Check for updates

Journal of the Korean Institute of Industrial Engineers
Vol. 52, No. 5, pp. 419-431, October 2026.
ISSN 1225-0988 | EISSN 2234-6457
https://doi.org/10.7232/JKIIE.2026.52.5.419
© 2026 KIIE

밀집 재배 환경 퓨샷 학습 기반 작물 개체 분할을 위한 방사형 손실함수

이혜승 · 김성수 · 김성범†

고려대학교 산업경영공학과

Radial Loss for Few-Shot Instance Segmentation of Densely Cultivated Crops

Hyeseung Lee · Sungsu Kim · Seoung Bum Kim

School of Industrial and Management Engineering, Korea University

Instance segmentation is a key technology for precision agriculture tasks such as per-plant growth monitoring and harvest-timing prediction. However, in real smart-farm environments where crops are densely cultivated, instance segmentation poses two practical challenges. First, because pixel-level annotation is highly costly, only a small number of labeled images are typically available. Second, severe leaf overlap between adjacent plants causes boundary ambiguity that degrades segmentation accuracy. To improve segmentation performance, prior studies have explored incorporating structural priors of the target object into the learning process. However, existing methods require redesigned activation functions or separate optimization modules, making them hard to integrate with foundation models and largely limited to single-object semantic segmentation. To address these limitations, we propose ‘radial loss,’ an auxiliary loss function that explicitly captures the radial convex structure of rosette-shaped crops by supervising predicted foreground probabilities along rays cast from each instance's center of mass toward its boundary. We further introduce ‘dynamic search,’ a threshold-triggered sampling mechanism that focuses supervision on boundary regions once the validation mAP exceeds a predefined threshold. Experiments on three lettuce datasets show that, under an extreme 10-shot setting, the proposed method improves mean mAP by 7.7% points over the baseline, demonstrating its effectiveness in label-scarce, densely cultivated environments.

Keywords: Instance Segmentation, Few-Shot Learning, Precision Agriculture, Deep Learning

1. 서 론

최근 인공지능(artificial intelligence, AI) 기술이 발전하면서, 스마트팜(smart farm) 환경에서 자동화된 작물 모니터링 시스템 도입이 빠르게 확대되고 있다(Kumar *et al.*, 2024; Li *et al.*, 2023). 특히 컴퓨터 비전 기반 영상 분석 기술은 작물의 생육 상태 파악, 수확량 예측, 병해 감지 등 다양한 목적으로 활용되며, 기존 수작업 기반 점검 방식을 대체하는 핵심 요소로 자리잡고 있다. 그중에서도 동일한 재배 공간 안에서 개별 작물 단위로 생육 상태를 분석하기 위해서는, 한 장의 이미지 안에서 개체를 픽셀 단위로 구별하는 개체 분할(instance segmentation) 기술이 요구된다(Weyler *et al.*, 2022). 개체 분할은 개별 작물의 무게 예측, 수확 시기 예측, 병해 발생 여부 판단 등 정밀 재배(precision agriculture) 응용 과제의 기반 기술로 사용되므로, 스마트팜 환경에서 활용도가 매우 높다. 그러나 실제 스마트팜 이미지에 개체 분할 모델을 적용하기 위해서는 두 가지 현

† 연락저자 : 김성범 교수, 02841 서울특별시 성북구 안암로 145 고려대학교 산업경영공학과, Tel : 02-3290-3397, Fax : 02-929-5888, E-mail : sbkim1@korea.ac.kr
2026년 5월 11일 접수; 2026년 8월 4일, 2026년 9월 7일 수정본 접수; 2026년 9월 8일 게재 확정.

실적인 어려움이 존재한다.

첫째, 픽셀 단위 레이블링(pixel-level labeling)에 막대한 시간과 비용이 든다. 작물 이미지와 같이 개체 수가 많고 경계가 복잡한 경우 이 어려움은 더욱 커지며, 실제 현장에서는 소수의 레이블 데이터만 확보 가능한 경우가 일반적이다. 이러한 데이터 부족 문제를 완화하기 위해, 최근에는 대규모 데이터로 사전학습된 파운데이션 모델(foundation model)을 활용하는 연구가 활발히 이루어지고 있다. 대표적으로 segment anything model(SAM)은 다양한 분할 과제에서 우수한 제로샷(zero-shot) 성능을 보였으며, 경량화된 fast segment anything(FastSAM)은 실시간 처리가 필요한 환경에서 실용적인 대안으로 활용되고 있다(Kirillov *et al.*, 2023; Zhao *et al.*, 2023). 그러나 이러한 파운데이션 모델은 일반 도메인 이미지를 기반으로 학습되었기 때문에, 스마트팜 이미지처럼 특수 도메인 데이터에 직접 적용할 경우 분할 성능이 급격히 저하된다(Chen *et al.*, 2023; Ma *et al.*, 2024). 이에 따라 소수의 레이블 데이터로 파운데이션 모델을 미세조정(fine-tuning)하는 방식이 활용되나, 학습 데이터가 극단적으로 부족한 상황에서는 모델이 도메인 지식을 충분히 학습하기 어려워 경계 모호성(boundary ambiguity) 문제가 여전히 남는다. 따라서 제한된 레이블 환경에서도 안정적으로 동작하는 퓨샷 학습(few-shot learning) 기반 개체 분할 방법론 개발이 요구된다.

둘째, 제한된 재배 공간에서 수확량을 극대화하기 위해 작물을 고밀도로 재배하는 경우가 많으며, 이로 인해 인접한 작물의 잎이 심하게 중첩되어 개체 간 경계가 불분명해진다(Buxbaum *et al.*, 2022). 이러한 경계 모호성은 모델의 분할 성능을 떨어뜨리는 원인이 된다. 이와 같은 경계 영역에서의 분할 성능을 안정적으로 확보하기 위한 방안으로, 대상 객체의 형태적 사전 정보(structural prior)를 손실 함수에 반영하는 접근법이 연구되어 왔다. 대표적으로 Zhao *et al.*(2025)는 다중 중심 별 모양 사전(multi-center star shape prior)을 변분 이론(variational theory)에 기반하여 신경망에 통합한 convex combination star shape prior를 제안하였다. 이 방법은 객체가 하나의 중심에서 방사형으로 볼록하게 퍼지는 구조를 가진다는 가정을 완화하여 복잡한 형태를 가지는 객체에도 적용 가능하도록 확장한 것이다. 그러나 이러한 연구들은 활성화 함수 재설계나 별도의 변분 최적화(variational optimization) 모듈을 요구하므로 기존 파운데이션 모델과 결합하기 위한 구조적 수정 비용이 크며, 대부분 단일 객체 대상의 의미 분할(semantic segmentation)을 전제로 하여 다수의 개체가 중첩된 밀집 작물 환경에 직접 적용하기 어렵다. 이처럼 기존 방법론들은 밀집된 작물 환경 특성을 충분히 반영하지 못하여, 인접한 상추 잎들이 겹쳐지는 경계면에서 분할 성능이 저하되는 문제가 발생한다.

이러한 한계를 해결하기 위해, 본 연구에서는 경량 파운데이션 모델인 FastSAM을 소수의 레이블 데이터로 미세조정하기 위한 방사형 손실 함수 radial loss를 제안한다. Radial loss는 상추가 식물학적으로 짧은 줄기에서 다수의 잎이 밀집하여 방사형으로 배열된 로제트(rosette) 형태를 가진다는 도메인 고유 특성에 기반한다. 먼저 정답 분할 마스크에서 무게중심을 계산하고, 해당 중심으로부터 여러 각도로 발사된 광선(ray)을 따라 객체 경계까지 도달하는 탐색을 수행한다. 이후 각 광선 위에서 샘플링된 지점에 대해 예측 값이 전경으로 일관되게 유지되도록 손실을 계산하여, 작물의 방사형 구조를 손실 함수에 명시적으로 반영한다. 나아가 학습이 진행되어 객체 중심부 예측이 충분히 수렴한 시점부터는 감독 신호를 경계 영역에 집중시키기 위해, 검증 mAP가 사전에 정의된 임계값을 초과하는 시점을 기준으로 광선의 탐색 범위를 전환하는 dynamic search 기법을 도입한다. Dynamic search는 검증 mean average precision (mAP)이 사전에 설정된 임계값을 초과하면 자동으로 활성화되며, 학습 후반부의 감독 신호를 경계 영역에 집중시킴으로써 밀집 환경에서의 경계 모호성 문제를 효과적으로 완화한다. 따라서 본 연구는 농업 현장에서 소수의 레이블 데이터만으로도 개별 작물을 정밀하게 구분할 수 있는 실용적인 방법론을 제시하며, 스마트팜 내 작물 생육 모니터링, 수확 시기 판단, 이상 개체 탐지 등 다양한 정밀 농업 응용 분야에 기여할 수 있을 것으로 기대된다(Lee *et al.*, 2026). 본 연구의 주요 기여점은 다음과 같다.

- 레이블 데이터 확보가 어려운 스마트팜 환경에서도 효과적인 개체 분할이 가능하도록, 경량 파운데이션 모델인 FastSAM을 기반으로 한 퓨샷 개체 분할 손실 함수 radial loss를 제안한다.
- 학습이 진행됨에 따라 작물 내부 영역 예측은 빠르게 수렴하는 반면 예측 오류는 점차 경계면에 집중되는 현상을 관찰하고, 이를 학습 과정에서 자동으로 감지하여 감독 신호를 경계 영역에 집중시키는 dynamic search 기법을 도입함으로써 경계 정밀도(boundary precision)를 효과적으로 향상시킨다.
- 실제 스마트팜 환경에서 수집된 상추 3품종 데이터셋을 활용하여 제안 방법론의 효과를 실증적으로 검증하였다. 레이블 데이터가 10장으로 제한된 극단적 환경에서 기존 star shape prior 기반 baseline 대비 평균 mAP를 7.7%p 향상시켜, 제한된 레이블 조건에서 제안 방법론의 효용성을 입증하였다.

