

# 산업 현장 드론 점검 자동화를 위한 U-Net 기반 부품 영역 분할 모델 연구

정태영, 유철우\*

명지대학교

jungtaeyeong420@gmail.com, \*cwyu@mju.ac.kr

## U-Net-Based Component Segmentation Model for Automated Drone Inspection in Industrial Environments

Taeyeong Jeong, Cheolwoo You\*

Myongji Univ.

### 요약

드론은 프로펠러, 날개, 본체 등 여러 부품으로 구성되어 있고 점검 시 각 부위의 상태를 따로 확인해야 한다. 하지만 사진이나 영상 속에서 사람이 매번 부위를 구분해 점검하기엔 시간이 오래 걸린다. 이러한 산업적 요구에 대응하기 위해, 기업과 대학이 IC-PBL(Industry-Coupled Problem Based Learning) 형태로 연계하여 실제 현장 문제 해결을 목표로 연구가 수행되었다. 본 연구는 기업의 요청에 따라 기존 오픈소스 이미지 분할 모델인 U-Net을 활용하여 드론 영상에서 부품별 영역을 추출하고, 입력된 드론 이미지에 대해 각 부위를 색상별로 자동 구분하는 딥러닝 기반 분할 모델을 학습·개발하는 것을 목적으로 한다. 특히 U-Net의 인코더-디코더 구조를 고려하여 특징 추출과 공간 정보 복원을 효과적으로 수행함으로써 과적합을 완화하고, IoU(Intersection over Union) 지표를 기반으로 모델의 분할 성능을 향상시키고자 한다.

### I. 서론

최근 드론은 다양한 산업 분야에서 활용 범위가 빠르게 확대되고 있으며 드론 운용 과정에서 발생하는 장비 상태 점검, 부품 관리, 정비 효율화가 필요하다. 특히 부품 단위의 상태를 확인하고 기록해야 하는데, 이 과정에서 사람의 수작업은 시간과 비용이 증가하는 한계가 있다. 이러한 기업의 문제를 IC-PBL(Industry - Coupled Problem Based Learning) 형태로 대학과 기업이 연계하여 실제 현장 문제 해결을 목표로 수행한다. 기업 측 요청사항은 드론 이미지를 입력으로 받아 드론의 각 부위를 자동으로 구분하는 segmentation 기반 AI 모델 개발이다.

그림 1의 U-Net 기반 대칭적인 인코더-디코더(encoder-decoder) 구조를 적용하여 의미론적 분할(semantic segmentation) 모델을 설계한다. 또한, 모델 학습을 위해 Kaggle에서 제공되는 공개 드론 이미지 데이터셋을 활용하고, 부품 단위의 학습 데이터를 구축하여 및 성능 평가를 수행한다.

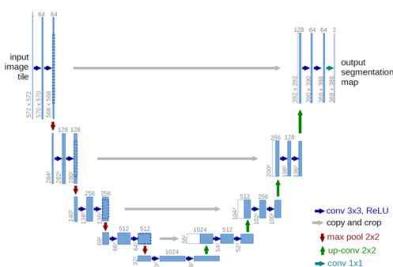


그림 1. U-NET의 기본 구조

### II. 본론

본 논문에서는 드론의 주요 부품을 기준으로 CVAT에서 라벨링 클래스를 정의하고 마스크 데이터를 생성한다. 그림 2는 드론의 부품별로 색깔을 매핑한 결과이다. 라벨링은 부품별 구분이 직관적으로 이루어지도록 색상 기준을 함께 적용하였고, 구체적으로 바디(body)는 빨강색, 암(arm)은 주황색, 프로펠러(propeller)는 노랑색, 랜딩기어(landing gear)는 파랑색, 카메라(camera)는 초록색으로 매핑한다.



그림 2. 부품별 색깔 라벨링

#### 가. ResNet34 모델

이미지 분류를 위해 개발된 영향력 있는 CNN(Conventional Neural Network)모델인 (Residual Network) ResNet34을 적용한다. ResNet34 기반 모델은 특징 추출 성능이 우수하여 학습 초기에 빠른 수렴을 보였다. 하지만 라벨링된 데이터셋이 약 800장으로 현저히 적어 그 결과 검증 데이터에서 성능이 저하되는 과적합(overfitting) 현상이 관찰되고, 제한된 데이터 환경에서 안정적인 일반화 성능을 확보하는 데 한계가 있음을 확인하였다.

