

## 특징 융합을 이용한 농작물 다중 분광 이미지의 의미론적 분할

# Semantic Segmentation of Agricultural Crop Multispectral Image Using Feature Fusion

문준열 · 박성준 · 백중환\*

한국항공대학교 항공전자정보공학부

Jun-Ryeol Moon · Sung-Jun Park · Joong-Hwan Baek\*

School of Electronics and Information Engineering, Korea Aerospace University, Gyeonggi-do, 10540 Korea

### [요 약]

본 논문에서는 농작물 다중 분광 이미지에 대해 특징 융합 기법을 이용하여 의미론적 분할 성능을 향상시키기 위한 프레임워크를 제안한다. 스마트팜 분야에서 연구 중인 딥러닝 기술 중 의미론적 분할 모델 대부분은 RGB(red-green-blue)로 학습을 진행하고 있고 성능을 높이기 위해 모델의 깊이와 복잡성을 증가시키는 데에 집중하고 있다. 본 연구는 기존 방식과 달리 다중 분광과 어텐션 메커니즘을 통해 모델을 최적화하여 설계한다. 제안하는 방식은 RGB 단일 이미지와 함께 UAV (unmanned aerial vehicle)에서 수집된 여러 채널의 특징을 융합하여 특징 추출 성능을 높이고 상호보완적인 특징을 인식하여 학습 효과를 증대시킨다. 특징 융합에 집중할 수 있도록 모델 구조를 개선하고, 작물 이미지에 유리한 채널 및 조합을 실험하여 다른 모델과의 성능을 비교한다. 실험 결과 RGB와 NDVI (normalized difference vegetation index)가 융합된 모델이 다른 채널과의 조합보다 성능이 우수함을 보였다.

### [Abstract]

In this paper, we propose a framework for improving the performance of semantic segmentation of agricultural multispectral image using feature fusion techniques. Most of the semantic segmentation models being studied in the field of smart farms are trained on RGB images and focus on increasing the depth and complexity of the model to improve performance. In this study, we go beyond the conventional approach and optimize and design a model with multispectral and attention mechanisms. The proposed method fuses features from multiple channels collected from a UAV along with a single RGB image to increase feature extraction performance and recognize complementary features to increase the learning effect. We study the model structure to focus on feature fusion and compare its performance with other models by experimenting with favorable channels and combinations for crop images. The experimental results show that the model combining RGB and NDVI performs better than combinations with other channels.

**Key words** : Semantic segmentation, Smart farm, Multispectral, Feature fusion, Attention mechanism.

<http://dx.doi.org/10.12673/jant.2024.28.2.238>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 4 April 2024; Revised 24 April 2024

Accepted (Publication) 27 April 2024 (30 April 2024)

\*Corresponding Author; Joong-Hwan Baek

Tel: +82-2-300-0125

E-mail: [jhbaek@kau.ac.kr](mailto:jhbaek@kau.ac.kr)

## 1. 서론

현대 사회가 빠르게 발전하며 전 세계의 인구 또한 계속하여 증가하고 있다. 동시에 의료의 발전도 이루어지고 평균 수명이 증가하며 인구노령화가 지속해서 진행되고 있다. 인구 대비 노동력은 부족해지는 상태이며 이에 가장 치명적인 주제 중 하나는 의식주에 속하는 농업이다. 근래 이 문제를 해결하기 위하여 스마트팜 (smart farm) 분야 관련 연구가 꾸준히 진행 중이다. 스마트팜은 인공지능(AI; artificial intelligence), 빅 데이터, 로봇 등의 기술을 농업에 접목하여 작물의 생육 환경을 최적화하여 농업의 효율성을 개선하는 기법을 말한다. 온도, 습도 등의 작물 성장에 영향을 미치는 환경을 실시간으로 모니터링하고 원하는 작물의 건강도를 계산하여 농작물을 최적의 환경에서 유지 및 관리하는 기술이다[1]-[3].

최근 연구 중 딥러닝 기술을 접목한 연구가 다양하게 진행되고 있는 가운데, 그중 대부분의 연구는 RGB 이미지만을 통하여 진행한다[4]. 그러나 이는 고수준의 객체 탐지나 분할을 하는 데에 분명한 한계가 존재한다.

본 논문에서는 RGB 이미지뿐만 아니라 다중 분광 이미지를 추가로 사용하여 작물 탐지와 분석을 위한 특징 융합 기반 의미론적 분할 모델을 제안한다. 데이터셋은 UAV를 통해 촬영된 작물 이미지로 청색광(blue), 녹색광(green), 적색광(red), 레드 엣지(RE; red edge), 근적외선(NIR; near infrared)으로 구성된 5개의 대역을 사용하고 전처리 작업을 거친다.

