환경영향평가 Vol. 34, No. 6(2025), pp.624-641 J. Environ. Impact Assess. 34(6): 624-641(2025) ISSN 1225-7184
https://doi.org/10.14249/eia.2025.34.6.624

Technical Article

농업지 모니터링을 위한 2단계 U-Net 기반 이진분할 정확도 향상 방법

안향식 · 신영규 · 최태봉
국립환경과학원 기후국토환경연구과

Improving Binary Segmentation Accuracy for Agricultural Land Monitoring through Two-Stage U-Net Training

HyangSig Ahn* · YoungKyu Shin* · TaeBong Choi
Division of Climate, Land, and Environment, National Institute of Environmental Research

요약: 이진 분할(binary segmentation)은 특히 작은 대상 영역을 놓치면 신뢰성을 크게 훼손할 수 있는 응용 분야에서 거짓 음성(False Negatives, FNs) 문제를 지속적으로 겪고 있다. 특히 희소한 토지 피복을 탐지하는 작업에서는 FN이 실제 대상을 모델이 검출하지 못한 것으로, 모델 성능의 신뢰성을 저해하는 핵심적인 요소로 간주된다. 이에 본 연구에서는 1단계 학습(Stage 1)에서 예측된 FP 및 FN 영역을 자동으로 축적하고, 이를 보조 감독 신호(auxiliary supervisory signals)로 활용하는 두 번째 출력 헤드를 통해 2단계 학습(Stage 2)에서 재학습하도록 하는 2단계 U-Net 학습 전략을 제안한다. 이 접근법은 모델 구조를 변경하지 않고도 불확실하거나 모호한 영역에 대한 학습 반영을 강화할 수 있다는 점에서 실용적이다. 제안된 방법은 대한민국의 고해상도 항공영상을 활용한 희소 분포 삼계(인삼) 재배지 탐지에 적용되었으며, FN을 감소시키고 FP의 소폭 증가를 허용하는 대신 평균 교차합집합(Intersection over Union, IoU)을 약 0.10 개선하는 결과를 보였다. 본 전략은 학습 방식만을 수정하는 방법으로, 다양한 U-Net 변형 모델과 호환 가능하며 FN 억제와 예측 안정성이 중요한 다른 도메인에도 일반화할 수 있다. 이러한 결과는 누적된 오차 영역으로부터의 학습이 복잡한 실제 환경에서 보다 신뢰할 수 있는 분할 성능을 달성하는 효율적이고 실용적인 경로가 될 수 있음을 보여준다.

주요어: 항공 영상, 거짓 음성, 인삼 재배지 탐지, 농업 토지 피복 분할

Abstract: Binary segmentation faces persistent challenges from false negatives (FNs), particularly in applications where missing small target regions can undermine reliability. FNs are critical in sparse land-cover mapping tasks because they represent true targets that models fail to detect. Here, we propose a two-stage U-Net training strategy that automatically accumulates FP and FN regions from Stage 1 predictions and reuses them as auxiliary supervisory signals through a second output head during Stage 2 learning. This enhances the model's ability to revisit uncertain or ambiguous areas

First & Corresponding Author: HyangSig Ahn, Tel: +82-32-560-7699, E-mail: hsahn0609@korea.kr, ORCID: 0000-0003-0651-6887
Co-Authors: YoungGyu Shin, Tel: +82-32-560-7555, E-mail: shingeo@korea.kr, ORCID: 0009-0007-5174-0358
TaeBong Choi, Tel: +82-32-560-7541, E-mail: tbochoi@korea.kr, ORCID: 0009-0007-4550-6805
Received: 4 December, 2025. Revised: 17 December, 2025. Accepted: 18 December, 2025.
www.kci.go.kr

without any architectural modification. The method was applied to mapping sparsely distributed ginseng fields from high-resolution aerial imagery from South Korea. The approach substantially reduced FNs, tolerated a modest increase in FPs, and improved Intersection over Union (IoU) by approximately 0.10 across various training conditions. Because the strategy only modifies the training scheme and not the model architecture, it is compatible with U-Net variants and generalizable to other domains where FN suppression and stability are essential. The findings demonstrate that learning from accumulated error regions offers an efficient and practical pathway toward more reliable segmentation in complex real world environments.

Keywords : aerial imagery, two-stage training, false negatives, ginseng field mapping, agricultural land cover segmentation

I. 서론

이진 분할(binary segmentation)은 토지 피복 분류를 비롯한 다양한 분야에서 널리 활용되는 핵심 기술이다(Chen et al., 2022; Cardama et al., 2023; Lee et at., 2024s). 대표적인 인코더-디코더 구조의 픽셀 단위 분류 모델인 U-Net은 다양한 이미지 분할 작업에서 폭넓게 사용되어 왔다(Ronneberger et al., 2015; Chen et al., 2024; Cheng et al., 2025; Pang et al., 2024). 그러나 클래스 불균형, 시각적으로 유사한 패턴, 이미지 노이즈 등과 같은 문제는 FP(False Positives) 및 FN(False Negatives) 오류를 초래하여 모델 성능을 저하시킨다. 이러한 이유로, 이진 분할에서 U-Net의 정확도를 향상시키기 위한 다양한 연구가 수행되어 왔다. U²-Net (Qin et al., 2020)은 중첩된 U자형 구조를 통해 분할 성능을 개선하였으며, BASNet (Qin et al., 2019)은 정교한 경계 검출을 위해 경계 인식(boundary-aware) 손실 함수를 도입하였다. PraNet (Fan et al., 2020)은 의료 영상에서 용종을 더욱 효과적으로 분할하기 위해 reverse-attention 메커니즘을 활용하였다. ResUNet++ (Jha et al., 2019)과 DoubleU-Net (Jha et al., 2020) 역시 소규모 객체의 안정적인 탐지를 목표로 개발된 구조적 개선 모델들이다.

FN 오류는 적용 분야에 따라 심각한 결과를 초래할 수 있다. 예를 들어 원격탐사에서는 불법 건축물, 산불 피해 지역 또는 특정 농경지의 탐지를 놓칠 경우 행정적 판단 오류나 비효율적인 재난 대응으로 이어질 수 있다. 따라서 FN을 억제하는 것은 원격탐사와 토지 피복 분류를 포함한 다양한 이미지 분할 작업에서 핵심적 과제이다.

FN 오류를 줄이기 위한 접근법은 크게 네 가지로 구분된다. 첫째, 손실 함수 측면에서 Focal Loss (Lin et al., 2017)나 Tversky Loss (Abraham and Khan, 2019)와 같이 양성 클래스에 더 높은 가중치를 부여하거나 어려운 샘플에 집중하는 방법이 있다. 둘째, 데이터 측면에서는 소수 클래스 데이터를 증강하거나 패치 추출 단계에서 양성 영역을 반드시 포함하도록 설계하는 전략이 활용된다(Shorten and Khoshgoftaar, 2019). 셋째, 다중 규모 특징 추출이나 attention 메커니즘을 활용한 구조적 개선을 통해 희소하거나 소규모 대상을 효과적으로 탐지할 수 있다(Zhou et al., 2018; Oktay et al., 2019). 넷째, 2단계 학습(two-stage) 또는 후처리 기반 방법은 초기 분할 결과를 정제하거나 누락된 영역을 추가 탐지하기 위해 별도의 정제 모듈을 도입한다(Chen at al., 2018; Jiang et al., 2019).

이러한 접근법들은 FN 감소에 기여해 왔으나, 경계가 불확실하거나 대상 패턴이 매우 희소한 영역에서는 안정적인 성능을 유지하는 데 한계를 보인다. 이러한 영역에서는 이미지나 학습 시점(epoch)에 따라 유사한 영역이 서로 다르게 분류되는 불안정성이 발생한다. 결과적으로 정확도 변동성이 커지고, 신뢰할 수 없는 분할 결과를 초래할 수 있다. 이 문제는 손실 함수 조정이나 구조적 개선만으로는 충분히 해결되지 않으며, FN 억제와 동시에 예측 안정성을 향상시킬 수 있는 새로운 전략이 요구된다.

본 연구는 1단계 학습에서 발생한 FP 및 FN 영역이 모델이 안정적으로 학습하지 못한 불확실하거나 모호

www.kci.go.kr

한 영역을 의미하며, 이러한 영역을 후속 학습 단계에서 명시적으로 감독하면 정확성과 안정성을 동시에 향상시킬 수 있다는 가정에 기반한다. 이를 바탕으로 본 연구는 다음과 같은 핵심 질문을 제기한다. 첫째, 네트워크 구조나 손실 함수를 변경하지 않고 학습 과정을 재구성하는 것만으로 의미 있는 성능 향상을 얻을 수 있는가? 둘째, 1단계에서 생성된 FP+FN 오류 마스크가 2단계에서 신뢰성이 낮은 영역을 보완하는 데 효과적으로 활용될 수 있는가? 셋째, FN이 성능을 좌우하는 희소하고 소규모 대상에서 이러한 전략이 실질적인 이점을 제공할 수 있는가?

이에 본 연구에서는 원래의 U-Net 구조를 유지하면서 학습 과정만을 개선하는 간단하지만 효과적인 2단계 학습(two-stage training) 전략을 제안한다. 기존 방법들은 주로 손실 함수 조정이나 구조적 변경에 집중했던 것과 달리, 제안된 전략은 학습 과정 자체를 재설계하여 불확실하고 모호한 영역을 명시적으로 감독한다. 반면 이 방법은 1단계 예측 결과를 기반으로 FP와 FN 영역을 자동으로 결합한 오류 마스크를 생성하고, 이를 2단계 학습에서 보조 감독 신호(auxiliary supervisory signal)로 활용하여 네트워크를 재학습시킨다. 이 방법은 네트워크 구조를 변경하지 않으며, 기존의 FN 억제 방법과 독립적으로 결합하여 상호 보완적 시너지와 추가적인 성능 향상을 기대할 수 있다.

