

2025 인재양성 프로그램
연구자 최종 결과보고서

시민과학자 데이터를 활용한 딥러닝 기반 국내 맞춤형 생태계교란 거북 종 분류 모델 개발

Deep Learning-Based Classification Model Development
for Invasive Turtle Species in Korea Using Citizen Science Data

2026.01.15

백종원(특정주제 연구자)

제 출 문

재단법인 숲과나눔 이사장 귀하

본 보고서를 “시민과학자 데이터를 활용한 딥러닝 기반 국내 맞춤 생태계교란
거북 종 분류 모델 개발”의 최종보고서로 제출합니다.

2026년 1월 15일

연구 원 : 백종원(특정주제연구자)

※ 본 보고서의 내용은 연구자의 의견이며, (재)숲과나눔의 공식적인 견해와는 다를 수
있습니다.

결과보고서 초록

이름	백 종 원	구분	특정주제 연구 장학생
연구제목	한글	시민과학자 데이터를 활용한 딥러닝 기반 국내 맞춤형 생태계교란 거북 종 분류 모델 개발	
	영문	Deep Learning-Based Classification Model Development for Invasive Turtle Species in Korea Using Citizen Science Data	
연구기간	2025.3.1 ~ 2025.12.31		
색인어	한글	침입종, 거북, 딥러닝, 이미지 분류, 시민과학	
	영문	Invasive species, Turtle, Deep learning, Image classification, Citizen science	

침입종은 전 세계적으로 생물다양성 손실의 주요 원인 중 하나로 인식되고 있다. 한국에서는 침입성 담수 거북에 의한 생태계 교란이 증가하면서 신속하고 정확한 종 식별이 가능한 실용적인 관리 기술이 시급히 요구되고 있다. 딥러닝 기반 객체 탐지 모델인 You Only Look Once(YOLO)는 높은 추론 속도와 향상된 분류 정확도를 제공하여 실시간 모니터링 및 애플리케이션 기반 배포에 적합하다. 본 연구에서는 YOLO 기반 접근법을 활용하여 침입성 담수거북 자동 분류 모델을 개발하고 시민과학 현장 데이터 적용 가능성을 평가하였다. 국내 생태 조건을 고려하여 생태계교란 거북 6종과 국내 자생 거북 2종을 대상 분류군으로 선정하였다. 시민과학자가 수집한 이미지를 훈련 데이터로 활용하였으며, 모델 훈련 중 손실 함수 변화와 훈련 후 평가 지표를 통해 모델 성능을 평가하였다. 모델은 과적합 증거 없이 안정적인 수렴을 보였다. 훈련 후 평가에서 추론 시간은 9.1ms, mean Average Precision (mAP)는 0.965, 평균 분류정확도는 92.5%를 기록했다. 현장 적용성을 평가하기 위해 실제 현장에서 수집한 시민과학 이미지에 모델을 적용한 결과 73.1%를 올바르게 분류하였다. 특히 상대적으로 낮은 출현 빈도를 보이는 종인 중국줄무늬목거북에 대해서는 100% 정확도를 달성했다. 이러한 결과는 딥러닝 기반 이미지 분류 모델이 시민과학자 및 형태 기반 종 분류 전문가의 종 식별 과정을 보조하는 기술적 도구로 활용 가능성을 제시한다. 본 연구에서 개발된 모델은 시민과학 기반 모니터링과 전문가 분석을 지원함으로써, 생태계교란 거북에 대한 조기 탐지 및 신속 대응(Early Detection and Rapid Response, EDRR) 전략의 성공률을 높일 수 있을 것으로 기대한다.

Abstract

Invasive species are recognized as one of the major drivers of biodiversity loss worldwide. In Korea, ecological disturbances have been increasingly caused by invasive freshwater turtles, creating an urgent need for practical management technologies that enable rapid and accurate species identification. You Only Look Once (YOLO), a deep learning (DL)-based object detection model, offers high inference speed and enhanced classification accuracy, making it suitable for real-time monitoring and for application-based deployment. In this study, we developed an automated classification model for invasive freshwater turtles using a YOLO-based approach and evaluated its applicability to citizen science-derived field data. Based on domestic ecological conditions, six invasive and two native turtle species were selected as target classes. Images collected by citizen scientists were used as training data. Model performance was evaluated using training-phase loss trends and post-training evaluation metrics. The model showed stable convergence without evidence of overfitting. In post-training evaluation, the inference time was 9.1 ms, mean Average Precision (mAP) reached 0.965, and average correct classification rates was 92.5%. To assess field applicability, the model was applied to real-world citizen science images, achieving a correct classification rate of 73.1%. Notably, 100% accuracy was obtained for *Mauremys sinensis*, a species with relatively low occurrence frequency. These results demonstrate that DL-based classification models using image data have strong potential as supportive technological tools for species identification by both citizen scientists and morphological classification experts. By facilitating citizen science-based monitoring and expert-driven analysis, the proposed model may contribute to Early Detection and Rapid Response (EDRR) management strategies for invasive freshwater turtles.

요약문

I. 제목

시민과학자 데이터를 활용한 딥러닝 기반 국내 맞춤 생태계교란 거북 종 분류 모델 개발

II. 연구 배경 및 목적

외래종 침입은 생물다양성 감소를 초래하는 주요 요인 중 하나로, 국내에서도 외래 민물거북을 중심으로 생태계 교란 문제가 지속적으로 보고되고 있다. 이에 따라 현장 적용이 가능하면서도 신속하고 정확한 종 식별을 지원할 수 있는 관리 기술의 개발이 요구된다. 딥러닝 기반 객체 탐지 모델인 You Only Look Once(YOLO)는 빠른 추론 속도와 높은 분류 정확도를 갖추고 있으며 애플리케이션 기반 현장 모니터링 환경에서 외래 민물거북의 자동 종 판별에 활용될 수 있다. 본 연구의 목적은 YOLO 모델을 활용하여 국내에서 지속적으로 발견되는 생태계교란 거북 6종과 국내 자생 거북 2종을 대상으로 자동 종 분류 모델을 개발하고, 시민과학 기반 현장 데이터에 대한 적용 가능성을 평가하는 데 있다.

III. 연구 결과 및 결론

개발된 모델은 학습 과정에서 과적합이 발생하지 않았으며 mAP0.5 기준 0.965, 평균 분류정확도 92.5%를 기록하여 민물거북 종 탐지 및 분류에서 높은 성능을 나타냈다. 특히 생태계교란 거북이 국내 자생 거북으로 오분류된 사례는 확인되지 않았으며, 이는 외래종 관리 측면에서 중요한 의미를 지닌다. 또한 개발된 모델을 실제 시민과학 데이터에 적용하여 현장 활용 가능성을 평가한 결과, 모델은 현장 데이터에서도 73.1%를 올바르게 분류하였다. 아울러 출현빈도가 낮은 중국줄무늬목거북에 대해서도 100% 올바르게 제시할 수 있음을 확인하였다. 이러한 결과는 딥러닝 기반 이미지 분류 모델이 시민과학자와 형태 기반 종 분류 전문가의 종 식별 과정을 보조하는 기술적 도구로 활용될 수 있음을 시사한다. 향후 데이터 불균형 문제를 완화, 다양한 전처리 기법 적용, 시민과학자 대상 가이드라인 제공 등을 체계적으로 개선하는 경우, 본 모델은 시민과학 기반 현장 조사와 전문가 중심의 대규모 데이터 분석 과정에서 발생하는 종 분류 및 식별 작업을 효율적으로 처리하는 데 활용될 수 있다. 나아가 생태계교란 거북에 대한 조기 탐지 및 신속한 관리 대응(Early Detection and Rapid Response, EDRR)을 기술적으로 지원하여 EDRR 전략의 성공률을 높일 수 있을 것으로 기대한다.

목 차

I. 연구 배경 및 목적	1
1. 연구 배경	1
2. 연구 목적	2
II. 연구 방법	3
1. 데이터 수집 및 선별	3
2. 모델 학습	4
3. 모델 성능 평가	6
III. 연구 결과	8
1. 모델 학습 과정 중 성능 평가	8
2. 모델 학습 후 성능 평가	9
3. 모델 학습 후 분류 결과	11
4. 현장 적용성 평가 결과	12
IV. 연구 논의	15
V. 연구 결론	18
VI. 인용문헌	19

그림 목 차

그림 1. YOLO11 모델의 구조도	4
그림 2. Stochastic gradient descent(SGD) optimizer 원리	5
그림 3. 모델 학습 기간 동안 평가된 train loss 분석	8
그림 4. 모델 학습 기간 동안 평가된 validation loss 분석	9
그림 5. 개발된 모델의 precision-recall curve	10
그림 6. 모델의 분류 결과를 타나낸 confusion matrix	11
그림 7. 시민과학자에 의해 수집된 이미지의 종 별 수집 위치 및 구성	12
그림 8. 본 연구에서 개발한 모델의 시민과학자 이미지 분류 결과 예시	13
그림 9. 시민과학자와 이미지분류 모델의 종 별 분류 정확도 그래프	14

표 목 차

표 1. 생태계교란 거북 6종 및 국내 자생 거북 2종 dataset 구축 결과	3
표 2. 개발된 모델의 성능 평가 결과	10

I. 연구 배경 및 목적

1. 연구 배경

생물다양성은 건강한 토양 유지, 영양소 순환, 기후 조절 등 생태계의 다양한 기능에 영향을 미치며 인간의 생활에도 직접적으로 관련되어 있다(Brusgaard et al., 2007; Mooney et al., 2009). 그러나 이러한 중요성에도 불구하고 최근 생물다양성의 급격한 감소가 관찰되고 있다 (Pimm et al., 2014). 생물다양성 감소의 원인 중 하나는 무역의 가속화에 따른 외래종 침입이며, 이는 생물다양성 감소의 주요 요인으로 지속적으로 대두되고 있다(Havel et al., 2015). 외래종은 토착종과의 서식지 경쟁을 통해 토착 생물의 서식에 부정적인 영향을 미치고 외래 병원균을 확산시켜 생태계 기능의 변화를 초래하는 것으로 알려져 있다(Ricciardi, 2007; Smith et al., 2009; Strayer et al., 2006). 이에 따라 우리나라를 비롯한 여러 국가들은 외래종 침입을 방지하기 위해 다각적인 노력을 기울이고 있다. 현재 우리나라는 환경부에서 "생태계교란생물"을 법적으로 지정하여 외래종을 관리하고 있다.