2. 관련 연구

2.1 작물 이미지 분할

작물 이미지 분할은 스마트팜 데이터 분석 연구에서 오랫동안 다루어져 왔으며, 초기에는 색 인덱스(color index)나 임계값 기반 규칙 방법으로 식물과 배경을 구분하는 연구가 주로 이루어졌다. 예를 들어 Riehle *et al.*(2020)은 HSV 및 CIELab 색 공간에서 계산된 색 인덱스와 Otsu 임계값을 결합하여 다양한

촬영 조건에서도 강건하게 작물과 배경을 분할하는 방법을 제시하였다. 이후 GrabCut과 같이 색 분포와 에너지 함수를 이용한 규칙 기반 분할 방법이 주로 활용되었다(Kim *et al.*, 2024). 규칙 기반 방법은 구현이 간단하다는 장점이 있으나, 조명 변화나 배경 잡음이 있는 환경에서 성능이 급격히 저하되는 한계가 있다. 이러한 한계를 극복하기 위해 딥러닝 기반 의미 분할 모델이 도입되었다. Chang *et al.*(2021)은 상추 분할에 U-Net 구조를 적용하였고, Jiang *et al.*(2024)은 DeepLabv3를 활용하여 픽셀 단위 정확도를 크게 향상시켰다. 그러나 의미 분할은 각 픽셀에 하나의 클래스 레이블만 부여하므로, 같은 품종 내 개별 작물을 구분하지 못한다는 근본적인 한계가 있다. 이러한 특성은 개별 작물 단위 생육 분석이나 수확 시기 예측이 요구되는 스마트팜 응용에 부적합하다.

개별 작물 단위 분석을 위해, 개체 분할(instance segmentation) 기반 접근이 연구되어 왔다. Islam *et al.*(2024)은 Mask R-CNN을 상추 이미지에 적용하여 개별 작물 단위 분할을 시도하였다. 그러나 Mask R-CNN과 같은 전통적인 개체 분할 모델은 대량의 픽셀 단위 레이블 데이터를 필요로 하며, 경계가 모호한 밀집 재배 환경에서는 충분한 품질의 레이블을 확보하기 어렵다(He *et al.*, 2017). 이러한 데이터 의존성 문제를 완화하기 위한 방향으로, 작물이라는 특정 도메인에 국한되지 않는 일반적인 few-shot 개체 분할 연구도 활발히 이루어져 왔다. Yan *et al.*(2020)은 메타러닝(meta-learning) 프레임워크를 도입한 Meta R-CNN을 제안하여, 소수의 support 이미지로부터 클래스별 표현을 학습함으로써 새로운 클래스에 대한 개체 분할 및 탐지를 가능하게 하였다. Fan *et al.*(2020)은 이를 확장한 FGN을 제안하여, support 특징을 활용해 탐지 및 분할의 각 단계를 안내(guide)하는 방식으로 few-shot 개체 분할 성능을 향상시켰다. 그러나 이러한 방법론들은 support-query 쌍으로 구성된 episodic 학습을 통해 새로운 클래스 일반화 능력을 확보하는 데 초점을 맞추고 있어, 대상 객체의 형태적 사전 정보를 학습 과정에 명시적으로 반영하지 않는다. 이로 인해 상추와 같이 뚜렷한 기하학적 구조(로제트 형태)를 가지는 도메인에서는 이러한 구조적 특성을 충분히 활용하지 못한다는 한계가 있다.

한편, 대량의 레이블 데이터 없이 도메인 지식을 확보하기 위한 또 다른 방안으로, 대규모 데이터로 사전학습된 파운데이션 모델을 활용하는 접근이 제안되었다. Bao *et al.*(2025)은 Grounding DINO와 SAM을 결합하고 식생(vegetation) 특화 prompt refinement 전략을 도입하여, 별도의 레이블링 없이 zero-shot 개체 분할을 수행하는 프레임워크를 제시하였다. 그러나 이 방법은 넓은 간격으로 재배된 작물을 주 대상으로 하였기 때문에, 밀집 환경에서 흔히 발생하는 심각한 가림 현상과 불분명한 개체 경계 문제를 효과적으로 해결하지 못한다. 이에 따라 최근 연구는 파운데이션 모델의 도메인 적응(domain adaptation)에 주목하고 있다. Lee *et al.*(2026)은 FastSAM을 상추 이미지로 미세조정(fine-tuning)하여 밀집 재배 환경에서의 개체 분할 성능을 개선하는 FLAsH 프레임워크를 제안하였다. 이는 도메인 특화 데이터로 파운데이션 모델을 적응시킴으로써 분할 성능을 향상시킬 수 있음을 보였다. 그러나 여전히 소수의 레이블만 확보 가능한 조건에서는 인접 작물 간 경계 영역에서 분할 오류가 집중적으로 발생하는 한계가 남아 있다. 본 연구는 이러한 한계를 극복하기 위해 작물의 구조적 특성을 손실 함수에 직접 반영함으로써, 제한된 레이블 조건에서도 안정적인 밀집 작물 개체 분할 성능을 달성할 수 있는 방법론을 제안한다.

2.2 구조적 사전 정보 기반 이미지 분할

대상 객체의 형태적 사전 정보(structural prior)를 개체 분할 학습에 반영하는 연구는 그래프 컷(graph cut) 기반 분할 방법에서 처음 도입되었다. Veksler(2008)는 객체가 중심점으로부터 모든 방향으로 볼록하게 퍼져 있는 별 모양(star shape) 구조를 가진다는 가정 아래 이 제약을 그래프 에너지 함수에 통합한 star shape prior를 제시하였다. 이후 Gulshan *et al.*(2010)은 단일 중심 기반 star shape 제약이 지나치게 엄격하다는 한계를 보완하고자, 복수의 중심점과 측지선(geodesic) 거리를 이용하는 multi-center star shape 구조로 확장하였다. 이러한 초기 연구들은 객체의 전역적 구조 정보를 분할 과정에 반영함으로써, 잡음이 존재하거나 경계가 모호한 환경에서도 안정적인 결과를 얻을 수 있음을 보였다. 그러나 그래프 컷 기반 방법은 전통적 변분 최적화(variational optimization) 프레임워크 안에서 동작하므로, 데이터 기반으로 특징을 학습하는 딥러닝 분할 모델에 직접 통합하기 어렵다.

이러한 한계를 극복하기 위해, 구조적 사전 정보를 딥러닝 손실 함수 형태로 변환하여 학습 과정에 통합하는 연구가 이루어졌다. Mirikharaji and Hamarneh(2018)는 피부 병변 분할 과제에서 별 모양 제약을 손실 함수로 확장한 star shape prior loss를 제안하였다. 이는 객체 중심점에서 이미지 내 각 픽셀로 이어지는 선분 위에서 동일 레이블 픽셀들의 예측이 일관되도록 강제함으로써, FCN 모델의 성능을 향상시켰다. 그러나 이 방법은 단일 객체를 대상으로 한 의미 분할을 전제로 설계되었기 때문에, 여러 개체가 중첩되어 존재하는 개체 분할 환경에 그대로 적용하기에는 한계가 있다. 또한 학습 전 과정에 걸쳐 supervision 범위가 고정된 정적 구조로 인해, 학습 후반부에 분할 오류가 경계 영역에 집중되는 시점에도 감독 신호를 적응적으로 재배분하는 메커니즘이 부재하다.

이후 연구들은 볼록성(convexity) 사전 정보를 네트워크 구조 수준에서 강제하는 방향으로 확장되어 왔다. Liu *et al.*(2020)는 soft threshold dynamics(STD) 기반 활성화 함수를 설계하여, 신경망 출력이 볼록 형태 사전 정보를 자동으로 만족하도록 하는 convex shape soft threshold dynamics(CS-STD) 블록을 제시하였다. Zhao *et al.*(2025)는 단일 중심 star shape의 표현력 한계를 지적하고, 다중 중심 star shape를 변분 이론에 기

반하여 신경망에 통합한 convex combination star shape prior를 제안하였다. 이러한 연구들은 형태 사전 정보를 네트워크 구조 수준에서 반영한다는 점에서 강력하지만, 별도의 변분 최적화 모듈이나 활성화 함수 재설계를 요구하므로 기존 파운데이션 모델에 곧바로 결합하기는 어렵다. Liu *et al.*(2020)의 CS-STD는 분할 헤드 내부에 별도로 설계된 활성화 함수 블록을 새로 삽입해야 하므로, 사전학습된 FastSAM의 가중치 구조와 호환되지 않아 처음부터 재학습에 준하는 구조 변경이 필요하다. Zhao *et al.*(2025)의 다중 중심 star shape prior 역시 순전파 과정에 변분 에너지 최소화 문제를 추가로 풀어야 하므로, 추론 시마다 별도의 최적화 단계가 요구되어 연산 비용이 늘어날 뿐 아니라 기존 네트워크의 계산 그래프 자체를 수정해야 한다. 반면 본 연구가 제안하는 radial loss는 손실 함수 항 하나를 추가하는 것만으로 동작하므로, 네트워크 구조나 추론 과정의 변경 없이 기존 파운데이션 모델의 미세조정 파이프라인에 그대로 결합할 수 있다.