제안된 전략의 실용성은 항공 영상을 활용한 인삼 재배지 자동 분할 실험을 통해 검증되었다. 인삼은 고부가가치 작물로서 상대적으로 작은 규모의 재배지에 산발적으로 분포한다. 정확한 분포도 제작은 농업 계획, 생태 모니터링, 정책 수립에 중요한 역할을 한다. 그러나 수작업 기반의 매핑은 많은 시간과 인력 자원이 소요되며, 원격탐사 영상 기반 자동 분할의 필요성이 증가하고 있다. 인삼 재배지를 고해상도 위성영상과 객체 기반 분류 방법으로 분류한 사례가 존재한다(Lee et al., 2021; Yang et al., 2021). 그러나 인삼 재배지는 일반적인 토지 피복 분할과 달리 극도로 희소하고 소규모라는 특성을 가지므로, 기존 단일 학습 방식에서는 FN으로 인해 정확도가 쉽게 저하된다. 제안된 FP+FN 기반 2단계 학습 전략은 인삼 재배지 분할의 정확성과 일관성을 개선하였으며, 기존 단일 단계 학습 방식의 한계를 극복하고 작물 특화 매핑 과제에서의 활용 가능성을 보여주었다.

II. 연구방법

1. 실험장비

실험은 Linux Ubuntu 24.04 운영체제와 Python 3.11.9 (Python Sofrware Foundation, 2024) 환경에서 수행되었다. 하드웨어 구성은 AMD Ryzen 1700X 중앙처리장치(CPU), 48 GB 메모리(RAM), NVIDIA RTX 3090 Ti 그래픽처리장치(GPU)로 구성되었다. 딥러닝 구현 및 GPU 가속을 위해 PyTorch 2.7.1(Paszke et al., 2019)과 NVIDIA CUDA Toolkit 12.6을 사용하였다.

2. 연구 지역

본 연구에서는 한국국토지리정보원(NGII)가 생산한 경기도 및 전라남도 지역의 고해상도 항공정사영상을 학습 및 평가 데이터로 활용하였다. 학습에는 경기도 지역의 영상 5개를 사용하였으며, 평가는 전라남도 지역의 영상 2개를 사용하였다(Figure 1). 테스트 데이터는 모델의 일반화 성능을 적절히 평가하기 위해 학습에 사용되지 않은 전라남도 지역에서 선정하였다. 학습 및 테스트 데이터는 겹치지 않도록 256 × 256 픽셀 단위로 타일화하였으며, 최종적으로 학습에 8,436개, 테스트에 3,496개의 타일을 사용하였다. 정답(GT) 마스크는 인삼 재배지의 반복적인 패턴과 텍스처를 시각적으로 판독하여 수작업으로 구축하였다. 후보 영역이 실제 인삼 재배지인지 확인하기 위해 공공 지도 서비스를 활용하여 교차 검증을 수행하였다

인삼 재배지는 반복적인 격자 형태의 차광시설(shade structure) 존재 여부로 식별할 수 있다(Figure 2). 그러나 이러한 구조물은 색상, 질감, 외형 등이 일정하지 않아 시각적 일관성이 낮다. 또한 유사한 반복 패턴은 다른 농업 지역에서도 나타나기 때문에 신뢰성 있는 구분이 어렵다. 이러한 변동성은 주로 계절 변화, 재료의 노후화, 조명 조건의 차이 등에 기인하며, 이는 동일 클래스 내부에서 큰 시각적 다양성(intra-class variation)을 발생시킨다. 더불어 온실, 배수로, 비닐하우스 등 다양한 농업 구조물 또한 유사한 반복 패턴을 가지므로,

www.kci.go.kr

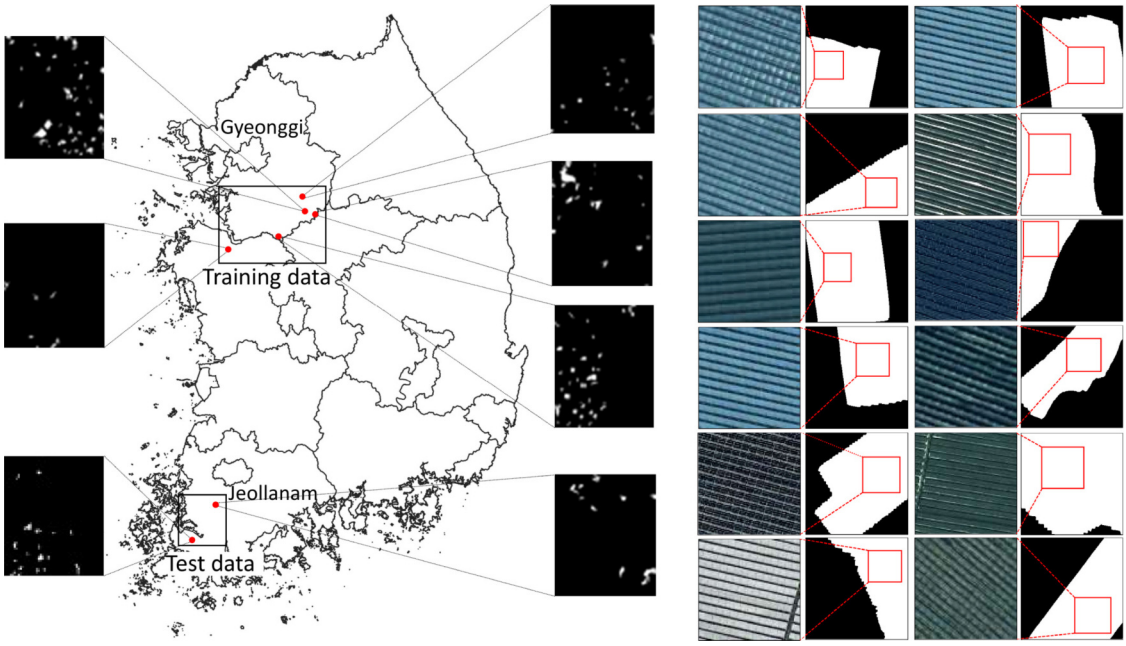

Figure 1. (a) Study area and corresponding ground-truth (GT) masks. White pixels represent ginseng fields, and black pixels denote the background used for training and testing. (b) Example 256 × 256 tiles GT label used for model training.

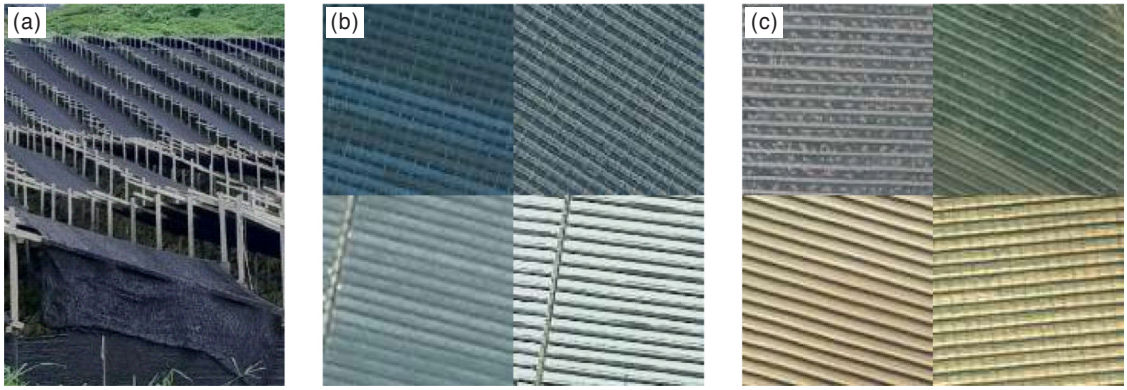

Figure 2. (a) Field photograph of ginseng farms; (b) representative shading patterns of ginseng cultivation structures; (c) examples of non-ginseng agricultural fields that visually resemble ginseng cultivation structures.

세그멘테이션 과정에서 클래스 간 혼동(inter-class confusion)의 가능성이 높아진다. 이러한 요인들은 고해상도 항공영상에서 인삼 재배지를 정확하게 분할하는 데 큰 어려움을 초래한다.

3. 모델 구성

본 연구에서는 인코더 단계 3개와 디코더 단계 3개로 구성된 U-Net 모델을 사용하였다. 각 단계는 두 개의 3 × 3 합성곱(convolution) 층으로 구성되며, 각 층 뒤에는 배치 정규화(batch normalization)와 ReLU 활성화 함수를 적용하였다. 인코더 단계 간 다운샘플링은 2 × 2 최대 풀링(max pooling)을 통해 수행되었으며, 디코더 단계 간 업샘플링은 2 × 2 전치합성곱(transposed convolution)을 사용하여 수행하였다. Skip connection

www.kci.go.kr

은 동일한 공간 해상도를 가진 인코더의 특성 맵을 채널 차원에서 디코더와 연결(concatenation)하는 방식으로 적용하였다. 최종 출력은 1 × 1 합성곱 층을 거쳐 생성되며, 시그모이드(sigmoid) 활성화 함수를 통해 이진 마스크로 변환된다. 드롭아웃(dropout)은 적용하지 않았으며, 입력과 출력의 공간 해상도를 동일하게 유지하기 위해 모든 합성곱 연산에서 패딩은 1로 설정하였다.

4. 학습 조건

학습에는 Adam 옵티마이저(Kingma and Ba, 2014)를 사용하였으며, 학습률은 0.001로 고정하고 배치 크기는 48로 설정하여 안정적이고 효율적인 수렴을 도모하였다. 손실 함수는 PyTorch 2.7.1에서 제공하는 BCEWithLogitsLoss를 사용하였으며, 이는 시그모이드 활성화 함수와 이진 교차 엔트로피(binary cross-entropy)를 통합하여 이진 분할에 효과적으로 활용된다. 제안된 2단계 학습 전략 외에도, SoftIoULoss (Huang et al., 2020), Focal Tversky Loss (Abraham and Khan, 2019), BCEIoULoss (Shorten and Khoshgoftaar, 2019)를 동일한 학습 조건에서 함께 평가하여 손실 함수 설계에 따른 성능 민감도를 비교하였다. 옵티마이저 구성, 가중치 초기화 방법, 손실 함수 등의 세부 하이퍼파라미터와 학습 설정은 Table 1에 요약하였다.

데이터셋 전체 픽셀 중 인삼 재배지에 해당하는 타깃 클래스의 비율은 2.2%에 불과하였다. 이로 인해, 학습 과정에서 하나의 배치(batch)에 타깃 픽셀이 포함된 타일이 선택될 확률이 매우 낮았으며, 아무런 조정 없이 학습을 진행할 경우 다수의 배치가 타깃 픽셀이 전혀 없는 상태로 구성되었다. 이러한 상황에서는 모델이 모든 픽셀을 비타깃으로 예측하더라도 전체 정확도가 높게 나타나는 기만적 현상이 발생하였다. 그 결과 FN 비율이 증가하고 IoU가 감소하며, 모델의 수렴이 제대로 이루어지지 않았다.