민물거북은 현재 늑대거북(*Chelydra serpentina*), 악어거북(*Macrochelys temminckii*), 중국줄무늬거북(*Mauremys sinensis*), 리버쿠테(*Pseudemys concinna*), 플로리다붉은배거북(*Pseudemys nelsoni*), 붉은귀거북속 전종(*Trachemys* spp.)이 생태계교란생물로 지정되어 있으며, 이는 척추동물 중 가장 많이 지정된 분류군이다(<https://kias.nie.re.kr/>). 민물거북은 수명이 길고 최적의 서식지에서도 생존할 수 있는 특성을 지니고 있어 다른 분류군에 비해 생태계에 미치는 영향이 클 수 있다(Gariboldi and Zuffi et al., 1994; Spinks et al., 2003). 또한 최근 기후변화로 인해 기온 및 수온이 상승하면서 온난한 기후에 적응한 민물거북들이 국내 토착종과 경쟁할 가능성이 높아지고 있다(Zhang et al., 2023). 대표적으로, 늑대거북과 악어거북은 과거 국내 야생에서 발견되지 않았으나 2020년 이후 지속적으로 발견되고 있다(Koo et al., 2020; Koo et al., 2021). 이러한 생태계교란거북을 효과적으로 모니터링하고 관리하기 위한 필요성이 점차 증대되고 있다.

생태계교란생물을 포함한 외래종 관리를 위한 방법 중 하나로 '조기 탐지와 신속 대응(Early Detection and Rapid Response, EDRR)' 전략이 많이 사용되고 있다(NISC, 2016). EDRR은 침입종이 돌이킬 수 없는 단계에 도달하기 전에 이를 봉쇄하고 근절하기 위해 신속한 대응을 목표로 한다. 그러나 이를 효과적으로 수행하기 위해서는 넓은 범위의 조사 지역을 면밀하게 모니터링해야 하지만, 소수의 연구 인력으로 광범위한 민물거북 서식지를 모두 모니터링하는 것은 한계가 있다. 최근 시민과학자들의 활동이 증가함에 따라 이들이 제공하는 데이터가 모니터링에 중요한 도움을 주고 있다(Suzuki-Ohno et al., 2022). 특히 시민과학자들이 수집한 방대한 양의 데이터는 국내의 다양한 조사 지역을 포함하고 있어, 국내 상황에 맞는 외래종 모니터링 결과 해석을 가능하게 한다. 다만, 이러한 데이터를 연구에 온전히 활용하기 위해서는 종 식별의 정확성이 보장되어야 하며, 따라서 현장 조사

시 시민과학자들이 효과적으로 종을 판별할 수 있는 방법이 필요하다.

딥러닝 기반 이미지 분류 기술은 이러한 문제를 완화하는데 유용할 수 있다. 딥러닝 기반 이미지 분류 기술은 이미 야생동물 모니터링에 성공적으로 적용되고 있다. 대표적으로 객체 탐지 모델은 이미지에서 생물의 객체 탐지와 분류를 동시에 수행할 수 있도록 개발되어 실시간으로 생물의 위치를 탐지하고 어떤 종인지 분류할 수 있다(Baek et al., 2023; Kim et al., 2022). 객체감지 모델 중 You Only Look Once(YOLO) 모델은 탐지 속도가 빠르고 경량화되어 있어 모바일 기기나 드론에 쉽게 탑재할 수 있는 장점이 있다(Li, 2023). 최근 이러한 장점을 기반으로 다양한 동물 및 식물을 대상으로 애플리케이션을 개발하고 시민과학자들에게 배포하여 모니터링을 진행하는 시도가 이루어지고 있다(Christakakis et al., 2024; Luong et al., 2023).

2. 연구 목적

본 연구의 목적은 시민과학자들이 수집한 데이터를 활용하여 현장에서 신속하고 정확하게 거북 종을 분류할 수 있는 YOLO 기반 생태계교란 거북 종 분류 모델을 개발하는 것이다. 연구의 대상종으로는 현재 국내에서 지속적으로 발견되는 생태계교란 거북 6종을 우선 선정하였으며, 국내 자생 거북 2종도 함께 학습하여 생태계교란 거북으로의 오동정을 방지하여 현장 대응성을 높이고자 하였다. 이를 통해 시민과학자들이 현장에서 보다 신속하고 정확한 종 분류가 가능하도록 지원하며, 생태계교란 거북의 조기 탐지와 신속한 대응에 기여하고자 한다.

II. 연구 방법

1. 데이터 수집 및 선별

민물 거북의 표준 dataset이 따로 존재하지 않기 때문에 8종에 대한 거북의 이미지는 iNaturalist(<https://www.inaturalist.org>)에서 수집되었다. 연구등급의 이미지는 iNaturalist의 Inat_images R 스크립트 패키지를 사용하여 수집되었다(Huerta-Ramos, 2021). 수집된 이미지는 학술 서적 및 형태학 기재 논문에 설명된 형태학적 특징을 추출하여 식별되었다(Baek et al., 2024a; Baek et al., 2024b; Farkas et al., 2019; Kim et al., 2021). 종이 명확하게 식별되지 않거나 품질이 낮은 이미지는 dataset에서 제외었다. 본 연구에서 사용된 모든 이미지는 500×500픽셀 이상 크기와 72dpi 이상 해상도를 포함하였다. 종을 구분하는데 사용되는 형태학적 특징이 몸 전체에 걸쳐 존재하기 때문에 몸 전체를 객체 영역으로 지정하였다. 객체 영역지정은 Labellmg를 사용하여 ground truth bounding boxes로 라벨링되었다(Tzutalin, 2015). 라벨링된 dataset는 무작위로 training set 4,743장(64%), validation set 1,181장(16%), test set 1,493장(20%)으로 분리되었다 (표 1).

표1. 생태계교란 거북 6종 및 국내 자생 거북 2종 dataset 구축 결과

No.	국명	학명	Training	Validation	Test
1	늑대거북	<i>Chelydra serpentina</i>	1,543	385	483
2	악어거북	<i>Macrochelys temminckii</i>	184	46	59
3	중국줄무늬목거북	<i>Mauremys sinensis</i>	474	118	149
4	리버쿠터	<i>Pseudemys concinna</i>	539	134	170
5	플로리다붉은배거북	<i>Pseudemys nelsoni</i>	643	160	203
6	붉은귀거북	<i>Trachemys scripta</i>	1,054	263	331
7	남생이	<i>Mauremys reevesii</i>	135	33	43
8	자라	<i>Pelodiscus maackii</i>	171	42	55
Total			4,743	1,181	1,493

2. 모델 학습

2.1 모델 구조

본 연구에서는 생태계교란 거북을 탐지 및 분류하기 위해 YOLO11 모델(Jocher and Qiu, 2024)이 적용되었다. YOLO11은 C3k2 블록을 C2PSA 모듈로 대체하고 구조를 전작과 유사하게 유지하면서 성능을 개선한 강화된 버전이다(Ghosh, 2024; Khanam and Hussain, 2024). C3k2 블록은 속도와 효율성을 높이기 위해 설계되어 전체적인 특징 집계 성능을 개선한다. 반면 C2PSA 모듈은 모델이 이미지 내 핵심 영역에 집중하도록 하여 작은 물체나 부분적으로 가려진 물체의 탐지 성능을 향상시켰다(Ghosh, 2024; Khanam and Hussain, 2024). 이러한 특징 통해 다양한 배경과 다양한 크기의 대상에 대한 모델의 적응성이 강화되어 정밀도와 연산 효율성 간의 최적의 균형을 달성한다. 그림 1은 YOLO11의 아키텍처를 보여준다. 이 모델은 네트워크 깊이와 너비에 따라 n , s , m , l , x 의 다섯 가지 변형으로 제공된다. 이 중 YOLO11x는 가장 높은 정확도를 보여주어 본 연구에서 사용되었다.

