

# U-Net 모델에 기반한 기간별 추출 소나무 고사목 데이터를 이용한 정사영상 탐지 정밀도 향상 연구

## A Study on Orthogonal Image Detection Precision Improvement Using Data of Dead Pine Trees Extracted by Period Based on U-Net model

김성훈<sup>1)</sup> · 권기욱<sup>2)</sup> · 김준현<sup>3)</sup>

Kim, Sung Hun · Kwon, Ki Wook · Kim, Jun Hyun

### Abstract

Although the number of trees affected by pine wilt disease is decreasing, the affected area is expanding across the country. Recently, with the development of deep learning technology, it is being rapidly applied to the detection study of pine wilt nematodes and dead trees. The purpose of this study is to efficiently acquire deep learning training data and acquire accurate true values to further improve the detection ability of U-Net models through learning. To achieve this purpose, by using a filtering method applying a step-by-step deep learning algorithm the ambiguous analysis basis of the deep learning model is minimized, enabling efficient analysis and judgment. As a result of the analysis the U-Net model using the true values analyzed by period in the detection and performance improvement of dead pine trees of wilt nematode using the U-Net algorithm had a recall rate of -0.5%p than the U-Net model using the previously provided true values, precision was 7.6%p and F-1 score was 4.1%p. In the future, it is judged that there is a possibility to increase the precision of wilt detection by applying various filtering techniques, and it is judged that the drone surveillance method using drone orthographic images and artificial intelligence can be used in the pine wilt nematode disaster prevention project.

Keywords : U-Net Models, Deep Learning Training Data, Dead Pine Trees, Pine Wilt Nematode

### 초 록

소나무 재선충 피해나무는 줄어들고 있으나, 피해 지역은 전국으로 확대되고 있다. 최근에 딥러닝 기술이 발전하면서 소나무재선충 고사목 탐지 연구에 적용이 빠르게 시도되고 있다. 본 연구의 목적은 딥러닝 학습데이터의 효과적인 취득과 정확한 참값을 확보하고, 학습을 통해 U-Net 모델의 탐지능력을 보다 향상시키기 위함이다. 이러한 목적달성을 위해 단계별 딥러닝 알고리즘을 적용한 필터링 방법을 이용하여 딥러닝 모델의 불명확한 분석 근거를 최소화하고, 효율적인 분석 및 판단을 할 수 있도록 하였다. 분석결과 U-Net 알고리즘을 이용한 소나무재선충 고사목 탐지 및 성능향상에 있어 기간별로 분석한 참값을 이용한 U-Net 모델이 기존에 제공하였던 참값을 이용한 U-Net 모델보다 재현율(Recall)은 -0.5%p, 정밀도(Precision)은 7.6%p, F-1 score는 4.1%p로 분석되었다. 향후 다양한 필터링 기법을 적용하여 재선충 탐지 정밀도를 높일 수 있는 가능성이 있을 것으로 판단되며, 드론 정사영상과 인공지능을 이용한 드론 예찰방법이 소나무재선충 방제 사업에 활용 가능할 것으로 판단된다.

핵심어 : U-Net 모델, 딥러닝 학습데이터, 소나무 고사목, 소나무재선충

Received 2022. 07. 20, Revised 2022. 08. 11, Accepted 2022. 08. 30

1) Smart Geo Co., Ltd., CEO (E-mail: ksh9232@smartgio.com)

2) Corresponding Author, Member, Professor, Dept. of Real Estate, Semyung University (E-mail: kwkwon69@semyung.ac.kr)

3) Member, Professor, Dept. of Geography, Kyungpook National University (E-mail: kamcho78@knu.ac.kr)

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.

## 1. 서론

최근 지구온난화로 인한 다양한 이슈를 해결하기 위해 재생 가능한 친환경적 에너지인 산림자원을 활용한 다양한 연구가 진행되고 있다. 산림은 온실가스 흡수, 산사태방지, 생물종의 다양성 보전, 휴양, 산소생산 등의 공익적 가치를 지니고 있다(Korea Forest Service, 2016). 이처럼 산림의 공익적 기능이 중시되면서 산림에 관한 연구가 지속해서 수행되고 있으며(Choi *et al.*, 2015; Lee and Park, 2019), 최근 산림의 혜택을 직간접적으로 방해하는 산림병 해충에 대한 방제의 필요성이 강조되고 있다(Jung *et al.*, 2015; Lee, and Seo, 2016). 산림병해 중 하나인 소나무재선충병은 우리나라 소나무림에 심각한 위협이 되는 질병이다.

소나무재선충(PWD: Pine Wilt Disease)은 1988년 부산에서 처음 발병되어 지금까지 전국으로 확대되어 소나무를 고사시키고 있는 심각한 질병이다. 산림청에서는 2005년 「소나무재선충병 방제특별법」을 제정하여 발병지역의 통제와 방제를 통해 피해를 최소화하려고 노력해 왔다. 소나무재선충병은 동아시아 지역에서 주로 발병하며, 한국과 중국, 일본 등의 소나무 목을 고사시켜 국가의 산림자원을 줄어든게 만드는 막대한 피해를 주는 병해충으로 단기간에 급속히 소나무를 고사 시키며, 한 번 감염되면 치료 회복이 불가능하며, 5월에서 8월까지 매개충의 우화 및 활동으로 인해 9월까지 감염목의 50% 이상이 고사목으로 발현하여 나타난다(Kim, 2022).