이 문제를 해결하기 위해, 전체 픽셀 대비 타깃 픽셀 비율이 0.1%를 초과하는 타일을 'positive tile'로 정의하였다. 이후 PyTorch 라이브러리의 WeightedRandomSampler 클래스를 활용하여 학습 시 positive tile의 선택 확률을 높이고, 하나의 배치를 구성할 때 positive tile과 negative tile의 비율이 약 2:1로 유지되도록 조정하였다. 이러한 샘플링 전략을 통해 매 학습 반복(iteration)마다 타깃 픽셀이 포함된 타일이 보다 안정적으로 배치에 포함되도록 하였다.

5. 학습 전략

본 연구에서는 거짓 음성(FN)을 억제하고 정확도를 향상시키기 위해 기존 U-Net 구조를 기반으로 한 2단계 학습(two-stage training) 전략을 적용하였다. 학습 과정은 기본 학습 및 오류 마스크 생성 단계인 Stage 1과, 오류 인지 기반 재학습 단계인 Stage 2로 구성된다. 먼저 Stage 1에서 충분한 학습이 진행된 후, 학습 데이터에서 발생한 오류 영역을 에폭(epoch)이 진행됨에 따라 추출하고 누적하였다. 각 에폭마다 누적된 마스크 영역의 면적을 계산하였다.

Stage 2에서는 원본 GT 마스크와 함께 누적된 오류 영역을 보조 감독 신호(auxiliary supervisory signals)로 사용하였으며, 이를 두 번째 출력 헤드(second output head)를 통해 학습에 활용하였다. Stage 1과 Stage 2 모두에서 매 에폭마다 학습 및 테스트 데이터셋에 대해 IoU, FP, FN 값을 계산하여 성능 변화를 모니터링하

Table 1. Training hyperparameters and model configuration.

Parameter	Value	Notes
Optimizer	Adam	
Learning rate	1e-3	
Beta1	0.9	Default
Beta2	0.999	Default
Epsilon	1e-8	Default
Batch size	48	
Weight initialization	He normal	Default

www.kci.go.kr

였다. IoU는 이미지 분할 성능을 평가하기 위한 대표적인 지표로, 식 (1)에 정의되어 있다(Rahman and Wang, 2016).

IoU는 다음과 같이 정의된다:

$$IOU = \frac{TP}{TP + FP + FN} \quad (1)$$

여기서,

- TP (True Positive): 실제 타깃을 타깃으로 정확히 예측한 픽셀 수
- FP (False Positive): 실제 배경을 타깃으로 잘못 예측한 픽셀 수
- FN (False Negative): 실제 타깃을 배경으로 잘못 예측한 픽셀 수

1) 1단계: 기본 학습(Baseline Training)

1단계에서는 기존 U-Net 구조를 활용하여 표준 이진 분할(binary segmentation) 학습을 수행하였다. 이 단계에 앞서 다양한 손실 함수를 비교하여 IoU 성능을 평가하였으며, 이후 오탐지 영역 추출 및 2단계 학습에서는 BCEWithLogitsLoss를 손실 함수로 사용하였다. Positive tile의 비율이 상대적으로 낮았기 때문에, 배치 내에서 positive tile과 negative tile의 비율을 약 2:1로 유지하기 위해 weighted random sampler를 적용하여 타깃 영역이 포함된 타일이 충분히 학습되도록 하였다. 학습은 100에폭(epoch) 동안 수행되었으며, 각 에폭마다 학습 데이터에 대해 loss, IoU, FP, FN 값을 기록하여 세부적인 모니터링이 가능하도록 하였다.

2) 1단계: 오류 마스크 생성(Error Mask Generation)

기본 학습을 100에폭 동안 수행한 후, 학습 데이터로부터 오탐지 영역을 추출하기 위해 추가적으로 100, 150, 200에폭의 확장 학습을 진행하였다. 이 단계에서는 FP 및 FN 영역을 반복적으로 탐지하고 누적하는 방식을 적용하였다. 매 에폭마다 전체 학습 데이터에 대해 FN 및 FP 영역을 세 차례 추출하여 누적하였으며, 누적된 오류 영역의 면적을 측정하였다. 또한, 학습의 진행에 따라 누적된 FP 및 FN 영역과 함께 FN, FP, IoU 값을 학습 및 테스트 데이터셋 각각에 대해 계산하여 기록하였다.

3) 2단계: 이중 채널 학습(Two-Channel Training)

2단계에서는 U-Net의 기본 구조를 그대로 유지하되 출력 채널 수를 두 개로 확장하였다. 첫 번째 출력 채널은 기존의 타깃 마스크를 예측하며, 두 번째 출력 채널은 1단계에서 생성된 FP 및 FN 결합 오류 마스크를 예측하도록 설계하였다. 각 출력 채널은 BCEWithLogitsLoss를 사용하여 독립적으로 학습되었다. 이 이중 채널 구조는 다중 클래스(multiclass) 학습 방식과는 다르며, 두 채널은 상호 독립적으로 학습되기 때문에 하나의 픽셀이 동시에 두 채널에서 타깃으로 분류될 수 있다. 반면, 다중 클래스 학습에서는 각 픽셀이 오직 하나의 클래스에 배타적으로 할당된다.

III. 결과

1. 1단계 학습

1단계 학습에 앞서 BCEWithLogitsLoss, SoftIoULoss, FocalTverskyLoss, BCEIoULoss 등 네 가지 손실 함수에 대해 성능을 비교하였다. 그 결과, BCEWithLogitsLoss가 가장 안정적인 IoU 성능을 보였다(Figure 3). 이에 따라, 오류 영역 마스킹 및 2단계 학습에서도 BCEWithLogitsLoss를 손실 함수로 선택하였다.

1단계 학습 동안 학습 데이터와 테스트 데이터에 대해 loss와 IoU 값을 기록하였다. 학습 loss는 약 30에서 시작하여 400에폭 이상 학습이 진행되는 동안 지속적

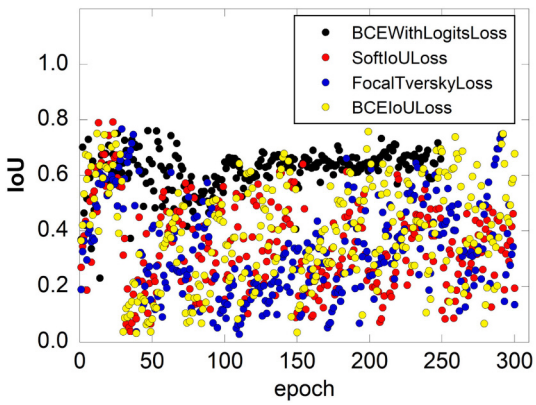

Figure 3. Prediction performance of a baseline U-Net trained with various loss functions, without applying the proposed two-stage strategy.

www.kci.go.kr

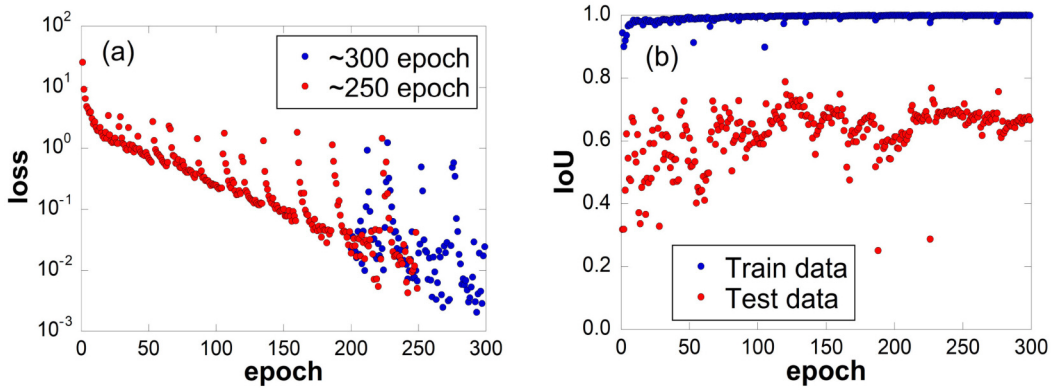

Figure 4. (a) Training and test loss on a logarithmic scale. Training loss steadily decreases with periodic fluctuations; (b) Training and testing the Intersection over Union (IoU). Training IoU exceeds 0.99 after approximately 100 epochs. Meanwhile, test IoU remains around 0.6 with vari-ation.

으로 감소하였으며, 약 100에폭 이후에는 1.0 이하로 떨어졌고 최종적으로 0.1 이하까지 감소하였다. 한편 학습 데이터의 IoU는 빠르게 증가하여 최종적으로 0.99를 초과하였으며, 이는 학습 데이터에 대해서 거의 완벽한 분할 성능을 보였음을 의미한다(Figure 4a).

반면, 테스트 데이터의 IoU는 이와 같은 향상을 보이지 않았으며, 전체 학습 과정에서 약 0.5에서 0.7 사이를 불규칙하게 변동하다가 최종적으로 약 0.7 수준으로 수렴하였다(Figure 4b). 이러한 학습 데이터와 테스트 데이터 간의 성능 격차는 일반화 능력이 제한적임을 나타낸다. 이는 모델이 학습 데이터의 공간적 패턴을 학습했다기보다는, 새로운 지역에서 적용 가능한 표현을 충분히 학습하지 못하고 단순히 훈련 데이터의 특정 패턴을 암기(memorization)했을 가능성을 시사한다. 즉, 기존의 단일 단계 학습 방식은 인삼 차광시설의 미세하고 반복적인 텍스처 특성을 충분히 학습하지 못했음을 의미한다. 따라서 더 높은 일반화 성능을 확보하고 모호한 패턴을 효과적으로 학습하기 위한 보다 강건한 학습 전략이 필요함을 시사한다.

