=1}^N v_i \quad (2)$$

Where, L_m is mean loss, N is the batch size, m_i is the mean of i^{th} sample, y_i is the label of i^{th} sample, L_v is variance loss, and v_i is the variance of i^{th} sample.

$$\text{Maturity} = \frac{1}{K-1} \sum_{j=0}^{K-1} j \cdot p_j \quad (3)$$

Where, K is number of the classes, j is class number from zero, and p_j is the probability of j^{th} class in softmax output.

학습방법 및 구현도구

토마토 성숙도 분류학습은 학습 및 검증데이터 각각 472장을 이용해서 300회 반복 실시하였다. 학습을 위한 파라미터(hyper-parameters)로 학습율은 데이터 규모 및 모델 크기를 고려하여 0.001로 설정하였으며, 배치 크기(batch size)는 64, 최적화기는 수렴 속도가 빠른 Adam optimizer를 이용하였다(Kim et al., 2020). 데이터의 다양성 확보를 위해 성숙도에 영향을 미치는 색상과 관련된 항목은 제외하고 반전(수직, 수평) 및 스케일 변환 등의 데이터 증강(data augmentation)을 수행하였다. 학습모델의 가중치 선정은 과적합(overfitting) 효과를 최소화하기 위해 검증(validation) 데이터의 분류 성능이 가장 좋을 때 가중치를 최종 모델의 값으로 선정하였다(Kim et al., 2021).

모델은 GPU (GeForce 1080ti, Nvidia, Santa Clara, CA, USA) 기반의 시스템으로 학습시켰으며, 학습모델 개발 및 기능 구현을 위한 프로그램은 python 3.7 및 pytorch 1.0을 이용하였다.

Results and Discussion

분류학습 결과

성숙도 분류 학습 결과 반복학습에 따른 전체손실(loss)과 분류 정확도(accuracy)는 Fig. 4와 같이 관찰되었다. 손실의 경우 소프트맥스 손실과 평균-분산 손실이 모두 포함된 총 손실로 표현되었으며, 학습데이터와 검증데이터 모두 유사한 경향을 보였다. 학습 및 검증데이터 모두 20회 반복시까지 손실이 급격히 감소하여 50회 이후 약 0.1에서 포화되었는데, 학습데이터의 경우 지속적으로 감소를 보인 반면, 검증데이터는 100회 이후 증가와 감소를 반복하며 향상된 결과가 관찰되지 않았다. 따라서 100회 반복 학습 시 모델의 가중치를 본 연구의 최종 분류 모델 가중치로 사용하였다.

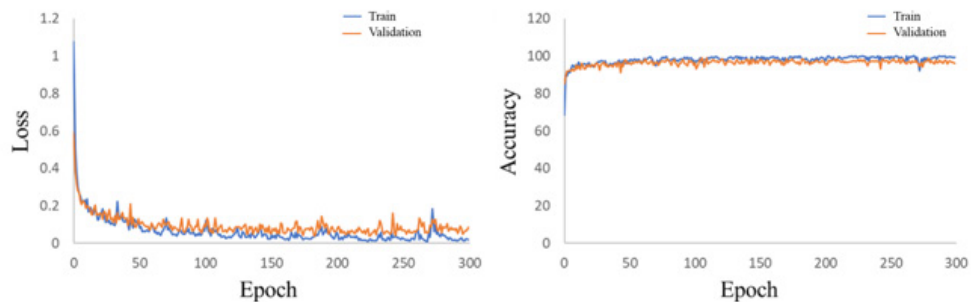


Fig. 4. Total loss and accuracy for tomato maturity classification by epoch.

학습된 합성곱신경망 모델을 이용한 시험데이터의 분류 성능은 일반적인 성능 지표인 재현율(recall), 정밀도(precision) 및 F1-score를 이용하여 Table 2와 같이 부류별로 계산되었다. 딥러닝 기반 농업 영상 내 객체 분류 및 검출 연구에서 F1-score가 약 0.92 수준임을 고려할 때(Kamilaris et al., 2017), 본 연구의 4개 층의 합성곱신경망이 기존 연구에 비해 비교적 가벼운 구조를 가짐에도 불구하고 유사한 성능을 보여줌을 알 수 있다. 부류별로 성능은 green, red의 경우 95% 수준으로 높게 나타난 반면 turning, pink의 경우 85 - 87%로 다른 부류에 비해 0.1 정도 낮게 관찰되었다. 이는 4단계로 분류할 경우 상대적으로 명확한 색상구성을 가지는 green, red에 비해 turning, pink는 모호한 색상구성을 가져 샘플 간의 차이가 크며 이로 인해 강인한 특징 표현이 어렵기 때문이다.

Table 2. Classification performance of tomato maturity estimation model.