본 연구에서 제안하는 radial loss는 이러한 한계를 극복하기 위해 개체별 볼록 사전 정보를 광선 단위로 반영하여 개체 분할에 직접 적용 가능하도록 설계되었으며, 손실 함수 수준에서만 동작하므로 기존 분할 모델 구조에 최소한의 수정으로 통합할 수 있다. 더불어, 분할 오류가 특히 경계 영역에 집중된다는 점에 착안하여, 학습 단계에 따라 광선의 샘플링 범위를 경계 영역으로 동적으로 축소하는 dynamic search 기법을 통해 전역 구조 학습과 경계 정밀도 향상을 동시에 달성한다.

3. 제안 방법론

3.1 문제 정의

본 연구는 소수의 레이블 데이터만으로 밀집 재배 작물의 개체 분할을 수행하는 문제를 다룬다. 전체 데이터셋 $D=\{x_i, M_i\}_{i=1}^{N}$이 주어졌을 때, 각 이미지 $x_i \in R^{H\times W\times 3}$는 다수의 작물 개체를 포함하며, $M_i=\{M_i^{(j)}\}_{j=1}^{K_i}$는 해당 이미지 내 K_i개의 개체 각각에 대한 이진 분할 마스크 $M_i^{(j)} \in \{0,1\}^{H\times W}$ 집합이다. 여기서 은 전체 데이터셋에 포함된 이미지의 수, H와 W는 각각 이미지의 세로 · 가로 픽셀 크기, 3은 RGB 3채널 입력임을 의미하며 K_i는 이미지 x_i 내에 존재하는 작물 개체의 수를 나타낸다. 소수 샘플 학습 조건을 모사하기 위해, 전체 데이터셋 중 극히 일부(10장 또는 20장)만을 학습용으로 활용한다. 학습의 목표는 사전학습된 파운데이션 분할 모델 f_ϕ을 제한된 학습 데이터로 미세조정하여, 테스트 이미지 x_{test}에 대해 개별 작물의 분할 마스크 집합 M_{test}를 정확하게 예측하는 파라미터 ϕ^*를 찾는 것이다.

본 연구의 핵심 가정은 분할 대상 객체가 별 모양 구조를 가진다는 것이다. 별 모양 객체란 객체 내부에 하나의 중심점 c가 존재하여, 임의의 내부 점 p에 대해 c와 p를 잇는 선분 위의 모든 점 q가 다시 객체 내부에 포함되는 형태를 의미한다(Veksler, 2008). <Figure 1>은 별 모양 객체의 정의를 시각화한 것이다. 상추, 배추 등 위에서 내려다본 잎채소는 짧은 줄기로부터 잎이 방사형으로 뻗어나가는 로제트 구조를 가지므로, 무게중심을 기준으로 한 별 모양 가정을 자연스럽게 만족한다(Minervini *et al.*, 2016). 본 연구는 이러한 도메인 고유의 기하학적 특성을 손실 함수에 직접 반영함으로써, 소수의 레이블 데이터만으로도 모델이 작물의 전역적 형태를 학습할 수 있도록 한다.

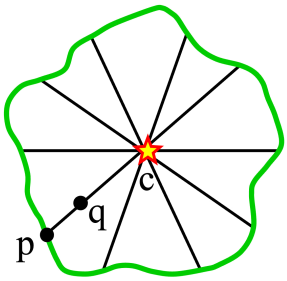


Figure 1. Definition of a star-shaped object. An object (green region) is star-shaped with respect to center c if, for any point p within the object, all points q on the line segment connecting c to p also lie within the object. Lettuce naturally satisfies this assumption when viewed from above because leaves radiate outward from a central stem in a rosette structure.

본 연구의 기반 모델로는 소수의 레이블 데이터로 도메인 적응이 가능하면서도 연산 효율이 뛰어난 FastSAM을 채택하였다. 상업적 규모의 스마트팜은 다수의 재배 베드에 설치된 여러 대의 카메라로 전체 재배 기간에 걸쳐 반복적으로 개체 분할을 수행해야 하므로, 모델 1회 추론에 소요되는 연산 비용과 메모리 요구량이 카메라 및 베드 수만큼 누적된다. 실제로 Lee *et al.*(2026)은 동일한 상추 스마트팜 환경에서 SAM과 FastSAM의 연산 효율을 비교하였으며, SAM은 이미지당 약 2,289ms의 추론 시간과 80GB의 GPU 메모리를 요구한 반면 FastSAM은 약 45ms, 15GB만으로 유사한 성능을 보였다. 이는 다수의 카메라를 상시 운용해야 하는 실제 배포 환경에서 경량 모델이 갖는 실용적 이점을 보여주며, 본 연구가 FastSAM을 기반 모델로 채택한 근거이다. FastSAM은 YOLOv8-seg 구조에서 가져온 네 가지 항의 가중합으로 정의된 손실 함수를 통해 학습되며 다음과 같이 나타낼 수 있다:

$$L_{FastSAM}=\lambda_{cls}L_{cls}+\lambda_{box}L_{box}+\lambda_{dfl}L_{dfl}+\frac{\lambda_{box}}{B}L_{mask} \quad (1)$$

이때 L_{cls}는 객체 존재 여부에 대한 분류 손실, L_{box}는 예측 경계 상자와 정답 간 일치도를 평가하는 회귀 손실, L_{dfl}은 경계 상자 회귀의 정밀도를 향상시키는 distribution focal loss이

며, L_{mask}는 예측 마스크와 정답 마스크 간의 픽셀 단위 이진 교차 엔트로피(binary cross-entropy) 손실이다. λ_{cls}, λ_{box}, λ_{dfl}은 각 손실 항의 가중치이며, 는 배치 크기이다. 기존 FastSAM 손실 함수는 각 픽셀을 독립적으로 평가하는 방식으로 설계되어, 개체의 전역적 형태나 작물 특유의 구조적 특성을 학습 과정에 반영하지 못한다. 이로 인해 소수의 레이블 데이터만으로 미세조정하는 경우, 인접한 작물 잎이 심하게 중첩된 밀집 환경에서 모델의 성능이 크게 떨어지는 문제가 발생한다. 이러한 한계를 극복하기 위해, 본 연구는 작물의 방사형 볼록 구조를 명시적으로 반영하는 radial loss를 기존 손실 함수에 추가하여 전체 학습 손실을 다음과 같이 정의한다:

$$L_{total} = L_{FastSAM} + \lambda_{radial} L_{radial} \quad (2)$$

L_{radial}은 본 연구에서 제안하는 방사형 손실 함수이며, λ_{radial}은 기존 손실과의 균형을 조정하는 가중치이다. Radial loss의 구체적 정의와 계산 과정은 3.2절에서 상세히 설명한다.

3.2 Radial Loss

본 연구에서 제안하는 radial loss는 기존 개체 분할 모델의 학습 손실 함수에 통합되는 부가 손실(auxiliary loss) 형태로 동작한다. <Figure 2>는 본 연구의 기반 모델인 FastSAM의 전체 구조를 보여준다. FastSAM은 SA-1B 데이터셋으로 사전학습된 파운데이션 분할 모델로, 일반 도메인에서 학습된 강건한 시각 표현을 보유하고 있어 소수의 레이블만으로 도메인 적응이 가능하다는 점에서 퓨샷 미세조정에 적합하다. 또한 SAM 대비 큰 폭의 추론 속도 향상을 달성한 경량화 구조로, 실시간 처리가 필요한 환경에 실용적으로 적용할 수 있다. 구조적으로 FastSAM은 YOLOv8-seg 기반 파운데이션 분할 모델로, 입력 이미지로부터 추출된 공유 특징을 탐지 헤드(detection head)와 프로토타입 헤드(prototype head)에 병렬로 전달한다. 탐지 헤드는 클래스 레이블·마스크 계수·경계 상자를 예측하고 프로토타입 헤드는 프로토타입 마스크를 생성하며, 탐지 헤드가 예측한 계수를 이용해 프로토타입들의 가중합을 계산한 뒤 이를 크롭 및 임계값 처리함으로써 최종 개체별 분할 마스크를 얻는다.

Radial loss는 상추와 같은 스마트팜 작물을 위에서 내려다본 형태가 근사적으로 볼록(convex) 구조를 가진다는 도메인 특성에 기반한다. 즉, 개체의 무게중심에서 출발하여 외곽 경계 방향으로 뻗어나가는 모든 직선 위의 지점은 전경(foreground)에 속한다는 기하학적 가정을 손실 함수에 직접 반영한다. Radial loss의 계산 과정은 다음 다섯 단계로 구성되며, <Figure 3>은 전체 흐름을 보여준다.

Step 1: Radial loss의 첫 번째 단계는 각 개체의 무게중심을 계산하는 과정이다. 개체 j의 정답 분할 마스크 $M_i^{(j)} \in \{0,1\}^{H \times W}$로부터 전경 픽셀 좌표의 평균을 계산하여 해당 개체의 무게중심 $c^{(j)}$를 다음과 같이 정의한다.

$$c^{(j)} = \frac{1}{|F^{(j)}|} \sum_{(x,y) \in F^{(j)}} (x, y),\ F^{(j)} = \{(x,y) : M^{(j)}(x,y) = 1\} \quad (3)$$

이때 $|F^{(j)}|$는 정답 마스크 $M^{(j)}$내 전경 픽셀의 총 수이고, (x,y)는 해당 전경 픽셀의 좌표이다.

