




# Development of a YOLO model-based program for estimating the population of hooded cranes in Suncheon Bay\*

## 순천만 흑두루미 개체수를 추정하는 YOLO 모델 기반의 프로그램의 개발

성홍주<sup>†</sup> · 전영국<sup>††</sup>   
 Hongju Seong<sup>†</sup> · Youngcook Jun<sup>††</sup>

### 요약

본 연구는 멸종 위기에 처한 두루미 군집의 개체수를 자동으로 추정하는 단일 이미지 학습 기반 YOLO 객체 탐지 시스템을 제시한다. 기존 객체 탐지 모델은 대량의 레이블링된 데이터를 필요로 하기 때문에 멸종 위기종의 보존 및 연구에 적합하지 않다. 이러한 문제를 해결하기 위해 본 연구는 고해상도 두루미 군집 이미지 한 장에서 사용자가 지정한 소수의 객체 및 배경 좌표를 기반으로 학습 데이터를 구축하고 YOLOv8n 모델을 사용하여 시스템을 학습시켰다. 우리는 또한 객체 탐지 과정에서 슬라이딩 윈도우 방식을 사용하여 이미지 전체를 유클리드 거리 기반 중복 제거 필터링을 통해 탐지된 객체의 고유성을 포착하였다. 이 논문은 시각화 이미지 파일과 생태학자들이 검토 및 수정할 수 있는 데이터로 흑두루미 개체수 탐지 결과를 제공한다. 제안된 시스템은 단일 이미지 내에서 개체수를 빠르고 정확하게 추정할 수 있는 잠재력과 대규모 데이터 세트 없이도 생태 정보를 분석하는 실용적 도구로서의 가능성을 보여준다.

**주제어** YOLOv8, 순천만 사진, 두루미, 객체 탐지, 개체수 추정

### ABSTRACT

This study presents a single-image learning-based YOLO object detection system that automatically estimates the population of a colony of the endangered hooded cranes. Existing object detection models require large amounts of labeled data, making them unsuitable for the conservation and research of this endangered species. To address this issue, this study constructs training data based on a small number of user-specified object and background coordinates from a single high-resolution hooded crane colony image and trains the system using the YOLOv8n model. The detection process utilizes a sliding window method to scan the entire image, and Euclidean distance-based deduplication filtering maintains the uniqueness of detected objects. The detection results are provided as visualization image files and data for review and revision by ecologists. The proposed system demonstrates its potential for quickly and accurately estimating the population of individuals within a single image and its potential as a practical ecological information analysis tool without the need for large datasets.

**Keywords** YOLOv8, Suncheon bay photo, hooded crane, object detection, population estimation

†정회원 IT educator  
 ††중신회원 Professor, Department of Computer Education, Suncheon National University (Corresponding Author)  
 논문투고 2025년 10월 13일  
 심사완료 2025년 11월 26일  
 게재확정 2026년 01월 05일  
 발행일자 2026년 02월 25일

\* 본 논문은 제1저자의 순천대학교 일반대학원 석사학위논문 일부를 발췌하여 요약, 정리한 것임.

## 1. Introduction

### 1.1 Research Background

Suncheon Bay's wetlands are renowned worldwide for their well-preserved status. Hooded cranes arrive from Siberia around late October, winter in Suncheon Bay, and depart for Siberia around late March of the following year. As of February 2025, the number of wintering cranes had increased to approximately 7,000 [1]. Tidal flat experts have been monitoring the population, using binoculars to observe hooded cranes two to three times a week as they nest in the tidal flats and fly to their feeding grounds, the farmlands.

Previous studies have shown that the number of hooded cranes, along with white-naped cranes and Japanese cranes, visiting wintering grounds in Korea fluctuates annually [2, 3]. In particular, the distribution of cranes traveling between nesting and farmland varies, making it crucial to collect and manage data through field surveys utilizing technology such as video recording and drones [4, 5].