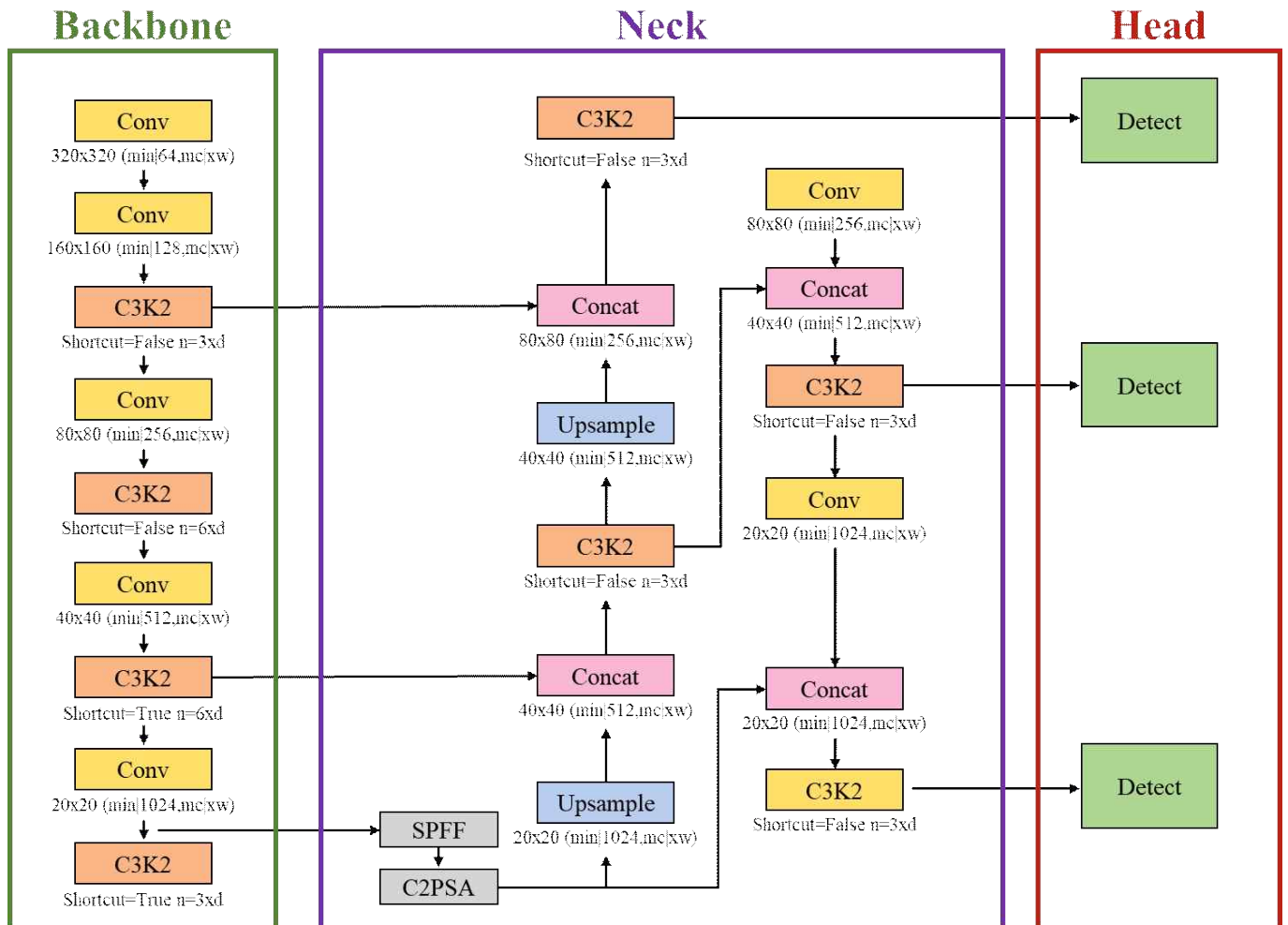


그림 1. YOLO11 모델의 구조도

2.2 모델 학습 설정값 및 학습 환경

학습 모델은 batch size 16, 입력 이미지 크기 640×640으로 최대 300 에포크 동안 훈련되었다. 에포크는 훈련 기간을 의미하며, batch size는 한 에포크를 구성하는 각 반복에서 사용되는 훈련 이미지 수를 나타낸다. 과적합을 방지하기 위해 조기 종료 기능과 데이터 증강이 적용되었다. 모델 훈련은 patience를 100으로 설정한 후 mAP가 100 에포크 동안 증가하지 않을 때 조기 종료되었다. 6개의 optimizer 중 랜덤 샘플링을 통해 local optima에 빠질 위험이 적고 빠른 수렴을 보이는 stochastic gradient descent (SGD) optimizer를 적용하여 학습하였다 (그림 2). 훈련 과정에서 두 가지 데이터 증강 방법인 Albumentation(Buslaev et al., 2020)과 Mosaic Augmentation(Austrheim et al., 2014)이 적용되었다. Albumentation 방법은 수평 반전, 번역, 확대/축소, HSV (Hue, Saturation, Value) 증강을 포함한다(Buslaev et al., 2020). 이 모델들의 실험 플랫폼은 Rocky Linux 8 운영 체제를 기반으로 하며, 두 개의 Intel Xeon Gold 6326 CPU, 8개의 64 GB REG.ECC DDR4 SDRAM, 및 24G 메모리를 갖춘 Nvidia RTX A5000 그래픽 카드를 사용한다. 실험 프로그램은 Python 3.11.3, CUDA 12.2, cuDNN 8.9.3, Pytorch 2.0.1을 기반으로 한다. 모든 실험 프로그램은 Anaconda 프롬프트를 사용한 가상 환경에서 실행되었다. 실험 프로그램에 필요한 패키지는 PyPI, SciPy, Pytorch에서 가져와 실험의 정상적인 작동과 실행이 되도록 설치하였다.

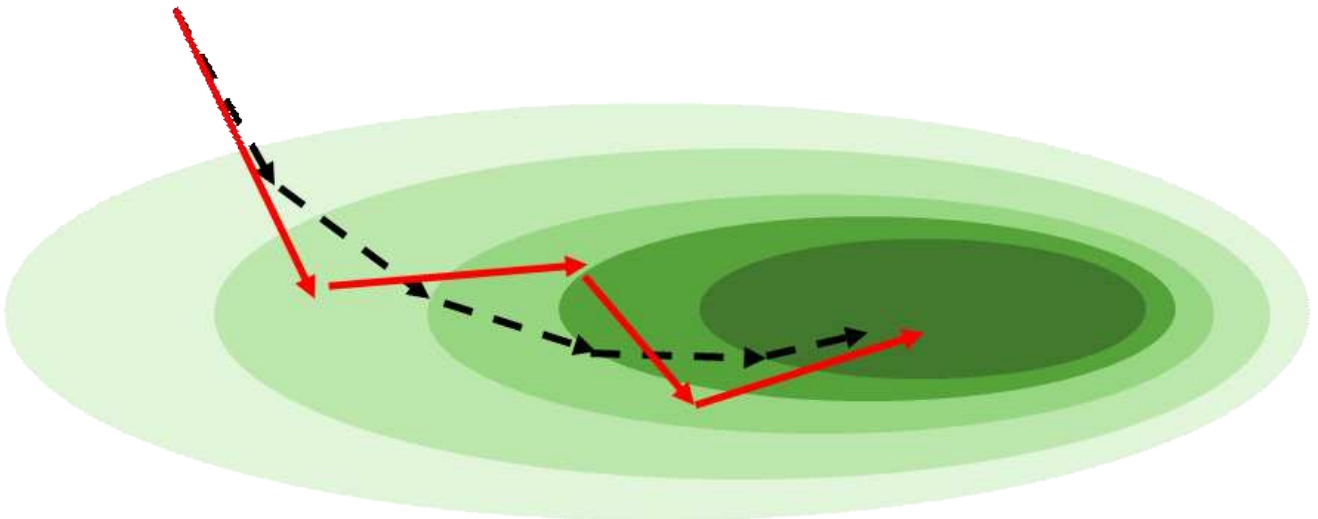


그림 2. Stochastic gradient descent(SGD) optimizer 원리. 빨간선: SGD optimizer 적용 학습 결과, 검정선: 일반적인 학습 결과

3. 모델 성능 평가

훈련 과정에서 모델의 성능은 세 가지 손실 함수를 사용하여 모니터링되었으며, 각 손실 함수는 validation set에서 평가되었다. 이 손실 함수는 예측된 결과와 실제 결과 간의 차이를 정량화하여 중요한 피드백을 제공하며, 이를 통해 모델은 매개변수를 반복적으로 조정하고 성능을 개선할 수 있다(Jocher and Qiu, 2024). Complete intersection over union(CIoU) loss ($LCIoU$)은 predicted bounding boxes와 ground truth bounding boxes 간의 불일치를 측정한다. 이는 IoU 지표를 기반으로 계산되며, bounding box의 중첩 면적, 상자 중심 간의 거리, 종횡비를 고려한다. 이 손실 함수는 위치 지정 부정확성과 기하학적 불일치 모두를 패널티로 적용하여 보다 정밀하고 안정적인 객체 위치 지정을 유도한다(Zheng et al., 2020). Classes loss ($Lcls$)은 탐지된 객체 분류와 관련된 평균 손실을 나타낸다. 이는 각 객체가 사전 정의된 클래스 중 하나에 정확히 할당되도록 보장한다. Distributed Focal loss ($Ldfl$)은 bounding box 회귀에 적용된다. bounding box 좌표를 직접 예측하는 기존 방법과 달리, $Ldfl$ 은 discrete box offset에 대한 확률 분포를 예측한다. 이 접근법은 모델이 위치 지정의 불확실성을 포착할 수 있게 하여 회귀 정확도를 향상시킨다. $Lcls$ 와 $Ldfl$ 모두 cross-entropy 공식에 기반하지만, 그 역할은 다르다. $Lcls$ 는 객체 범주에 대한 표준 다중 클래스 분류를 수행하는 반면 (Jocher and Qiu, 2024), $Ldfl$ 은 가중 cross-entropy 손실을 사용하여 discrete distribution 모델링을 통해 연속적인 bounding box 좌표를 회귀한다(Li et al., 2020). $LCIoU$ 의 계산은 방정식 (1)에 정의된 intersection over union(IoU) 지표에서 파생되었으며, 여기서 G 와 P 는 각각 ground-truth bounding box 및 predicted bounding box를 나타낸다. ν 는 두 박스 간의 종횡비 일관성 정도를 설명하며, 방정식 (2)에서 정의됩니다. 여기서 w^{gt} 와 h^{gt} 는 ground-truth bounding box의 너비와 높이를 나타내고, w 와 h 는 predicted bounding box의 너비와 높이를 각각 나타낸다. 매개변수 α 는 방정식 (3)을 사용하여 결정된 균형 조정 요소이다. $LCIoU$ 는 방정식 (4)를 사용하여 계산되었으며, 여기서 d 는 두 박스의 중심점 사이의 거리를, c 는 두 박스를 모두 포함하는 가장 작은 외곽 박스의 대각선 길이를 나타낸다.