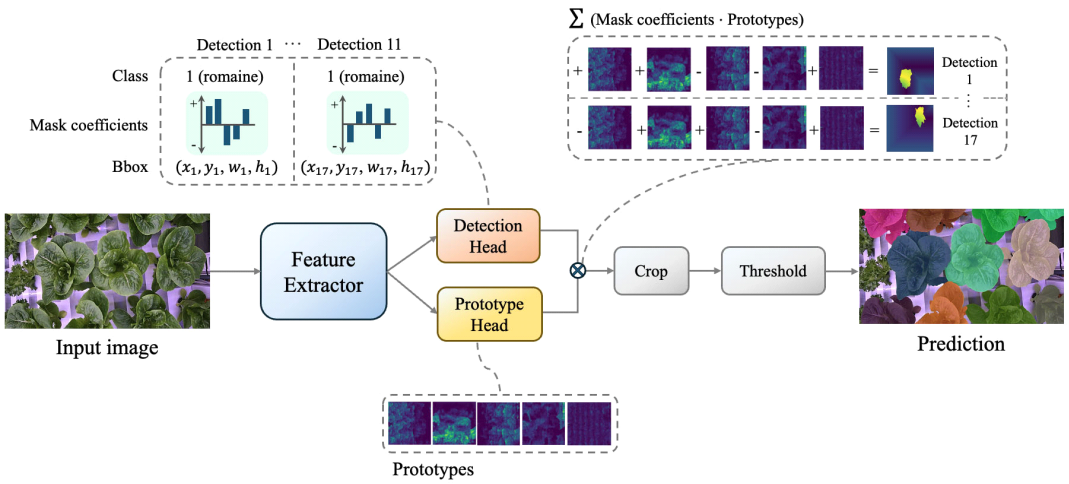


Figure 2. Overview of the FastSAM architecture. A feature extractor feeds shared representations into a detection head and a prototype head in parallel. The detection head outputs class labels, mask coefficients, and bounding boxes, while the prototype head generates prototype masks. Final instance masks are obtained by computing a weighted sum of prototypes using the predicted coefficients, followed by cropping and thresholding.

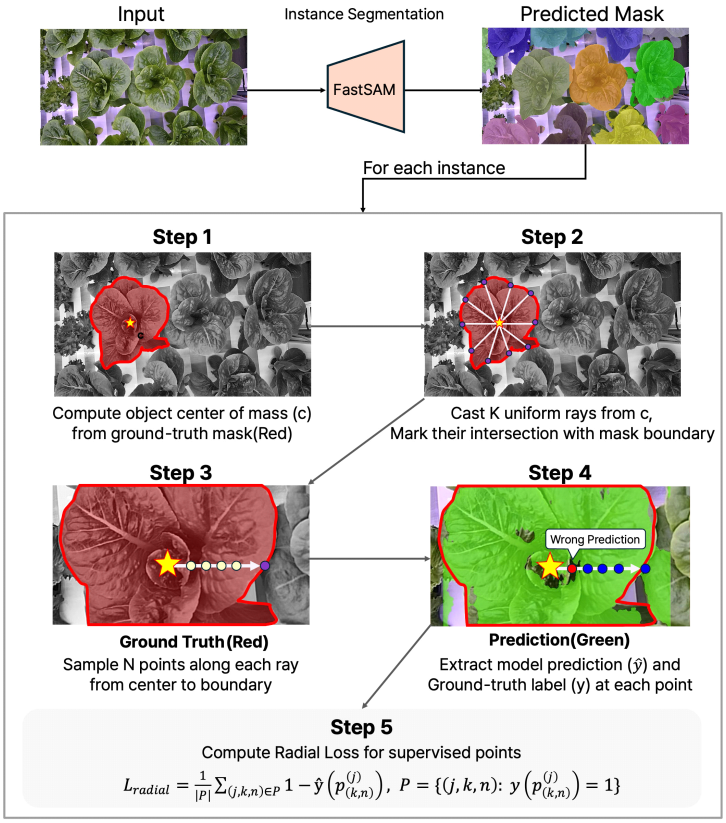


Figure 3. Overview of the radial loss computation pipeline. For each instance, the pipeline follows five steps: (1) computing the center of mass c from the ground-truth mask, (2) casting K uniformly spaced rays and searching for boundary points, (3) sampling N points along each ray, (4) extracting the predicted foreground probabilities at sampled locations, and (5) computing the loss at foreground-labeled points.

Step 2: 다음으로, 계산된 무게중심을 기준으로 외곽 방향으로 광선을 발사하여 개체의 경계점을 탐색한다. $c^{(j)}$를 기준으로 K개의 동일 간격 각도 $\theta_k = 2\pi k/K (k=0,1,\cdots,K-1)$에 대해 광선 $r_k^{(j)}$를 발사하며, 각 광선은 중심에서 외곽 방향으로 뻗어나가 정답 마스크의 경계와 처음 만나는 지점을 해당 방향의 경계점 $b_k^{(j)}$로 정의한다. 경계점은 다음과 같이 표현된다:

$$b_k^{(j)} = c^{(j)} + L_k^{(j)} \cdot (\cos\theta_k,\ \sin\theta_k) \tag{4}$$

이때 $L_k^{(j)}$는 무게중심 $c^{(j)}$에서 경계점 $b_k^{(j)}$까지의 거리이다.

$$L_k^{(j)} = min\{r>0 : c^{(j)} + r\cdot(\cos\theta_k,\ \sin\theta_k) \notin F^{(j)}\} \tag{5}$$

Step 3: 이렇게 정의된 광선을 따라 균일한 간격으로 N개의 지점을 샘플링함으로써 손실 계산의 대상이 되는 지점 집합을 구성한다. n번째 샘플링 지점 $p_{(k,n)}^{(j)}$는 다음과 같이 정의된다:

$$p_{(k,n)}^{(j)} = c^{(j)} + \frac{n}{N}\cdot L_k^{(j)} \cdot (\cos\theta_k,\ \sin\theta_k),\ n=1,2,\cdots,N \tag{6}$$

Step 4: 생성된 샘플링 지점은 중심에서 경계까지 이어지는 방사형 구조를 명시적으로 반영한다. 이후 각 샘플링 지점 $p_{(k,n)}^{(j)}$에서 모델이 예측한 전경 확률 $\hat{y}\left(p_{(k,n)}^{(j)}\right)\in[0,1]$과 해당 위치의 정답 레이블 $y\left(p_{(k,n)}^{(j)}\right)\in\{0,1\}$을 추출한다.

Step 5: 마지막으로, 정답 레이블이 전경인 샘플링 지점에 한하여 예측 확률이 1에 가까워지도록 감독 신호를 부여한다. 이에 따라 radial loss는 다음과 같이 정의된다:

$$L_{radial} = \frac{1}{|P|}\sum_{(j,k,n)\in P}\left(1-\hat{y}\left(p_{(k,n)}^{(j)}\right)\right), \tag{7}$$
$$P=\{(j,k,n) : y(p_{(k,n)}^{(j)})=1\}$$

여기서 P는 정답 레이블이 전경인 샘플링 지점의 인덱스 집합, $|P|$는 해당 집합의 크기이다. Radial loss는 광선을 따라 샘플링된 개별 전경 지점을 직접 감독함으로써 작물의 방사형

볼록 구조를 각 개체 단위로 반영한다. 이러한 설계는 다수 개체가 중첩된 밀집 작물 환경에서 개체별 경계 학습을 용이하게 하며, 기존 픽셀 단위 손실 함수가 포착하지 못하는 작물의 기하학적 구조를 학습 과정에 직접 주입한다.

3.3 Dynamic Search

학습 초기에는 모델이 작물의 전역적 형태를 학습해야 하므로, radial loss가 광선 전체 길이에 걸친 감독 신호를 제공하는 것이 효과적이다. 그러나 <Figure 4>와 같이, 학습이 진행되어 개체 중심부 예측이 충분히 수렴한 이후에는, 남은 예측 오류가 주로 경계 영역에 집중되는 현상이 나타난다. 이 시점에서도 광선 전체에 감독 신호가 균일하게 분산되면, 이미 잘 예측되는 내부 영역에 불필요한 감독이 할당되어 경계 영역 학습이 늦어진다.

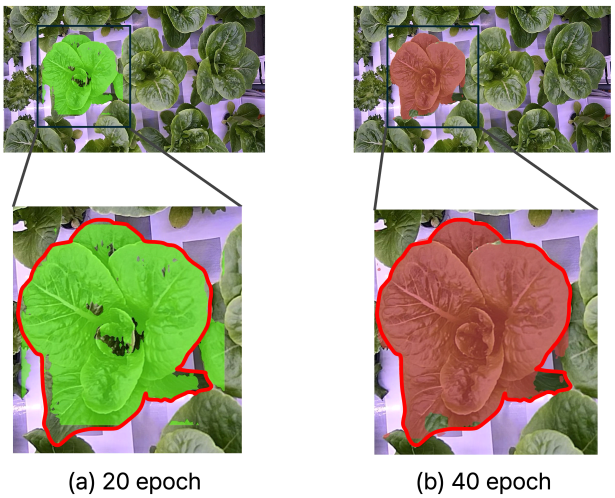


Figure 4. Visualization of prediction error concentration over training epochs. (a) At epoch 20, errors are distributed across both interior and boundary regions. (b) At epoch 40, interior predictions converge while residual errors concentrate along the boundary, motivating boundary-focused supervision in later training stages.

이러한 문제를 해결하기 위해, 본 연구는 검증 mAP가 사전에 정의된 임계값을 초과하는 시점을 기준으로 광선 샘플링 범위를 전환하는 dynamic search 기법을 도입한다. <Figure 5>와 같이, 매 학습 에폭 종료 시점에 검증 세트에 대한 mAP를 측정하고, 사전에 설정된 임계값 τ를 처음 초과한 시점부터 샘플링 시작 위치를 경계 방향으로 이동시킨다.