손실 값을 로그 스케일로 나타냈을 때, 100에폭 이후를 포함하여 간헐적으로 상대적으로 높은 값의 이상치(outlier)가 관찰되었다(Figure 4a). 또한 학습 데이터의 IoU가 1.0에 근접했음에도 불구하고, 일시적인 감소가 반복적으로 발생하는 현상이 확인되었다(Figure 4b). 이는 100에폭 이후에도 오류 영역이 지속적으로 발생하고 있음을 의미한다. 따라서 100에폭 이후에 생성된 오류 영역을 누적하여 2단계 학습을 위한 추가 데이터로 활용하였다.

먼저 0~100에폭 구간에서는 오류 영역을 누적하지 않고 기본 학습을 수행하였다. 이후 100~200에폭 구간에서는 오류 영역을 누적하면서 단일 채널 구조의 1단계 학습을 지속하였다. 이때 누적된 오류 영역은 2단계 이중 채널 학습의 보조 입력 데이터로 사용되었으며, 이를 Case 1: A_100-200이라 정의하였다. 이어서 저장된 200에폭의 모델을 불러와 200~250에폭까지 1단계 학습을 재개하였으며, 이 구간에서 오류 영역은 기존(100~200에폭) 누적 영역에 추가로 누적되었다. 이렇게 100~250에폭 동안 누적된 오류 영역은 Case 2: A_100-250으로 정의되며, 이후 이 데이터를 활용하여 2단계 학

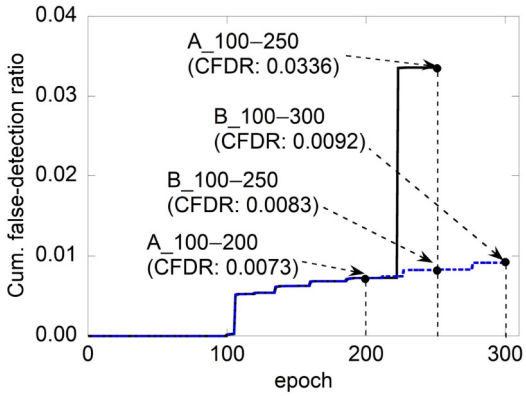

Figure 5. Increase in cumulative false-detection ratio over epochs during first-stage training. The second stage of training was conducted at the four points.

www.kci.go.kr

습을 수행하였다.

또한 동일한 200에폭 모델을 불러온 뒤 다시 200~250에폭 구간에서 학습을 재개하여 오류 영역을 추가로 누적하였고, 이를 Case 3: B_100-250로 정의하였다. 이후 250~300에폭 구간에서 오류 영역을 더욱 누적하고, 해당 누적 영역을 활용하여 2단계 학습을 수행하였으며, 이를 Case 4: B_100-300으로 명명하였다(Figure 5).

네 가지 조건에서 서로 다른 누적 오류 영역을 활용하여 각각 독립적인 2단계 학습을 수행하였다. 전체 데이터셋 대비 누적 오류 영역의 비율은 누적 오탐지 비율(Cumulative False-Detection Ratio, CFDR)로 정의되며 Figure 4에 나타내었다. CFDR은 전체 데이터셋 중 FP 또는 FN 오류 영역이 차지하는 면적 비중을 의미하며, 에폭이 진행됨에 따라 점진적으로 증가하였으나 간헐적으로 급격한 상승이 발생하는 구간도 확인되었다. 특히 100~250에폭 구간에서 오류를 누적한 Case A_100-250에서는 Epoch 223에서 CFDR이 급격히 증가하는 양상이 관찰되었다(Figure 5).

Figure 6은 Case A_100-250에서 누적된 오류 영역의 타일 예시를 제시한다. Figure 6a는 GT 상 positive로 이미 라벨링된 타일에서 발생한 누적 FN 영역을 나타낸다. 이 FN 영역은 실제 타깃 영역과 강하게 중첩되므로, 학습 성능 향상에 미치는 영향은 상대적으로 제한적일 수 있으며, 이에 대해서는 추가적인 분석이 필요하다.

반면 Figure 6b의 누적 FP 영역은 실제 인삼 재배지는 아니지만, 시각적으로 인삼 재배지와 유사한 패턴을 가진 지역을 강조한다. 모델이 이러한 모호한 영역을 비타깃(non-ginseng)으로 과도하게 학습할 경우, 실제 인삼 재배지에 대해서도 예측을 억제할 수 있어 FN 증가로 이어질 가능성이 있다. 이러한 작용은 학습 과정에서 관찰된 FN 값의 변동에도 영향을 미쳤을 것으로 판단된다.

2. 2단계 학습

2단계 학습에서는 1단계와 동일한 U-Net 구조를 사용하되, 이중 출력 채널 구조(dual-output channel)를 도입하여 학습을 수행하였다. 첫 번째 출력 채널은 기존의 GT 마스크(타깃 및 배경)를 예측하며, 두 번째 출력 채널은 1단계에서 생성된 누적 오류 마스크(FP + FN 및 배경)를 예측하도록 설계하였다. 각 채널은 BCEWithLogitsLoss를 사용하여 독립적으로 학습되었으며, 이를 통해 GT 기반 감독 신호와 오류 기반 감독 신호가 모두 적절히 반영되도록 하였다. 이러한 구성은 기존에 잘못 분류되었거나 모호한 영역을 별도의 채널에 할당하여 재학습하는 과정을 통해 해당 영역의 표현 학습을 강화할 수 있도록 지원한다. 최종 예측 결과는 두 채널에서 예측된 타깃 영역의 합집합(union)에 의해 결정되었다.

누적 오류 마스크를 보조 입력 채널로 활용하여 2단계 학습을 수행한 결과는 Figure 7에 나타내었다. 모든 실험 사례에서 정확도는 향상되었으며, 에폭 진행에

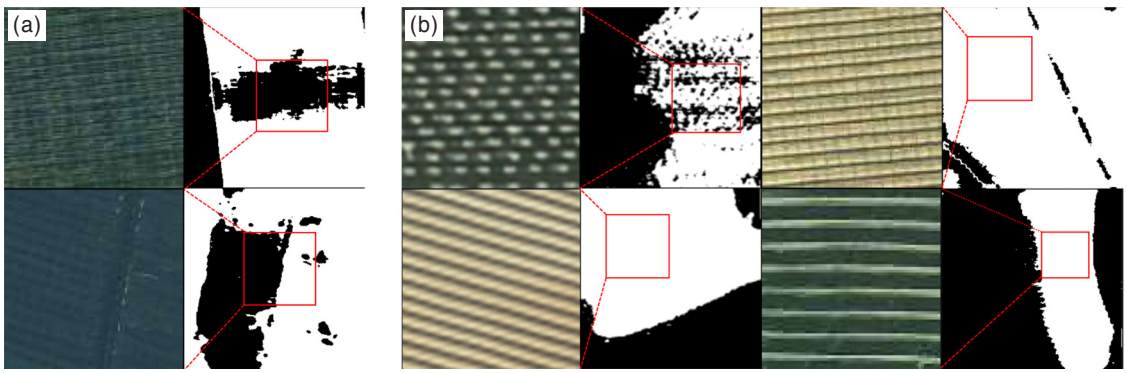

Figure 6. Accumulated error tiles from the training data. Color images show aerial photographs. Grayscale images highlight accumulated errors, with white representing error regions: (a) FN areas that overlap with ground-truth ginseng fields; (b) FP areas outside ginseng fields show visually similar patterns.

따른 변동성도 감소하여 수렴 안정성이 향상되었음을 확인할 수 있었다. 네 가지 사례 중 Case B_100-250이 가장 높은 최종 정확도(약 0.79)에 도달하였다. 학습 에폭이 증가할수록 정확도 변동 폭이 점차 줄어들었으며, 이는 향상된 학습 안정성과 성능 흔들림에 대한 저항성이 강화되었음을 의미한다.

1단계 학습에서 관찰된 IoU의 진동(oscillation)은 오류 영역의 불안정한 분류로 인해 발생한 것으로 판단된다. 이는 모델이 시각적으로 모호한 영역에 대해 반복적으로 결정 경계를 변경했기 때문으로 해석된다. 반면 2단계 학습에서는 이러한 오류 영역이 추가 출력 채널을 통해 보조 타깃(auxiliary targets)으로 재정의되면서 FN의 수가 지속적으로 감소하였다. 동시에 FP의 수는 증가하였는데, 이는 모델이 타깃 클래스와 시각적으로 유사한 미세 패턴에 대해 더욱 민감하게 반응하기 시작했음을 시사한다. 이러한 변화는 IoU의 성능 향상과 변동성 감소에 기여하여 보다 안정적인 수렴을 가능하게 한 것으로 판단된다.

딥러닝 모델에서는 과적합(overfitting)이 일반적으로 발생하는 문제이다(Salman and Liu, 2019; Srivastava et al., 2024). 그러나 400에폭에 걸친 본 실험에서는 과적합으로 인한 성능 저하가 뚜렷하게 나타나지 않았다. 이는 제안된 방법이 실험 조건에서 안정적인 일반화 성능을 유지했음을 의미한다.