최근 4차 산업혁명의 핵심기술 가운데 무인항공기(UAV: Unmanned Aerial Vehicle)가 적극 활용되면서 한국임업진흥원에서도 2017년부터 본격적으로 소나무재선충병 감염 의심목의 위치를 찾아내기 시작하였고, 재선충고사목 발현 시기 중 9월 한 달간 전국의 선단지(재선충병 발생지역과 그 외곽의 확산 우려 지역)를 촬영한 정사영상을 판독하는 데에는 많은 인력과 시간이 소요되고 있다. 특히 전문가의 경험과 능력에 따라 소나무재선충 의심목 검출 정밀도가 차이가 있어 완전한 검출에 한계가 있으며, 소나무재선충 피해목은 줄어들고 있으나 피해지역은 전국으로 확대되고 있는 상황으로 방제에 어려움이 있다.

현재는 딥러닝 기술이 발전되면서 소나무재선충 의심 고사목 탐지 연구에 그 적용이 활발히 시도되고 있다. 본 연구와 관련하여 휴대용 근적외선 카메라로부터 얻어진 수치영상을 이용한 소나무재선충 피해목의 조기감별을 시도한 연구가 있으며(Kim *et al.*, 2011), 무인항공기를 이용하여 정사영상을 구축한 후 재선충선단지 육안관찰을 통하여 231본의 피해 의심목을 추출한 바 있다(Kim *et al.*, 2017). 또한 딥러닝 기법을 이용한 연구로는 RGB 정사영상을 이용하여 객체기

반 옴로(YOLOv2) 모델과 이미지 분할방법(SegNet)을 이용하여 소나무재선충 감염 탐지성능을 비교하였고(Lim *et al.*, 2021), ResNet18 패치기반 분류기를 구성하여 RGB 정사영상을 이용한 히트맵을 통한 재선충 피해목의 분포를 탐색하였다(Zhang *et al.*, 2020).

본 연구의 목적은 드론 정사영상을 이용하여 이미지 객체인 식 딥러닝 방법을 적용하여 소나무재선충 위치정보를 신속하게 탐지하는 방법을 제시하는 것으로, U-Net 모델을 이용하여 딥러닝 탐지 정밀도를 보다 향상시키기 위해서 시기별 정사영상을 이용하여 영상을 분석하고, 정도 높은 학습데이터의 참값을 활용하여 이미지 경계 부분의 오분류 문제를 해결하기 위해 이미지 세그먼테이션 기반 딥러닝 모델을 이용하여 단계별 필터링을 통해 탐지 정밀도를 향상 방안을 연구하고자 한다.

### 1.2 연구의 내용 및 범위

U-Net 모델을 이용하여 소나무재선충 의심목을 효과적으로 추출하고 정밀도를 높이기 위해 한국임업진흥원의 기간별 동일위치의 전·후 정사영상 약 80m × 80m 크기로 고사목 위치데이터를 이용하였다. 이러한 정사영상의 크기는 대용량 데이터의 효율을 위해 적정규모의 사이즈를 테스트 한 결과 약 80m 보다 작을 경우 데이터 처리 효율이 떨어지는 경향이 있었고, 약 80m 보다 큰 경우 데이터 분석에서 속도가 늦어져 본 연구를 위한 서버사양에서 적합한 처리 규격으로 결정하였다.

연구의 시간적 범위는 2017년, 2018년 5월과 9월이며, 공간적 범위는 대상지역 및 구축 데이터 수는 김천시, 남양주시, 김해시, 구미시, 안동시, 밀양시, 포항시, 제주시, 경주시를 대상으로 68장의 정사영상에서 2,846개의 고사목 위치 데이터를 추출 및 구축하였다.

본 연구에서는 학습데이터에 따른 정밀도 비교 분석을 위해 소나무재선충 고사목 위치정보 데이터와 동일지역의 일정기간에 대한 고사목이 영상에서 베어진 것을 확인한 고사목 위치 데이터를 각각 세그먼테이션 하여 학습데이터로 사용하였다. U-Net 모델을 이용하여 기존 소나무재선충고사목 데이터와 재분류한 소나무재선충 고사목 데이터를 각각 세그먼테이션 하여 학습 및 훈련한 후 검증대상지역을 선정하여 정밀도를 확인하였다. U-Net 모델 분석과정중 1단계는 소나무재선충 고사목을 학습데이터로 학습시키고 분석을 수행한다. 2단계에서는 CNN (Convolutional Neural Network)의 Sequential 모델을 이용하였다. CNN Sequential 모델은 딥러닝 모델의 한종류로 이미지 분류에 많이 쓰이는 모델이며, 기존 영상처리의 필터역할을 하는 합성곱(convolution)과 신

경망을 결합한 구조이다. 합성곱 레이어(convolution layer)가 필터를 통해 이미지의 특징을 추출하면 풀링 레이어(pooling layer)에서 특징을 강화하고 이미지의 크기를 줄인다. 이러한 과정을 반복하며 이미지의 특징(feature)을 추출하는 방법이다(Lee *et al.*, 2021). 그래서 2단계에서는 산지와 산지가 아닌 지역으로 나누고 이미지를 학습하여 소나무, 고사목 외 오분류 된 데이터를 필터링 하였다. 3단계에서는 산지 내 오분류 된 데이터를 종류별로 세분화하고 세그멘테이션 하여 소나무 재선충 고사목이 아닌 데이터로 학습시켰다. 이를 통해 딥러닝 모델의 정밀도 향상 방안을 제시하고, 객체탐지 딥러닝 모델 성과와 결과를 비교 평가하였다.

