Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 6, December 2025, Page. 5435-5449
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

Detection of Endangered Indonesian Species Across Multiple Taxonomic Classes Using Faster R-CNN

Moh. Jabir Mubarok*[1], Rizky Fitria Haya[2], Eka Fitria[3], Brilian Surya Budi[4]

[1]Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[2,3]Informatics, Universitas Syiah Kuala, Indonesia
[4]Computer Science, University of Tsukuba, Japan

Email: [1]jabirmubarok@gmail.com

Received : Jun 2, 2025; Revised : Jun 17, 2025; Accepted : Jun 18, 2025; Published : Dec 22, 2025

Abstract

Indonesia's rich biodiversity includes many endangered species across various taxonomic groups. This study presents a Faster R-CNN deep learning model to detect ten endangered Indonesian species, covering birds, reptiles, mammals, and fishes. A custom dataset with diverse images was annotated and used to train the model with transfer learning on the Detectron2 framework. Evaluation using COCO metrics yielded an average precision (AP) of 54.93%, with the Komodo Dragon achieving the highest AP (82.57%) and Wallace's Standardwing the lowest (30.82%). The model excels at detecting larger, distinct species but has difficulty with smaller or camouflaged ones in complex environments. Training results confirm that transfer learning aids performance despite limited data. Analysis of misclassifications suggests the need for additional data modalities or context to improve accuracy. This work highlights the potential of Faster R-CNN for automated endangered species monitoring in Indonesia and recommends dataset expansion, data augmentation, and model refinement to enhance detection, particularly for challenging species. This study contributes to computer vision applications in conservation, particularly within low-resource biodiversity contexts.

Keywords : *Deep Learning, Detection, Endangered Species, Faster R-CNN, FPN, Transfer Learning*

1. INTRODUCTION

Biodiversity in Indonesia is among the richest in the world, containing a vast number of endemic species and those at risk of extinction [1]. The country features diverse ecosystems, from tropical forests to marine environments, each supporting various organisms [2]. However, human activities like habitat destruction, poaching, and the illegal wildlife trade pose threats to many species [3]. Due of the extensive areas that need to be covered, traditional biodiversity monitoring techniques like field surveys and manual observations are costly, time-consuming, and limited [4]. Advances in technology, especially deep learning and computer vision, have created opportunities to automate species detection and monitoring [5]. This automation offers faster, more scalable, and more accurate tools for conservationists to track endangered species [6].

Region-based Convolutional Neural Networks (R-CNNs), one of the many deep learning methods, have shown remarkable efficacy in object detection tasks [7]. The two-stage deep learning architecture known as the Faster R-CNN model has demonstrated special promise because of its capacity to precisely produce region proposals and identify objects in intricate, cluttered environments [8]. These capabilities make Faster R-CNN an ideal tool for detecting endangered species in the wild, where environmental conditions are often less controlled and highly variable.

Recent advancements in computer vision and remote sensing technologies have significantly enhanced biodiversity monitoring and the detection of endangered species. To increase the effectiveness

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

and precision of wildlife monitoring systems, numerous research have used deep learning techniques such Faster R-CNN and YOLO to detect species using satellite and drone imagery. For instance, research [9] demonstrated that convolutional neural networks (CNNs) are effective at detecting large species with distinct features but face difficulties with smaller or camouflaged species. Their results demonstrated how crucial high-quality datasets and data augmentation are to enhancing model performance, which is consistent with the study's methodology. Additionally, study [10] combined remote sensing data from multiple sensors, including optical and radar imagery, with machine learning to monitor biodiversity in complex landscapes, highlighting the necessity of multimodal data integration for tracking endangered species obscured by dense vegetation or environmental factors.

Further investigations have focused on enhancing detection accuracy by incorporating spatial and temporal features into CNN models, addressing challenges related to species movement and changing habitats [11]. The use of high-resolution aerial imagery combined with data augmentation and transfer learning techniques was emphasized in study [12], improving detection of small or difficult-to-spot species and enabling adaptation of models to new environments. Study [13] explored the influence of agricultural practices on endangered species distribution, advocating for the integration of environmental variables such as soil and climate data to strengthen biodiversity monitoring frameworks. Most relevant to the current research, study [14] conducted a thorough evaluation of Faster R-CNN with a ResNet backbone for endangered species detection in remote sensing data, providing critical insights into model performance in cluttered and complex backgrounds and guiding the evaluation strategy and focus areas for model improvement in this work.

This work is among the first to apply Faster R-CNN for multi-class detection of Indonesian endangered species, including avian, reptilian, amphibian, and marine taxa. Specifically, it aims to detect ten species at risk of extinction from images captured in their natural habitats. Utilizing deep learning, this study suggests a dependable and effective method for monitoring biodiversity that may greatly improve conservation initiatives in tropical areas.