$$IoU = \frac{G \cap P}{G \cup P} \quad (1)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (2)$$

$$\alpha = \frac{\nu}{(1 - IoU) + \nu} \quad (3)$$

$$LCIoU = 1 - IOU + \frac{d^2}{c^2} + \alpha \nu \quad (4)$$

훈련 과정을 완료한 후, test set을 사용하여 모델 성능을 precision, recall, 및 mean Average Precision (mAP)로 평가했다. precision는 방정식(5)로 계산되며 모델이 정확히 식별한 양성 예측의 비율을 측정한다. recall은 방정식(6)으로 계산되며 모델이 실제 양성 사례 중 정확히 탐지한 비율을 나타내는 한편, 방정식(7)과 같이 계산되는 Average precision(AP)는 방정식(8)에 따라 계산되며, n은 dataset의 총 객체 수를 의미한다. AP는 precision과 recall을 균형 있게 반영하는 지표로, precision-recall 곡선 아래 면적을 측정하여 객체 탐지 및 분류 평가를 모두 지원한다. mAP는 방정식(8)를 사용하여 계산되었다. 여기서 Q는 dataset의 총 쿼리 수를 나타내며, AP(q)는 각 쿼리 q에 대한 AP 점수를 의미한다. 본 연구에서는 두 가지 버전의 mAP를 평가했다: mAP0.5와 mAP0.5-0.95. 전자는 고정된 IoU 임계값 0.5를 사용하며, 후자는 0.5에서 0.95까지의 여러 IoU 임계값에 대한 평균을 계산한다. IoU는 ground truth bounding boxes와 모델이 예측한 predicted bounding boxes 간의 중첩 비율을 비교하여 true positive를 결정하는 데 사용되었다. 또한 모델의 분류 성능을 시각화하기 위해 confusion matrix로 나타내었다.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (5)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (6)$$

$$\text{Average Precision (AP)} = \sum_{x=0}^{x=n-1} \{\text{Recall}(x) - \text{Recall}(x+1)\} \times \text{Precision}(x) \quad (7)$$

$$\text{mean Average Precision (mAP)} = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (8)$$

Ⅲ. 연구 결과

1. 모델 학습 과정 중 성능 평가

그림 3는 훈련 과정 동안의 train loss에 대한 값을 보여준다. $LCIoU$, $Lcls$, $Ldff$ 의 값 모두 지속적으로 감소하고 안정적으로 수렴하였다. $LCIoU$ 값은 3.15955에서 시작하여 0.24891의 값으로 수렴하였다. $Lcls$ 값은 4.78063에서 시작하여 0.1491의 값으로 수렴하였다. $Ldff$ 값은 4.20184에서 시작하여 0.84615의 값으로 수렴하였다. 300번의 epoch 중 295번째 epoch 값이 가장 최적의 성능을 보였다. 최적의 epoch에서 train loss값은 $LCIoU$ 에서 0.25938, $Lcls$ 에서 0.15562, $Ldff$ 에서 0.84876의 값을 보였다.

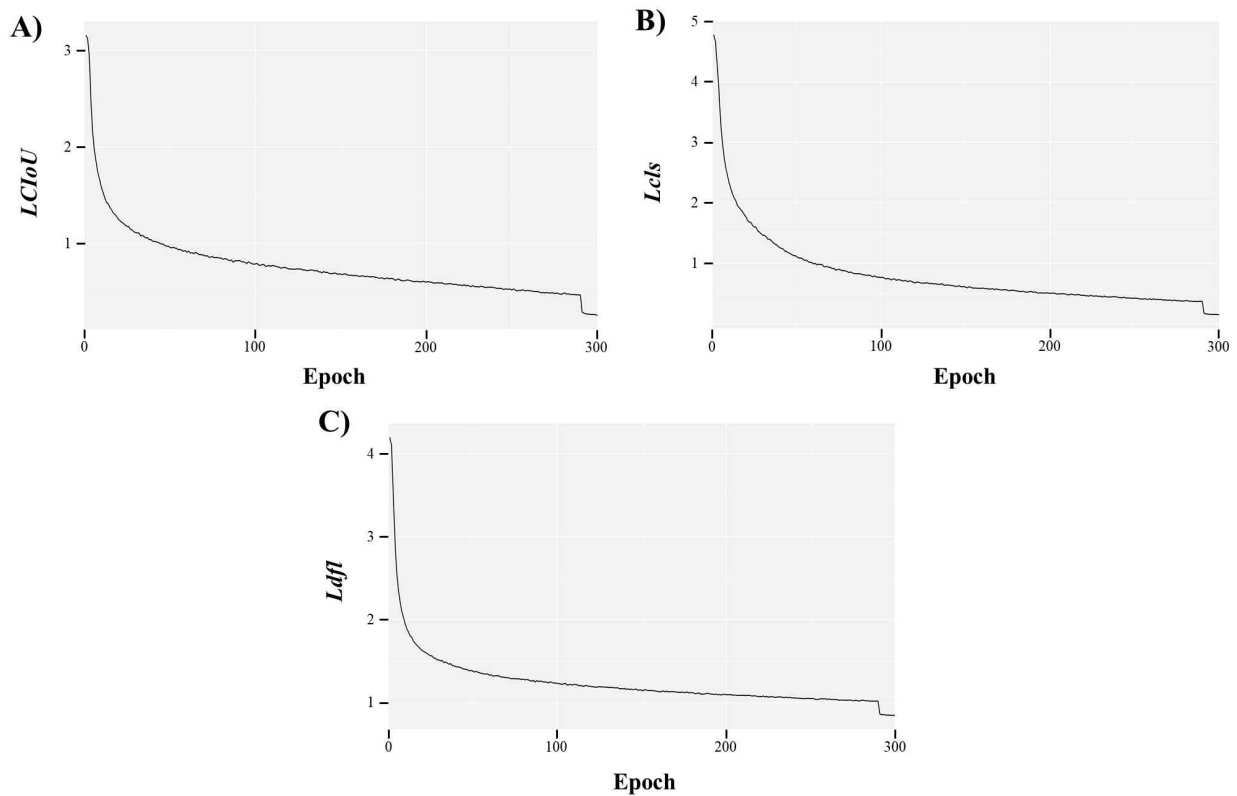


그림 3. 모델 학습 기간 동안 평가된 train loss 분석. A) Complete intersection over union loss ($LCIoU$), B) Cls loss ($Lcls$), C) Distributed Focal loss ($Ldff$).

그림 4은 훈련 과정 동안의 validation loss에 대한 값을 보여준다. $LCIoU$, $Lcls$, Ldf 3개의 값 모두 지속적으로 감소하고 안정적으로 수렴하였다. $LCIoU$ 값은 2.92817에서 시작하여 0.54488의 값으로 수렴하였다. $Lcls$ 값은 6.07941에서 시작하여 0.31429의 값으로 수렴하였다. Ldf 값은 4.19394에서 시작하여 1.04375의 값으로 수렴하였다. 최적의 epoch에서 validation loss값은 $LCIoU$ 에서 0.5471, $Lcls$ 에서 0.31597, Ldf 에서 1.06277의 값을 보였다.

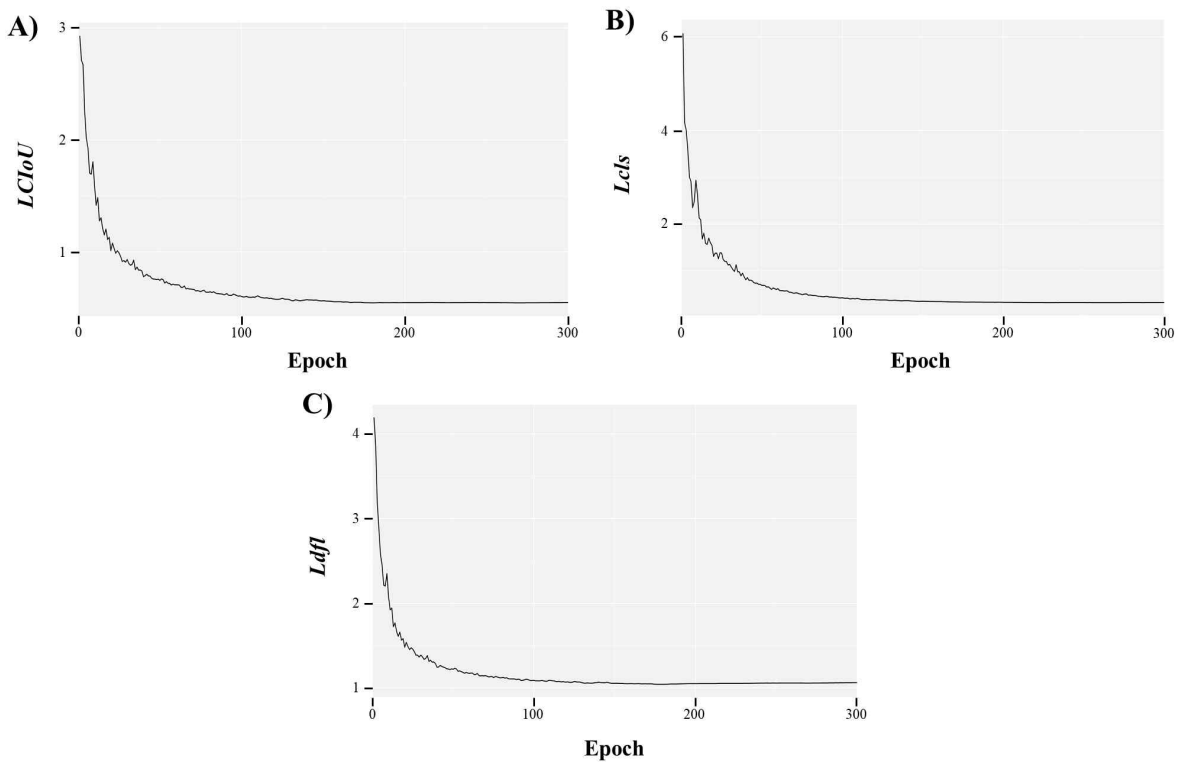


그림 4. 모델 학습 기간 동안 평가된 validation loss 분석. A) Complete intersection over union loss ($LCIoU$), B) Cls loss ($Lcls$), C) Distributed Focal loss (Ldf).