## 2. 딥러닝 모델

### 2.1 U-Net 모델

FCN모델 계열인 U-Net 모델은 시맨틱 분할의 인코더-디코더 아키텍처를 기반으로 인코더는 다운샘플링(downsampling)을 과정을 거치며, 디코더는 업샘플링(upsampling)을 거치는 구조로 되어있다. Fig. 1은 FCN 모델에 마지막 layer인 fully connected layer를 1×1 convolution layer로 교체하고, transposed convolution을 이용하여 업샘플링을 진행하여 입력 데이터와 같은 크기의 맵을 생성하였다.

U-Net 모델은 contract되는 레이어와 업스케일링이 반복되면서 U자 형태의 구조를 가진 네트워크로 contract 과정에서 이미지의 크기를 점차 줄여나가면서 여러 개의 이미지로 나누어 분할된 이미지를 얻기 위해 업샘플링 과정을 거치게 된다. Fig. 2는 입력 영상의 특징을 추출하면서 차원을 축소해가는 컨볼루션 처리와 contract 과정과 위치정보에 대한 손실 값이 원래의 크기로 복원하면서 정답 레이블 영상에 가장 부합하도록 영상필터의 가중치를 결정하는 역으로 응답하는 디콘볼루션 과정을 거쳐 영상인식 정확도를 향상시킨 것을 나타낸 것이다(Olaf *et al.*, 2015).

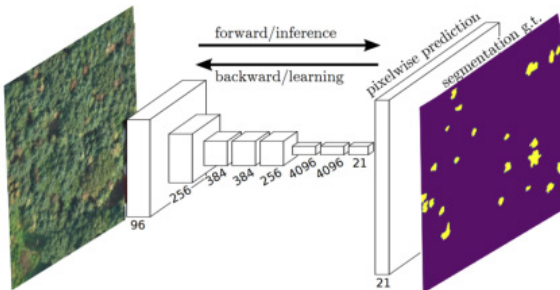


Fig. 1. Transposed convolution upsampling

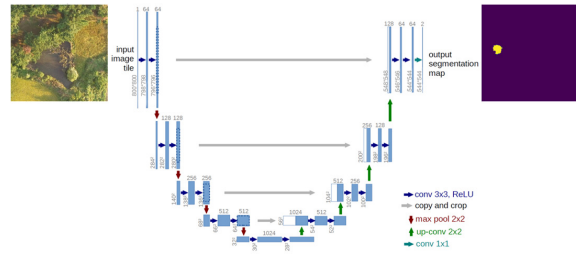


Fig. 2. U-Net model architecture

위치정보 속성을 가진 정사영상 크기는 면적이 넓고 용량이 커서 분석속도가 느리며 컴퓨터 성능이 높아야 한다. 따라서 분석 속도와 데이터 저리 효율을 높이기 위해 정사영상을 영상편집 프로그램을 이용하여 약 80m × 80m 크기로 잘라 영상을 분석하였다. 또한, 공간정보 데이터를 공유하고 편집할 수 있는 지오서버 프로그램을 이용하여 정사영상을 저장하고, 자동으로 일정 크기로 잘라 영상을 분석할 수 있도록 하였다. 지오서버의 줌레벨(zoom Level)은 19~20으로 설정하고 약 80m × 80m 이하로 잘라 영상분석을 실행하였다. U-Net 모델 분석을 위해 학습데이터는 37회(epoch)의 훈련을 수행하여 정확도는 0.996의 성능을 갖추었다.

### 2.2 객체검출 정확도 평가

이미지 객체인식에 있어 객체검출 정확도를 평가하는 방법으로 오차행렬(confusion matrix)을 적용하였다. 오차행렬은 학습된 분류모델의 성능을 나타내는 지표로, 어떠한 유형의 예측오류가 발생했는지 나타내는 지표이며, TP (True Positive: 맞는 것을 맞다고 예측한 것), TN (True Negative: 아닌 것을 틀리다고 예측한 것), FP (False Positive: 아닌 것을 맞다고 예측한 것), FN (False Negative: 맞는 것을 틀리다고 예측한 것)의 개념을 가지고 있으며, 이를 통해 재현율(recall)은 Eq. (1), 정밀도(precision)는 Eq. (2), F-1 score는 Eq. (3), 정확도(accuracy)는 Eq. (4)로 산출할 수 있다(Kotsiantis *et al.*, 2006). 이때 각각의 의미는 재현율은 실제로 true인 데이터를 학습모델이 true라고 인식한 데이터의 수를 의미하며, 정밀도는 학습모델이 true로 예측한 데이터 중 실제로 true인 데이터의 수를 의미한다. 그리고 학습모델의 성능을 측정하는데 있어서 여전히 학습모델이 얼마나 효과적인지를 설명할 수 있는 한 가지 지표를 더 필요로 하는데 이 때 F-1 score가 사용되며, F-1 score는 정밀도와 재현율을 조합하여 하나의 통계치로 계산하는데 사용된다. 그리고 정확도는 올바르게 예측된 데이터의 수를 전체 데이터의 수로 나눈 값이다. 위에서 언급한 내용에 대한 계산식은 다음과 같으며, 전체적인 연산과

정은 Table 1과 같다.

$$\text{재현율 } Recall = \frac{TP}{TP + FN} \quad (1)$$

$$\text{정밀도 } Precision = \frac{TP}{TP + FP} \quad (2)$$

$$F1 - score = \frac{2 \times recall \times precision}{recall + precision} \quad (3)$$

$$\text{정확도 } Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4)$$

Table 1. Confusion matrix

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

### 3. 학습데이터 구축 및 딥러닝 모델 구축

#### 3.1 정사영상을 이용한 참값 학습데이터 구축

학습데이터는 Table 2와 같이 정사영상 총 213장과 소나무 재선충 고사 의심목 위치데이터 6,189개를 이용하였다.