This study makes the following primary contributions:

- We explore how Faster R-CNN can be used to identify endangered species across various types of taxonomic groups, showcasing its resilience in challenging environments.
- We present a framework for multi-class species detection that can be easily adapted to monitor different species in Indonesia and other biodiversity hotspots.
- We evaluate the model’s performance based on a dataset of endangered species, providing awareness on the advantages and disadvantages of the prevailing deep learning-based conservation strategies.

2. RESEARCH METHODS

This study identifies endangered species in Indonesia using a deep learning technique and the Faster R-CNN model. The research process starts with collecting and labeling species images, followed by preprocessing and data augmentation to improve model learning. Then, the model is trained using transfer learning with a ResNet-101 backbone. The block diagram in Figure 1 summarizes these main steps.

2.1. Datasets

The dataset utilized for this research was gathered from iNaturalist [15], a reputable, community-driven bio-diversity platform that hosts millions of wildlife images along with rich taxonomic and geospatial metadata. To ensure ethical usage, only images licensed under Creative Commons (CC BY and CC0) were selected, allowing academic and non-commercial use.

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 6, December 2025, Page. 5435-5449
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

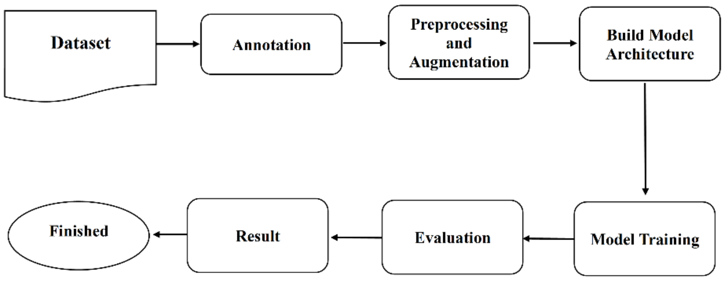


Figure 1. Overview of the Research Methodology

This study focuses on ten endangered species native to Indonesia, representing diverse taxonomic classes including Aves (birds), Reptilia (reptiles), Mammalia (mammals), and Actinopterygii (ray-finned fishes). Each species contributes approximately 40 samples, resulting in a total of 400 curated images. Although limited in size, this dataset is sufficient for prototyping and evaluating the proposed method.

Images were filtered using the following quality criteria:

- High visual clarity, with sharp focus on the subject.
- Varied natural backgrounds (e.g., forests, coral reefs, open sea, inland waters) to simulate real-world ecological conditions.
- Pose and orientation diversity, including frontal, lateral, and oblique views.
- Geographical diversity across Indonesian regions and islands.

These selection criteria were designed to maximize data diversity and ensure robust model training under realistic deployment scenarios [11].

Table I lists ten endangered species used in this study, including their common names, scientific names, and taxonomic classifications. These species were chosen based on conservation status, data availability, and ecological importance to Indonesian biodiversity [26]-[35].

Table 1. List of Endangered Target Species

No.	Common Name	Scientific Name	Taxonomic Class
1	Bali Myna	*Leucopsar rothschildi*	Aves
2	Banggai Cardinalfish	*Pterapogon kauderni*	Actinopterygii
3	Green Sea Turtle	*Chelonia mydas*	Reptilia
4	Java Green Peafowl	*Pavo muticus*	Aves
5	Komodo Dragon	*Varanus komodoensis*	Reptilia
6	Maleo	*Macrocephalon maleo*	Aves
7	Napoleon Wrasse	*Cheilinus undulatus*	Actinopterygii
8	Orangutan	*Pongo spp.*	Mammalia
9	Pygmy Seahorse	*Hippocampus bargibanti*	Actinopterygii
10	Wallace's Standardwing	*Semioptera wallacii*	Aves

The inclusion of species from multiple taxonomic groups presents a complex multi-class detection problem. This diversity mirrors real-world wildlife monitoring challenges, where systems must generalize across varying morphologies, sizes, and environmental contexts [36].

2.2. Preprocessing and Augmentation

At this stage, each image in the dataset was meticulously annotated using the Label Studio platform, a flexible and open-source tool tailored for computer vision tasks [37]. Annotators manually drew bounding boxes around visible instances of the target species and assigned them to their respective

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

class labels [38]. To ensure the integrity and consistency of annotations, each image was reviewed by two independent reviewers [39]. Discrepancies, if any, were resolved through consensus to maintain annotation reliability [40].