Figure 8은 1단계와 2단계 학습에서 각각의 에폭 단위 IoU 값에 대해 이후 10에폭 구간을 기준으로 계산된 이동 표준편차(moving standard deviation)를 나타낸다. 이 값은 학습 과정에서 IoU 변동의 지역적(local) 안정성을 정량적으로 확인하기 위한 지표로, 각 슬라이딩 윈도우 내에서의 성능 변동 정도를 보여준다. 그림에서 확인되듯이, 2단계 학습의 IoU 값은 1단계 학습보다 더 밀집되어 있으며 변동 폭도 작아, 학습 안정성 및 수렴 일관성이 크게 향상된 것으로 나타났다. 이는 제안된 2단계 학습 전략이 분할 정확도를 향상시킬 뿐만 아

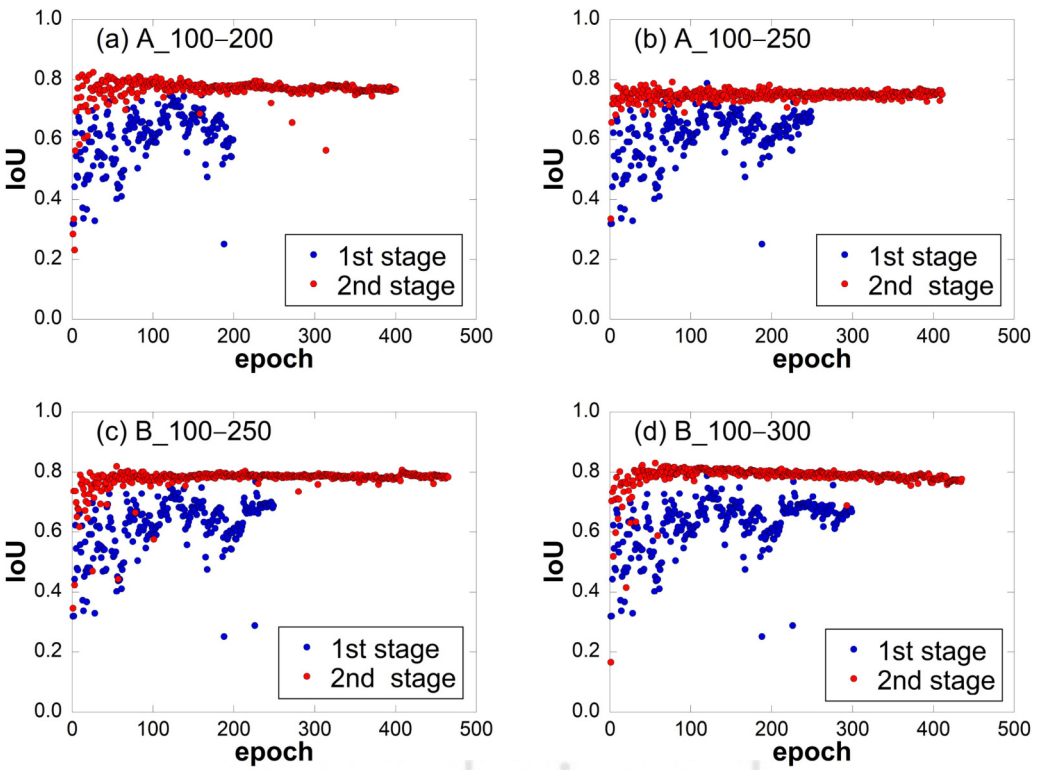

Figure 7. Prediction performance in the second-stage training: (a) Case 1: A_100-200; (b) Case 2: A_100-250; (c) Case 3: B_100-250; and (d) Case 4: B_100-300.

www.kei.ge.kr

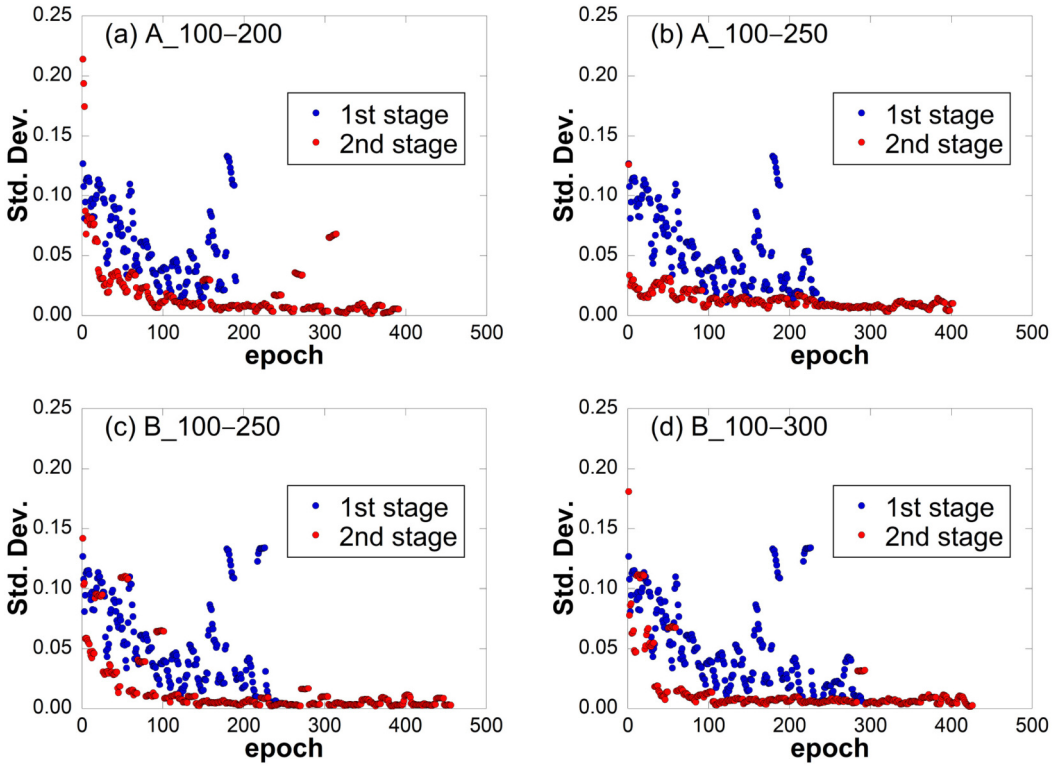

Figure 8. Standard deviation of IoU of the next ten epochs (a) Case 1: A_100-200; (b) Case 2: A_100-250; (c) Case 3: B_100-250; and (d) Case 4: B_100-300.

니라, 에폭 간 성능 변동을 줄여 보다 신뢰할 수 있고 일관된 최적화 경로를 형성함을 보여준다.

Figure 9는 학습 에폭에 따른 FN 비율의 변화를 나타낸다. 각 에폭에서 FN 및 FP 면적을 TP 면적과 비교하여 그 비율을 계산함으로써, 학습 과정에서 오류 분포가 어떻게 변화하는지를 평가하였다. 이 지표는 모델이 시간에 따라 올바른 탐지와 오분류 감소 간의 균형을 어떻게 조정해 나가는지를 파악하는 데 유용하다.

2단계 학습에서는 1단계에 비해 FN 비율이 감소하였으며, 이는 제안된 접근 방법이 누락된 영역을 효과적으로 복원(recovering missed regions)하고 FN을 억제하는 데 효과적임을 보여준다(Figure 9). 이러한 결과는 2단계 학습 메커니즘이 1단계에서 간과되었던 영역을 성공적으로 재학습하고, 인삼 재배지와 관련된 미세한 공간적 패턴(spatial patterns)을 포착하는 모델의 능력을 강화했음을 나타낸다. 이는 희소 타깃 분할 문제에서 특히 도전적인 과제라는 점에서 의미가 있다.

이러한 효과는 1단계에서 충분한 오류 정보가 누적된 경우에 가장 두드러지게 나타난다. 누적된 오류 정보는 이후 학습 과정에서 모호한 타깃 특성을 보다 효과적으로 학습하도록 지원한다. 특히 1단계에서 가장 높은 누적 오탐지 비율(CFDR = 0.32)을 기록한 Case A_100-250은 모든 경우 중 가장 낮은 FN 비율을 보였으며, 약 0.06 수준까지 수렴하였다. 이러한 감소는 2단계 학습 전략이 부분적으로 관측되거나 약하게 표현된 차광 구조(shade structure) 패턴에 대한 모델의 민감도를 정교하게 향상시킬 수 있음을 보여준다.

또한 전반적인 FN 감소 추세는, 누적된 오류 정보가 학습 안정성을 해치지 않으면서도 재검출 능력(recall)을 향상시킬 수 있는 유효한 감독 신호(supervisory signal)로 작용했음을 의미한다. 이러한 결과는 제안된 2단계 학습 방법이 희소하고 시각적으로 모호한 농업 객체를 포함한 이진 분할 문제에서 실용적이고 강건한 개선 효과를 제공함을 뒷받침한다.

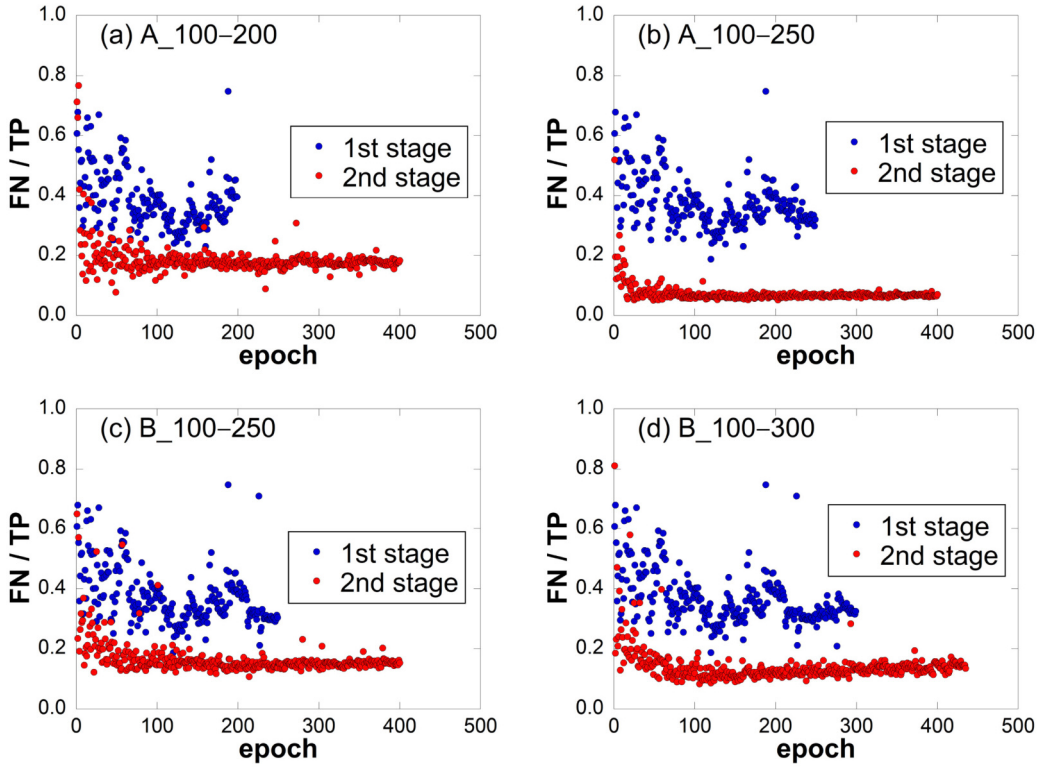

Figure 9. Difference in FN between the first and second stages: (a) Case 1: A_100-200; (b) Case 2: A_100-250; (c) Case 3: B_100-250; and (d) Case 4: B_100-300

Figure 10은 학습 과정에서의 FP 비율 변화를 나타낸다. Figure 9에서 보인 FN 감소 경향과는 달리, FP 비율은 2단계 학습에서 일관적으로 증가하는 모습을 보였다. 이는 제안된 접근법의 본질적인 트레이드오프(trade-off)를 반영하는 것으로, 모델이 이전에 놓쳤거나 모호한 영역을 복원하기 위해 더욱 적극적으로 탐지하려는 방향으로 민감도가 조정되었기 때문이다. 모델이 불확실하거나 약하게 표현된 시각적 단서를 더 넓게 포착하게 되면서, 실제 인삼 재배지가 아닌 영역에서도 타깃과 유사한 패턴에 반응하여 FP가 증가하는 현상이 나타났다.