The hooded crane is an internationally protected endangered species, and monitoring its migration paths and habitats is crucial for ecological conservation research. Aerial photography using drones is a useful tool for observing entire populations at high resolution. However, a single image often contains hundreds of individuals, making manual identification and counting of individuals a time-consuming and labor-intensive process. In these environments, object detection technologies, particularly deep learning-based models like YOLO (You Only Look Once), can offer an alternative with fast and accurate detection performance.

However, most deep learning models require large amounts of labeled training data, which poses a practical limitation for species like the hooded crane, where observational data is limited and expertise is required. Furthermore, ecological fields often require precise analysis based on a single image, making existing approaches focused on large-scale learning limited in practicality.



Figure 1. Manual work to determine the number of hooded cranes in Suncheon Bay.

Accordingly, there is a need to develop a system that can perform effective object detection in a single image with only a limited number of manual labels, which can contribute to supporting the counting work of ecological experts and establishing a foundation for data accumulation.

### 1.2 Research Purpose and Necessity

The purpose of this study was to develop a software tool that automatically detects the locations of hooded crane individuals in a single high-resolution ecological image, thereby enabling ecological experts to quickly and accurately determine population numbers. Hooded cranes occur in dense colonies, and further research requires a detection-based approach that clearly identifies the locations of individual individuals, rather than a density-based estimation method. However, due to the nature of this endangered species, securing large-scale training data is difficult, and initially, only a small number of individuals can be manually labeled.

This study is based on the following needs:

First, we address the issue of data insufficiency by designing a model capable of training with a small number of user labels. Second, we utilize lightweight models such as YOLOv8n to ensure both learning speed and practicality. Third, we visualize detection results that can be directly modified and reviewed by experts, providing a tool that can be immediately applied in actual ecological settings.

## 2. Related Research

Object detection technology is effectively utilized for tracking and counting wildlife. In particular, the YOLO (You Only Look Once) family of models is widely used in ecological monitoring due to its high detection speed and accuracy.

Chen and his colleagues developed an advanced and accurate method for automatically calculating the number of cranes (*Grus grus*), using thermal cameras at night and visible light (RGB) cameras during the day onboard unmanned aerial vehicles (UAVs), based on image analysis and computer vision from the YOLO v3 platform [6]. Chen et al. developed the YOLO-SAG model, which incorporates the Softplus Attention technique, enabling reliable wildlife detection even in complex environments [7]. Noguchi et al. applied YOLOv7 to automatically estimate the population of green sea turtles based on coastal images captured by drones [8], and Tripathi et al. implemented a lightweight system for real-time detection of swamp deer using YOLOv8 [9].

Furthermore, Yang et al. proposed a lightweight model based on YOLOv5s in a forest environment [10], and Chalmers et al. developed a system for real-time detection and warning of curlew nests using YOLOv10 [11]. Ma et al. proposed the YOLOv8-bird model, an improved version of YOLOv8, for migratory bird monitoring in Poyang Lake, improving detection performance for small objects [12, 13].

Studies targeting social birds are also increasingly emerging, and the Appsilon research team developed a single-class YOLOv5 model to detect Antarctic cormorant nests [14]. Furthermore, Kim S and Kim M proposed an approach utilizing a Density Activation Map (DAM) to accurately estimate the population of migratory bird colonies, such as the Hooded Crane [15].

While the effectiveness of YOLO-based detection models has been demonstrated across diverse environments and species, studies on detecting endangered species like the Hooded Crane, which occurs in high density in a confined area, from a single image are rare, highlighting the importance of this study [16, 17].

### 3. Program Design and Development

#### 3.1 Development of a User Label Collection Interface

In this study, we developed a user interface-based label collection tool to build training data with minimal user labels for single-image YOLO detection model training. This tool is designed to enable ecological experts to visually distinguish between objects (hooded cranes) and background in high-resolution images, allowing them to easily register object and background regions in rectangular units by clicking or dragging.