Maturity stage (class)	Recall	Precision	F1-score
Green	0.98	0.91	0.94
Turning	0.77	0.94	0.85
Pink	0.83	0.92	0.87
Red	0.98	0.95	0.97
Total	0.94	0.94	0.94

토마토 성숙도 평가

본 연구에서는 평균-분산 손실을 이용한 데이터 분포를 내재적으로 학습하였으며, 분류 학습 가중치를 이용하여 최종적으로 입력영상에 대한 토마토 성숙도를 0-1 사이의 값으로 평가하였다.

Fig. 5는 4단계의 성숙도 분류와 본 연구에서 제안된 성숙도 평가를 비교한 그림으로 왼쪽의 경우 일반적인 분류 학습의 결과로 성숙도를 4개의 단계로 분류한 경우이며, 오른쪽 그림은 학습결과를 이용하여 성숙도를 0-1 사이의 값으로 표현한 것이다. 검정색 테두리의 원은 올바르게 검출된 샘플, 주황색 원은 잘못 검출된 샘플을 의미하며, 분류 결과에서는 전체 성숙도 영역에 분포를 표현하지 못하고 등분된 값에 각각 맵핑되어 연속적인 색상의 분포를 가지는 성숙도의 평가가 어려움을 알 수 있다. 이에 반면, 본 연구에서의 평균-분산 손실 기반의 분포 학습 결과를 활용했을 때, 오른쪽 그림과 같이 토마토 색상의 정도 및 영역에 따라 동일한 부류 내에서도 분포의 표현이 가능함을 알 수 있다. 성숙 단계별로 분포를 보면 중간 성숙 단계인 turning, pink 단계에서 전반적으로 넓은 범위를 가지고 있어 명확한 분류가 어려웠으며 green, red 단계는 상대적으로 밀집된 영역을 보여주고 있어 이는 위에서 언급한 분류결과와 일치함을 알 수 있다. 또한, 올바르게 검출되지 않은 샘플의 경우 학습모델의 성능저하 뿐만 아니라 데이터 라벨 작업 시 작업자의 시각적 정보와 경험에 크게 의존될 수 있는 것을 알 수 있으며, 중간 단계의 경우 이러한 샘플이 다른 영역과의 경계에 많이 분포되어 있어 일반적인 객체분류가 아닌 본 연구와 같은 분포에 대한 고려가 적합한 것으로 판단된다. 하지만 본 연구에서 제시된 결과는 토마토 성숙 단계 구분의 모호함과 관련해 이를 연속적인 값으로 표현할 수 있는 학습방법을 보여주고 있지만 실제 성숙도에 대한 측정값 부재로 성능에 대한 정확한 검증이 어려우며 분포가 고려된다고 해도 수동적인 성숙도 학습이 이루어진다는 점에서 활용의 한계점을 가진다. 이러한 한계점에도 불구하고 4개의 분류학습을 통해 데이터들이 가지고 있는 분포를 내재적으로 학습하여 연속적인 지표화가 가능하다는 점에서 수확로봇 연구에 기여가 가능하다고 판단되며, 데이터 수집 시 현장의 정량적 성숙도 평가와 이를 전문적으로 분류할 수 있는 기술 개발을 통해 본 연구의 결과를 실용적으로 확장 가능할 것으로 판단된다. Fig. 6은 현장에서 수집된 원 영상을 이용하여 본 연구의 성숙도를 평가한 대표적인 결과로 다양한 성숙 단계를 이산적으로 분류하는 것이 아닌 연속적인 값으로 평가 가능하며, 정량적인 검증은 어렵지만 직관적으로 활용 가능성을 보여줌을 알 수 있다.

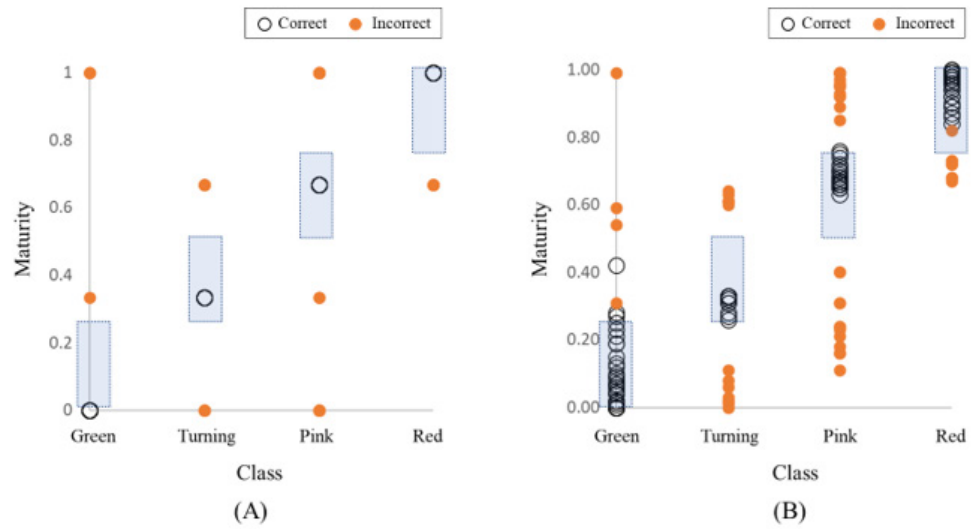


Fig. 5. Tomato maturity estimation by classification (A) and distribution (B).