이 현상은 모델의 결정 경계가 누락을 최소화하는 방향으로 조정되었음을 시사한다. 그 결과, 모델은 도전적인 환경에서도 더 높은 재검출 성능(recall)을 확보하도록 민감도가 증가한 것으로 이해할 수 있다.

네 가지 실험 사례 간 FP 증가의 정도는 서로 다르게 나타났다. 특히 1단계 학습에서 더 많은 오류 정보가 누적된 경우, 2단계 학습에서 보다 강한 보정 신호(corrective signal)가 작용하면서 FP 증가 현상이 더욱 뚜렷하게 나타났다. 특히 Case A_100-250에서는 2단계 학습에서 FP 비율이 0.20을 초과하며 가장 큰 증가폭을 보였다. 이 사례는 FP 증가가 가장 컸음에도 불구하고 FN 감소 또한 가장 두드러졌으며, 두 지표 간 균형 역시 가장 양호하게 유지되었다.

제안된 2단계 학습 전략은 FP의 통제된 증가를 유발하지만, FN의 현저한 감소 효과가 이 단점을 충분히 상쇄한다. 그 결과 IoU 향상으로 이어지며, 희소하게 분포된 타깃을 복원(recall)하는 데 더 유리한 성능을 나타낸다. 이는 복잡한 농업 환경에서 드물게 나타나는 타깃 객체의 탐지를 우선시할 경우, 일정 수준의 FP 증가는 불가피하지만 수용 가능한 비용(trade-off)임을 시사한다. 즉, 실제 타깃을 놓치는 것(FN)이 일시적인 오탐지(FP)보다 더 큰 위험을 초래한다는 점에서, FN 감소는 분할 성능 신뢰도를 결정하는 핵심 요소로 간주

www.kci.go.kr

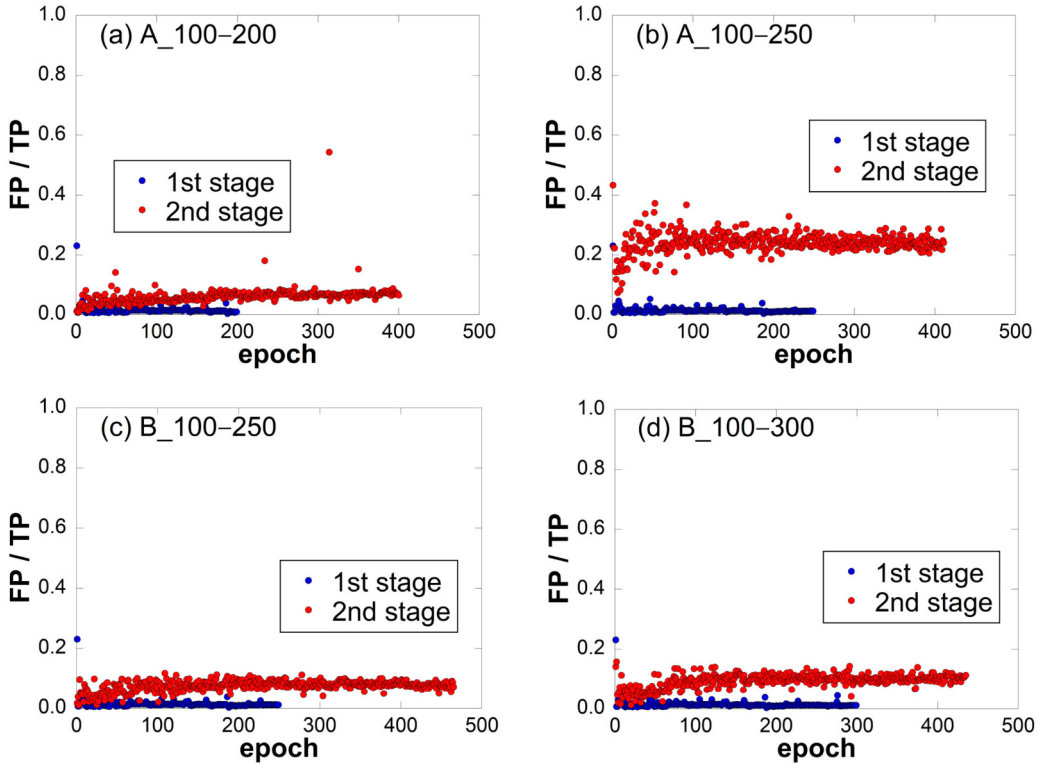

Figure 10. Difference in FP between the first and second stages. (a) Case 1: A_100-200; (b) Case 2: A_100-250; (c) Case 3: B_100-250; and (d) Case 4: B_100-300.

될 수 있다.

Figure 11은 제안된 2단계 학습 프레임워크를 통해 도출된 인삼 재배지 분류 결과를 나타낸다. Figure 11a는 GT 마스크이며, Figure 11b와 11c는 각각 1단계 및 2단계 모델의 예측 결과이다. Figure 11c에서는 GT 기반 예측 영역을 흰색, 누적 오류 마스크에 의해 예측된 영역을 빨간색, 두 영역이 중첩된 부분을 파란색으로 표시하였다. Figure 11d에서는 전체 영역이 흰색으로 표현되어 있으며, 2단계 학습 결과가 1단계 모델보다 GT에 더 근접한 예측을 생성함을 확인할 수 있다.

2단계 모델은 1단계 학습에서 탐지하지 못했던 인삼 재배지 영역을 성공적으로 식별하며, 희소하게 분포된 농업 타깃에 대한 민감도가 향상되었음을 보여준다. 이는 누적 오류 정보를 추가 학습에 활용함으로써, 기존 모델이 쉽게 간과하던 미세한 패턴을 효과적으로 포착하도록 네트워크의 표현 학습 능력이 강화되었음

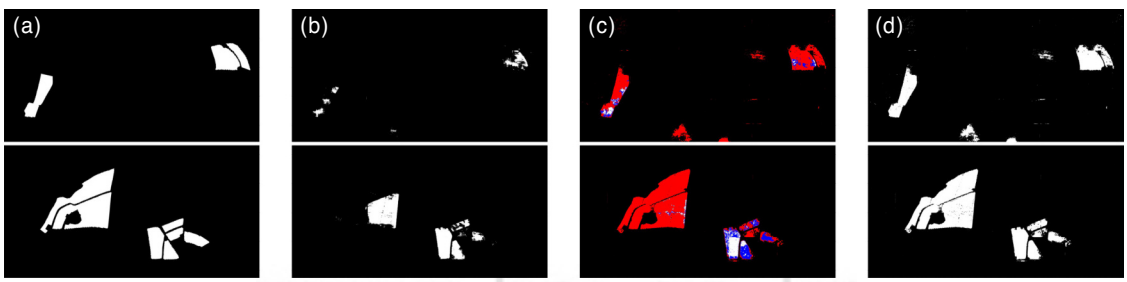

Figure 11. Difference in FP between the first and second stages. (a) Case 1: A_100-200; (b) Case 2: A_100-250; (c) Case 3: B_100-250; and (d) Case 4: B_100-300.

www.kci.go.kr

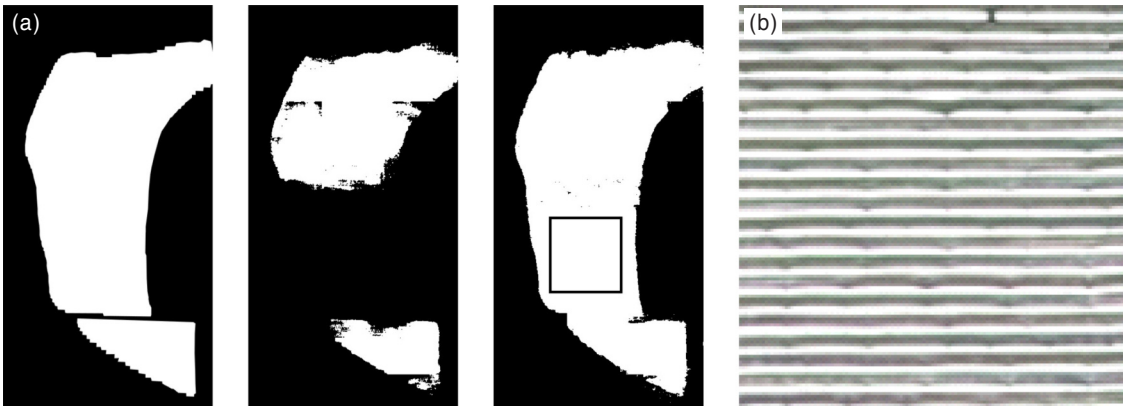

Figure 12. GT regions that were missed by the Stage 1 model but successfully detected after Stage 2 training: (a) ground truth (GT); (b) Stage 1 prediction; (c) Stage 2 prediction; and (d) corre-sponding zoomed area from the aerial image, showing visually ambiguous shading structures where the vertical grid pattern is subtle.

을 의미한다. 그러나 이러한 개선은 모호하거나 약하게 표현된 특징을 복원하려는 과정에서 FP의 증가를 동반하는 트레이드오프(trade-off)를 가져온다. 이는 희소 타깃 분할 문제에서 재검출(recall) 향상과 정밀도(precision) 유지 간 균형이 필요함을 보여주며, 향후 FP를 억제하면서도 향상된 민감도를 유지하기 위한 정교한 개선 전략의 필요성을 시사한다.