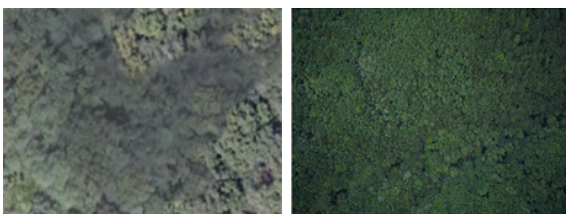


Fig. 3. The blurring and vignetting images

Fig. 3과 같이 위의 학습데이터 중 영상이 번져 흐리게 보이는 블러링(blurring) 현상과 구름 및 날씨의 밝기 변화 원인에 따른 명암으로 발생하는 비네팅(vignetting) 현상 등으로 학습데이터로 사용 불가능한 위치 데이터는 제외 하였고, 추가적으로 제거목이 확인된 나무는 별도 레이블링을 통해 구축 하였다. 또한 정사영상이 어두운 이미지의 경우 소나무재선

충 탐지 정밀도가 떨어질 수 있어 정사영상을 검수 및 색상보정을 진행하였다. 소나무재선충 고사의심목 데이터 중 참값이 아닌 오류 학습데이터로 인하여 탐지모델의 정밀도가 낮아지는 문제를 방지하기 위해서 정사영상 상에서 소나무재선충 고사목이 방제되어 제거된 이미지를 추출하여 레이블링한 것만을 학습데이터로 사용하였다.

Table 2. Pine wilt disease dead wood orthoimage and learning materials

Location of image	Shooting time (Y/M)	Number of images	Number of objects	Species of trees.
Uljin	2018/10	19	88	pine
Yeongyang	2018/9	14	273	pine
Bonghwa	2018/10	7	167	pine
Cheongsong	2018/10	11	317	pine
Sangju	2018/9	14	229	pine
Gimcheon	2018/10	11	155	pine
Hapcheon	2018/10	9	186	pine
Sancheong	2018/10	14	217	pine
Hadong	2018/10	12	314	pine
Seocheon	2017/5	13	69	pine
Gyeonggi area	2017	17	63	pine
<b>Subtotal</b>		<b>141</b>	<b>2,078</b>	
Gyeonggi area	2017	17	885	oak
Anyang	2018	3	127	oak
Incheon	2018	1	253	oak
<b>Subtotal</b>		<b>21</b>	<b>1,265</b>	
Gwangju	2017/2018	5	124	pine
Namyangju	2017/2018	6	71	pine
Gimhae	2017/2018	8	282	pine
Gumi	2017/2018	8	628	pine
Andong	2017/2018	4	15	pine
Yeongju	2017/2018	3	3	pine
Ulju/Miryang	2017/2018	8	624	pine
Pohang	2017/2018	6	323	pine
Jeju	2017/2018	12	418	pine
Gyeongju	2017/2018	8	358	pine
<b>Subtotal</b>		<b>68</b>	<b>2,846</b>	
<b>Sum</b>		<b>213</b>	<b>6,189</b>	<b>Pine : 4,924</b>

추출방법은 같은 위치의 정사영상을 QGIS에서 중첩하고,

재선충고사목의심목 위치데이터를 참조하여 제거목을 확인하고 레이블링하였다. Fig. 4는 시기별 정사영상을 이용하여 소나무재선충 고사목을 추출한 것이다. 제거목 확인 및 레이블링 작업을 진행하여 기간별 정사영상 총 68개에서 추출한 소나무재선충 고사목 위치데이터 학습자료수 2,846개 중에서 별채목 2,573개를 참값으로 구축하였다. 또한 구축한 참값 데이터 중 딥러닝 모델의 탐색성능 결과를 검증하기 위해 경주시 양월리 참값 데이터 450개는 학습에서 제외하였다. 이때 한국임업진흥원에서 제공받은 경주시 지역의 학습자료수는 358개였으나 별채목 확인결과 소나무재선충 고사목의 실제 수량은 450개로 확인되었다.



Fig. 4. Extraction of true value of pine wilt nematode dead tree using orthographic images by period

소나무재선충 고사목은 감염초기 점차 연노랑으로 소나무

상부부터 변하기 시작하며, 중기에는 검붉은색으로 완전히 갈변하며, 말기에는 회백색을 띠다가 잎이 탈리된다. 따라서 재선충고사목의 다양한 형태와 색깔을 인식 및 탐지하기 위해서 별채목 확인결과를 Table 3과 같이 코드로 구별하여 기간분석 추출 참값을 학습데이터로 구축하였다.

기간분석 추출 참값 데이터의 초기, 중기, 말기로 구분하기 위해 Fig. 5와 같이 이미지를 세분화하여 데이터 속성정보를 각각 Code 1, Code 2, Code 3으로 구분하여 입력하였으며, 검증지역을 제외한 구축데이터 수는 초기 19개, 중기 1,967개, 말기 137개로 총 2,123개의 데이터를 학습데이터로 사용하였다.