As shown in <Figure 2>, the user selects approximately 50 hooded crane objects from the entire image and individually marks them in a red rectangular window. Furthermore, approximately 50 non-hooded crane objects, such as mudflats, are designated in blue windows, preparing for label-based machine learning. A YOLO-based auxiliary program then performs object detection by slicing the entire image at regular intervals.

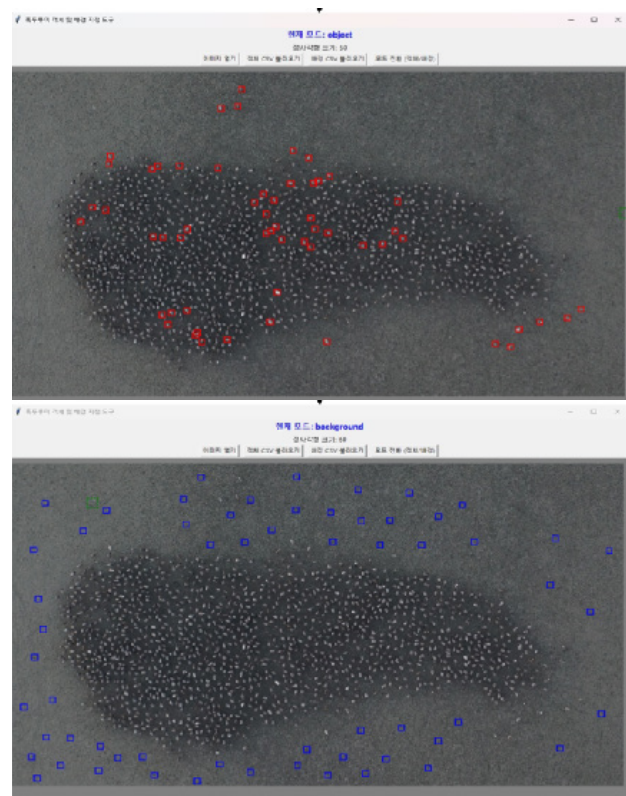


Figure 2. Designating hooded crane objects in a red window (top) & designating non-objects in a blue window (bottom)

Furthermore, the following features were designed to ensure ease of use and review efficiency for actual field experts:

1) Object/Background Mode Switching and Deletion: Keyboard shortcuts allow for easy switching between object and background

designation modes, and incorrectly designated areas can be deleted or corrected.

2) Previous Work Retention and Save Function: Designated object/background information can be saved and retrieved as a file, ensuring continuity during extended work.

3) Counting Mode: The system automatically counts and displays the number of hooded cranes in selected areas, either predicted by the model or user-specified objects, allowing experts to quickly review and refine model predictions (Figure 3).

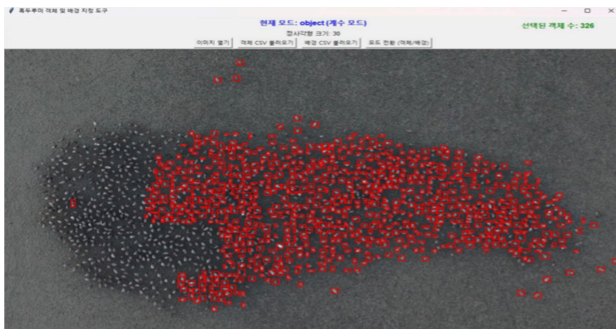


Figure 3. Example of using the hooded cranel population mode.

Users can cancel recently selected objects or background rectangles using the space key, and quickly adjust the size of a selected rectangle using the '+' and '-' keys. The developed auxiliary tool is designed to be operable by users without any separate coding knowledge, providing an interface that allows ecological experts to directly identify objects within an image in real time and make necessary corrections. These tools support the rapid and accurate construction of label data required for YOLO model training, and are therefore suitable for playing a key role in the development of high-precision detection models based on single images.