Figure 12는 테스트 데이터에서 1단계 모델이 탐지하지 못했지만, 제안된 2단계 학습을 통해 성공적으로 검출된 인삼 재배지 영역을 보여준다. Figure 12a는 GT, Figure 12b는 1단계 모델의 예측 결과, Figure 12c는 2단계 모델의 예측 결과이다. Figure 12d에서 볼 수 있듯이, 2단계에서 새롭게 검출된 영역은 육안으로도 인식하기 어려울 정도로 시각적 표현이 약한 특성을 가지며, 이는 제안된 학습 전략이 미세하고 불분명한 재배지 패턴에 대한 탐지 민감도를 효과적으로 향상시켰음을 보여준다.

Figure 13은 테스트 데이터 예측 과정에서 관찰된 FP 사례를 나타낸다. Figure 13a에서는 항공영상이 인삼 재배지와 유사한 패턴을 보이지만, 자세히 확인해 보면 실제로는 개방된 농경지(open farmland)임을 알 수 있다. 반면 Figure 13b의 영역은 색상과 질감이 인삼 차광시설의 특성과 매우 유사하여 실제 인삼 재배지와 시각적으로 구분하기 어렵다. 이 영역은 육안으로 정밀하게 관찰하더라도 모호한 시각적 특성을 보인다. 이는 해당 영역이 스펙트럴 및 구조적 특성이 상대적으로 약하게 표현된 작은 규모의 인삼 재배지일 가능성이 있으며, 이러한 미세한 표현 특성으로 인해 수작업 라벨링 과정에서 누락되었을 가능성을 시사한다.

Figure 11에서 제시한 시각적 비교 결과와 함께, Table

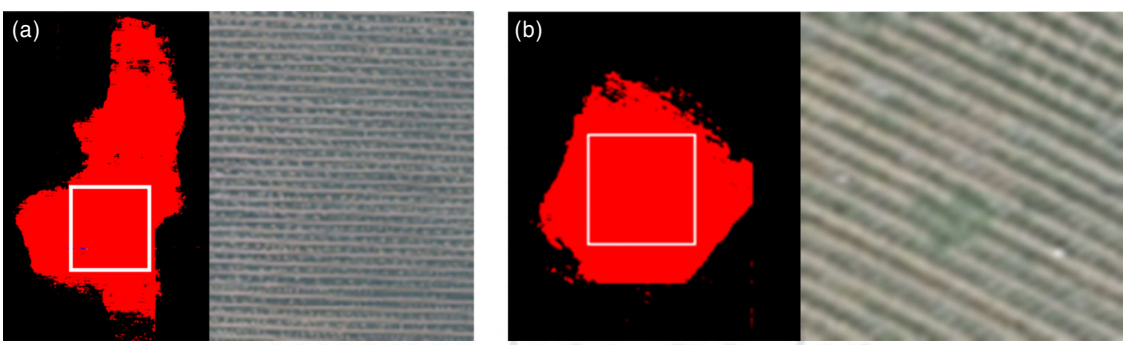

Figure 13. FP regions in the test data and magnified patterns: (a) A farmland area where the vertical grid lines are not continuous. (b) A visually ambiguous region with faint vertical striping.

www.kci.go.kr

Table 2. First- and second-stage training results under different error accumulation ranges

Case	Accu. range (epochs)	CFDR	Avg. IoU (1st stage)	Avg. IoU (2nd stage)	ΔIoU	Avg. FN/TP (1st stage)	Avg. FN/TP (2nd stage)	ΔFN/TP
A_100-200	100-200	0.0073	0.61	0.77	0.16	0.38	0.19	0.19
A_100-250	100-250	0.0336	0.62	0.75	0.13	0.38	0.07	0.31
B_100-250	100-250	0.0083	0.62	0.78	0.16	0.37	0.16	0.31
B_100-300	100-300	0.0092	0.63	0.79	0.16	0.37	0.14	0.23

2는 서로 다른 오류 누적 범위에 대해 1단계 및 2단계 학습의 성능을 요약하여 보여준다. CFDR은 누적 에폭 수가 증가함에 따라 함께 증가하였는데, 이는 학습 IoU가 높은 수준에 도달한 이후에도 FP 및 FN 영역이 지속적으로 발생했음을 의미한다. 그러나 이러한 오류 영역을 2단계 학습에 반영함으로써 분할 정확도는 모든 경우에서 일관되게 개선되었다. 평균 IoU는 약 0.13-0.16 향상되었으며, 평균 FN/TP 비율은 최대 0.31 감소하여 제안된 2단계 학습 전략이 FN을 효과적으로 억제하고 학습 안정성을 높였음을 보여준다. 네 가지 실험 사례 중 B_100-300은 가장 높은 IoU(0.79)와 낮은 FN/TP 비율 중 하나를 기록하였다. 이는 오류 누적 기간을 확장함으로써 모델의 강건성과 일반화 성능을 더욱 향상시킬 수 있음을 시사한다.

네 가지 실험 사례 중 B_100-250이 수렴 시점에서 가장 높은 IoU(약 0.79)를 기록하였다. 한편, B_100-300은 초기 학습 구간에서 비교적 높은 정확도를 유지하여 평균 IoU가 소폭 더 높았으나, 학습 후반으로 갈수록 성능이 점진적으로 감소하는 경향을 보였다. 이러한 감소 현상의 원인은 명확하지 않지만, 오류 영역의 과도한 누적 또는 후기 학습 단계에서의 과적합(overfitting)과 관련되었을 가능성이 있다. 결과적으로, B_100-250이 최종 수렴 성능 측면에서 가장 우수한 성능을 달성하였다.

IV. 논의

1. 손실 함수 기반 FN 억제와의 비교

본 연구에서는 초기 단계에서 FN 억제를 위해 Focal TverskyLoss (Abraham and Khan, 2019)와 같은 손실 함수 기반 전략을 적용하였다(Figure 3). 이 방법은 간헐적으로 IoU가 0.77까지 도달하기도 했으나, 전체적인 학습 과정은 수렴 안정성이 낮았고 에폭 간 정확도의 변동 폭이 상당히 컸다. 이에 반해, 제안된 2단계 학습 전략은 손실 함수를 변경하지 않고 감독 신호를 강화함으로써 더욱 일관된 성능을 보였으며, FN을 효과적으로 억제하면서도 IoU의 안정성을 유지하였다.

FocalTverskyLoss는 검출되지 않은 타깃 영역(FN)에 높은 패널티를 부여하여 FN을 감소시키는 데 효과적이다. 그러나 이러한 접근만으로는 모호한 경계 영역에서의 비균질한 분류 문제를 해결하기 어렵다. 특히 원격탐사 토지피복 분류에서는 다양한 공간적 특성과 시계열 변화를 포함하는 대규모 이질적 데이터가 활용되기 때문에, 학습의 안정성과 일관성을 확보하는 것이 중요하다. 따라서 본 연구는 단순히 손실 기반 FN 억제 성능을 넘어서, 학습 안정성과 신뢰성을 향상하는 새로운 방향을 제시한다는 점에서 의의가 있다.

2. 오류 누적 특성과 모호성 처리 성능

학습 데이터에서 누적된 오류 영역은 테스트 데이터에서도 유용하게 작용하였다. 단일 에폭에서 발생하는 예측 오류는 매우 미미하지만, 수백 개의 에폭을 거치며 누적된 작은 오류들이 상당한 의미 있는 영역으로 집적되었고, 이를 보조 감독 신호로 활용했을 때 테스트 성능이 향상되었다. 일반적으로 학습 IoU가 0.99 이상 도달하면 과적합(overfitting)으로 간주되지만(Srivastava et al., 2014), 본 연구에서는 이러한 과적합 상태에서 발생한 잔여 오류(error residual)가 오히려 일반화에 긍정적인 영향을 미친다는 점을 실험적으로 확인하였다. 이는 누적된 오분류 결과가 학습에 유용한 정보들을 포함하고 있음을 시사한다.

고해상도 항공영상에서는 불명확한 경계 영역이나 육안 식별이 어려운 모호한 텍스처 패턴이 빈번하게 나타난다. 이 경우 사람의 수작업 라벨링과 딥러닝 모델

www.kci.go.kr

모두에서 분류 불확실성이 존재할 수 있다. 제안된 듀얼 채널 기반 피드백 구조는 이러한 영역을 보조 학습 대상으로 명시적으로 재학습함으로써, 모호한 패턴에 대한 강건성을 향상시킨다. 이 접근은 겉보기에는 멀티클래스 학습과 유사하지만, 실제로는 각 채널에서 동일한 픽셀이 동시에 참여할 수 있으며, 공유된 가중치 구조를 통해 역전파(backpropagation) 과정에서 상호 보완적 특징 학습이 이루어진다는 점에서 근본적으로 다르다(Gooefellow et al., 2016; Rumelhart et al., 1986; Caruana, 1997). 이러한 구조는 과적합을 완화하고 표현 일반화를 강화하는 데 기여한다.

3. 한계점

본 연구에서 사용된 데이터는 극도로 희소한 농업 타깃을 포함하는 특성을 갖는다. 추후 연구에서는 객체가 더 넓은 공간적 범위에 분포하는 INRIA Aerial Image Labeling Dataset (Maggiori et al., 2017)과 같은 복합 환경에서의 성능을 검증할 필요가 있다. 다른 데이터셋에서도 IoU가 0.99 이상으로 수렴하는지, 오류 누적이 FN 감소에 계속 기여하는지, 아니면 라벨 노이즈(label noise)가 지배적인 영향을 미치며 성능 저하를 유발하는지도 분석해야 한다. 한편, 제안된 방법은 FN을 상당히 억제했지만 FP는 증가하였기 때문에, FP를 최소화하기 위한 후처리(post-processing) 기법의 추가 적용 가능성도 고려해야 한다.

분할 학습에서의 라벨 오류 감지 연구는 불확실성 추정이나 손실 기반 이상치 탐색을 통해 주로 데이터셋 품질 개선을 목표로 해왔다. Rottmann and Reese (2023)는 불확실성 기반 라벨 오류 탐지 방법을 제안했고, Schubert et al. (2024)는 손실 신호를 분석하여 수동 라벨링 오류를 찾아냈다. Suárez et al. (2025)는 농업 영상에서 자동 라벨 수정 기법을 적용하였다. 그러나 이러한 연구들은 공통적으로 '데이터 품질 개선'을 목표로 하였으며, 누적된 예측 오류를 학습 과정에 직접 활용하는 접근은 아니었다.