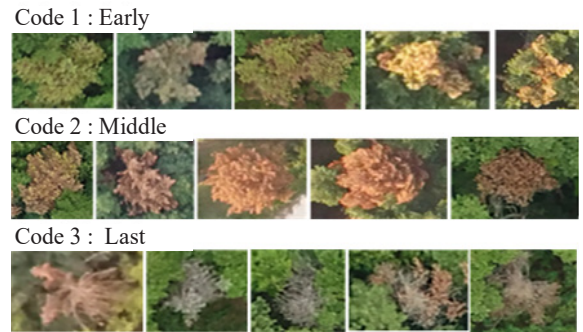


Fig. 5. Criteria for classification of pine wilt nematodes in the early, middle, and last stages

Table 3. Period analysis extraction true value

Shooting area	Code 1 (early)	Code 2 (middle)	Code 3 (last)	Sum	Note
Jeoji-ri, Hangeong-myeon, Jeju	18	252	20	290	
Yongpyeong-dong, Miryang		397	15	412	
Geomam-ri, Andong		190	10	200	
Doyori, Saengnim-myeon, Gimhae		108	2	110	
Siuri, Joan-myeon, Namyangju		9	4	13	
Ipsok-ri, Nongso-myeon, Gimcheon		27	2	29	
Seohari, Duseo-myeon, Ulju-gun		89	9	98	
Jungmyeong-ri 1, Yeonil-eup, Nam-gu, Pohang		27	14	41	
Jungmyeong-ri 2, Yeonil-eup, Nam-gu, Pohang		293	29	322	
Geogeo-dong, Gumi, Gyeongsangbuk-do		224	22	246	
Bonghan-ri, Goa-cup, Gumi	1	343	7	351	
Taejang-ri, Seohu-myeon, Andong		8	3	11	
<b>Sum</b>	<b>19</b>	<b>1,967</b>	<b>137</b>	<b>2,123</b>	
Yangwol-ri, Angang-eup, Gyeongju	12	419	19	450	Verification Area
<b>Total sum</b>	<b>31</b>	<b>2,386</b>	<b>156</b>	<b>2,573</b>	

### 3.2 U-Net 모델 구축

소나무재선충 고사목 탐감 데이터와 U-Net 모델을 이용하여 분석시행 결과 오분류 문제와 미분류 문제가 발생한다. 이것은 정밀도(precision)와 재현율(recall) 저하의 요인이 된다. 따라서 성능을 높이기 위해서는 일반적으로 학습데이터의 종류와 양을 늘리는 것이 일반적이다. 그러나 탐지 결과에 대한 문제 발생 시 딥러닝 모델들은 결과에 대한 원인을 알 수가 없기 때문에 이러한 문제점을 최소화하고 성능을 높이는 방안으로 단계별 딥러닝 필터링을 적용하였다.

특히 색상과 형태가 비슷한 잡풀, 건물지붕, 묘지, 밭, 나무그림자 땅 등 다양한 이미지가 주로 오분류되어 정밀도 저하의 요인이 된다. 이러한 오분류 데이터를 필터링하여 정밀도를 높이기 위해서 CNN Sequential 모델을 이용하여 필터링 분석을 시행하고 오분류 데이터를 제외하였다.

Sequential 모델의 학습데이터는 정사영상을 256m × 256m로 그리드 형태로 잘라 산과 기타 파일이름으로 구별하여 Table 2와 같이 총 1,237개 산 이미지와 563개의 기타 이미지로 학습을 시켰다. 학습 파라미터 설정을 위해 epoch는 100번 수행하여 훈련 시켰다. Fig. 6은 모델의 파라미터 설정값을 나타낸 것으로 Fig. 7과 같이 오분류 데이터 제거를 수행하였다.

```

train_datagen = ImageDataGenerator(
    rotation_range=15,
    rescale=1. / 255,
    shear_range=0.1,
    zoom_range=0.2,
    horizontal_flip=True,
    width_shift_range=0.1,
    height_shift_range=0.1)
    
```

Fig. 6. Training data setting parameters of sequential model

학습한 Sequential 모델을 이용하여 안동지역 정사영상 검증 데이터 (9,629m × 15,104m)를 이용하여 정밀도를 분석한 결과 Table. 4와 같이 적중률은 산지의 경우 약 85.9%, 기타 약 94.3%로 나타났다.

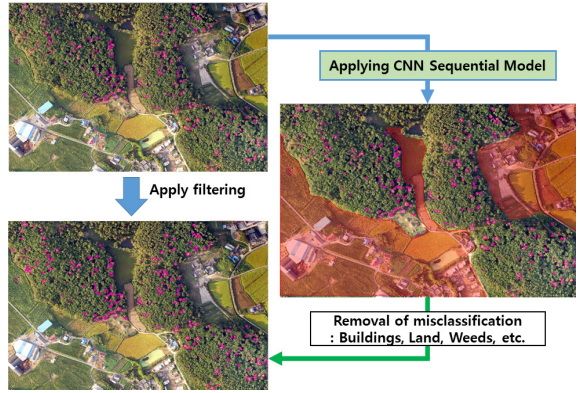


Fig. 7. Removal of misclassification outside of mountains using CNN Sequential model

학습한 Sequential 모델을 이용하여 안동지역 정사영상 검증 데이터 (9,629m × 15,104m)를 이용하여 정밀도를 분석한 결과 Table. 4와 같이 적중률은 산지의 경우 약 85.9%, 기타 약 94.3%로 나타났다.