학습 모델은 기본적으로 CNN (convolutional neural network) 모델을 기반으로 한다. 입력 개수에 따라 인코더의 수를 늘리고 다중 분광 간의 특징 융합을 위한 여러 모듈을 적재적소에 추가하여 다중 분광을 이용한 최적의 모델을 제안한다. 본 연구에서는 단일 RGB 이미지만을 사용하는 기존 모델과 제안된 모델을 학습 및 평가하여 성능 비교를 진행한다.

## II. 의미론적 분할 관련 모듈 연구

### 2-1 EAEF

본 논문에서 응용할 EAEF (explicit attention enhanced fusion) 모듈을 사용하는 EAEFNet은 ResNet Backbone을 사용하는 두 개의 인코더와 디코더로 구성되며 인코더에 서로 다른 모달리티(modality)의 입력을 받는다. EAEF 모듈은 두 인코더 사이에 배치되어 두 입력의 서로 다른 특징을 효과적으로 결합하도록 설계된 융합 방식으로 그림 1과 같이 구성된다[5].

먼저, EAEF는 두 인코더를 통해  $E_X$ 와  $E_Y$ 의 입력을 받아서 전역 평균 풀링(GAP; global average pooling)과 다층 퍼셉트론(MLP; multi layer perceptron)을 적용하여 주목해야 할 가중치를 생성한다. 여기서  $E_X$ 와  $E_Y$ 는 서로 다른 채널의 특징 맵을 말한다.

계산된 가중치는 EAEF 프레임워크에서 각각 AIB (attention interaction branch)와 ACB (attention complement branch)의 두 가지 분기로 나누어 처리된다. AIB는 두 모달리티의 유의미한 특징을 모두 인식하는 경우, 또는 모두 인식하지 못하는 경우를 상호작용하여 특징을 강화하는 방식이다. 가중치끼리 요소별 곱셈을 하고 결과를  $E_{X,Y}$ 와 채널별 곱셈하여 각 인코더의 특징 맵에 적용한다.

이렇게 생성된  $I_X, I_Y$ 는 AIB를 통해 깊이별 컨볼루션(DWC; depth-wise convolution), 전역 맥스 풀링(GMP; global max pooling), 다중 퍼셉트론을 통해 새로운 특징 맵을 형성한다. 이때,  $I_X, I_Y$ 는 동시에 연산되어 모달리티 간의 상호작용을 더욱 구체화하며 특징 추출을 효과적으로 할 수 있게 된다.

반면, ACB는 하나의 모달리티만을 유용한 특징으로 인식하여 추출하는 경우를 상호보완적으로 처리하여 서로의 결여된 특징을 개선하는 역할을 한다.

AIB와 ACB를 통해 얻은 출력인  $\hat{I}_X, \hat{I}_Y$ 와  $C_X, C_Y$ 를 더한  $O_X, O_Y$ 를 최종적으로 1x1 컨볼루션 계층과 소프트맥스 함수를 적용하여 특징 융합을 완수한다. 그림 1에 EAEF의 구조도를 보인다.

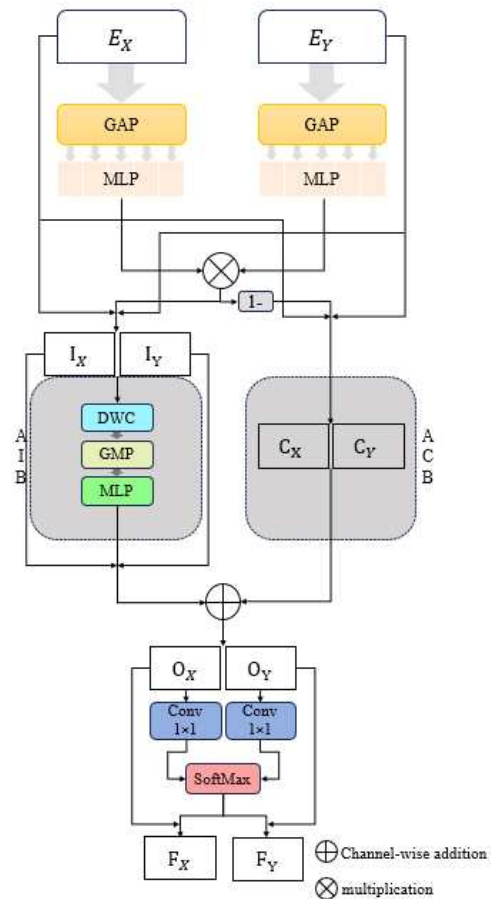


그림 1. EAEF 모듈  
Fig. 1. EAEF module.