이에 비해 본 연구는 누적된 FP 및 FN 영역을 보조 감독 신호로 활용하여 모델 학습 과정 자체를 개선한 최초의 시도라는 점에서 차별성을 가진다. 특히 고해상도 항공영상과 같이 희소 타깃 및 불분명한 시각적 패턴이 존재하는 상황에서, 구조 변경 없이 학습 안정성과 FN 억제를 동시에 달성한 간단하고 범용적인 방법이라는 점에서 실용성이 높다.

V. 결론

본 연구에서는 U-Net 기반 이진 분할 성능을 향상하기 위한 아키텍처 비의존적 2단계 학습 전략을 제안하였다. 제안된 방법은 1단계 학습 과정에서 누적된 오분류 영역(FP + FN)을 보조 감독 신호로 활용함으로써, 모델이 모호한 영역을 더욱 정확하게 학습하고 FN을 효과적으로 감소시키도록 유도한다. 이를 희소하게 분포된 인삼 재배지를 대상으로 하는 고해상도 항공영상 분할 작업에 적용한 결과, 기존 U-Net 대비 IoU가 향상되었으며, 수렴 안정성 또한 크게 개선되었다.

본 연구는 단일 손실 함수와 단일 아키텍처, 그리고 희소 타깃 기반 데이터셋을 중심으로 수행되었으므로, 향후 연구에서는 보다 복합적인 벤치마크 데이터셋과 다양한 모델에 대해 일반화 가능성을 검증할 필요가 있다. 또한 오류 누적 조건의 체계적인 최적화와 후처리 기법의 적용을 통해 성능 향상 여지가 남아 있다.

이러한 결과는 반복적인 오류 기반 감독(iterative error-guided supervision)이 아키텍처 변경 없이도 분할 모델의 안정성과 정확도를 효과적으로 향상시킬 수 있음을 시사한다. 제안된 전략은 기존 손실 함수 및 모델 구조와 쉽게 결합 가능한 유연하고 계산 효율적인 학습 프레임워크로서, 원격탐사, 의료 영상, 농업 모니터링 등 FN 억제와 경계 모호성 해결이 중요한 다양한 분야에 활용될 잠재력을 지닌다.

사사

본 논문은 환경부의 재원으로 국립환경과학원의 지원을 받아 수행하였습니다(NIER-2023- 01-01-103).

References

Abraham, S., Khan, N. M., (2019). A Novel Focal Tversky Loss Function with Improved Attention U-Net

www.kci.go.kr

for Lesion Segmentation. In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2019), 2019 May 28-June 1; Venice, Italy. https://doi.org/10.1109/ISBI.2019.8759329

Cardama, F. J., Heras, D. B., Argüello, F. (2023). Consensus Techniques for Unsupervised Binary Change Detection Using Multi-Scale Segmentation Detectors for Land Cover Vegetation Images. *Remote Sensing, 15*, 2889. https://doi.org/10.3390/rs15112889

Caruana, R. (1997). Multitask Learning. *Machine Learning, 28*, 41-75. https://doi.org/10.1023/A:1007379606734

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A. L., (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*, 834-848. https://doi.org/10.1109/TPAMI.2017.2699184

Cheng, H., Zhang, Y., Xu, H., Li, D., Zhong, Z., Zhao, Y., Yan, Z. (2025). MSGU-Net: A Lightweight Multi-Scale Ghost U-Net for Image Segmentation. *Frontier Neurorobot, 18*, 1480055. https://doi.org/10.3389/fnbot.2024.1480055

Chen, H., Han, Y., Yao, L., Wu, X., Li, K., Yin, J. (2024), MS-UNet: Multi-Scale Nested UNet for Medical Image Segmentation with Few Training Data Based on an ELoss and Adaptive Denoising Method. *Mathematics, 12*, 2996. https://doi.org/10.3390/math12192996

Chen, Z., Wu, Z., Gao, J., Cai, M., Yang, X., Chen, P.; Li, Q. (2022). A Convolutional Neural Network for Large-Scale Greenhouse Extraction from Satellite Images Considering Spatial Features. *Remote Sensing, 14*, 4908. https://doi.org/10.3390/rs14194908

Fan, D. P., Ji, G. P., Zhou, T., Chen, G., Fu, H., Shen, J., Shao, L. (2020). PraNet: Parallel Reverse Attention Network for Polyp Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2020), 2020 October 4-8; Lima, Peru (Virtual Conference). https://doi.org/10.48550/arXiv.2006.11392

Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning; MIT Press: Cambridge, MA, USA.

Huang, Y., Tang, Z., Chen, D., Su, K. (2020). Batching Soft IoU for Training Semantic Segmentation Networks. *IEEE Signal Processing Letters, 27*, 66-70. https://doi.org/10.1109/LSP.2019.2956367

Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P., Johansen, H. D. (2020). Double U-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. In Proceedings of the 33rd IEEE International Symposium on Computer-Based Medical Systems (CBMS 2020), 2020 July 28-30; Rochester, MN, USA. https://doi.org/10.1109/CBMS49503.2020.00111

Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., De Lange, T., Halvorsen, P., Johansen, H. D. (2019). ResUNet++: An Advanced Architecture for Medical Image Segmentation. In Proceedings of the 2019 IEEE International Symposium on Multimedia (ISM 2019), 2019 December 4-6; San Diego, CA, USA. https://doi.org/10.48550/arXiv.1911.07067

Jiang, Z., Ding, C., Liu, M., Tao, D. (2019). Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In Proceedings of the International MICCAI BrainLesion Workshop (BrainLes 2019), Shenzhen, China, 13-15 October 2019. https://doi.org/10.1007/978-3-030-46640-4_22

Kingma, D. P., Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv. https://doi.org/10.48550/arXiv.1412.6980

Lee, D. H.; Cho, M. -A.; Park, J. -M.; Kang, M. -S.; Kim, S. -H. (2021). UAV, a Farm Map, and Machine

www.kci.go.kr

Learning Technology Convergence Classification Method of a Corn Cultivation Area. *Agronomy, 11*, 1554. https://doi.org/10.3390/agronomy11081554

Lee, K., Lee, H., Park, J., Hwang, J. Y. (2024). Fine-Grained Binary Segmentation for Geospatial Objects in Remote Sensing Imagery via Path-Selective Test-Time Adaptation. IEEE Trans. Geosci. *Remote Sensing, 62*, 5623216. https://doi.org/10.1109/TGRS.2024.3378311

Lin, T. Y., Goyal, P., Girshick, R., He, K., Dollár, P. (2017). Focal Loss for Dense Object Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017), 2017 October 22-29; Venice, Italy. https://doi.org/10.1109/ICCV.2017.324

Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P. (2017). Can semantic labeling methods generalize to any city? The INRIA aerial image labeling benchmark. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS); 2017 July 23-28; Fort Worth, TX, USA. https://doi.org/10.1109/IGARSS.2017.8127684

Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K. McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., Ruecherk, D. (2018). Attention U-Net: Learning Where to Look for the Pancreas. arXiv. https://doi.org/10.48550/arXiv.1804.03999

Pang, B., Chen, L., Tao, Q., Wang, E., Yu, Y. (2024). GA-UNet: A Lightweight Ghost and Attention U-Net for Medical Image Segmentation. *Journal of Imaging Informatics in Medicine, 37*, 1874-1888. https://doi.org/10.1007/s10278-024-01070-5

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems (NeurIPS); Curran Associates, Inc.: Red Hook, NY, USA, pp. 8024-8035. https://doi.org/10.48550/arXiv.1912.01703

Python Software Foundation. (2024). Python: A Programming Language; Python Software Foundation: Wilmington, DE, USA, Available online: https://www.python.org/

Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., (2020). Jagersand, M. U^2-Net: Going Deeper with Nested U-Structure for Salient Object Detection. *Pattern Recognition, 106*, 107404. https://doi.org/10.48550/arXiv.2005.09007

Qin, X., Zhang, Z., Huang, C., Gao, C., Dehghan, M. (2019) Jagersand, M. BASNet: Boundary-Aware Salient Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019), 2019 June 16-20; Long Beach, CA, USA.

Rahman, M. A. & Wang, Y. (2016). Optimizing Intersection-over-Union in Deep Neural Networks for Image Segmentation. In Proceedings of the International Symposium on Visual Computing (ISVC 2016), 2016 December 12-14; Las Vegas, NV, USA. https://doi.org/10.1007/978-3-319-50835-1_22

Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), 2015 October 5-9; Munich, Germany. https://doi.org/10.48550/arXiv.1505.04597

Rottmann, M., Reese, M. (2023). Automated Detection of Label Errors in Semantic Segmentation Datasets via Deep Learning and Uncertainty Quantification. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2023 January 3-7; Waikoloa, HI, USA. https://

www.kci.go.kr

doi.org/10.1109/WACV56688.2023.00323

Rumelhart, D. E., Hinton, G. E., Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature, 323*, 533-536. https://doi.org/10.1038/323533a0

Salman, M. S., Liu, Y. (2019). Overfitting Mechanism and Avoidance in Deep Neural Networks. arXiv. https://doi.org/10.48550/arXiv.1901.06566

Schubert, M., Riedlinger, T., Kahl, K., (2024). Kröll, D.; Schoenen, S.; Šegvić, S.; Rottmann, M. Identifying label errors in object detection datasets by loss inspection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2024 January 3-7; Waikoloa, HI, USA.

Shorten, C. & Khoshgoftaar, T. M. A (2019). Survey on Image Data Augmentation for Deep Learning. *Jounal of Big Data, 6*, 1-48. https://doi.org/10.1186/s40537-019-0197-0

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning. Research, 15*, 1929-1958.

Yang, Z., Dong, J., Kou, W., Qin, Y., Xiao, X. (2021). Mapping Panax notoginseng Plantations by Using an Integrated Pixel- and Object-Based (IPOB) Approach and ZY-3 Imagery. *Remote Sensing, 13*, 2184. https://doi.org/10.3390/rs13112184

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In Proceedings of the International Workshop on Deep Learning in Medical Image Analysis, 2018 September 20; Granada, Spain. https://doi.org/10.1007/978-3-030-00889-5_1

www.kci.go.kr