#### 나. SAM 모델

YOLOv8n 기반의 SAM(Segment Anything Model) 계열 접근도 검토하였다는데, 해당 방식은 객체의 경계 및 형태를 비교적 잘 추출하여 이미지 기반 분할 결과 자체는 양호하게 도출되는 경향이 있다. 그러나 본 연구의 목적은 드론의 부품을 정의된 클래스(바디, 암, 프로펠러, 랜딩기어, 카메라) 단위로 일관되게 분리하는 것이고, SAM 기반 결과는 클래스별 라벨의 정합성 및 일관성을 유지하기 어렵거나, 다중 부품이 인접한 구간에서 부품 간 구분이 불명확해지는 문제가 발생한다. 즉, 분할 성능 자체와 별개로 학습용 마스크 데이터로 활용하기 위한 라벨링 체계 적용에 한계가 있음을 확인하였다.

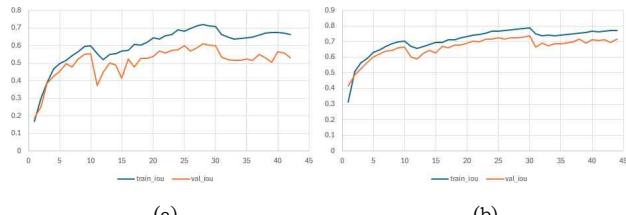


그림 3. 훈련, 검증 데이터 IoU값 (a) ResNet34, (b) VGG-16

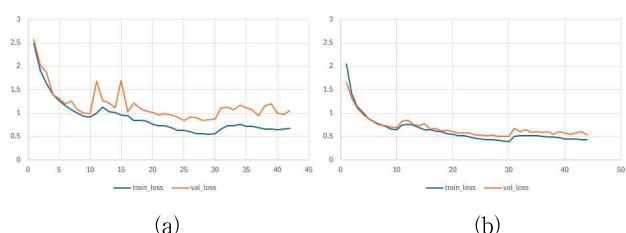


그림 4. 훈련, 검증 데이터 Loss 값 (a) ResNet34, (b) VGG-16

다. VGG-16 모델

VGG-16 기반 모델은 비교적 단순하고 일관된 특징 추출 구조를 갖기 때문에 부품 간 형태적 차이를 반영한 클래스 분할을 수행하는 데 적합하다. 또한, 제한된 데이터셋 환경에서도 안정적으로 학습이 가능하고, 클래스 기반 분할 목적에 적합한 구조를 고려하여 최종적으로 VGG-16을 인코더-디코더 모델로 채택한다. 과적합(overfitting) 방지를 위해 얼리 스탑핑(early stopping)을 추가하여 학습 과정에서 시간을 단축한다.

그림 3은 ResNet34 모델과 VGG-16 모델의 학습 특성을 비교하기 위해, 학습 과정에서의 훈련 데이터와 검증 데이터의 IoU(Intersection over Union) 변화를 epoch에 따라 시각화하여 분석하였다. 훈련 데이터와 검증 데이터는 7:3의 비율을 나타낸다. IoU 그래프에서는 두 모델 모두 초기 epoch 구간에서 빠른 성능 향상을 보였으나, 데이터 수가 제한된 조건에서는 ResNet34 모델의 경우 학습이 진행될수록 훈련 성능과 검증 성능 간 격차가 커지는 경향이 관찰된다. 이는 훈련 데이터에 대한 적합이 과도해지면서 검증 IoU의 개선 폭이 제한되거나 변동성이 증가하는 형태로 나타나, 과적합 가능성을 보여준다. 반면 VGG-16 모델은 epoch가 증가함에 따라 검증 IoU의 변동 폭이 상대적으로 작고 안정적으로 수렴하는 경향을 보인다. ResNet34의 가장 높은 IoU값은 0.61, VGG-16의 값은 0.73으로 약 19.7%로 상승한 결과를 확인하였다.

그림 4는 epoch에 따른 Loss 그래프이다. ResNet34 모델은 훈련 Loss가 지속적으로 감소하는 반면 검증 Loss가 일정 구간 이후 감소하지 않거나 증가하는 양상이 나타날 수 있어 일반화 성능 확보에 한계가 있다. 이에 비해 VGG-16 모델은 훈련/검증 Loss가 함께 감소하며 수렴하는 형태를 보여, 제한된 데이터셋 환경에서 학습 안정성과 일반화 측면에서 상대적으로 유리한 특성을 확인하였다. ResNet34의 가장 높은 Loss값은 0.83, VGG-16의 값은 0.49로 약 41% 감소한 결과를 확인하였다. 본 연구의 결과는 단순한 모델 성능 비교에 그치지 않고, 실사용을 고려한 형태로 웹페이지에 적용한다. 그림 5는 원본 이미지와 분할 결과를 동시에 출력하고, 부품별로 감지 결과를 전달한다. 또한, 데이터셋 특성상 팔(arm)이 존재하는 드론과 존재하지 않는 드론이 혼재되어 있기 때문에, 팔이 없는 기체에 대해서는 결과 화면에서도 이상 탐지 결과를 보여주며 ‘arm’이 감지되지 않음을 보인다. 이렇게 산출된 결과 보고서는 pdf로 다운 받을 수 있도록 구현한다.