## 2-2 어텐션 모듈

Channel attention과 Spatial attention은 특징 추출 수행 능력을 향상하기 위한 효과적인 모듈이다[6]. Channel attention은 입력 특징 맵의 각 채널이 얼마나 중요한지를 학습하여, 중요한 채널의 특징은 강조하고 중요도가 낮은 경우에는 억제한다. 전역 평균 풀링과 전역 맥스 풀링을 통해 특징을 요약하고, 다중 퍼셉트론을 통해 가중치를 도출한다. 이를 시그모이드 함수를 통하여 정규화하고 원본 특징 맵에 적용하여 입력의 중요 특징에 더 집중할 수 있게 한다. 식은 다음과 같고,  $E$ 는 특징 맵을 나타낸다.

$$CA(E) = \sigma(MLP(GAP(E)) + MLP(GMP(E))) \quad (1)$$

Spatial attention은 이미지 내의 중요한 공간적 정보를 식별 및 강조하는 메커니즘이다. 입력 특징 맵에서 위치별로 중요도를 파악하고 특징 맵을 재조정한다. Channel attention은 채널을, spatial attention은 공간의 중요도를 파악하여 정보를 효율적으로 처리한다. 전역 평균 풀링과 전역 맥스 풀링을 진행하고 채널 방향으로 연결하여 합친다.  $7 \times 7$  필터를 통해 연산하여 시그모이드 함수로 중요한 공간 위치의 특징을 강조한다. 자세한 수식은 식 (2)로 나타낸다. 그림 2에 Channel attention과 spatial attention 구조도를 보인다.

$$SA(E) = \sigma(f^{7 \times 7}([GAP(E); GMP(E)])) \quad (2)$$

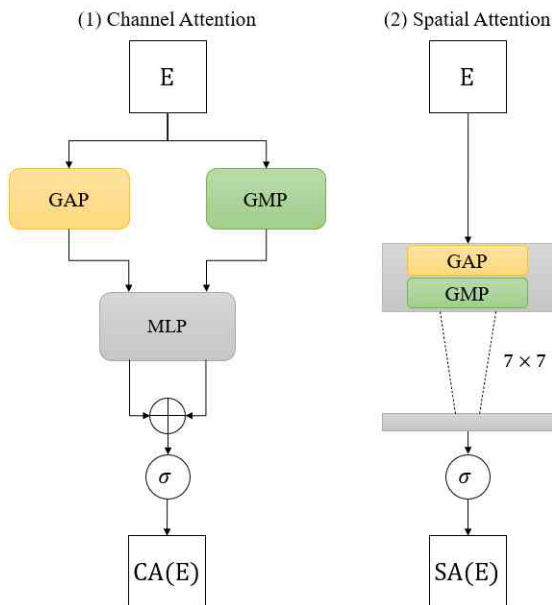


그림 2. (1) Channel attention 구조도 (2) Spatial attention 구조도

Fig. 2. (1) Structure of Channel attention (2) Structure of Spatial attention.

## III. 특징융합 기반 의미론적 분할

### 3-1 다중 분광 데이터 전처리

본 논문에서는 스마트팜 관련 작물 영상의 의미론적 분할을 위한 다중 분광 데이터 확보를 위해 DJI사의 P4 Multispectral 기체(그림 3)를 사용한다. 다중 분광 이미지를 촬영할 수 있도록 RGB 1개 센서와 5 밴드의 센서가 존재하는 고정밀 UAV로 스마트팜 연구에 최적화 되어있다. P4 Multispectral 기체의 제원은 표 1과 같다[7].

데이터 수집을 위해 화분에 작물을 심어서 폭넓게 배치하였다. 시간이 흐름에 따라 작물이 자라는 모습을 주기적으로 촬영하여 다양한 작물의 데이터셋을 갖추어 실험을 진행하였다. 기체에 탑재된 6개의 센서를 이용하여 동시에 촬영하며 RGB, 청색광(blue), 녹색광(green), 적색광(red), 레드엣지(red edge), 근적외선(NIR)의 5가지 대역이 그림 4와 같이 촬영된다.



그림 3. DJI P4 Multispectral

Fig. 3. DJI P4 Multispectral.

표 1. DJI P4 Multispectral 제원

Table. 1. Specifications of DJI P4 Multispectral.

Size	350 mm
Weight	1487 g
Max. altitude	6000 m
Max. speed	50-58 km/h
Max. flight time	27 min
Accuracy of hovering	±0.1 m
Camera sensor	six 1/2.9" CMOS sensors (1 RGB sensor, 5 monochrome sensors)
FOV of lens	62.7 °
Image size	1600×1300