이를 정식화하기 위해, t번째 에폭의 검증 mAP를 $mAP_{val(t)}$, 해당 에폭에서의 광선 샘플링 시작 위치 비율을 $S^{(t)}$, 광선 전체 길이 대비 샘플링 구간의 비율을 $\beta \in (0,1]$로 정의한다. 이때, 광선 위 샘플링 지점은 다음과 같이 조정된다:

$$p_{(k,n)}^{(j)} = c^{(j)} + \left(s^{(t)} + (1-s^{(t)})\cdot\frac{n}{N}\right)\cdot L_k^{(j)} \cdot (\cos\theta_k, \sin\theta_k),\ n=1,2,\ldots,N \tag{8}$$

$s^{(t)}$는 검증 mAP가 임계값 τ를 초과하기 전까지 $s^{(t)}=0$을 유지하여 광선 전체 길이에 걸쳐 샘플링이 이루어지도록 하고, τ를 처음 초과한 에폭부터 $s^{(t)}=1-\beta$로 전환하여 광선의 외곽 β구간에만 감독 신호가 집중된다. 본 연구에서는 실험을 통해 $\tau=0.5$, $\beta=0.6$으로 설정하였으며, 해당 설정의 타당성은 4절에서 탐색 비율에 대한 추가 분석으로 검증한다.

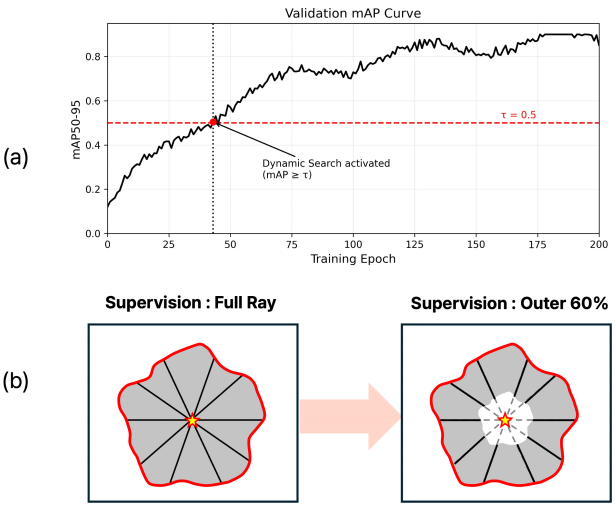


Figure 5. Overview of the dynamic search mechanism. (a) Validation mAP curve during training. Dynamic search is activated when the validation mAP first exceeds the threshold τ. (b) Transition of the ray sampling range from full-ray supervision (left) to outer 60% supervision (right) upon activation, concentrating the supervisory signal on boundary regions in the later stage of training.

이러한 동적 조정을 통해, 학습 초기에는 작물의 전역 볼록 구조를 학습하고 학습 후반부에는 경계 정밀도 향상에 감독 신호를 집중함으로써, 전역 형태 보존과 경계 세밀도 향상을 단계적으로 달성한다. 제안 방법론의 전체 학습 절차는 <Algorithm 1>에 요약되어 있다.

Algorithm 1: Few-Shot Instance Segmentation Training with Radial Loss and Dynamic Search

Require: train dataset D_{train}, validation dataset D_{val}, pre-trained FastSAM model f_ϕ, total epochs T, ray count K, samples per ray N, loss weight λ_{radial}, activation threshold τ, search ratio β

Output: Fine-tuned model parameters ϕ^*

1: $\leftarrow s0$ // sampling start ratio
2: activated $\leftarrow$ False
3: for $t=1$ to T do
4: for each mini-batch (X, Y) in D_{train} do
5: Forward pass through $f_\phi(X) \rightarrow$ predictions $\hat{Y}$
6: Compute $L_{FastSAM}$ via Eq. (1)
7: $\leftarrow 0$, $|P| \leftarrow 0$
8: for each instance mask $M^{(j)}$ in Y do
9: Compute center of mass $c^{(j)}$ via Eq. (3)
10: for $k=0$ to $K-1$ do
11: $\theta_k \leftarrow 2\pi k/K$
12: Compute boundary distance $L_k^{(j)}$ via Eq. (5)
13: for $n=1$ to N do
14: $p_{(k,n)}^{(j)} \leftarrow$
$c^{(j)} + (s + (1-s) \times n/N) \times L_k^{(j)} \times (\cos\theta_k, \sin\theta_k)$,
15: if $y(p_{(k,n)}^{(j)}) = 1$ then
16: $L_{radial} \leftarrow L_{radial+}(1 - \hat{y}(p_{(k,n)}^{(j)}))$
17: $|P| \leftarrow |P| + 1$
18: end if
19: end for
20: end for
21: end for
22: $L_{radial} \leftarrow L_{radial}/|P|$
23: $L_{total} \leftarrow L_{FastSAM} + \lambda_{radial} \times L_{radial}$
24: Update ϕ via backpropagation on L_{total}
25: end for
26: Compute $mAP_{val(t)}$ on D_{val}
27: if not activated and $mAP_{val(t)} \geq \tau$ then
28: $s \leftarrow 1-\beta$ // activate Dynamic Search
29: activated $\leftarrow$ True
30: end if
31: end for

Return: $\phi^* \leftarrow \phi$

4. 실험 및 결과

4.1 실험셋팅

본 연구는 제안 방법론의 효과를 검증하기 위해 국내 식품업체 스마트팜에서 수집된 상추 3품종(butterhead, romaine, red lettuce) 데이터셋을 활용하였다. 각 품종은 밀집 재배 환경에서 촬영된 이미지로 구성되어 있으며, 한 이미지 내에 다수의 개체가 중첩되어 분포한다. <Table 1>은 각 품종별 학습, 검증, 테스트 데이터의 구성을 보여준다. 전체 데이터셋은 학습 70%, 검증 20%, 테스트 10%의 비율로 분할하였으며, 검증 및 테스트 분할은 모든 실험에 걸쳐 동일하게 유지된다. 학습 데이터는 별도의 학습 풀로 두고, 퓨샷 시나리오에 따라 그중 일부만을 선별하여 사용하였다. 학습 풀을 실제 사용하는 10장, 20장보다 충분한 규모로 구축함으로써, 향후 40장과 같이 더 큰 규모의 few-shot 조건에서도 곧바로 실험이 가능하도록 하였다. 이에 본 연구에서는 학습 풀을 상대적으로 크게 유지하는 대신, 검증 및 테스트 분할은 모든 학습 조건에서 동일하게 고정하여 조건 간 성능 비교의 공정성을 확보하였다. 실제 농업 현장에서 흔히 발생하는 극단적 데이터 부족 상황을 모사하기 위해, 본 연구는 학습 풀에서 10장 및 20장의 이미지를 무작위로 선별하는 두 가지 퓨샷 시나리오를 구성하였다. 두 시나리오 간 비교의 신뢰성을 확보하기 위해, 10장 학습 세트는 20장 학습 세트의 부분집합으로 구성하는 계층적 샘플링(hierarchical sampling) 방식을 채택하였다. 이를 통해 두 조건 간 성능 차이가 데이터 분포의 우연한 차이가 아니라 학습 데이터 규모의 차이에서 기인함을 보장하였다. 결과의 신뢰성을 추가로 확보하기 위해 모든 실험은 서로 다른 3개의 무작위 시드(random seed)에 대해 반복 수행하였으며, 평균 성능을 제시하였다.

Table 1. Number of training, validation, and test samples for each lettuce dataset.

	Butterhead	Romaine	Red lettuce
# of Training Pool	44	49	48
↳ 10-shot subset	10	10	10
↳ 20-shot subset	20	20	20
# of Validation Data	14	16	15
# of Test Data	11	13	12
Total size	69	78	75

기반 모델로는 SA-1B 데이터셋으로 사전학습된 FastSAM을 사용하였다. 모든 실험은 NVIDIA RTX 4090 GPU 환경에서 수행하였으며, 초기 학습률(initial learning rate) 1×10^{-2}, 최종 학습률 비율(final learning rate factor) 1×10^{-2}, 배치 크기 8, 학습 에폭 300으로 설정하였다. 첫 3 에폭 동안 warmup을 적용한 후 선형 학습률 감소(linear learning rate decay) 스케줄을 사용하였으며, 옵티마이저는 Ultralytics 기본 설정(momentum 0.937, weight decay 5×10^{-4})을 따랐다. Radial loss의 하이퍼파라미터는 광선 개수 $K=36$, 광선당 샘플링 지점 수 $N=8$, 손실 가중치 $\lambda_{radial}=0.5$로 설정하였으며, dynamic search의 활성화 임계값은 $\tau=0.5$, 광선 탐색 비율 $\beta=0.6$으로 설정하였다. 평가 지표로는 개체 분할 분야의 표준 지표인 mean average precision (mAP)와 intersection over union(IoU) 임계값 0.5 및 0.75에서 측정한 AP50, AP75를 사용하였다.

비교 baseline은 다음 두 가지를 사용하였다. 첫째, 기존 FastSAM 손실 함수 $L_{FastSAM}$만을 사용하여 미세조정한 lower bound은 구조적 사전 정보 없이 픽셀 단위 손실만으로 학습한 경우의 성능을 나타낸다. 둘째, star loss는 본 연구와 동일한 형태로 구조적 사전 정보를 손실 함수 수준에서 반영한 baseline으로 $L_{FastSAM}$에 star shape prior loss를 부가 손실로 추가하여 구성된다(Mirikharaji and Hamarneh, 2018). 제안 방법론 radial loss는 $L_{FastSAM}$에 L_{radial}를 부가 손실로 추가한 형태로, 세 방법은 모두 FastSAM 기존 손실을 공통 기반으로 하였으며 추가 손실 항의 차이만으로 비교를 수행하였다. Radial loss는 star loss와 마찬가지로 중심점 기반의 별 모양 가정을 손실 함수에 반영한다는 점에서 출발점을 공유하지만, 구조적으로는 뚜렷이 구분된다. 첫째, star shape prior loss는 이미지 내 모든 픽셀 p에 대해 p와 중심을 잇는 선분 위의 픽셀 q간의 예측값 불일치를 벌점화하는 조밀한 쌍별(pairwise) 정규화 항인 반면, radial loss는 중심에서 균일한 각도로 발사된 K개의 광선을 따라 샘플링된 N의 지점에서 예측값과 정답 레이블 간의 오차를 직접 감독하는 희소한(sparse) 점별(point-wise) 지도 방식이다. 둘째, star shape prior loss는 전체 선분을 모두 고려할 경우의 높은 연산 비용 때문에 실제 구현에서는 근접 m개 픽셀과 이산화된 d개 방향만으로 근사하는 반면, radial loss는 처음부터 K개의 고정된 광선과 광선당 N개의 샘플링 지점만을 대상으로 설계되어 근사 없이 정확히 정의된 범위에서 계산이 이루어진다. 아울러 star shape prior loss는 학습 전 과정에 걸쳐 고정된 정적 구조를 가지는 반면, radial loss는 3.3절에서 설명한 dynamic search를 통해 학습 진행 단계에 따라 샘플링 범위를 조정할 수 있는 학습 단계 의존적 구조를 가진다. 이러한 구조적 차이는 바로 다음 절의 실험 결과에서 radial loss가 star loss 대비 일관된 성능 우위를 보이는 것과도 부합한다.