### 3.2 YOLOv8n-Based Detection and Deduplication

Based on object and background information obtained through a user label collection tool, this study trained the YOLOv8n (nano) model to build a detection model specialized for a single image. YOLOv8n boasts advantages in model weight reduction and learning speed, and its architecture allows for stable performance even in environments with small training data volumes, such as drone photos of individual hooded cranes. The data training process proceeded as follows:

Training Data Composition: Training images of

a certain size and label files in YOLO format were generated based on user-specified object (positive) and background (negative) regions. If background images were insufficient, augmentation was performed through position shifting, etc., to balance the data ratio between the object and background.

YOLOv8n Training and Application: The YOLOv8n model was trained using the constructed training data, and the trained model was applied to the original image to perform full detection. Detection was performed by dividing the image into equal intervals, and the location and size information of the object predicted by the model was collected from each region.

Deduplication Filtering: Since detection results may contain duplicate detections of the same object across multiple windows, Euclidean distance-based filtering was applied to remove duplicate objects from adjacent areas. Predictions within a certain distance are considered the same object and merged into one, thereby increasing the accuracy of the final detection results.

Through this structure, the researchers developed an auxiliary program that enables reliable object detection in a single image with only a small amount of manual labeling. In particular, the detection results for the hooded crane colony were reviewed by ecological field experts, demonstrating the ability to secure additional high-quality data and refine the object recognition model.

## 4. Experiments and Results

### 4.1 Experimental Environment and Dataset

This study utilized NVIDIA T4 GPUs in a Google Colab environment for training and detection experiments of the YOLOv8n model. The model was implemented using Python and PyTorch, based on the Ultralytics YOLOv8 package. YOLOv8n is the most lightweight version of the YOLO series, boasting fast training speeds and a low risk of overfitting in environments with small data sets, making it ideal for this study.

The data used in the experiment were high-resolution vertical images captured using a drone

around Suncheon Bay, South Jeolla Province, provided by the Suncheon City Ecological Conservation Division and a professional ecological surveyor. Hooded cranes, which winter in Suncheon Bay from late October, were photographed from approximately 50 meters above the tidal flats between 6 and 7 a.m. The researchers conducted detection experiments on 10 randomly selected images. Each image contained hundreds of hooded cranes, providing ideal conditions for verifying detection accuracy.

**Table 1.** Experimental environment

Item	Description
Platform	Google Colab
Runtime Settings	GPU(NVIDIA T4)
Framework	Pytorch-based YOLOv8
Usage Model	YOLOv8n
Programming Language	Python 3.10
Major Libraries	OpenCV, Pandas etc.
Test Target Images	10 high-resolution photos of hooded crane colony

The researchers also compared the model's detection results against the manual counts (actual population counts) performed by ecological experts for the images. Some of the images contained over 1,000 individuals, effectively validating the detection model's practicality.

## 4.2 Results of Hooded crane population recognition

The following factors were collected to evaluate the model's detection performance. First, the number of true objects (TP+FN) represents the number of true objects in the error matrix. This is based on the results of manual counting by experts on high-resolution colony images provided by the Suncheon City Ecological Conservation Division. Second, the number of detected objects (TP+FP) represents the number of objects that the model detected as hooded cranes. This is based on the total number of objects detected in each image using the YOLO-based detection model. Third, the number of false positives (FP) represents the number of objects that the model detected that were not true hooded cranes. This is based on the value manually counted in collaboration with experts.

Meanwhile, the metrics used to evaluate the

performance of estimating the hooded crane population are precision, recall, and the F-1 score.

Precision: The percentage of objects detected by the model that are identified as real.

"Precision" =  $\frac{\text{Number of objects true positives (TP)}}{\text{Number of objects detected (TP+FP)}}$

Recall: The percentage of objects that actually exist that the model correctly detects.

"Recall" =  $\frac{\text{Number of objects true positives (TP)}}{\text{Number of actual objects (TP+FN)}}$

F1-score: The harmonic mean of precision and recall, calculated as follows.