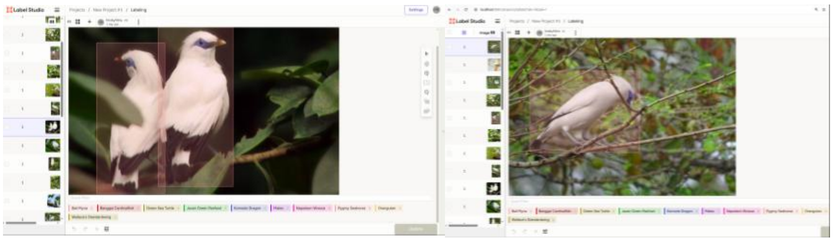


Figure 2. Example of image annotation and labeling procedure in Label Studio

A stratified splitting strategy was employed to prepare the dataset for training and validation, dividing it into two subsets: 80% for training and 20% for validation. Stratification was based on class labels to ensure that each subset maintained the original distribution of species classes. This approach was essential to prevent data imbalance, especially considering the limited number of samples per class, and to present a more accurate assessment of the model's performance in every class [14].

During training, a number of data augmentation techniques were used to improve the model's performance under different visual conditions and increase its capacity for generalization. By simulating real-world situations, these enhancements make the model more resilient to variations in scale, color, lighting, and position. The techniques implemented include:

- Horizontal Flipping: Images are randomly flipped horizontally so the model can recognize objects from different directions, especially if they are symmetrical.
- Random Brightness and Contrast: Brightness and contrast are randomly adjusted to mimic different lighting conditions like bright sunlight or shadows.
- Color Jittering (Hue and Saturation): Colors are randomly changed in hue and saturation to help the model handle variations caused by cameras or lighting.
- Scaling and Resizing: Images are scaled and resized so the model can detect objects at different sizes and distances.
- Random Cropping: Random parts of images are cropped so the model learns to recognize partially visible objects, as often happens in real-world scenes.

All augmentations were dynamically applied during training using a custom data mapper in the Detectron2 framework. By performing these transformations on-the-fly, the training dataset diversity was effectively increased without requiring additional storage [21].

To ensure input consistency and improve training efficiency, every image was downsized to 640×640 pixels, which was the fixed resolution. This resolution offers a balance between preserving essential visual details and maintaining computational efficiency for accelerated training and inference [21].

2.3. Model Architecture

The Faster Region-based Convolutional Neural Network (Faster R-CNN) architecture is used in this study to identify endangered species from different taxonomic groups. There are two primary parts to the Faster R-CNN:

1) Region Proposal Network (RPN) creates region proposals that are likely to contain objects.
2) Fast R-CNN head, which classifies the proposed regions and refines bounding box coordinates.

ResNet-101 integrated with a Feature Pyramid Network (FPN) serves as the foundation for feature extraction. ResNet-101 was chosen due to its ability to produce deep visual representations, whereas

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 6, December 2025, Page. 5435-5449
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

FPN enhances multi-scale object detection, which is essential for recognizing wild species of different sizes and positions. Model implementation is conducted using the Detectron2 frame-work, which provides modular pipelines for efficient object detection training [8].

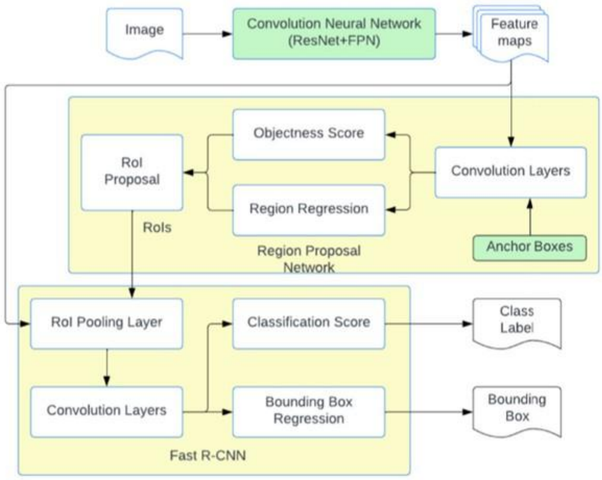


Figure 3. The structure and flowchart of a Faster R-CNN architecture in this paper

2.4. Model Training

The objective of the training process was to maximize item detection capabilities with constrained data. It involved transfer learning, carefully tuned hyperparameters, and a robust training setup using modern deep learning frameworks.