2. 모델 학습 후 성능 평가

개발된 모델의 inference time은 9.1ms로 측정되었다. 훈련 완료 후 모델 성능은 표 2에 제시되어 있다. 두 개의 mAP 값 ($mAP_{0.5}$, $mAP_{0.5-0.95}$)은 각각 0.965와 0.851이었다. Precision은 붉은귀거북과 남생이의 0.922부터 자라의 0.982까지의 범위를 보였다. Recall은 남생이의 0.833부터 늑대거북의 0.985까지의 범위를 보였다. precision과 recall을 기반으로 계산된 precision-recall curve는 그림 5를 통해 확인할 수 있다. $mAP_{0.5}$ 에서는 늑대거북이 0.992로 가장 높은 값을 보였고 붉은귀거북이 0.943으로 가장 낮은 값을 보였다. $mAP_{0.5-0.95}$ 는 늑대거북이 0.908로 가장 높은 값을 보였고 남생이가 0.818으로 가장 낮은 값을 보였다. 생태계교란 거북 6종의 경우 $mAP_{0.5}$ 에서 0.967을 보였고 $mAP_{0.5-0.95}$ 에서 0.967의 값을 보였다. 국내 자생 거북 2종의 경우 $mAP_{0.5}$ 에서 0.961을 보였고 $mAP_{0.5-0.95}$ 에서 0.846의 값을 보였다.

표 2. 개발된 모델의 성능 평가 결과

No.	국명	Precision	Recall	mAP0.5	mAP0.5-0.95
1	늑대거북	0.969	0.985	0.992	0.908
2	악어거북	0.954	0.899	0.949	0.832
3	중국줄무늬목거북	0.955	0.941	0.969	0.828
4	리버쿠터	0.951	0.918	0.967	0.880
5	플로리다붉은배거북	0.977	0.952	0.979	0.844
6	붉은귀거북	0.922	0.901	0.943	0.821
7	남생이	0.922	0.833	0.954	0.818
8	자라	0.982	0.909	0.968	0.874
Total		0.954	0.917	0.965	0.851

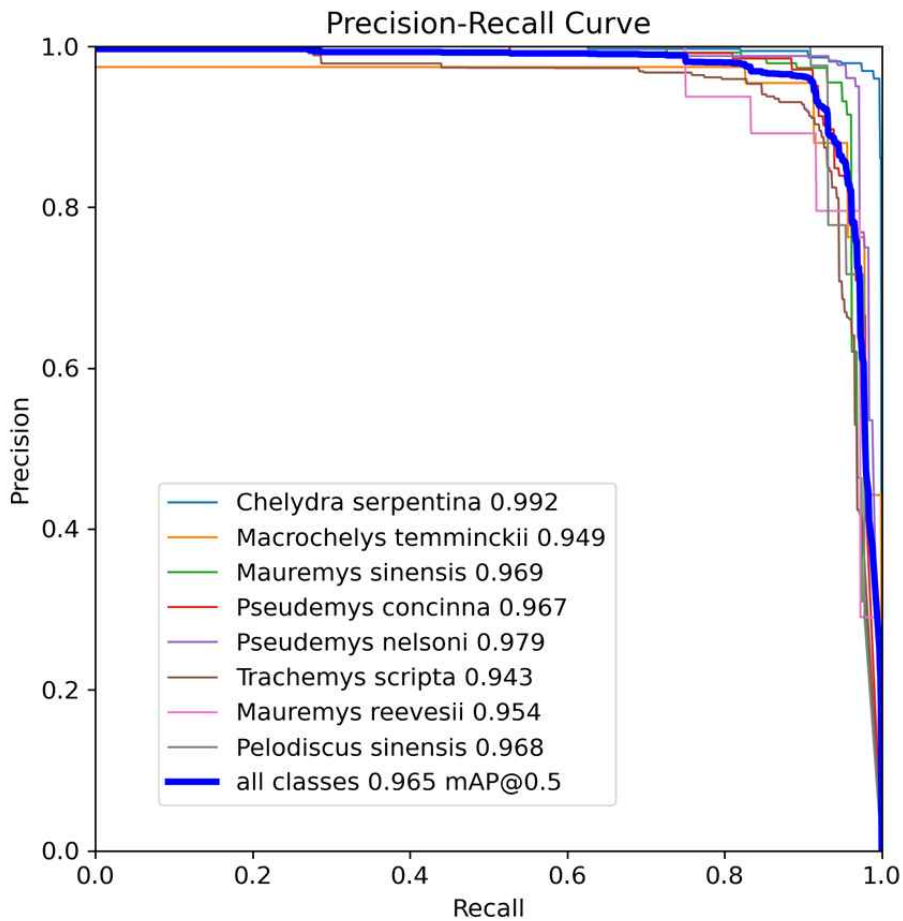


그림 5. 개발된 모델의 precision-recall curve

3. 모델 학습 후 분류 결과

개발된 모델의 분류 결과는 confusion matrix를 통해 나타낸다 (그림 6). 평균 분류 정확도는 92.5%였다. 가장 낮은 분류 정확도는 남생이로 86.1%였고 가장 높은 분류 정확도는 99.7%로 늑대거북이었다. 남생이는 늑대거북으로 5.6%만큼 오분류되었으며 중국줄무늬목거북, 붉은귀거북, 배경으로 각각 2.8%만큼 오분류되었다. 가장 오분류 비율이 높은 결과는 악어거북은 늑대거북으로 가장 높은 비율인 13%이었다. 다음으로 리버쿠터가 붉은귀거북으로 오분류 된 비율이 7.4%였다. 생태계교란 거북 6종의 경우 평균 93.4%의 평균 분류 정확도를 보였으며 국내 자생 거북 2종의 경우 89.7%의 평균 분류 정확도를 보였다. 생태계교란 거북에서 국내 자생 거북으로 오분류된 경우는 없었으나 국내 자생 거북은 남생이가 생태계교란 거북으로 11.2%만큼 오분류 되었으며 자라는 생태계교란 거북으로 4.6%만큼 오분류 되었다.

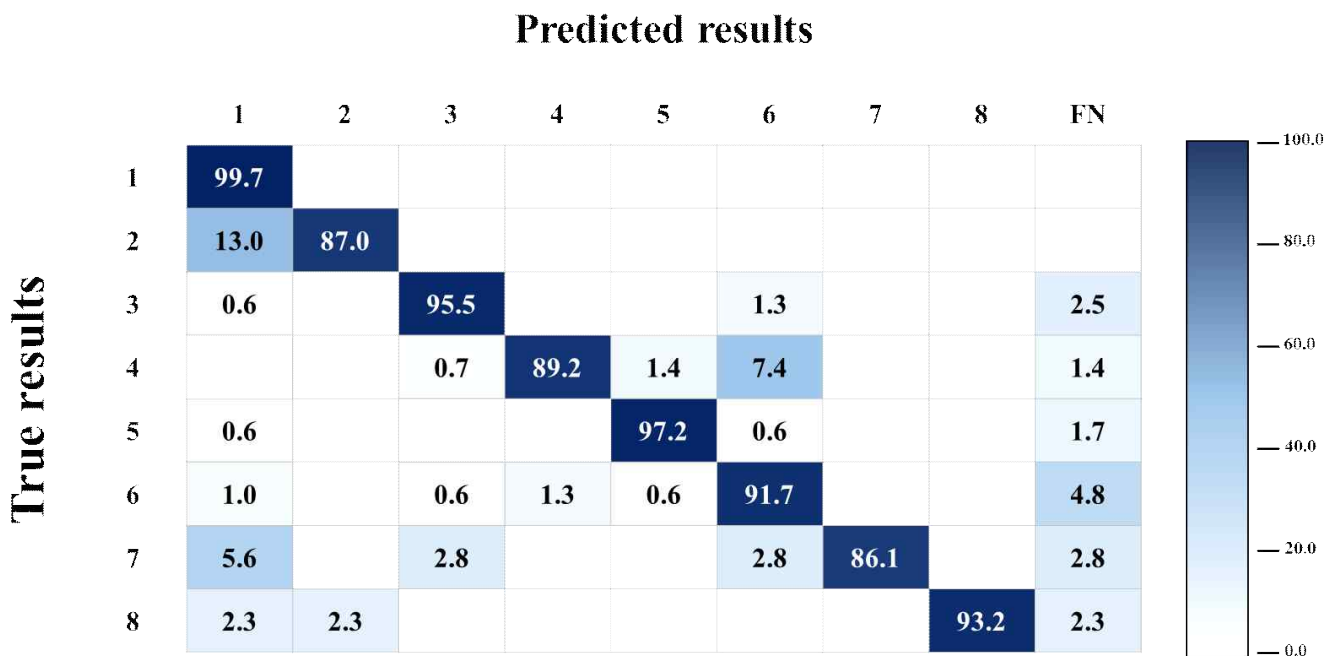
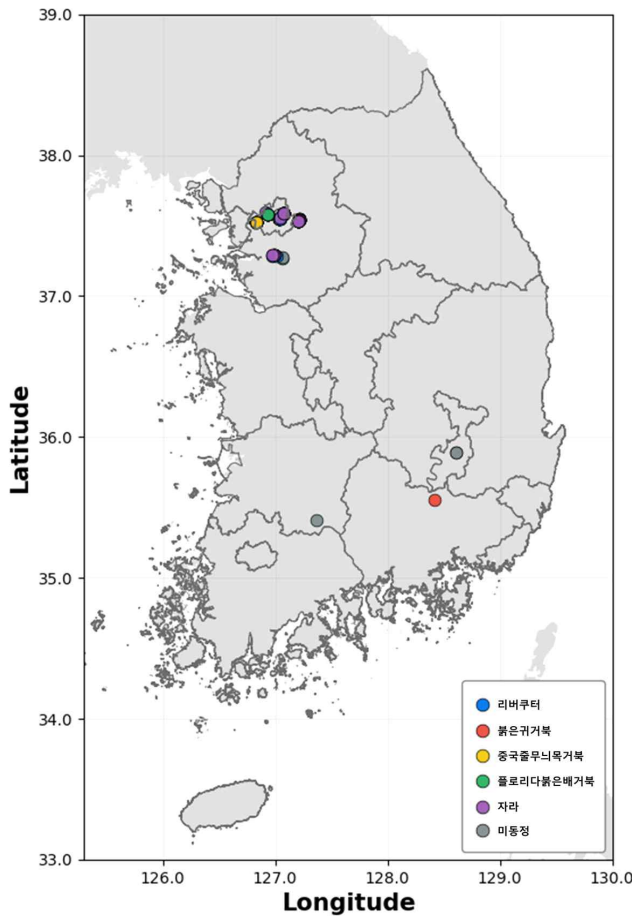


그림 6. 모델의 분류 결과를 나타낸 confusion matrix. 1: 늑대거북, 2: 악어거북, 3: 중국줄무늬목거북, 4: 리버쿠터, 5: 플로리다붉은배거북, 6: 붉은귀거북, 7: 남생이, 8: 자라. FN은 거북류가 배경으로 잘못 분류한 경우의 확률을 나타낸다. 행은 실제 종을 나타내며, 열은 모델이 예측한 종을 나타낸다. 예측 결과는 백분율 값으로 표시된다. 대각선 값은 정확한 예측을 나타내며, 대각선 외의 값은 잘못된 예측을 나타낸다.