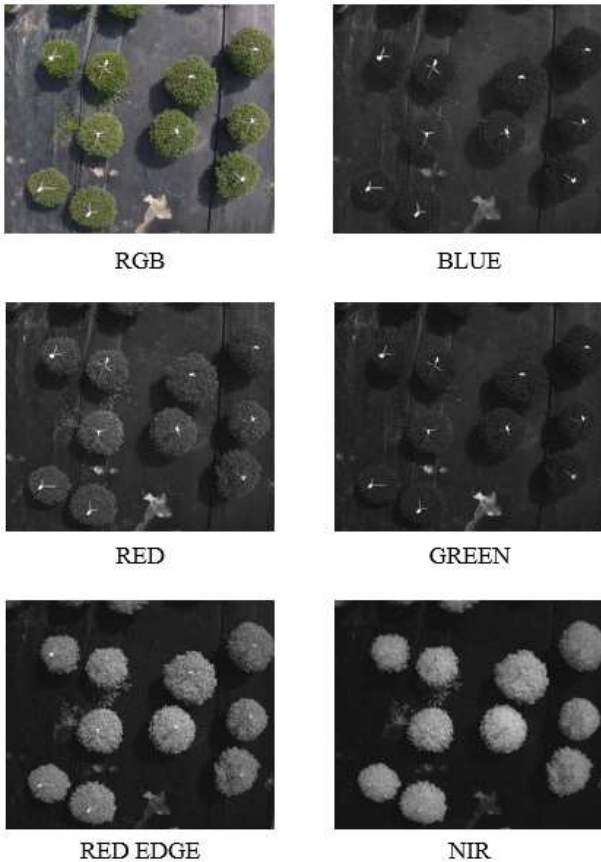


그림 4. DJI P4 Multispectral로 수집된 RGB와 다중 분광 이미지 예시  
 Fig. 4. Example of RGB and Multispectral images using DJI P4 Multispectral.

DJI P4 Multispectral을 통해 촬영된 다중 분광 이미지는 5 채널이지만 본 논문의 실험에서는 식생지수라고 불리는 NDVI 채널을 추가한다. NDVI는 식생의 밀도와 상태를 평가하기 위한 지표로 태양으로부터 받는 빛 중 일부를 흡수하고 반사하는 특성을 기반으로 작물의 건강 상태 추정에 도움을 준다. 식생지수는 근적외선과 적색광의 반사도 차이로 계산되고 -1에서 1 사이의 값을 가지며 식 (3)과 같다. 일반적으로 0.3 이상이면 밀도가 높은 식생을, 0.3 미만은 식생이 거의 없음을 나타낸다. 0에 가깝거나 음수 값일 경우 물이나 맨땅일 가능성이 높다[8].

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3)$$

식생지수는 스마트팜 관련 연구에 널리 사용되고 있으며 본 논문에서도 채널 간의 특징 융합 시 강점을 드러낼 것을 기대한다. 그림 5는 근적외선과 적색광으로 계산한 NDVI 이미지 예시이다.



그림 5. NIR과 RED 이미지로 계산된 NDVI 이미지 예시  
 Fig. 5. Example of NDVI images using NIR and RED images.

DJI P4 Multispectral 기체는 호버링 정확도가 ±0.1 m이고 아래 그림 6과 같이 물리적인 렌즈 위치 또한 다르며 카메라 렌즈의 왜곡에 대한 보정이 이루어지지 않는다.

그 결과 똑같은 조건에서 동시에 촬영된 이미지에 대해서도 채널별로 오차가 발생한다. 이러한 문제는 본 논문에서 진행할 의미론적 분할에 치명적인 문제로 작용하며 캘리브레이션 작업을 필수적으로 진행해야 한다.

캘리브레이션을 하기 위한 방법으로 본 논문에서는 SuperGlue를 채택하여 진행한다. SuperGlue는 그래프 신경망(GNN; graph neural network)과 어텐션 메커니즘을 활용하여 두 이미지 세트의 지역 특징점 사이의 대응 관계를 정교하게 찾아내고 매칭할 수 없는 점을 걸러내는 역할의 딥러닝 아키텍처이다[9]. RGB 이미지를 기준으로 이미지 사이의 적합한 특징점을 매칭하고 이를 기반으로 기하학적 변환을 추정한다. 매개변수 조정 및 정확도를 검증하여 캘리브레이션의 신뢰성을 높인다. 전체 데이터에 대하여 검토를 거친 뒤, 아직 오차가 존재하는 데이터에만 픽셀 단위의 후보정 작업을 직접 진행한다.

캘리브레이션 과정 중에 수행된 기하학적 변환에 의해 이미지의 가장자리 부분에서 시프트 현상이 발생하는데, 이에 따라 원본 이미지 외부의 픽셀에 대해 패딩을 적용하여 정제하게 된다. 이후 이미지의 중심을 기준으로 해상도를 기존 1600×1300에서 1300×1100 픽셀로 크기를 조정하여, 관심 영역의 데이터를 보존하고 불필요한 외곽 데이터를 제거한다. 이 과정은 실험의 정확성 및 통일성을 제공한다.