4.2 퓨샷 미세조정 실험 결과

<Table 2>는 두 가지 학습 데이터 규모(10장, 20장)에서 비교 방법론과 제안 radial loss의 분할 성능을 보여준다. 제안 방법론은 모든 학습 규모에서 lower bound(FastSAM loss)와 baseline(Star loss)대비 일관된 성능 향상을 달성하였다. 특히 학습 데이터가 10장으로 극단적으로 제한된 환경에서 평균 mAP 0.609를 기록하여, FastSAM loss (0.513) 및 star loss (0.532) 대비 각각 9.6%p, 7.7%p의 큰 폭의 성능 향상을 보였다. 학습 데이터가 20장으로 증가한 경우에도 평균 mAP 0.800을 달성하여, baseline 대비 2.2%p의 향상을 확인하였다. 이는 레이블 데이터가 부족할수록 제안 방법론의 효과가 더욱 뚜렷해짐을 시사하며, 작물의 방사형 볼록 구조를 명시적으로 반영함으로써 모델이 제한된 데이터로부터 효율적으로 도메인 지식을 학습할 수 있음을 입증한다.

품종별로 살펴보면, butterhead 품종에서 가장 큰 성능 향상이 관찰되었다. 10장 학습 조건에서 baseline 대비 mAP 20.5%p의 향상을 보였으며, 이는 butterhead가 다른 품종 대비 로제트 구조가 명확하여 제안 방법론의 형태적 가정을 가장 잘 만족하기 때문으로 분석된다. Romaine 품종에서도 일관된 성능 향

Table 2. Performance comparison of FastSAM across three lettuce datasets using different loss functions. The best performance among compared methods is highlighted in bold, and values in parentheses indicate standard deviations. An asterisk (*) indicates a statistically significant improvement of the proposed method over the baseline (paired t-test, $p < 0.05$).

Fine-tuning method	Butterhead			Romaine			Red lettuce			Average		
	AP50	AP75	mAP	AP50	AP75	mAP	AP50	AP75	mAP	AP50	AP75	mAP
Zero-shot												
-	0.044	0.010	0.016	0.101	0.031	0.039	0.467	0.429	0.403	0.204	0.157	0.153
Fine-tuned with 10 labeled images												
FastSAM loss	0.798 (0.075)	0.480 (0.306)	0.493 (0.196)	0.630 (0.063)	0.320 (0.130)	0.337 (0.085)	0.879 (0.009)	0.810 (0.019)	0.710 (0.013)	0.769 (0.127)	0.537 (0.250)	0.513 (0.187)
Star loss	0.828 (0.061)	0.515 (0.242)	0.506 (0.134)	0.720 (0.027)	0.329 (0.009)	0.381 (0.005)	**0.888** (0.018)	0.798 (0.015)	0.708 (0.013)	0.812 (0.085)	0.547 (0.236)	0.532 (0.165)
Radial loss (Proposed)	**0.860** (0.024)	**0.812*** (0.159)	**0.711*** (0.112)	**0.733** (0.043)	**0.550*** (0.095)	**0.493** (0.108)	0.875 (0.034)	**0.830*** (0.016)	**0.721*** (0.012)	**0.821** (0.078)	**0.670** (0.157)	**0.609** (0.129)
Fine-tuned with 20 labeled images												
FastSAM loss	0.917 (0.047)	0.845 (0.149)	0.751 (0.165)	0.845 (0.026)	0.733 (0.038)	0.695 (0.033)	0.900 (0.017)	0.851 (0.020)	0.756 (0.025)	0.887 (0.038)	0.810 (0.066)	0.734 (0.034)
Star loss	0.930 (0.008)	0.904 (0.004)	0.830 (0.006)	0.842 (0.017)	0.756 (0.002)	0.705 (0.021)	0.915 (0.008)	0.872 (0.008)	0.799 (0.005)	0.896 (0.047)	0.844 (0.078)	0.778 (0.065)
Radial loss (Proposed)	**0.961*** (0.006)	**0.961*** (0.007)	**0.865*** (0.006)	**0.866** (0.002)	**0.772*** (0.001)	**0.735** (0.013)	**0.919** (0.013)	**0.874** (0.015)	**0.801** (0.009)	**0.916** (0.048)	**0.869** (0.095)	**0.800** (0.065)

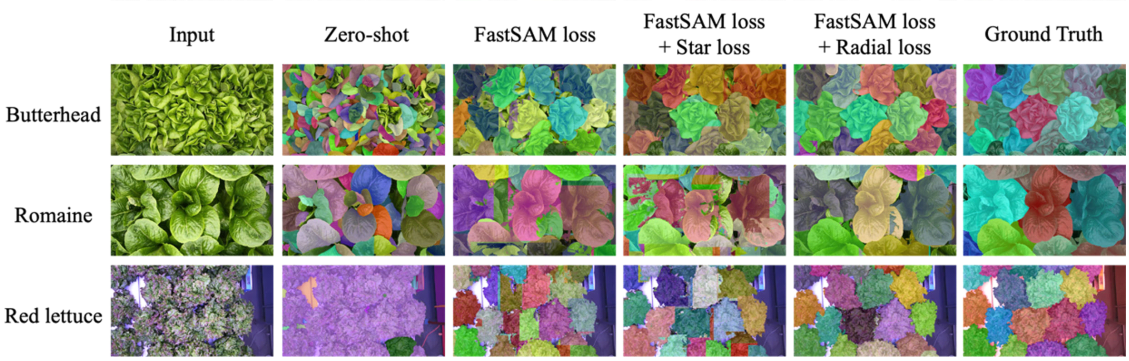


Figure 6. Qualitative comparison of segmentation results under the 10-shot setting. From left to right: input image, zero-shot prediction, FastSAM loss baseline, FastSAM loss + Star loss, FastSAM loss + Radial loss (proposed), and ground truth. The proposed method produces cleaner instance boundaries in densely overlapping regions, whereas baselines show frequent over-segmentation and merging errors near contact zones.

상이 확인되었으나, 잎이 길게 뻗는 형태 특성으로 인해 butterhead 대비 향상 폭은 다소 작게 나타났다. Red lettuce의 경우 zero-shot 성능이 이미 상대적으로 높아(평균 mAP 0.403) baseline 간 격차가 작았으며, 그 결과 제안 방법론의 향상 폭도 다른 품종 대비 작게 나타났다. 다만 10-shot 조건의 mAP에서는 통계적으로 유의한 향상을 확인하였다.

<Figure 6>은 학습 데이터 10장 조건에서의 분할 결과를 시각적으로 비교한 것이다. Lower bound(FastSAM loss)의 경우 인접한 작물 잎의 경계를 정확히 구분하지 못하여 단일 개체를 여러 부분으로 과분할하는 오류가 빈번하게 관찰되었다. 비교 방법론은 star loss를 학습에 함께 활용함으로써 lower bound 대비 분할 안정성이 다소 향상되었으나, 여전히 인접 작물 간 경계 영역에서 부정확한 분할이 발생하였다. 반면 제안 radial loss는 개체의 무게중심으로부터 방사형으로 뻗어나가는 구조를 명시적으로 학습함으로써, 인접 작물 간 경계가 모호한 영역에서도 개별 작물을 일관성 있는 단일 개체로 분할하였다. 특히 작물 외곽의 경계선 추적 정확도가 baseline 대비 두드러지게 향상되었으며, 이는 dynamic search 기법을 통해 학습 후반부에 감독 신호가 경계 영역으로 집중된 결과로 해석된다.

4.3 광선 탐색 비율 실험 결과

<Table 3>은 광선 샘플링 범위를 학습 전 기간에 걸쳐 고정한 경우, 탐색 비율 β가 분할 성능에 미치는 영향을 분석한 결과이다. $\beta = 0.5$(외곽 50% 구간만 감독), $\beta = 0.6$, $\beta = 0.7$, $\beta = 1.0$(전체 구간 감독) 네 가지 설정을 비교한 결과, $\beta = 0.6$ 설정이 평균 mAP 0.602로 가장 우수한 성능을 달성하였다. $\beta = 0.5$ 설정은 감독 영역이 지나치게 좁아 객체 내부 구조 학습이 불충분하였으며, $\beta = 0.7$ 및 $\beta = 1.0$ 설정은 경계 집중 효과가 약화되어 분할 성능이 저하되었다. 이러한 결과는 외곽 60% 구간이 전역 형태 학습과 경계 정밀도 향상 간의 균형을 가장 효과적으로 달성하는 탐색 범위임을 시사하며, 3.3절에서 제안하는 dynamic search는 이 비율($\beta = 0.6$)을 학습 후반부에 적응적으로 적용함으로써 추가적인 성능 향상을 도모한다.