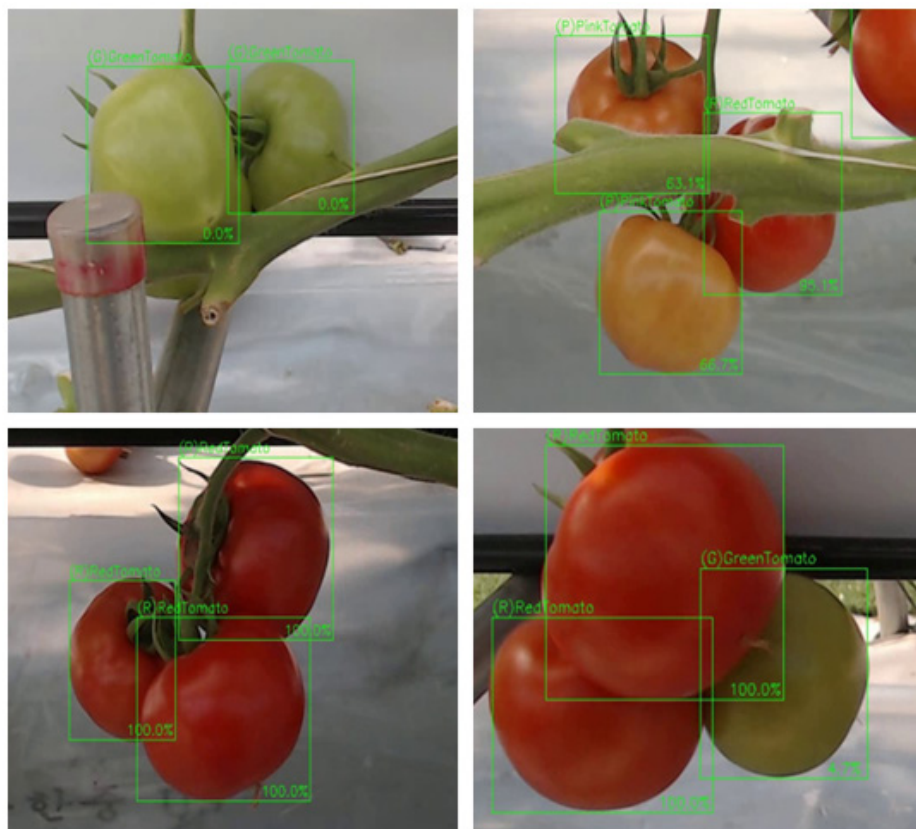


Fig. 6. Representative tomato maturity estimation using field images with background.

Conclusion

본 연구에서는 시설 토마토 수확로봇 개발을 위한 기초연구로서 딥러닝 기술을 이용하여 토마토의 성숙도를 평가하였으며, 부류별 분포 학습을 통해 성숙도를 연속적인 값으로 계산할 수 있는 방법을 제안하였다. 현장 토마토 영상을 수집 및 학습을 위한 데이터 세트를 구축하였으며, 4단계의 성숙도 분류학습을 위해 심층신경망을 이용한 학습모델을 설계하였다. 연속적인 특징을 가지는 토마토 성숙도를 평균-분산 손실을 이용하여 분류학습 시 내재적으로 부류별 분포의 학습을 가능하게 하였으며, 학습된 모델을 이용하여 최종 성숙도 평가 시 출력되는 부류별 확률값을 이용하여 라벨의 기댓값을 계산하였으며, 정규화를 통해 0 - 1 사이의 성숙도로 변환하였다. 분류학습의 결과 94%의 토마토 성숙도 분류 정확도를 보여주었으며, 성숙 단계별로는 green, red 단계에서는 90% 이상의 높은 성능을 보여주었지만 중간 단계의 경우 85% 수준으로 다소 낮게 관찰되었다. 평균-분산 손실 기반의 성숙도 평가 결과, 토마토 색상의 정도 및 영역에 따라 동일한 부류 내에서도 분포의 표현이 가능하였으며 중간 단계의 경우 샘플이 다른 영역과의 경계에 많이 분포되어 있어 성숙도 평가의 경우 일반적인 객체분류가 아닌 분포에 대한 고려가 적합한 것으로 판단된다. 본 연구의 결과를 통해 검출된 토마토의 이산적인 성숙 단계뿐만 아니라 연속적인 성숙도 평가도 가능한 것을 알 수 있었다. 실제 수확로봇 적용을 위해, 데이터 수집 시 정확한 성숙도의 측정 및 조도 정보와의 융합을 통해 평가된 연속적인 성숙도 자체의 신뢰성을 향상하고 토마토 개체별 영역 검출 기술과 융합 등 강인한 알고리즘으로 고도화를 여러 개체의 겹침 등으로 인해 발생하는 한 성숙도 오차 해결이 필요할 것으로 판단된다.