최종 데이터셋은 객체 분할 모델 학습을 위하여 픽셀 단위의 어노테이션을 수행한다. 클래스는 작물 영역과 배경 영역으로 구성된 이진 클래스로 구성되며 그림 7과 같이 명확하게 구분되어 객체 분할 훈련에 활용된다.





그림 6. DJI P4 Multispectral의 렌즈  
Fig. 6. Lens of DJI P4 Multispectral.

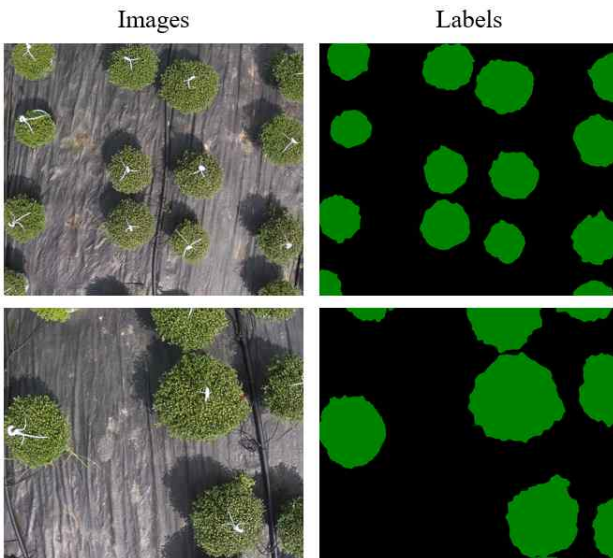


그림 7. RGB 이미지와 라벨링 이미지 예시  
Fig. 7. Example of RGB and labeling images.

### 3-2 특징 융합 기반 제안 모델

본 논문은 앞선 관련 연구에서 소개한 모델과 다중 분광 이미지를 활용하여 중요한 특징만을 추출함으로써 비교적 가벼운 모델을 사용하더라도 정확성과 신뢰도를 향상시킬 수 있는 아키텍처를 제안한다. Attention 모듈은 이미 많은 딥러닝 모델에서 적용하고 성과를 이루고 있다. 그중 channel attention과 spatial attention을 이용한 성능 개선은 충분히 입증된 효과적인 모듈이다. 이를 서로 다른 채널의 특징을 융합하는 EAEF 모듈 기반의 CNN 모델에 배치하여 여러 실험을 진행하고 가장 적합한 모델 구조를 찾아낸다. 그림 8에는 본 논문에서 제안하는 최적의 특징 융합 기반 의미론적 분할 모델의 구조를 보인다.

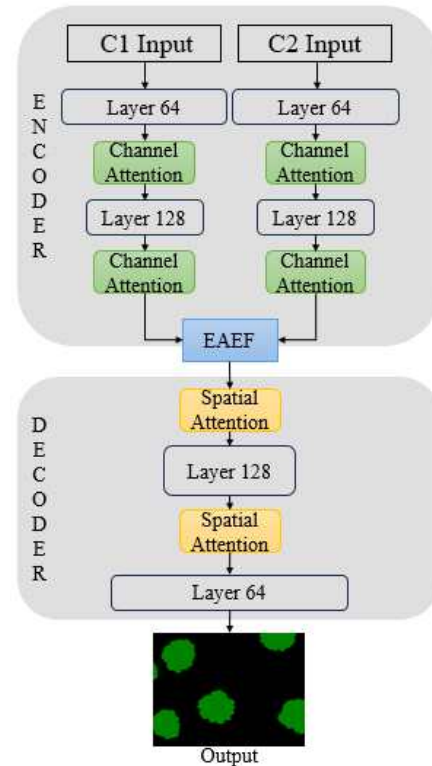


그림 8. 제안 모델의 구조도  
Fig. 8. Proposed model architecture.

제안 모델은 C1 인코더와 C2 인코더를 사용하면서 서로 다른 채널을 입력으로 받는다. 인코더의 Layer는 기본적인 컨볼루션 계층, 배치 정규화, Relu 함수를 포함하는 구조로 다운샘플링 후 Layer 뒤의 숫자만큼의 채널 수를 갖는다. Layer를 거친 뒤 중요한 채널의 특징을 channel attention을 통해 강조된다. 그렇게 강조된 특징 맵은 EAEF 모듈을 통해 채널 간의 상호작용이 이루어져 주요 정보가 증진된다. 디코더에서는 Layer 뒤의 숫자만큼의 채널을 입력받고, 업샘플링과 spatial attention을 연속적으로 진행하면서 공간적 중요도를 기반으로 특정 영역을 강조하며 최종 의미론적 분할 맵을 생성한다. 제안된 모델은 다중 분광 이미지의 잠재력을 이용하여 RGB 이미지만으로는 얻을 수 없는 중요한 정보를 다양한 모듈로써 강조하여 성능을 높이기 위한 새로운 접근 방식을 제안한다.