Table 3. Performance comparison of fixed ray search ratios under the 10-shot setting. β denotes the ratio of each ray to which the supervisory signal is applied, measured from the boundary inward (e.g., β = 0.6 supervises the outer 60% of each ray; β = 1.0 corresponds to full-ray supervision). The best performance among compared methods is highlighted in bold, and values in parentheses indicate standard deviations.

Method	Butterhead			Romaine			Red lettuce			Average		
	AP50	AP75	mAP	AP50	AP75	mAP	AP50	AP75	mAP	AP50	AP75	mAP
$\beta = 0.5$	0.760 (0.041)	0.616 (0.073)	0.531 (0.064)	0.890 (0.019)	0.822 (0.025)	0.723 (0.031)	0.703 (0.052)	0.354 (0.081)	0.376 (0.066)	0.784 (0.038)	0.597 (0.060)	0.543 (0.054)
$\beta = 0.6$	**0.887** (0.026)	**0.835** (0.038)	**0.725** (0.035)	**0.899** (0.017)	0.820 (0.022)	0.720 (0.027)	0.672 (0.047)	0.333 (0.069)	0.362 (0.058)	**0.819** (0.030)	**0.663** (0.043)	**0.602** (0.040)
$\beta = 0.7$	0.859 (0.034)	0.722 (0.061)	0.644 (0.052)	0.875 (0.021)	0.822 (0.024)	0.730 (0.026)	0.646 (0.055)	0.409 (0.077)	0.383 (0.064)	0.793 (0.037)	0.651 (0.054)	0.586 (0.047)
$\beta = 1.0$	0.843 (0.037)	0.687 (0.066)	0.620 (0.056)	0.877 (0.020)	**0.825** (0.023)	**0.730** (0.025)	**0.717** (0.049)	**0.452** (0.073)	**0.419** (0.060)	0.812 (0.035)	0.655 (0.052)	0.590 (0.045)

4.4 Dynamic Search 효과 분석 실험 결과

<Table 4>는 dynamic search 기법의 효과를 분석한 결과이다. Dynamic search를 제거한 경우 10-shot 조건에서 평균 mAP가 0.609에서 0.590으로, 20-shot 조건에서 0.800에서 0.765로 각각 1.9%p, 3.5%p 감소하였으며, 모든 지표에서 통계적으로 유의한 성능 저하가 관찰되었다. 이는 두 학습 규모 모두에서 dynamic search 기법이 일관된 성능 향상에 기여함을 입증한다. 특히 레이블 데이터가 가장 제한적인 10-shot 조건에서는, 엄격한 IoU 임계값을 요구하여 경계 정밀도에 민감한 AP75 지표에서 가장 큰 폭의 향상(2.5%p)이 관찰되었으며, 이는 AP50 향상(0.9%p) 및 mAP 향상(1.9%p)을 모두 웃도는 수치이다. 이러한 결과는 dynamic search 기법이 학습 후반부에 감독 신호를 경계 영역에 집중시킴으로써, 인접 작물 간 경계 분할 품질을 효과적으로 향상시킨다는 본 연구의 설계 의도를 뒷받침한다.

Table 4. Effect of dynamic search. We compare the average performance of the proposed method with and without the dynamic search mechanism across the three lettuce datasets under both 10-shot and 20-shot fine-tuning settings. The best performance in each column is highlighted in bold, and values in parentheses indicate standard deviations. An asterisk (*) indicates a statistically significant improvement of the proposed method over the without-Dynamic-Search variant (paired t-test, $p < 0.05$).

Dynamic Search	AP50	AP75	mAP
Fine-tuned with 10 labeled images			
✓	**0.821*** (0.078)	**0.670*** (0.157)	**0.609*** (0.129)
✗	0.812 (0.134)	0.645 (0.147)	0.590 (0.141)
Fine-tuned with 20 labeled images			
✓	**0.916*** (0.048)	**0.869*** (0.095)	**0.800*** (0.065)
✗	0.896 (0.127)	0.848 (0.135)	0.765 (0.138)

4.5 광선 개수 실험 결과

광선 개수 K는 radial loss가 작물의 방사형 구조를 표현하는 해상도를 결정하는 핵심 하이퍼파라미터이다. K가 작으면 인접 광선 사이의 각도 간격이 넓어져 외곽 경계의 세부 형태를 충분히 포착하지 못하는 반면, K가 과도하게 크면 인접 광선이 거의 동일한 영역을 중복 감독하여 손실 신호의 다양성이 감소하고 에폭당 학습 비용이 선형적으로 증가하는 trade-off가 존재한다. 이에 본 절에서는 작물의 로제트 구조를 충실히 반영하면서도 학습 효율을 유지할 수 있는 적정 광선 밀도가 존재할 것이라는 가설 아래, 레이블 데이터가 가장 제한적인 10-shot 조건에서 $K \in \{24,\ 30,\ 36,\ 42\}$ 네 가지 설정의 분할 성능과 에폭당 학습 시간을 함께 측정하여 최적 광선 개수를 결정하였다.

<Table 5>는 그 결과를 보여준다. K=36 설정이 평균 mAP 0.609로 가장 우수한 성능을 달성하였으며, AP50 (0.821) 및 AP75 (0.670) 지표에서도 모두 최고 성능을 기록하였다. K를 24, 30으로 줄인 경우 평균 mAP가 각각 0.564, 0.593으로 하락하였는데, 이는 광선 밀도가 낮을 경우 인접 광선 사이의 각도 간격이 벌어져 작물 외곽의 굴곡을 충분히 감독하지 못하고, 결과적으로 방사형 볼록 구조에 대한 사전 정보가 학습 신호로 충분히 전달되지 못한 것으로 해석된다. 한편 K=42로 광선 수를 더 증가시킨 경우, 평균 mAP가 소폭 감소하였으며 특히 경계 정밀도 지표인 AP75에서 더 큰 폭의 하락이 관찰되었다. 이는 광선 수가 작물의 로제트 구조가 요구하는 해상도를 초과하면 인접 광선이 거의 중복된 경로를 감독하게 되어 손실 신호가 특정 방향으로 편향되고, 한정된 학습 데이터 조건에서 경계 정밀도가 우선적으로 저하됨을 시사한다.

학습 비용 측면에서는 K가 24에서 42로 증가함에 따라 에폭당 학습 시간이 25.63초에서 49.51초로 약 1.93배 선형적으로 증가하였다. 특히 K=42 설정은 K=36 대비 약 17%의 추가 학습 비용을 요구함에도 분할 성능에서 일관된 향상을 보이지 못하였으며, 오히려 경계 정밀도 지표인 AP75에서 성능 저하를 동반하였다. 이러한 결과는 작물의 방사형 구조를 효과적으로 학습하기 위해서는 적정 수준의 광선 밀도가 필요하며, 본 연구에서 채택한 K=36이 분할 성능과 학습 효율 간의 균형 측면에서 가장 합리적인 선택임을 입증한다.

Table 5. Effect of the number of rays K on segmentation performance and per-epoch training time under the 10-shot fine-tuning setting. Performance metrics (AP50, AP75, mAP) are averaged across the three lettuce varieties, and values in parentheses indicate standard deviations. Training time per epoch is measured in seconds on a single NVIDIA RTX 4090 GPU. The best performance among compared settings is highlighted in bold.

# of Rays	Accuracy			Wall-clock time per epoch (s)
	AP50	AP75	mAP	
K=24	0.796 (0.112)	0.592 (0.232)	0.564 (0.174)	25.63
K=30	0.814 (0.135)	0.625 (0.221)	0.593 (0.171)	37.40
K=36	**0.821** (0.078)	**0.670** (0.157)	**0.609** (0.129)	42.49
K=42	0.802 (0.153)	0.618 (0.312)	0.589 (0.222)	49.51

5. 결 론

본 연구는 작물의 방사형 볼록 구조를 손실 함수에 직접 반영함으로써, 레이블 데이터가 극단적으로 부족한 밀집 재배 환경에서도 안정적인 작물 개체 분할 성능을 확보할 수 있는 방사형 손실 함수 radial loss를 제안하였다. Radial loss는 정답 마스크의 무게중심으로부터 방사형으로 뻗어나가는 광선을 따라 전경 예측을 감독함으로써, 작물의 방사형 볼록 구조를 손실 함수 수준에서 명시적으로 학습하도록 유도한다. 또한, 검증 mAP가 사전에 정의된 임계값을 초과하는 시점을 기준으로 광선의 샘플링 범위를 경계 영역으로 전환하여 축소하는 dynamic search 기법을 도입하여, 학습 후반부에 발생하는 경계 영역 집중 오류를 효과적으로 완화하고 분할 정밀도를 추가적으로 개선하였다. 실제 스마트팜에서 수집된 상추 3품종 데이터셋에 대한 실험 결과, 제안 방법론은 레이블 데이터 10장만 사용하는 극단적 퓨샷 조건에서 기존 star shape prior 기반 baseline 대비 평균 mAP를 7.7%p 향상시켰으며, 레이블 20장 조건에서도 2.2%p의 성능 개선을 달성하였다. 더불어 제안 방법론은 손실 함수 수준에서만 동작하므로 별도의 네트워크 구조 수정 없이 기존 파운데이션 분할 모델에 통합 가능하다는 실용적 강점을 가진다. 이러한 결과는 제안 방법론이 레이블 확보가 현실적으로 어려운 실제 스마트팜 환경에서 높은 실용성을 가질 수 있음을 시사하며, 나아가 생육 단계 분류나 수확량 예측과 같은 후속 downstream 과제의 정확도 향상에도 기여할 수 있을 것으로 기대된다.