4. 현장 적용성 평가 결과

본 연구에서 개발한 이미지분류모델의 현장 적용성을 평가하기 위해 시민과학 데이터의 분류에 모델을 적용하였다. 시민과학 데이터는 동아사이언스 지구사랑탐사대의 리버헬(시민과학풀씨 4기)팀에서 수집한 이미지 649장으로 구성되었다 (그림 7). 시민과학 이미지 중 거북이 존재하는 이미지는 325장이었으며 나머지 324장은 배경을 촬영한 이미지로 확인되었다. 325장의 거북 이미지를 전문가가 분류한 결과, 분류 형질이 나타나있어 분류가 가능한 것은 103장이었다. 나머지 222장의 이미지는 촬영거리가 멀어 낮은 해상도 및 화질로 인해 정확한 종 분류가 어렵거나, 촬영 각도, 빛 번짐 또는 물체에 의해 가려져 분류 형질을 확인하기 어려웠다. 분류가 가능한 103장의 이미지에는 리버쿠터, 붉은귀거북, 중국줄무늬목거북, 플로리다붉은배거북, 자라가 포함되었다. 시민과학 이미지에 대해 모델의 분류 시, detection 하이퍼파라미터는 모델 train 및 test 조건과 동일하게 설정하였다.



No.	국명	이미지 수
1	리버쿠터	26
2	붉은귀거북	7
3	중국줄무늬목거북	3
4	플로리다붉은배거북	9
5	자라	58
6	동정 불가	222
거북 이미지 수		325
배경 이미지 수		324
이미지 수 합계		649

그림 7. 시민과학자에 의해 수집된 이미지의 종 별 수집 위치 및 구성.

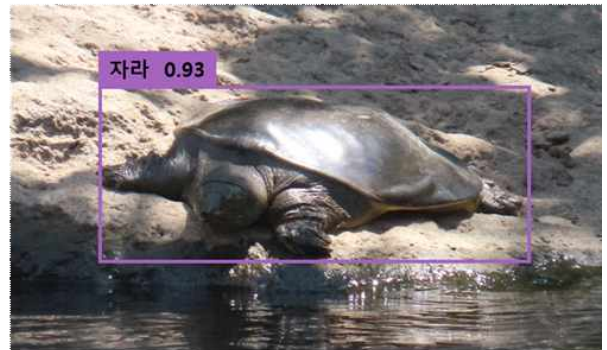
(좌) 종 별 위치 분포; (우) 종 별 이미지 수

시민과학자 이미지를 모델이 분류한 결과, 모델은 거북이 존재하는 325장의 이미지 중 241장(74.1%)에서 거북을 탐지 및 분류하였다. 모델이 거북으로 탐지한 241장의 이미지에는 분류 형질이 드러나는 103장의 이미지가 모두 포함되었다 (그림 8). 모델은 분류 형질이 드러나는 이미지 103장 중 75장을 올바르게 분류하였으며, 이미지에 포함된 거북 5종에 대해 73.1%의 평균 분류정확도를 보였다. 종 별 분류정확도는 중국줄무늬목거북이 100%로 가장 높았으며, 리버쿠터는 34.6%로 가장 낮은 정확도를 보였다 (그림 9). 반면 시민과학자는 103장의 거북 이미지 중 73장을 올바르게 분류하였으며, 48.2%의 평균 분류 정확도를 보였다. 시민과학자의 종 별 분류 정확도는 자라가 94.8%로 가장 높았으며, 중국줄무늬목거북은 모두 분류하지 못하여 0%의 정확도를 보였다.

A)



B)



C)



D)



그림 8. 본 연구에서 개발한 모델의 시민과학자 이미지 분류 결과 예시. A) 붉은귀거북, B) 자라, C) 리버쿠터, D) 중국줄무늬목거북. 예측된 증명 우측의 숫자는 해당 종으로 분류한 모델의 신뢰도(confidence score)를 나타낸다. Photo credit: A) 과학초보팀, B) 넓적이를 찾아서팀, C) 임형제가족팀, D) 임형제가족팀

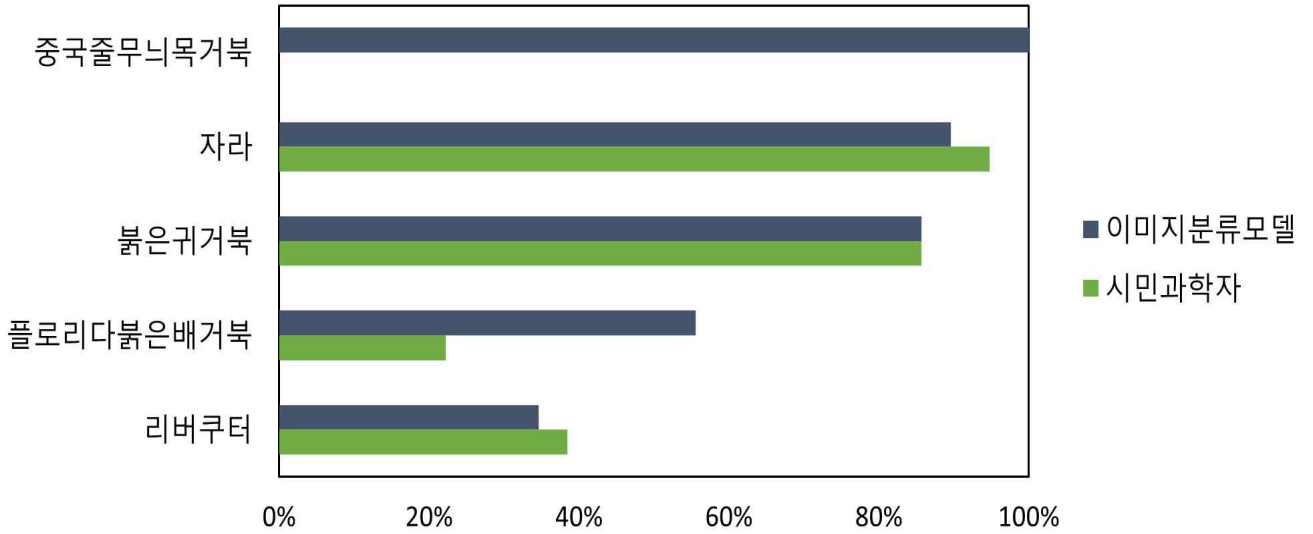


그림 9. 시민과학자와 이미지분류 모델의 종 별 분류 정확도 그래프. 청색은 본 연구에서 개발한 이미지분류 모델의 분류 정확도를 나타내며, 녹색은 시민과학자의 분류 정확도를 나타낸다.

IV. 연구 논의

모델 학습 과정에서의 성능 평가 결과, train loss와 validation loss 모두에서 세 가지 loss값($LClou$, $Lcls$, Ldf)이 안정적으로 감소하며 수렴하는 양상을 보였다(그림 3 및 4). 또한 손실 값이 더 이상 감소하지 않은 이후에도 상승하는 경향은 관찰되지 않았으며, 이는 본 연구에서 학습된 모델에서 과적합이 발생하지 않았음을 시사한다.

모델 학습 후 성능 평가 결과, 탐지와 분류 성능을 종합적으로 평가하는 지표인 mAP0.5 및 mAP0.5-0.95에서 늑대거북이 각각 0.992와 0.908로 가장 높은 성능을 나타냈다. 일반적으로 mAP 값이 높을수록 학습 완료 후 모델의 탐지 및 분류 성능이 우수함을 의미한다. 이러한 결과는 본 연구에서 대상으로 설정한 8종 중 늑대거북이 가장 많은 학습 데이터를 확보한 종이었기 때문일 가능성이 있다(표 2). 충분한 양의 학습 데이터는 딥러닝 모델의 성능 향상에 있어 가장 기본적이면서도 핵심적인 요소로 알려져 있으며(Sun et al., 2017), 이에 따라 늑대거북을 탐지하고 분류하기 위한 형태적 특징이 모델에 충분히 학습되었을 가능성이 있다. 반면, 가장 낮은 성능을 보인 종은 mAP0.5 기준에서는 붉은귀거북, mAP0.5-0.95 기준에서는 남생이로 나타났다. mAP0.5는 상대적으로 탐지 난이도가 낮은 조건에서의 모델 성능을 반영하는 지표인 반면, mAP0.5-0.95는 다양한 IoU 기준을 종합하여 여러 탐지 난이도 수준에서의 성능을 평가한다 (Jocher et al., 2023). 붉은귀거북의 경우, 붉은 귀덮개 유무 확인을 통한 상대적으로 쉬운 형태분류 특징을 지니고 있기 때문에 수면 아래 위치하거나 이미지 내에서 객체 크기가 작은 등 탐지 난이도가 높은 이미지가 다수 포함되어 있었다. 이러한 데이터 특성으로 인해, 상대적으로 쉬운 탐지 조건보다는 다양한 난이도의 탐지 조건을 포함하는 평가 지표에서 더 안정적인 성능을 보였을 가능성이 있다. 남생이의 경우, 본 연구 대상종 중 가장 적은 양의 데이터로 학습이 수행되었으며 (표 2), 이로 인해 다양한 객체 크기와 촬영 거리 조건을 학습에 충분히 반영하지 못했을 가능성이 있다. 특히 남생이 이미지 중 다수는 이미지 내에서 거북 객체가 차지하는 비율이 큰 경우가 많아, 상대적으로 작은 객체에 대한 탐지 성능을 학습하는 데에는 한계가 있었을 수 있다. 이러한 데이터 특성은 모델이 소형 객체를 탐지하는 데 불리하게 작용했을 가능성이 있다. 또한 두 가지 mAP 지표를 기준으로 생태계교란 거북 6종과 국내 자생 거북 2종의 성능을 비교한 결과, 전반적으로 생태계교란 거북 종에서 더 높은 탐지 성능이 나타났다. 이는 종별 데이터 수집량의 차이로 인한 데이터 불균형에 기인한 결과일 가능성이 크다 (Chan et al., 2023). 동일한 수집 절차를 적용하였음에도 불구하고, 국내 자생 거북 2종은 생태계교란 거북 종에 비해 종별 데이터의 절대적인 양이 부족하였다. 이러한 결과는 해외 데이터에 비해 국내 자생 거북 종에 대한 공개 이미지 데이터가 상대적으로 제한적임을 시사하며, 향후 국내 자생 생물종을 대상으로 한 체계적인 데이터 수집 및 dataset 구축의 필요성을 강조한다.