## IV. 실험

### 4-1 실험 환경

본 논문에서 진행한 실험 환경은 Ubuntu 18.04이다. 프로세서는 Intel Core I9-10900KF를, 그래픽카드는 NVIDIA GeForce RTX 3080을 사용하고 메모리는 31.3 GB이다. 데이터셋의 클

래스는 배경과 작물, 이진 클래스로 구분하며 총 데이터셋은 16,975장(2,425세트)으로 RGB와 NDVI를 포함한 총 7개의 채널이 존재한다. 학습 데이터셋과 테스트 데이터셋은 4:1의 비율로 나누고 epoch은 30으로 한다.

성능 평가 지표는 의미론적 분할에 보편적으로 사용되는 IoU (intersection over union)를 사용한다. IoU는 객체에 대한 예측이 정확히 이루어졌는지를 결정하기 위한 지표로 분모는 실제와 예측의 합집합, 분자는 실제와 예측의 교집합이다. 식은 아래와 같다.

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + FN_i} \quad (4)$$

여기서  $TP$ 는 true positive,  $TN$ 는 true negative,  $FN$ 는 false negative,  $FP$ 는 false positive이고  $N$ 은 샘플 수이다. 인코더 입력단에 7개의 채널을 조합하여 가장 성능이 우수한 경우를 찾기 위한 실험을 진행한다. 이를 위해 인코더를 추가한 3개의 입력을 받는 특징 융합 기반 모델과 1개의 입력을 받는 모델에 실험에 추가하여 사용한다. 추가로 RGB만을 이용하는 기존 모델과 제안 모델의 성능과 파라미터를 함께 비교하여 최적의 모델을 제안한다.

#### 4-2 실험 결과 및 분석

우선, 1개의 입력을 받는 경우를 훈련하여 각 채널에 대한 실험을 진행한다. 입력이 하나이기 때문에 그림 8의 제안한 모델에서 불필요한 특징 융합을 위한 모듈을 제외하여 기본적인 CNN 구조로 변환한다. 실험 결과는 그림 9와 같다. 1개의 입력을 통한 작물 영역 분할 결과, RGB 채널이 가장 높은 성능을 보인다. 단일 채널에 비해 풍부한 정보를 제공하며 정확한 결과를 달성할 수 있음을 말한다. 또한, NDVI와 같은 특수 채널의 성능도 높은 것으로 보아 식물의 건강이나 밀도 등의 정보를 효과적으로 반영하여 학습하고 있음을 알 수 있다.

다음으로 본 논문에서 제안하는 모델로 학습한 작물 의미론적 분할 결과를 그림 10과 같이 보인다. 1개의 입력을 통한 실험과 마찬가지로 동향을 가지며 RGB 기반의 조합이 높은 성능을 보인다. 그중 RGB+NDVI의 조합이 가장 높은 성능인 96.67%의 mIoU를 달성함으로써 NDVI가 제공하는 추가적인 정보가 RGB 정보와 함께 정확도 개선에 도움을 주고 있음을 파악한다.

마지막으로 제안된 모델에 똑같은 인코더를 추가하여 3개의 채널이 서로 특징 융합하여 좋은 성능을 보일 수 있도록 한다. 앞서 진행한 실험을 근거로 조합을 고려하고 실험을 진행한다. 결과는 다음 그림 11과 같다. RGB+RE+NIR의 조합이 가장 높은 mIoU를 나타냈지만, 이것은 직전에 실험한 RGB+NDVI의 결과에 비해서는 좋은 성능을 보이는 것이 아니다. 즉, 입력을 늘려 정보의 양이 증가하는 것이 득이 되는 것이 아니라 오히려 채널 간의 집중도를 떨어뜨리는 결과를 보인다.

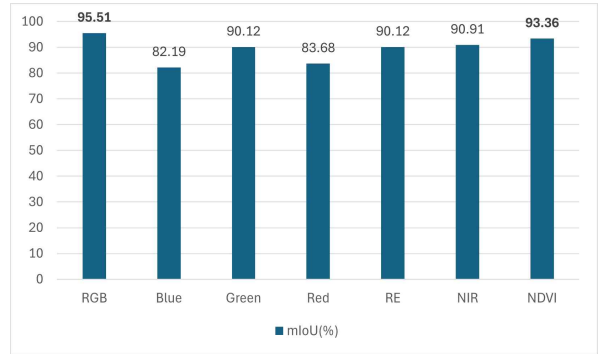


그림 9. 1개의 입력을 통한 분할 결과  
Fig. 9. 1 Channel segmentation evaluation metric.