또한 오분류 및 미분류율을 낮추기 위해 Table 2의 참나무 시들음병 고사목 이미지 1,265개를 필터링을 위해 추가로 학습시켰다. Fig. 8은 학습시킨 참나무 시들음병 고사목 이미지 예시이다. 또한 Fig. 9는 오분류를 방지하여 탐지 성능을 높이기 위해 필터링 및 산출결과를 학습데이터로 활용하는 처리 흐름도를 나타낸 것이다.



Fig. 8. Oak wilt disease dead tree learning data

Table 4. CNN Sequential model performance evaluation

Division	Total number of images	Number of correct answers	Number of incorrect answers	Hit rate(%)	Note
Mountain	1,237	1,063	174	85.9	
Etc	563	531	32	94.3	

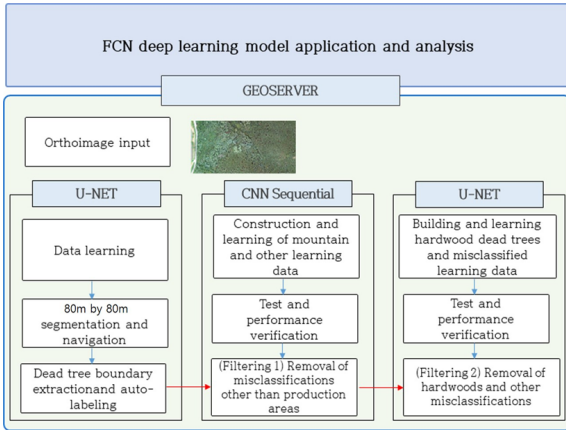


Fig. 9. Use of filtering and calculation result learning data

### 3.4 U-Net 모델 검증

Table 5와 같이 구축된 U-Net 모델의 검증 지역은 경주시 양월리 지역 전체 정사영상 중 603m × 518m 크기의 대상지를 선정하였다.

Table 5. Establishment of true value of pine wilt nematode dead wood in verification area

Area	Division	Number
Yangwol-ri, Angang-eup, Gyeongju-si (603m × 518m)	Period analysis extraction true	228
	Existing provided true value	199

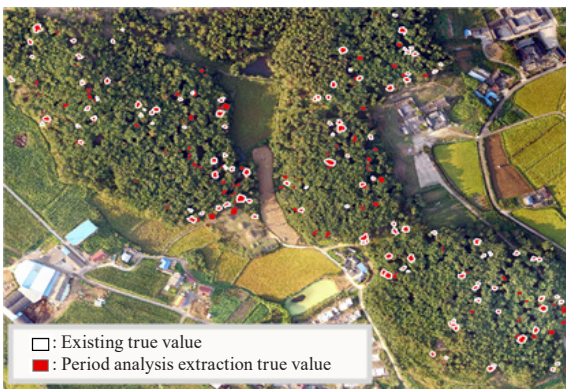


Fig. 10. Compare the original provided true value and the period analysis extracted true value

검증지역 단일 정사영상 분석 고사목 참값과 기간 분석 추

출 참값은 각각 228개, 199개가 추출되었으며, 추출결과 객관적인 제거목 검출이 가능하며, 제거된 고사목의 명확한 확인으로 참값 데이터의 수량이 기존 제공 참값 보다 증가하였다. Fig. 10은 기존 제공 참값과 기간분석 추출 참값을 비교한 것이다.

#### 3.4.1 기존 제공 참값을 이용한 U-Net 모델 분석

한국임업진흥원에서 제공받은 기존 제공 참값 4,924개와 참나무시들음병 1,265개를 학습시킨 U-Net 모델을 이용하여 검중 정사영상에 적용하여 분석한 결과 Fig. 11과 같이 재선충고사목이 탐지되었으며, Table 6과 같이 272개의 소나무재선충 의심목이 집계되었다. 그러나 일부 소나무재선충 의심 고사목 외 건물, 묘지, 땅, 풀, 밭 등과 같이 색상과 부분 형상이 유사한 것들이 오분류 되었다. 오분류된 이미지는 CNN Squential 딥러닝 모델을 이용하여 필터링을 수행하였다. 그러나 산지의 경계부 일부에서 오분류 데이터가 나타났다.



Fig. 11. U-Net model analysis result using the existing provided true value

분석결과 재선충 고사목이 아닌 것을 재선충 고사목으로 오인식한 수량이 57개가 검지되었으며, 찾아내지 못한 것은 13개가 있었다. 분석에 대한 검증을 위해 오차행렬에서 정밀도와 재현율을 이용하여 산정하고, 모델의 성능 평가에 있어서 정밀도만으로 평가할 경우 클래스별 데이터의 균형이 비슷할 경우 적용 가능하나, 데이터가 불균형한 경우 정밀도가 높으면 재현율이 낮고, 정밀도가 낮으면 재현율이 높은 경향이 나타남으로 정밀도에 따른 클래스별 불균형으로 인한 정밀도의 영향을 평가하기 위한 지표로 F-1 score를 사용하였다. 본 연구에서는 F-1 score는 소나무재선충병 감염목 데이터보다 반대로 정상 소나무 및 임목의 데이터 값인 TN의 개수가 정상적인 소나무 뿐 아니라 임목의 수량이 너무 많아서 카운팅하는 것이 불가능하기 때문에 딥러닝 알고리즘의 최적화된 성

능을 검증하기 위해 불균형한 데이터에 대해서도 성능 측정이 가능한 F-1 score를 사용하였다(Derczynski, 2016). 정밀도와 재현율, F-1 score를 계산한 결과는 Table 6과 같다.