다만, radial loss는 작물이 방사형 볼록 구조를 가진다는 가정에 기반하고 있어 비볼록 구조 작물에는 직접 적용하기 어려운 한계가 있으며, dynamic search의 활성화 임계값과 광선 탐색 비율이 사전에 고정된 하이퍼파라미터로 설정되어 재배 환경 변화에 따라 재탐색이 필요한 한계가 있다. 본 연구에서 검증한 butterhead, romaine, red lettuce 세 품종은 로제트 형태를 공유하면서도 잎의 곡률과 결구 형태에서 뚜렷한 차이를 보이는데, 세 품종 모두에서 radial loss가 일관된 성능 향상을 보였다는 결과는 본 방법론이 특정 품종에 국한되지 않고 로제트형 작물이 공유하는 구조적 특성을 학습하고 있음을 뒷받침한다. 이에 따라 배추, 케일, 청경채 등 위에서 내려다볼 때 짧은 줄기로부터 잎이 방사형으로 퍼져나가는 다른 로제트형 엽채류에도 별도의 구조 변경 없이 적용 가능할 것으로 판단된다. 반면 덩굴성 작물이나 줄기가 길게 뻗는 작물과 같이 뚜렷한 단일 중심을 갖지 않는 비방사형 작물에는 현재 형태의 radial loss를 직접 적용하기 어려우며, 이 경우 복수의 중심점을 두고 각 중심으로부터 국소적으로 방사형 볼록성을 강제하는 다중 중심(multi-center) 구조로 확장한다면 적용 범위를 넓힐 수 있을 것으로 기대되나, 중심점 수가 늘어날수록 경계 영역에서의 감독 신호가 분산되어 성능이 저하될 가능성이 있어 그 민감도는 향후 추가 검증이 필요하다. 아울러 본 연구에서 사용한 이미지는 상용 스마트팜에서 실제로 널리 채택되는 상단 고정형 카메라 배치로 촬영된 것으로, 별 모양 가정은 완전한 대칭을 요구하지 않으므로 다소 기울어진 각도에서도 원리적으로 성립할 수 있다. 다만 카메라 각도가 극단적으로 기울어져 잎들이 서로를 가리는 심각한 가림 현상이 발생하는 경우에는 객체의 투영 형태가 더 이상 단일 중심 기준의 별 모양 구조를 만족하지 못할 수 있어, 본 연구의 적용 범위는 상단 또는 이에 가까운 각도에서 촬영이 이루어지는 환경으로 한정된다. 본 연구에서 활용한 데이터셋은 실제 스마트팜 현장에서 수집, 라벨링된 것으로 품종별 전체 규모가 69~78장 수준에 불과하며, 이 중 학습 풀을 제외한 나머지를 검증과 테스트에 배분함으로 인하여 테스트 표본 수 역시 11~13장으로 제한되었다. 학습 풀에 남은 여분의 이미지를 테스트셋으로 재배분하면 표본 수를 다소 늘릴 수 있으나, 이는 본 연구가 다루는 few-shot 조건(10장, 20장)보다 더 큰 규모로 실험을 확장할 수 있는 여지를 줄이는 결과로 이어진다. 따라서 본 연구는 현재의 분할을 유지하는 대신, 이러한 표본 크기의 한계를 향후 대규모 데이터 수집을 통해 보완할 필요가 있는 과제로 남겨둔다. 따라서 향후 연구에서는 복수의 무게중심을 활용하는 다중 중심 광선 설계를 통해 비볼록 작물로 적용 범위를 확장하고, 검증 성능 추이에 기반하여 하이퍼파라미터를 학습 중 자동으로 조정하는 적응형 dynamic search 기법을 개발하며, 보다 대규모의 데이터셋을 확보하여 통계적 신뢰도를 높이는 방안을 모색하고자 한다.

참고문헌

Bao, Q. Z., Yang, Y. X., Li, Q., and Yang, H. C. (2025), Zero-shot instance segmentation for plant phenotyping in vertical farming with foundation models and VC-NMS, *Frontiers in Plant Science*, **16**, 1536226.

Buxbaum, N., Lieth, J. H., and Earles, M. (2022), Non-destructive plant biomass monitoring with high spatio-temporal resolution via proximal RGB-D imagery and end-to-end deep learning, *Frontiers in Plant Science*, **13**, 758818.

Chang, S., Lee, U., Hong, M. J., Jo, Y. D., and Jin-Baek, K. (2021), Lettuce growth pattern analysis using U-net pre-trained with Arabidopsis, *Agriculture*, **11**(9), 890.

Chen, T., Zhu, L., Deng, C., Cao, R., Wang, Y., Zhang, S., ... and Mao, P. (2023), Sam-adapter: Adapting segment anything in underperformed scenes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3367-3375.

Fan, Z., Yu, J. G., Liang, Z., Ou, J., Gao, C., Xia, G. S., and Li, Y. (2020), Fgn: Fully guided network for few-shot instance segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9172-9181.

Gulshan, V., Rother, C., Criminisi, A., Blake, A., and Zisserman, A. (2010), Geodesic star convexity for interactive image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3129-3136.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017), Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*,

2961-2969.
Islam, S., Reza, M. N., Chowdhury, M., Ahmed, S., Lee, K. H., Ali, M., ... and Chung, S. O. (2024), Detection and segmentation of lettuce seedlings from seedling-growing tray imagery using an improved mask R-CNN method, *Smart Agricultural Technology*, **8**, 100455.
Jiang, Y., Tan, Y., Ji, F., Su, D., Wang, S., Zhang, L., and Zhou, Q. (2024), CFIHL: A variety of chlorophyll a fluorescence transient image datasets of hydroponic lettuce, *Frontiers in Plant Science*, **15**, 1414324.
Kim, J. S. G., Moon, S., Park, J., Kim, T., and Chung, S. (2024). Development of a machine vision-based weight prediction system of butterhead lettuce (Lactuca sativa L.) using deep learning models for industrial plant factory, *Frontiers in Plant Science*, **15**, 1365266.
Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... and Girshick, R. (2023), Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4015-4026.
Kumar, V., Sharma, K. V., Kedam, N., Patel, A., Kate, T. R., and Rathnayake, U. (2024). A comprehensive review on smart and sustainable agriculture using IoT technologies, *Smart Agricultural Technology*, **8**, 100487.
Lee, H., Kim, S., and Kim, S. B. (2026), FLAsH: FastSAM-Based Lettuce Analysis for High-Density Cultivation, *Smart Agricultural Technology*, **14**, 102020.
Li, Y., Wang, D., Yuan, C., Li, H., and Hu, J. (2023), Enhancing agricultural image segmentation with an agricultural segment anything model adapter, *Sensors*, **23**(18), 7884.
Liu, J., Tai, X. C., and Luo, S. (2020), Convex shape prior for deep neural convolution network based eye fundus images segmentation, arXiv preprint arXiv:2005.07476.
Ma, J., He, Y., Li, F., Han, L., You, C., and Wang, B. (2024), Segment anything in medical images, *Nature Communications*, **15**(1), 654.
Minervini, M., Fischbach, A., Scharr, H., and Tsaftaris, S. A. (2016), Finely-grained annotated datasets for image-based plant phenotyping, *Pattern Recognition Letters*, **81**, 80-89.
Mirikharaji, Z. and Hamarneh, G. (2018), Star Shape Prior in Fully Convolutional Networks for Skin Lesion Segmentation, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11073 LNCS, 737-745. https://doi.org/10.1007/978-3-030-00937-3_84.
Riehle, D., Reiser, D., and Griepentrog, H. W. (2020), Robust index-based semantic plant/background segmentation for RGB-images, *Computers and Electronics in Agriculture*, **169**, 105201.
Veksler, O. (2008, October), Star shape prior for graph-cut image segmentation. In *European Conference on Computer Vision*, Berlin, Heidelberg: Springer Berlin Heidelberg, 454-467.
Weyler, J., Magistri, F., Seitz, P., Behley, J., and Stachniss, C. (2022), In-field phenotyping based on crop leaf and plant instance segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2725-2734.
Yan, X., Chen, Z., Xu, A., Wang, X., Liang, X., and Lin, L. (2019). Meta r-cnn: Towards general solver for instance-level low-shot learning. In *Proceedings of the IEEE/CVF international conference on computer vision*, 9577-9586.
Zhao, X., Ding, W., An, Y., Du, Y., Yu, T., Li, M., ... and Wang, J. (2023), Fast segment anything, arXiv preprint arXiv:2306.12156.
Zhao, X., Xie, J., Chen, S., and Liu, J. (2025), Convex Combination Star Shape Prior for Data-driven Image Semantic Segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 14068-14077.

저자소개

이혜승: 건국대학교 식품유통공학과에서 2024년 학사학위를 취득하고 고려대학교 산업경영공학과에서 석사과정에 재학 중이다. 연구분야는 image segmentation이다.

김성수: 경희대학교 산업공학과에서 2022년 학사학위를 취득하고 고려대학교 산업경영공학과에서 석박통합과정에 재학 중이다. 연구분야는 computer vision, vision-language model이다.

김성범: 고려대학교 산업경영공학부 교수로 2009년부터 재직하고 있으며, 인공지능공학연구소 소장, 기업산학연협력센터 센터장, 한국데이터마이닝학회 회장을 역임했다. 미국 University of Texas at Arlington 산업공학과에서 교수를 역임하였으며, 한양대학교 산업공학과에서 학사학위를 미국 Georgia Institute of Technology에서 산업시스템공학 석사 및 박사학위를 취득하였다. 인공지능, 머신러닝, 최적화 방법론을 개발하고 이를 다양한 공학, 자연과학, 사회과학 분야에 응용하는 연구를 수행하고 있다.