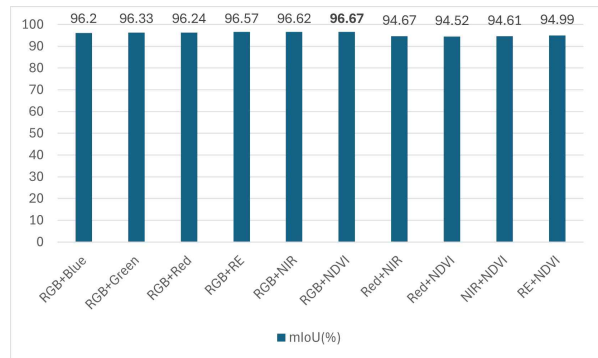


그림 10. 2개의 입력을 통한 분할 결과  
Fig. 10. 2 Channel segmentation evaluation metric.

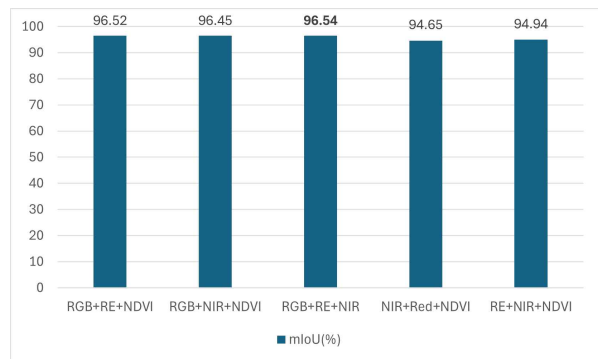


그림 11. 3개의 입력을 통한 분할 결과  
Fig. 11. 3 Channel segmentation evaluation metric.

기존 RGB 모델을 이용한 학습 결과와 제안 모델의 결과를 비교하여 그림 12에 보인다[10]-[13]. U-Net의 mIoU가 98%가 가장 높은 것을 확인할 수 있으나 각 파라미터 수에 주목해 보면 제안 모델과 약 126배가 차이 난다. RGB 단일 입력을 통해 성능 개선을 하고자 모델의 복잡도를 높인 결과이다. 반면, 제안 방식을 사용하면 모델을 무겁게 하지 않아도 성능 차이 대비 모델이 훨씬 가벼워지면서 학습의 효율성을 극대화할 수 있음을 보인다. 추가로 그 차이로 인하여 제안 모델은 두 개의 입력으로 학습함에도 불구하고, 기존 모델보다 학습 시간 또한 개선 되는 것을 확인하였다.

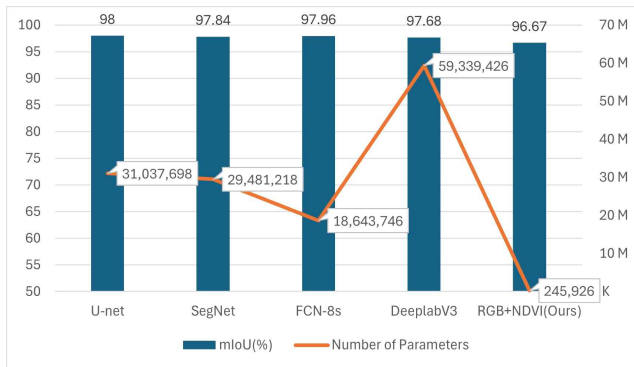


그림 12. 기존 모델과 비교  
 Fig. 12. Comparison with other models.

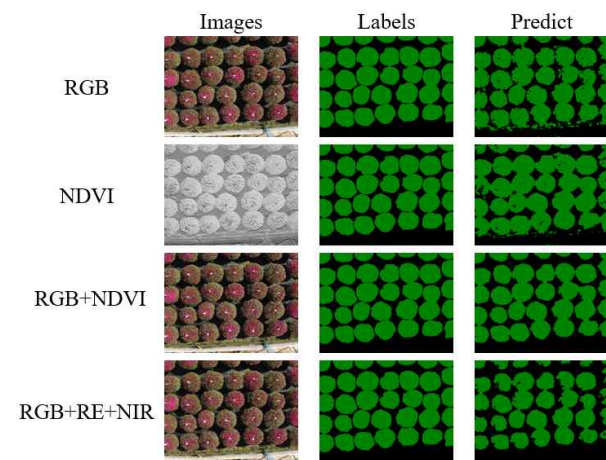


그림 13. 객체 분할 결과 예시  
 Fig. 13. Case of segmentation results.

그림 13은 실험을 통하여 얻은 유의미한 조합에 대한 의미론적 분할 예측 이미지이다. RGB만을 입력한 결과보다 RGB+NDVI를 입력한 훈련 결과가 정성적으로도 개선되었음을 보인다.