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Using 10 randomly selected high-resolution drone photos, the researchers assigned approximately 50 labels to each of hooded cranes and tidal flats, then ran the program to recognize hooded cranes, as shown below. After removing duplicated objects using an auxiliary program, the number of hooded cranes actually identified, the number of falsely detected hooded cranes, and the number of undetected hooded cranes were confirmed through key operations. Through this process, the number of accurately recognized hooded cranes was finally confirmed, and precision, recall, and F1 score were calculated according to the indices introduced above.

**Table 2.** Results of the hooded crane population recognition experiment

Image Number	Actual Object	Detected Object	False Detection Object	Precision	Recall	F1 score
1	1,337	1,114	13	0.988	0.823	0.898
2	621	661	41	0.937	0.998	0.967
3	707	536	13	0.975	0.739	0.841
4	181	219	38	0.826	1	0.905
5	460	430	3	0.993	0.928	0.959
6	455	436	22	0.950	0.910	0.929
7	829	796	2	0.997	0.958	0.977
8	318	316	7	0.978	0.972	0.975
9	815	800	20	0.975	0.957	0.966
10	495	511	26	0.949	0.980	0.964
Avg.	621.8	581.9	18.5	0.957	0.927	0.938

As shown in Table 2, the experimental results

for estimating the number of hooded cranes using a single image demonstrate that the single-image-based YOLOv8n model proposed in this study achieved high detection performance even with a small number of user labels. Specifically, it achieved an average precision of 95.7%, a recall of 92.7%, and an F1-score of 93.8%, experimentally demonstrating the feasibility of reliable object detection without the need for large-scale training.

The process of recognizing hooded cranes in each image took approximately 1-2 hours. However, this approach is inherently fundamental, requiring manual verification of the entire hooded crane population and calculation of each metric. Future work remains to increase the number of images and obtain statistical data on the metric calculations as parameters (e.g., window size, stride, etc.) change.

### 4.3 Model Comparison

To verify the effectiveness of the proposed single-image-based model, this study conducted comparative experiments on the following three learning methods using the same YOLOv8n architecture.

**Single-Image-Based Model (Model 1):** This method collects user labels only for a specific image, trains a model specialized for that image, and performs detection on the same image. This is identical to the proposed system's architecture.

**Multi-Image-Based Integrated Model (Model 2):** This method collects user labels from multiple images, integrates the entire data set, trains, and then performs detection on a specific image. The goal is to improve general performance.

**Generalization Model (Model 3):** This method performs detection on new images not used in training. This model is intended to evaluate generalization performance in real-world ecological surveys.

The experimental results showed that the single-image-based model performed similarly to the multi-image-based model even with a small number of user labels. Its high detection accuracy, especially for specific images, confirms its suitability for supporting expert review and counting tasks. On the other hand, the generalized model had high precision but low recall, resulting in many missed detections. This suggests that although it may be advantageous for universal detection across a variety

of images, it is inefficient for estimating the number of objects within a single image.

**Table 3.** Model comparison and results

Item	Model 1	Model 2	Model 3
Number of training images	1	1, 2, 3, 4, 5	2, 3, 4, 5, 6
Number of target images to predict	1		
Precision	0.988	0.998	1.0
Recall	0.823	0.850	0.614
F1-score	0.898	0.918	0.760

Comparative experiments between the three models showed that the single-image-based model performed similarly to the multi-image-based model and had significantly fewer missed detections than the generalized model, making it more suitable for assisting field counting tasks.

These results suggest that the proposed system can be utilized as a practical auxiliary tool in data-limited endangered species ecological monitoring sites, and demonstrate that its structure allows for long-term data accumulation and model performance improvement through iterative learning.

## 5. Conclusion

This study developed an object detection-based auxiliary system capable of automatically estimating the population of the endangered hooded crane (Old Crane) by analyzing high-resolution ecological images. Using drone images of hooded cranes in Suncheon Bay, we trained a YOLOv8n model with only a small number of user labels and built a detection model specialized for a single image. This experimentally demonstrated high detection performance without the need for large-scale training data.