Conflict of Interests

No potential conflict of interest relevant to this article was reported.

Acknowledgements

본 결과물은 농림축산식품부 및 과학기술정보통신부, 농촌진흥청의 재원으로 농림식품기술기획평가원과 재단법인 스마트팜연구개발사업단의 스마트팜다부처 패키지혁신기술개발사업의 지원을 받아 연구되었으며(421031-04), 2022년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과임(2021RIS-004).

Authors Information

Taehyeong Kim, <https://orcid.org/0000-0002-9608-4362>

Dae-Hyun Lee, <https://orcid.org/0000-0001-9544-5974>

Seung-Woo Kang, <https://orcid.org/0000-0001-6318-2006>

Soo-Hyun Cho, <https://orcid.org/0000-0003-0553-7440>

Kyoung-Chul Kim, <https://orcid.org/0000-0001-6699-881X>

References

- Baek SM, Kim WS, Kim YJ, Chung SO, Nam KC, Lee DH. 2020. Development of a real-time crop recognition system using a stereo camera. *Korean Journal of Agricultural Science* 47:315-326. [in Korean]
- Begum N, Hazarika MK. 2022. Maturity detection of tomatoes using transfer learning. *Measurement: Food* 7:100038.
- Benavides M, Canton-Garbin M, Sanchez-Molina JA, Rodriguez F. 2020. Automatic tomato and peduncle location system based on computer vision for use in robotized harvesting. *Applied Sciences* 10:5887.
- Chun CJ, Shim SB, Kang SM, Ryu SK. 2018. Development of evaluation of automatic pothole detection using fully convolutional neural networks. *The Journal of the Korea Institute of Intelligent Transport Systems* 17:55-64. [in Korean]
- Fu L, Majeed Y, Zhang X, Karkee M, Zhang Q. 2020. Faster R-CNN-based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. *Biosystems Engineering* 197:245-256.
- Kamilaris A, Kartakoullis A, Prenafeta-Boldú FX. 2017. A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture* 143:23-37.
- Kang H, Zhou H, Wang X, Chen C. 2020. Real-time fruit recognition and grasping estimation for robotic apple harvesting. *Sensors* 20:5670.
- Kim G, Seo D, Kim KC, Hong Y, Lee M, Lee S, Kim H, Ryu H, Kim YJ, Chung SO, et al. 2020. Tillage boundary detection based on RGB imagery classification for an autonomous tractor. *Korean Journal of Agricultural Science* 47:205-217. [in Korean]
- Kim WS, Lee DH, Kim T, Kim G, Kim H, Sim T, Kim YJ. 2021. One-shot classification-based tilled soil region segmentation for boundary guidance in autonomous tillage. *Computers and Electronics in Agriculture* 189:106371.
- Li H, Zhu Q, Huang M, Gui Y, Qin J. 2018. Pose estimation of sweet pepper through symmetry axis detection. *Sensors* 18:3083.
- Li J, Tang Y, Zou X, Lin G, Wang H. 2020. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots. *IEEE Access* 8:117746-117758.
- Lin G, Tang Y, Zou X, Xiong J, Li J. 2019. Guava detection and pose estimation using a low-cost RGB-D sensor in the field. *Sensors* 19:428.
- Pan H, Han H, Shan S, Chen X. 2018. Mean-variance loss for deep age estimation from a face. pp. 5285-5294. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
- Rong J, Dai G, Wang P. 2021. A peduncle detection method of tomato for autonomous harvesting. *Complex & Intelligent Systems* 8:2955-2969.
- Seo D, Cho B, Kim KC. 2021. Development of monitoring robot system for tomato fruits in hydroponic greenhouses. *Agronomy* 11:2211.
- Stefas N, Bayram H, Isler V. 2019. Vision-based monitoring of orchards with UAVs. *Computers and Electronics in Agriculture* 163:104814.
- Zhang F, Gao J, Zhou H, Zhang J, Zou K, Yuan T. 2022. Three-dimensional pose detection method based on keypoints detection network for tomato bunch. *Computers and Electronics in Agriculture* 195:106824.