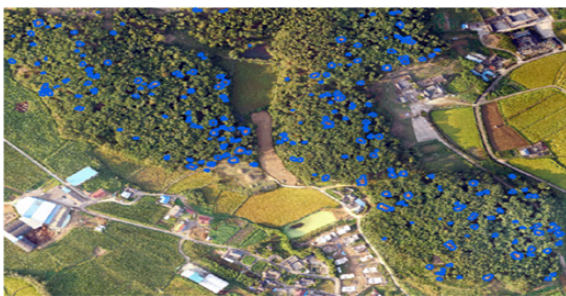
**Table 6. Result of error matrix analysis using existing true values**

Classification	Image of pine wilt disease damage	Other images	Total
Real			
Image of pine wilt disease damage	TP 215	FN 13	228
Other images	FP 57	TN -	57
Total	272	13	
Recall	$\frac{TP}{TP+FN} = 0.943$		
Precision	$\frac{TP}{TP+FP} = 0.790$		
F-1 score	$\frac{2 \times Recall \times Precision}{Recall + Precision} = 0.860$		

3.4.2 기간 분석 추출 참값을 이용한 U-Net 모델 분석

시기별 영상을 분석하여 제거된 고사목 참값 데이터를 추출하여 학습시킨 U-Net 모델을 이용하여 검증지역 정사영상에 적용하여 Fig. 12와 같이 재선충고사목 분석을 진행하였다.

분석결과 건물, 모지, 땅, 풀, 밭과 같은 오분류가 나타났으며, CNN Sequential 딥러닝 모델을 이용하여 필터링을 수행하여 Table 7와 같이 247개의 소나무재선충 의심목이 탐지되었다. 또한 재선충 고사목이 아닌 것을 재선충 고사목으로 오인식한 수량이 33개가 검지되었으며, 찾아내지 못한 것은 14개가 있었다. 이를 바탕으로 오차행렬을 이용하여 정밀도와 재현율, F-1 score를 계산한 결과이다.



**Fig. 12. Results of U-Net model analysis using periodic analysis extract true values**

**Table 7. Result of error matrix analysis using period analysis extraction true value**

Classification	Image of pine wilt disease damage	Other images	Total
Real			
Image of pine wilt disease damage	TP 214	FN 14	228
Other images	FP 33	TN -	33
Total	247	14	
Recall	$\frac{TP}{TP+FN} = 0.938$		
Precision	$\frac{TP}{TP+FP} = 0.866$		
F-1 score	$\frac{2 \times Recall \times Precision}{Recall + Precision} = 0.901$		

Table 6과 7을 통해 학습데이터 참값에 따른 U-Net 모델을 분석한 결과 Table. 8과 같이 기존 제공 참값 데이터보다 기간별 고사목 확인을 통한 재선충 고사목 데이터 구축 방법이 정밀도는 약 7.6%p 높았으며, 재현율은 약 0.5%p가 낮게 분석되었다.

**Table 8. Comparison of deep learning model analysis according to the true value of training data**

Division	U-Net result using the existing provided true value	U-Net result using period analysis extraction true value	Note
Recall	0.943	0.938	- 0.5%p
Precision	0.790	0.866	7.6%p
F-1 Score	0.860	0.901	4.1%p

5. 결론 및 시사점

본 연구는 딥러닝 학습데이터의 효과적인 취득과 정확한 참값을 확보하고, 학습을 통해 U-Net 모델의 탐지능력을 보다 향상시키는데 목적이 있다. 이를 위해 단계별 딥러닝 알고리즘을 적용한 필터링 방법을 이용하여 딥러닝 모델의 불명확한 분석 근거를 최소화하고, 효율적인 분석 및 판단을 할 수 있도록 하였고 분석결과 U-Net 알고리즘을 이용한 소나무 재선충 고사목 탐지 및 성능향상에 있어 기간별로 분석한 참값을 이용한 U-Net 모델이 기존에 제공하였던 참값을 이용한 U-Net 모델보다 재현율은 -0.5%p, 정밀도는 7.6%p, F-1 score



는 4.1%p로 분석되었으며 다음과 같은 결론을 얻었다.

첫째, 기존 제공 참값을 이용한 U-Net 모델의 탐지 성능이 떨어지는 원인으로는 기존에 제공된 참값은 2017년과 2018년 수집된 데이터로 드론 도입 및 활용 초기로 영상촬영의 해상도와 관련된 품질과 분석기술이 현재의 촬영 및 분석 기술보다 낮아 참값의 정도가 떨어졌으며, 특히 드론 촬영 정사영상의 해상도 차(10cm~20cm)와 블러링 및 비네팅 등에 의한 영상분석의 정확성 저하가 가장 큰 원인인 것으로 판단된다.

둘째, 정사영상 상의 수목의 수직 이미지 형상으로는 침엽수와 활엽수, 고사목과 계절에 따른 단풍, 잎의 갈변목들과의 식별이 어려운 경우가 많아 영상분류 시 오판 문제가 발생할 수 있다. 따라서 정확한 소나무재선충 고사목 참값을 이용하는 것이 탐지 정도를 높일 수 있으며, 기간별 분석 추출 참값(기간에 따른 소나무재선충 고사 제거목을 분석한 참값) 데이터 수집 방법을 통해 성능을 개선할 수 있을 것으로 판단된다.