However, the developed program does not finalize hooded crane detection results; rather, it serves as an auxiliary tool to support expert review, enabling rapid and efficient population census in ecological survey fields [1, 3, 6, 13]. Furthermore, this study focused on flat terrain (rice paddies and tidal flats) and a single species, the hooded crane. Further experimental validation is needed for the diverse terrain of Suncheon Bay, including tidal

flats, rice paddies, and mountainous terrain, as well as for other bird species.

The proposed system demonstrated promising results in practical research and field applications in the following ways.

**Reduced data burden:** Effective model training is possible with ecological experts manually designating a small number of individuals and backgrounds, eliminating the need for extensive labeling required in previous studies.

**High-precision detection performance:** With an average F1-score of 93.8%, single-image-based models achieved detection precision and recall levels sufficient for practical population estimation.

**Practical, expert-centered design:** A user interface designed to visualize prediction results and allow for direct review and modification can simultaneously improve the reliability of detection results and work efficiency for field experts.

As previously discussed in previous studies, ecological research utilizing object detection technology is becoming increasingly active. In particular, YOLO-based models have been effectively applied to the detection and population estimation of various species. However, most of these studies typically train models based on a large number of training images and large, labeled datasets. While leveraging a sufficient number of samples and class information to improve generalized detection performance is advantageous in environments where data acquisition is readily available, it is difficult to apply to subjects with limited data collection, such as endangered species or grouping birds.

Furthermore, while learning common patterns from multiple images can enhance the model's versatility, it can lead to missed detections or inaccurate location predictions in single-image analysis situations where the identification of objects within a specific image is crucial for rapid and accurate identification. In contrast, this study presents the following advantages:

The program achieves practical detection performance even without large-scale data by learning and applying a YOLOv8n detection model tailored to a single high-resolution image and a small number of user labels. Further, it enhances field applicability through designing a visualization-based auxiliary system that allows direct review and modification. It turns out to complement the

limitations of existing density-based estimation methods by clearly identifying the individual locations of the hooded cranes in Suncheon bay area.

In conclusion, this study presents a practical ecological monitoring tool that can be immediately applied in the field, while demonstrating the potential of a learning structure capable of reliable detection even in data-limited environments. The developed auxiliary program is expected to assist field experts who monitor hooded cranes by manually counting them weekly from late October, the wintering season, to late March of the following year. By automating some of the manual work required to count individuals in clusters on tidal flats, even with drone photography, it can have the additional benefit of systematically storing hooded crane data and building a big data base in the long term.

During this process, it was determined that variations in height and angle, such as those observed when photographing flocks of hooded cranes from an aerial drone, did not significantly affect the preprocessing process when applying the YOLOv8n model.

In the future, this research will train on a variety of data to ensure high detection performance not only for hooded cranes and Suncheon Bay, but also for a variety of species and environments. Furthermore, a follow-up task is to develop an iterative improvement system that retrains the hooded crane detection model, incorporating the findings of tidal flat and bird experts.

---

## References

---

- [1] Son, J. (2024). Distribution changes of wintering hooded cranes (*Grus monacha*) in Suncheon Bay and behavioral responses against potential threats. Master thesis at Chonnam National University (in Korean).
- [2] Kim, H. (2018). Land use changes due to the release of civilian control zone and trends in population fluctuations of cranes. Master thesis at Gyeongnam National University of Science and Technology (in Korean).
- [3] Yoo, S., Ham, J., Yang, M., & Kwon, I. (2024). Distribution and use of roosting site by wintering Red-crowned Cranes (*Grus japonensis*) in Ganghwa tidal flat, Korea. *Korean Journal of Ornithology*, 31(2), 200-207.