그림 6은 제안한 모델 성능을 최고 정확도 epoch, mIoU, mDice, Best Loss의 지표를 시각화해서 보여준다. 본 연구는 기업의 실제 현장의 문제를 해결하기 위함이므로 기업에도 사용된 모델 성능을 보여준다. 또한, 부품별 세그멘테이션 성능을 Accuracy, IoU, F1 Score를 주요 지표로 선정하고 각 부품 클래스에 대한 지표값을 그래프로 나타낸다. 모델의 저반전이

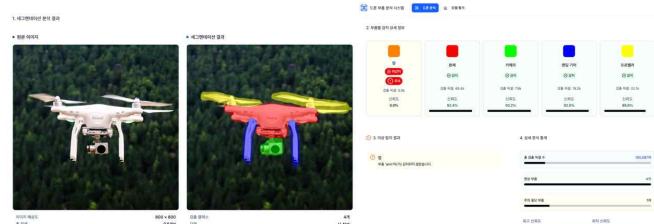


그림 5. 드론 부품별 segmentation, 이상 탐지 결과

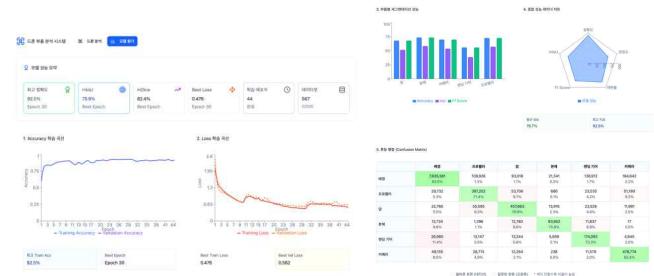


그림 6. VGG-16 모델 성능지표, 부품별 핵심 정확도 레이더 차트 및 혼동 행렬

균형 성능과 강점을 직관적으로 확인할 수 있는 종합 성능 레이더 차트를 나타낸 것이다. 이후, 행을 실제 클래스 열을 예측 클래스로 구성하며, 피셀 개수를 지표로 설정하였다. 전체 피셀 10,485,760에서 7,835,581 피셀인 배경을 정확도 93.5%로 가장 높았고 다음 478,774피셀인 카메라가 정확도 82.4%로 분류 성능이 뛰어난 것을 확인하였다.

### III. 결론

본 연구에서는 기업 요구사항을 기반으로 드론 점검을 위한 AI 기반 Segmentation 모델을 개발하였다. 모델 선정 과정을 통해 우선 채택된 34개의 층을 가지고 있는 ResNet34 모델은 데이터셋의 크기가 적을수록 Overfitting으로 이어지는 문제가 발생하였다. 이러한 문제를 완화하기 위해 우리는 가벼운 모델인 VGG-16 모델을 적용하였다. 실험 결과, 제안된 모델이 IoU(Intersection over Union)과 Loss를 효과적으로 개선한다는 것을 확인하였다. 또한, Early stopping 기술을 추가하여 Overfitting 방지를 효과적으로 수행할 수 있음을 확인하였다.

#### ACKNOWLEDGMENT

본 과제(결과물)는 교육부와 경기도의 재원으로 지원을 받아 수행된 경기 지역혁신중심 대학지원사업(경기RISE사업)의 연구결과임(2025-RISE-09-A15) 또한, 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2024-00335012).

참 고 문 헌

- [1] Ronneberger, O., Fischer, P., & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation pp. 234-241, 2015
  - [2] Simonyan, K., & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014
  - [3] He, K., Zhang, X., Ren, S., & Sun, J. Deep Residual Learning for Image Recognition. In CVPR 2016, pp. 280-285, Aug. 2016.
  - [4] N. H. Huy, N. D. Binh, T. Q. Quy, Q. X. Truong, and N. V. Hai, "Comparative Performance of ResNet50 and VGG16 in Lung Infection Detection," in Advances in Information and Communication Technology, pp. 733-744, 2025