## V. 결론

본 논문에서는 UAV로 촬영된 다중 분광 작물 이미지를 통해 특징 융합 기반 의미론적 분할 기법을 제안하였다. 제안된 방식을 검증하기 위해 농작물과 관련된 NDVI 채널을 추가하고 캘리브레이션을 적용하여 데이터 전처리를 수행했다. 의미론적 분할은 특징 융합 모듈인 EAEF와 channel attention, spatial attention 모듈을 기반으로 하는 제안 모델과 다중 분광을 이용하여 진행하였다. 그 결과 RGB와 NDVI에서 강점을 드러내고, 이 둘의 조합인 RGB+NDVI의 실험 결과가 기존 RGB의 실험 결과에 비하여 mIoU가 1.16% 향상됨을 보였다. 모델의 깊이도 중요하지만 다중 분광 이미지를 통해 부족한 특징을 보충하고,

이를 집중시키고 효과적으로 융합시키기 위한 모듈로도 충분한 성능 개선을 이루어낼 수 있었다. 또한, 파라미터를 줄임으로써 가벼우면서도 정확한 모델을 제안할 수 있었다.

추후에는 본 논문에서 실험하지 않은 다른 다중 분광에 대해서도 연구하며 작물 이미지의 의미론적 분할 성능을 높일 수 있도록 한다. 추가로 각 특징에 집중하고 융합하기 위한 또 다른 기법들을 고안하여 기존 모델보다는 가볍지만 정확도도 높일 수 있는 연구를 진행할 예정이다.

## Acknowledgments

본 연구는 경기도기술개발 사업의 사업비 지원 (과제번호 #D2323011)과 경기도 지역협력 연구센터 사업의 일환으로 수행하였음. [GRRC항공2023-B02, 지능형 인터랙티브 미디어 및 공간 융합 응용 서비스 개발]

## References

- [1] M. Dhanaraju, P. Chenniappan, K. Ramalingam, S. Pazhanivelan, and R. Kaliaperumal, "Smart farming: internet of things (IoT)-based sustainable agriculture," *Agriculture*, Vol. 12, No. 10, pp. 1745, Oct. 2022.
- [2] R. R. Samantaray, A. Azeez, and N. Hegde, "Efficient smart farm system using machine learning," in *Proceedings of the 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS)*, Bangalore: India, pp. 576-581, 2023.
- [3] V. R. Saraswathi, J. Sridharani, S. P. Chowdary, K. Nikhil, S. M. Harshitha, and M. K. Sai, "Smart farming: the IoT based future agriculture," in *Proceedings of the 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli: India, pp. 150-155, 2022.
- [4] A. Khan, T. Ilyas, M. Umraiz, Z. I. Mannan, and H. Kim, "CED-Net: Crops and weeds segmentation for smart farming Using a small cascaded encoder-decoder architecture," *Electronics*, Vol. 9, No. 10, pp. 1602, Oct. 2020.
- [5] M. Liang, J. Hu, C. Bao, H. Feng, F. Deng, and T. L. Lam, "Explicit attention-enhanced fusion for RGB-thermal perception tasks," arXiv:2303.15710, 2023. Available: <https://arxiv.org/abs/2303.15710>
- [6] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich: Germany, pp. 3-19, 2018.
- [7] DJI. DJI P4 Multispectral information [Internet]. available: <https://www.dji.com/kr/p4-multispectral/specs>

- [8] S. Huang, L. Tang, J. P. Hupy, et al., “A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing,” *Journal of Forestry Research*, Vol. 32, pp. 1-6, 2021.
- [9] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning feature matching with graph neural networks,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle: WA, pp. 4938-4947, 2020.
- [10] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Munich: Germany, Vol. 9351, pp. 234-241, 2015.
- [11] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481-2495, 2017.
- [12] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston: MA, pp. 3431-3440, 2015.
- [13] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, “Encoder-decoder with Atrous separable convolution for semantic image segmentation”, in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich: Germany, pp. 801-818, 2018.



**문 준 렬 (Jun-Ryeol Moon)**

2020년 ~ 현재 : 한국항공대학교 전자 및 항공전자공학 학사과정  
 ※ 관심분야 : 컴퓨터비전, 영상처리



**박 성 준 (Sung-Jun Park)**

2019년 2월 : 한국항공대학교 전자 및 항공전자공학 (공학사)  
 2021년 2월 : 한국항공대학교 일반대학원 항공전자정보공학 (공학석사)  
 2021년 ~ 현재 : 한국항공대학교 일반대학원 항공전자정보공학 박사과정  
 ※ 관심분야 : 컴퓨터비전, 영상처리



**백 중 환 (Joong-Hwan Baek)**

1981년 2월 : 한국항공대학교 항공통신공학 (공학사)  
 1987년 7월 : 오클라호마주립대학원 전기 및 컴퓨터공학 (공학석사)  
 1991년 7월 : 오클라호마주립대학원 전기 및 컴퓨터공학 (공학박사)  
 1992년 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수  
 ※ 관심분야 : 영상처리, 패턴인식, 멀티미디어, 가상현실