- [4] Choi, M. & Park, B. (2021). Technological DMZ at <https://www.anthropocene-curriculum.org/case-study/technological-dmz-digital-technologies-and-the-conservation-of-cranes>
- [5] Lee, N. (2022). Spatial distributions of shorebirds and their hourly changes in tidal flats observed with unmanned aerial vehicles. Master thesis at Anyang University (in Korean).
- [6] Chen, A., Jacob, M., Shoshani, G., et al. (2023). Using computer vision, image analysis and UAVs for the automatic recognition and counting of common cranes (*Grus grus*). *Journal of Environmental Management*, 328, 116948. <https://doi.org/10.1016/j.jenvman.2022.116948>
- [7] Chen, L., Li, G., Zhang, S., et al. (2024). YOLO-SAG: An improved wildlife object detection algorithm based on YOLOv8n. *Ecological Informatics*, 83, 102791. <https://doi.org/10.1016/j.ecoinf.2024.102791>
- [8] Noguchi, N., Nishizawa, H., Shimizu, T., et al. (2025). Efficient wildlife monitoring: Deep learning-based detection and counting of green turtles in coastal areas. *Ecological Informatics*, Volume 86, 103009. <https://doi.org/10.1016/j.ecoinf.2025.103009>
- [9] Tripathi, R., Agarwal, K., Tripathi, V., et al. (2025). Conservation in action: Cost-effective UAVs and real-time detection of the globally threatened swamp deer (*Rucervus duvaucelii*). *Ecological Informatics*, 85, 102913. <https://doi.org/10.1016/j.ecoinf.2024.102913>
- [10] Yang, W., Liu T., Jiang P., et al. (2023). A forest wildlife detection algorithm based on improved YOLOv5s. *Animals*, 2023, 13(19), 3134. <https://doi.org/10.3390/ani13193134>
- [11] Chalmers, C., Fergus, P., Wich, S., et al. (2025). AI-driven real-time monitoring of ground-nesting birds: A case study on curlew detection using YOLOv10. *Remote Sensing*, 17(5), 769. <https://doi.org/10.3390/rs17050769>
- [12] Wang, L., Zhang, H., Yang, T., et al. (2021). Optimized detection method for Siberian crane (*Grus leucogeranus*) Based on Yolov5. 11th International Conference on Information Technology in Medicine and Education (ITME), Wuyishan, Fujian, China, 2021, pp. 01-06, <https://doi.org/10.1109/ITME53901.2021.00031>
- [13] Ma, J., Guo, J., Zheng, X., & Fang, C. (2024). An improved bird detection method using surveillance videos from Poyang Lake based on YOLOv8. *Animals*, 14(23), 3353. <https://doi.org/10.3390/ani14233353>
- [14] Cusick, A., Fudala, K., Storozenko, P., et al. (2024). Using machine learning to count Antarctic shag (*Leucocarbo bransfieldensis*) nests on images captured by remotely piloted aircraft systems. *Ecological Informatics*, Volume 82, 2024, 102707 <https://doi.org/10.1016/j.ecoinf.2024.102707>
- [15] Kim, S., & Kim, M. (2020). Learning of counting crowded birds of various scales via novel density activation maps. *IEEE Access*, 8, 155296-155305. <https://doi.org/10.1109/ACCESS.2020.3019069>
- [16] Ferrante, G., Nakamura, L., Sampaio, S., et al. (2024). Evaluating YOLO architectures for detecting road killed endangered Brazilian animals. *Scientific Reports*, 14, Article 1353. <https://doi.org/10.1038/s41598-024-52054-y>
- [17] Seong, H. (2025). *Development of a Program for Estimating Hooded Crane Population Using the YOLO Model*. Master's thesis at Suncheon National University.



**Hongju Seong**

- 2021 Suncheon National University, Department of Computer Education (B.S.)
- 2025 Suncheon National University, Department of Computer Education & Information (M.S.)

✦ Research Interest : graph theory, artificial intelligence, deep learning, education, gifted education

✉ labmen42@naver.com



**Youngcook Jun**

- 1986 Suwon University, Department of Mathematics (B.S.)
- 1986 Chicago State University, Department of Mathematics (M.S.)
- 1995 University of Illinois at Urbana-Champaign, Ph.D.
- 1996–Present Professor, Department of Computer Education, Suncheon National University

✦ Research Interest : symbolic computation, agent modeling, robotics, case studies

✉ ycjun@suncheon.ac.kr