셋째, 또한 정사영상 상의 식별이 어려운 문제를 해결하고 성능을 높이기 위해 소나무재선충 의심 고사목 외 건물, 묘지, 땅, 풀, 밭 등과 같이 색상과 부분 형상이 유사한 것들이 오분류 되어 성능이 저하되는 문제를 CNN Sequential 모델을 적용하여 산림과 비산림 지역으로 구분하여 식별하고 그에 따른 오분류 데이터를 다시 필터링 함으로써 오분류를 현저히 줄일 수 있도록 성능을 개선할 수 있었다.

본 연구를 통해 9월달 한 달 내의 정사영상을 이용하여 소나무재선충 의심 고사목을 탐지하여야 하는 계절적 특수성 때문에 10월달 이후 촬영 및 제작된 정사영상의 경우 단풍, 활엽수 및 침엽수의 갈변 등에 의해 탐지 성능이 떨어지는 경우가 발생되었다. 이는 해상도를 높이고 소면적으로 분석할 경우 탐지 성능을 높일 수는 있으나 전국적으로 대면적의 촬영을 통해 신속히 분석해야 하는 상황을 고려할 경우 활용성에 한계가 있음을 확인할 수 있었다.

향후 분석 PC의 성능이 높아지고 영상 촬영 및 실시간 데이터 전송 기술이 발달하여 상용화된다면 높은 해상도에서 탐지 성능 개선과 활용이 가능할 것으로 판단된다. 또한 본 연구의 딥러닝 알고리즘을 적용한 필터링 방법의 성능 우수성의 검증을 통해 향후 다양한 필터링 기법을 적용하여 재선충 탐지 정밀도를 높일 수 있는 가능성이 있을 것으로 판단되며, 드론 정사영상과 인공지능을 이용한 드론 예찰방법이 소나무재선충 방제 사업에 활용 가능할 것으로 판단된다.

## References

Choi, Y.R., Lee, J.S., and Yun, H.C. (2015), Extraction

of forest resources using high density LiDAR data, *Journal of the Korean Society of Survey Geodesy Photogrammetry*, Vol. 33, No. 2, pp. 73-81. (in Korean with English abstract).

<https://doi.org/10.7848/ksgpc.2015.33.2.73>

Derczynski, L. (2016), Complementarity, F-score, and NLP evaluation, *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, European Language Resources Association (ELRA), pp. 261-266.

Jung, Y.J. (2015), Pine wood nematode, *The Magazine of The Korean Society of Hazard Mitigation*, Vol. 15, pp. 124-128.

Kim, M.I., Lee, W.K., and Kwon, T.H. (2011), Early detecting damaged trees by pine wilt disease using DI(Detection Index) from portable near infrared camera, *Journal of Korean Society of Forest Science*, Vol. 100, No.3, pp. 374~381. (in Korean with English abstract). <https://doi.org/10.14578/jkfs.2011.100.3.8>

Kim, M.J., Bang, H.S., and Lee, J.W. (2017), Accredited use of unmanned aerial vehicle for forecasting pine wood nematode in boundary area: A case study of Sejong metropolitan autonomous city, *Korean Society of Forest Science*, Vol. 106, No. 1, pp. 100~109. (in Korean with English abstract)

<https://doi.org/10.14578/jkfs.2017.106.1.100>

Kim, S.H. (2022), *A Study on the Improvement of Orthoimage Detection by the U-Net model : Focusing on the Case of Pine Wilt Disease*, Ph.D. dissertation, Youngnam University, Gyeongsan, Korea, 91p.

Korea Forest Service. (2016), *National Forest Inventory Report*, Korea Forest Service, Daejeon, Korea, pp. 1-192 (in Korean).

Kotsiantis, S., Kanellopoulos, D. and Pintelas, P. (2006). "Handling imbalanced datasets: A review." *GESTS International Transactions on Computer Science and Engineering*, Vol. 30, No. 1, pp. 25-36.

Lee, K.W. and Park, J.K. (2019), Economic evaluation of unmanned aerial vehicle for forest pest monitoring, *Journal of the Korea Academia Industrial cooperation Society*, Vol. 20, No. 1, pp. 440-446 (in Korean with English abstract).

<https://doi.org/10.5762/KAIS.2019.20.1.440>

Lee, N.K., Kim, J.Y., and Shim, J.H. (2021), Empirical study on analyzing training data for CNN-based product classification deep learning model, *The Journal of Society for e-Business Studies*, Vol. 26, No. 1, pp. 107-126. (in Korean with English abstract)

<https://doi.org/10.7838/jsebs.2021.26.1.107>

Lee, S.K. and Seo, S.T. (2016), First report of Armillaria root disease caused by Armillaria tabescens on carpinus tschonoskii in south Korea. *Plant Disease*. Vol. 100, No. 1, pp. 213.

Lim, E.T. and Do, M.S. (2021), Pine wilt disease detection based on deep learning using an unmanned aerial vehicle. *Journal of Civil and Environmental Engineering Research*, Vol. 41, No. 3, pp. 317–325. (in Korean with English abstract).

<https://doi.org/10.12652/Ksce.2021.41.3.0317>

Olaf, R., Philipp, F., and Thomas B. (2015), U-Net : Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, 9351, pp. 234–241.

Zhang, R.R., You, J., Kim, B.J., Sun, J.N., and Lee, J.H. (2020), Searching the damaged pine trees from wilt disease based on deep learning, *Korean Institute of Smart Media*, Vol. 9, No. 3, pp. 46-51. (in Korean with English abstract).

<https://dx.doi.org/10.30693/SMJ.2020.9.3.46>