Confusion matrix를 기반으로 종별 분류 결과를 분석한 결과, 늑대거북이 가장 높은 분류 정확도를 보인 반면, 남생이가 가장 낮은 분류 정확도를 나타냈다. 이러한 결과는 종별 학습 데이터의 양적 차이에 기인했을 가능성이 있다. 늑대거북은 본 연구에서 대상으로 설정한 8종 중 가장 많은 수의 dataset로 학습이 수행된 반면, 남생이는 가장 적은 양의 데이터로 학습되었다 (표 2). 특히 남생이의 경우, 오분류된 대상 종이 늑대거북, 중국줄무늬목거북, 붉은귀거북 등 특정 한 종에 국한되지 않고 여러 종으로 분산되어 나타났다 (그림 6). 이는 남생이를 분류하기 위한 형태적 특징이 모델에 충분히 학습되지 못했음을 시사하며, 종 특이적인 형태 정보를 효과적으로 반영하기에는 데이터가 부족했을 가능성을 보여준다. 따라서 향후 남생이의 분류 정확도를 향상시키기 위해서는 남생이의 형태적 특징이 명확하게 드러나는 이미지를 추가적으로 확보하는 것이 필요하다. 오분류가 가장 빈번하게 발생한 사례는 악어거북이 늑대거북으로 분류된 경우였다. 늑대거북과 악어거북은 양측 costal scutes와 marginal scutes 사이에 scute의 존재 여부로 구분될 수 있으나, carapace에 세 개의 keels가 존재하는 등 전반적으로 유사한 형태학적 특성을 공유하고 있어 형태 기반 분류가 어려운 종으로 알려져 있다 (Baek et al., 2024a; Kim et al., 2021). 이러한 형태학적 유사성은 딥러닝 기반 분류 모델에서 종 오분류를 유발하는 주요 요인으로 선행 연구들에서도 보고된 바 있다 (Baek et al., 2024; Kim et al., 2025; Spiesman et al., 2021). 더불어 악어거북은 멸종위기종으로, 본 연구에서 포함된 생태계교란 거북 중에서 상대적으로 가장 적은 양의 데이터로 학습이 이루어졌다 (표 2). 이와 같은 데이터 불균형은 형태적 차이를 충분히 학습하는 데 한계로 작용했을 가능성이 있다 (Chan et al., 2023). 따라서 향후 악어거북의 분류 정확도를 개선하기 위해서는 종 특이적인 형태 특징이 명확히 표현된 고품질 이미지를 추가적으로 확보하는 것이 필요할 것으로 판단된다. 두 번째로 높은 오분류 사례는 리버쿠터가 붉은귀거북으로 분류된 경우였다. 붉은귀거북은 일반적으로 눈 뒤와 귀덮개 부근에 나타나는 붉은색 또는 주황색 띠의 유무를 통해 구분될 수 있다. 그러나 붉은귀거북 dataset에 일부 포함된 아종인 노란배거북의 경우, 눈 뒤와 귀덮개 부근에 띠가 존재하지 않으며 대신 눈 뒤쪽에 노란색 얼룩이 나타나는 것으로 보고되어 있다 (Baek et al., 2024a; Kim et al., 2021). 이러한 형태적 특징은 눈 사이에 다섯 개의 밝은 노란색 줄무늬가 존재하는 리버쿠터와 유사하게 인식될 수 있다 (Baek et al., 2024a; Kim et al., 2021). 이러한 결과는 본 연구에서 적용된 모델이 두 종 간 눈 뒤에 나타나는 노란색 무늬를 구분할 만큼의 형태적 특징을 충분히 학습하지 못했음을 시사한다. 따라서 향후 리버쿠터의 분류 성능을 향상시키기 위해서는 형태 특징이 명확히 드러나는 이미지를 추가로 확보하는 한편, 하이퍼파라미터 최적화와 같은 학습 전략을 적용하여 dataset 특성에 보다 적합한 학습 설정을 도입할 필요가 있다.

시민과학자 이미지의 분류에 모델을 적용한 결과, 모델의 평균 분류 정확도는 test set에 비해 상대적으로 낮게 나타났으며, 이는 현장 데이터의 특징으로부터 기인한 것으로 사료된다. 모델의 train 및 test에는 사진 내 거북의 사이즈 및 해상도를 고려하여 선별한 고품질의 이미지 dataset가 이용된 반면, 시민과학자의 이미지는 촬영 거리가 멀어 해상도가 낮거나, 거북의 일부가 물체에 의해 가려지거나, 수중에서 촬영되는 등 분류형질을 확인하기 어렵게 하는 다양한 요인이 존재하였다. 해당 요인들은 모델의 탐지와 분류 성능을 낮추는 요인으로 작용할 수 있다고 알려져 있다 (Baek et al., 2023; Pegorara and Pflugfelder, 2020; Redmon et al., 2016). 따라서 향후 현장 적용 가능성을 향상시키기 위해선 시민과학자의 현장 데이터를 기반으로 본 연구의 모델을 개선할 필요가 있다. 이를 위해 모델의 학습 전 또는 과정 중에서 객체 크기 조절, 흐림 효과, 부분 가림, 수중 색 편향 등 현장 상황을 고려한 다양한 전처리 기법 적용을 고려하거나 하이퍼파라미터 최적화를 통해 성능을 향상시킬 수 있다 (Dahiya et al., 2023; Maharana et al., 2022; Tan et al., 2023). 그럼에도 불구하고, 개발한 모델은 시민과학자보다 높은 평균 분류정확도를 보였을 뿐 아니라, 관측 빈도가 낮아 시민과학자에게 덜 알려졌을 가능성이 있는 종에 대한 보조 도구로서의 잠재력을 보여주었다. 예를 들어 모델은 중국줄무늬목거북의 이미지를 모두 올바르게 분류한 반면, 시민과학자의 경우 해당 이미지를 분류하지 못하였다. 중국줄무늬목거북은 2020년 생태계교란 생물로 지정되었으며 (MCEE, 2025), 다른 거북류에 비해 낮은 관측빈도를 보이므로 (NIE, 2023), 시민과학자에게 상대적으로 덜 알려져 제한된 정보로 인해 분류가 어려울 수 있다. 이러한 결과는 본 연구에서 개발한 모델이 현장에서 관측빈도가 낮은 생태계교란생물의 분류 및 조기 탐지를 지원하는 방식으로 활용될 수 있음을 시사한다. 다만 국내 출현 빈도가 낮은 종은 검증 이미지 수가 제한적이므로 향후 추가적인 현장 데이터를 확보하여 일관성을 재확인할 필요가 있다.

본 연구에서 개발한 모델은 분류 형질이 드러나는 이미지에서는 거북을 안정적으로 탐지하였으며, 분류 형질이 비교적 명확한 경우에는 높은 분류 성능을 보여주었다. 따라서 현장 데이터에서도 일정 수준의 이미지 품질이 확보되면 모델의 활용 가능성이 높아질 수 있다. 현장 데이터의 품질 변동성이 큰 시민과학의 특성을 고려하여, 시민과학자 대상 촬영 가이드를 제공하는 등 데이터 수집 단계에서부터 품질을 향상한다면 데이터의 활용 가능성과 분류 성능을 향상시킬 수 있을 것으로 기대된다.

V. 연구 결론

본 연구에서는 생태계교란 거북 6종과 국내 자생 거북 2종을 대상으로, YOLO11 모델을 기반으로 한 딥러닝 종 분류 모델을 개발하였다. 개발된 모델은 mAP0.5 기준 0.965, 분류 정확도 92.5%의 성능을 달성하여, 거북 종 탐지 및 분류에 있어 우수한 성능을 보였다. 특히 생태계교란 거북이 국내 자생 거북으로 오분류된 사례는 관찰되지 않았으며, 이는 관리 측면에서 중요한 의미를 갖는다. 또한 개발한 모델을 실제 시민과학 데이터에 적용하여 현장 활용 가능성을 평가한 결과, 현장 데이터에서도 73.1%를 올바르게 분류하였다. 아울러 출현빈도가 낮은 종에 대해서도 올바르게 분류할 수 있음을 확인하였다. 이러한 결과는 딥러닝 기반 이미지 분류 모델이 시민과학자 및 형태 분류 전문가가 신속한 종 식별을 필요로 하는 상황에서 효과적인 보조 도구로 활용될 수 있음을 시사한다. 한편, 본 연구에서는 종별 데이터 수의 불균형 문제가 존재하였으며, 특히 멸종위기종의 경우 이러한 문제가 더욱 두드러질 수 있다. 따라서 향후 연구에서는 generative adversarial network 기반 데이터 증강, transfer learning과 같은 기법을 적용하여 데이터 불균형으로 인한 영향을 완화하고 모델의 일반화 성능을 향상시킬 필요가 있다. 또한 실제 현장 환경에서는 객체 크기 조절, 흐림 효과, 부분 가림, 수중 색 편향 등으로 인해 데이터 품질 저하가 발생하는 문제가 확인되었다. 이에 따라 다양한 전처리 기법의 적용이나 하이퍼파라미터 최적화와 함께, 시민과학자를 대상으로 한 촬영 가이드를 제공하는 등 데이터 수집 단계에서부터 품질을 체계적인 개선이 필요하다. 이러한 체계적인 개선을 통해 고도화된 모델은 시민과학자의 현장 조사와 전문가의 대규모 데이터 분석을 동시에 지원함으로써, 생태계교란 거북의 조기 탐지 및 신속한 관리 대응(EDRR)에 기여할 수 있을 것으로 기대된다.

VI. 인용문헌

Austrheim, G., Speed, J. D., Martinsen, V., Mulder, J., Mysterud, A. (2014). Experimental effects of herbivore density on aboveground plant biomass in an alpine grassland ecosystem. *Arctic, Antarctic, and Alpine Research*, 46(3), 535-541.

Baek, H.J., Min, J.S., Lee, S.I., Yoo, H.R., (2024a). Non-native turtles in South Korea. *Korea*, 20–113.

Baek, H. J., Cheong, E., Kim, Y., Koo, K. S., Kim, S. H., Park, C. D., Yoon, J. D. (2024b). Geographical distribution of *Mauremys sinensis*, *Mauremys reevesii*, and their hybrids in South Korea. *Animals*, 14(18), 2626.

Baek, J.W., Kim, J.I., Kim, C.B. (2024). Deep learning-based image classification of sea turtles using object detection and instance segmentation models. *PLoS One* 19(11), e0313323. <https://doi.org/10.1371/journal.pone.0313323>.

Brussaard, L., De Ruiter, P. C., Brown, G. G. (2007). Soil biodiversity for agricultural sustainability. *Agriculture, ecosystems & environment*, 121(3), 233-244.

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A. A. (2020). Albumentations: fast and flexible image augmentations. *Information*, 11(2), 125.

Chan, W. H., Fung, B. S., Tsang, D. H., Lo, I. M. (2023). A freshwater algae classification system based on machine learning with StyleGAN2-ADA augmentation for limited and imbalanced datasets. *Water Research*, 243, 120409.

Christakakis, P., Papadopoulou, G., Mikos, G., Kalogiannidis, N., Ioannidis, D., Tzovaras, D., Pechlivani, E. M. (2024). Smartphone-Based Citizen Science Tool for Plant Disease and Insect Pest Detection Using Artificial Intelligence. *Technologies*, 12(7), 101.

Dahiya, N., Sharma, Y. K., Rani, U., Hussain, S., Nabilal, K. V., Mohan, A., Nuristani, N. (2023). Hyper-parameter tuned deep learning approach for effective human monkeypox disease detection. *Scientific Reports*, 13(1), 15930.

Farkas, B., Ziegler, T., Pham, C. T., Ong, A. V., Fritz, U. (2019). A new species of *Pelodiscus* from northeastern Indochina (Testudines, Trionychidae). *ZooKeys*, (824), 71.

Gariboldi, A. Zuffi, M. A. (1994). Notes on the population reinforcement project for *Emys orbicularis* (Linnaeus, 1758) in a natural park of northwestern Italy. *Herpetozoa* 7, 83–89.

Ghosh, A. (2024). YOLO11: Redefining Real-Time Object Detection [online]. <https://learnopencv.com/yolo11/>. Accessed 15 March 2025.

Havel, J. E., Kovalenko, K. E., Thomaz, S. M., Amalfitano, S., Kats, L. B. (2015). Aquatic

invasive species: challenges for the future. *Hydrobiologia*, 750, 147-170.

Huerta-Ramos, G., Luštrik, R. (2021) Inat_Images [dataset]. Zenodo.

Jocher, G., Derrenger, P., Yasin, M., Munawar, M., R., (2023). Assessed 21 December 2025. <https://docs.ultralytics.com/guides/yolo-performance-metrics/>

Jocher, G., Qiu, J., Ultralytics yolo11 (2024). Assessed 3 April 2025.

Khanam, R., Hussain, M., Yolov11: an overview of the key architectural enhancements (2024).

Kim, B. K., Sung, H. C. (2020). New record of the non-native snapping turtle *Chelydra serpentina* (Linnaeus, 1758) in the wild of the Republic of Korea. *BiolInvasions Records*, 9(2), 444-449.

Kim, J.I., Baek, J.W., Kim, C.B. (2025). Hierarchical image classification using transfer learning to improve deep learning model performance for amazon parrots. *Scientific reports*, 15(1), 3790. <https://doi.org/10.1038/s41598-025-88103-3>.

Kim, S.H., Lee, H.H., Kim, D.E., Park, J.S., Lee, H.J., Kim, N.H., Kim, Y.H., Kim, P.J., Park, J.E., Baek, H.J., Jang, H.J., Jeong, Y.J., Choi, D.H., Woo, S.H., Lee, S.I., Cho, S.Y., Jo, A.R., Kim, D.Y., Kim, H.S., Park, E.H., Ban, Y.G., Son, S.H., Lee, M.J., Jang, B.J., Chae, D.Y., Shin, J.W., Song, H.R., (2021). Information for the Field Management of Invasive Alien Species in Korea. *Korea*, 36–65.

Koo, K. S., Park, S. M., Kang, H. J., Park, H. R., Choi, J. H., Lee, J. S., Kim, B. K., Sung, H. C. (2020). New record of the non-native snapping turtle *Chelydra serpentina* (Linnaeus, 1758) in the wild of the Republic of Korea. *BiolInvasions Records*, 9(2), 444-449.

Koo, K. S., Park, S. M., Choi, J. H., Sung, H. C. (2021). New report of an alligator snapping turtle (*Macrochelys temminckii* Troost, 1835) introduced into the wild in the Republic of Korea. *BiolInvasions Record*, 10(1).

Li, X. Wild Animals Detection Based on YOLOv5. (2023). *Applied and Computational Engineering* 8(1):630-640.

Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J., Yang, J., (2020). Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection. *Advances in neural information processing systems*, 33, 21002–21012.

Luong, C. T., Farhan, A., Vasquez, R. D., Roldan, M. J. M., Lin, Y. K., Hsu, S. Y., Lin, M. D. Hsiao, C. D., Hung, C. H. (2023). SpiderID_APP: A User-Friendly APP for Spider Identification in Taiwan Using YOLO-Based Deep Learning Models. *Inventions*, 8(6), 153.

Maharana, K., Mondal, S., Nemade, B. (2022). A review: Data pre-processing and data

augmentation techniques. *Global Transitions Proceedings*, 3(1), 91-99.

Ministry of Climate, Energy and Environment. (2025). 생태계교란 생물 지정 고시. 환경부고시 제2025-165호.

Mooney, H., Larigauerie, A., Cesario, M., Elmquist, T., Guldberg, O. H., Lavorel, S., Mace, M. G., Palmer, M., Scholes, R., Yahara, T. (2009). Biodiversity, climate change, and ecosystem services. *Current opinion in environmental sustainability*, 1(1), 46-54.

National Invasive Species Council (NISC) 2016–2018 NISC Management Plan. NISC Secretariat (Washington, DC 2016).

National Institute of Ecology. (2023). Nationwide Survey of Non-native Species in Korea.

Pegoraro, J., Pflugfelder, R. (2020). The problem of fragmented occlusion in object detection. arXiv preprint arXiv:2004.13076.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roverts, C. M., Rexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *science*, 344(6187), 1246752.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Ricciardi, A. (2007). Are modern biological invasions an unprecedented form of global change?. *Conservation Biology*, 21, 329–336.

Smith, K. F., Behrens, M., Schloegel, L. M., Marano, N., Burgiel, S., Daszak, P. (2009). Reducing the risks of the wildlife trade. *Science*, 324(5927), 594-595.

Spiesman, B. J., Gratton, C., Hatfield, R. G., Hsu, W. H., Jepsen, S., McCornack, B., Patel, K., Wang, G. (2021). Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Scientific reports*, 11(1), 7580. <https://doi.org/10.1038/s41598-021-87210-1>.

Spinks, P. Q., Pauly, G. B., Crayon, J. J. Shaffer, H. B. (2003). Survival of the western pond turtle (*Emys marmorata*) in an urban California environment. *Conservation Biology*, 113, 257–267.

Strayer, D. L., Eviner, V. T., Jeschke, J. M. Pace, M. L. (2006). Understanding the long-term effects of species invasions. *Trends in ecology & evolution*, 21, 645–651.

Sun, C., Shrivastava, A., Singh, S., Gupta, A. (2017). Revisiting unreasonable effectiveness of data in deep learning era. In *Proceedings of the IEEE international conference on computer vision* (pp. 843-852).

Suzuki-Ohno, Y., Westfechtel, T., Yokoyama, J., Ohno, K., Nakashizuka, T., Kawata, M., Okatani, T. (2022). Deep learning increases the availability of organism photographs taken by citizens in citizen science programs. *Scientific Reports*, 12(1), 1210.

Tan, F. G., Yuksel, A. S., Aksoy, B. (2023, November). Deep Learning-Based Hyperparameter Tuning and Performance Comparison. In *The International Conference on Artificial Intelligence and Applied Mathematics in Engineering* (pp. 128-140). Cham: Springer Nature Switzerland.

Tzotalin, D., (2015). Labellmg [software]. GitHub. <https://github.com/tzotalin/labellmg>.

Zhang, C., Xu, Y., Hua, Y., Li, P., Storey, K. B., Zhang, W. (2023). Greater physiological resistance to heat may favour an invasive freshwater turtle, enabling it to outcompete native species in a changing climate. *Freshwater Biology*, 68(9), 1588-1601.

Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., Ren, D. (2020). Distance-IoU loss: faster and better learning for bounding box regression. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34(7), pp. 12993–13000.

사사

본 연구는 숲과나눔 장학생 지원을 받아 연구를 진행하였습니다. 현장 적용성 평가에 동아사이언스 지구사랑탐사대의 리버셀 [시민과학풀씨 4기] 팀의 위치 데이터와 이미지가 사용되었습니다. 이원재 팀장님, 임형제가족팀, 과학초보팀, 넓적이를 찾아서 팀의 기여에 감사드립니다.

본 연구보고서의 내용은 연구자의 의견이며,
(재)숲과나눔의 공식적인 견해와는 다를 수 있습니다.