1) *Transfer Learning Strategy*: A transfer learning approach was adopted by initializing the model with COCO-pretrained weights of the Faster R-CNN architecture [8]. This strategy allows the model to leverage feature representations learned from a large-scale dataset (MS COCO) [16], enabling better convergence and performance even with a smaller target dataset. By using pre-trained weights, the network starts with generalized object features such as edges, shapes, and textures, thus reducing the need to learn from scratch.
2) *Hyperparameters:* The training was configured with the following hyperparameters, optimized for stability and performance:

Table 2. Training Hyperparameters

Parameter	**Value**
Backbone	ResNet-101 + FPN
Optimizer	Stochastic Gradient Descent (SGD)
Learning Rate	$1e^{-4}$
Momentum	0.9
Weight Decay	$1e^{-4}$
Batch Size	2 per GPU
Training Iteration	~ 10,000
Learning Rate Scheduler	Warmup + Step Decay
Loss Function	RPN + ROI Classification+ Bbox Regression

3) *Loss Function:* To evaluate the model's optimal performance during training, a loss function is used to compute the RPN loss and the classifier loss [8]. Equation (1) provides a definition of the loss function.

(1)

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

$$L(p_i, t_i) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* \, L_{\text{reg}}(t_i, t_i^*)$$

The total loss function used during training combines multiple components from the Faster R-CNN architecture:

$$\mathcal{L}_{total} = \mathcal{L}_{RPN} + \mathcal{L}_{cls} + \mathcal{L}_{bbox} \quad (2)$$

$\mathcal{L}_{RPN}$ represents the Region Proposal Network loss, which includes classification and bounding box regression for anchor boxes. $\mathcal{L}_{cls}$ is the classification loss for the final region of interest (ROI) predictions. $\mathcal{L}_{bbox}$ is the bounding box regression loss, typically calculated using smooth L1 loss. This multi-task loss formulation helps the network learn to propose regions, classify objects, and refine bounding box coordinates simultaneously.

4) *Training Environment:* Training was conducted using the PyTorch 2.6 framework with the Detectron2 library [17]. The hardware specifications used are as follows:

Table 3. Training Environment Specifications

Component	**Specification**
GPU	NVIDIA RTX 4060
RAM	16 GB
Framework	PyTorch 2.6 + Detectron2

2.5. Model Evaluation

The performance of the object detection model was assessed using the standard COCO Evaluation Metrics [16], which provide a comprehensive measure of model accuracy across varying thresholds and object sizes.

1) *Mean Average Precision (mAP):* Computed over multiple Intersection over Union (IoU) thresholds and ranges from 0.5 to 0.95 in increments of 0.05, is the main metric that is employed. This gives a more balanced view of the model's detection quality.

$$AP = \frac{1}{T} \sum_{t=1}^{T} AP_{\text{IoU}=t} \quad (3)$$

Where T = 10 for the 10 IoU thresholds {0.50, 0.55, . . ., 0.95}, and AP IoU=t is the average precision at a specific IoU threshold.

2) P Metrics at Specific Thresholds:
- AP50: Average Precision at IoU = 0.50, representing more lenient matching criteria.
- AP75: Average Precision at IoU = 0.75, representing stricter criteria with tighter bounding box alignment.

These metrics help distinguish how well the model balances precision and recall under both relaxed and strict matching conditions.

3) *AP Based on Object Sizes:* To further analyze model performance, Average Precision is also reported based on object scale:
- APS (Small): AP for objects with area less than 32x32 pixels.
- APM (Medium): AP for objects with area between 32x32 and 96x96 pixels.

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

- APL (Large): AP for objects with area greater than 96x96 pixels.

A confusion matrix can be used to examine the model's classification performance in addition to Average Precision (AP) metrics. A table that breaks down the number of correct and incorrect predictions by class is called a confusion matrix. It is useful to determine which classes the model frequently confuses. The degree of success for the classification task was determined by examining test performance using the confusion matrix. The true positive (TP), true negative (TN), false positive (FP), and false negative (FN) variables in the confusion matrix corresponded to the results of the classification process [24]. One key indicator of a performance assessment's accuracy is the degree of correspondence between predicted and actual data [25]. In the context of multi-class detection for endangered species, the confusion matrix helps to pinpoint specific species pairs that are frequently misclassified, guiding further improvements in data annotation and model training [23].

3. RESULT

After training the model using Faster R-CNN with a ResNet-101 backbone, the evaluation was performed using the prepared test dataset. The model was evaluated using standard COCO metrics for detecting endangered species. Table 4 are the evaluation results based on several metrics, measured using the COCO API.

3.1. Evaluation Metrics

The following are the Average Precision (AP) values calculated based on various criteria.

Table 4. Average Precision (AP) Values For Different Criteria

Metric	**Value**
AP (mean)	54.93%
AP50	85.99%
AP75	60.13%
AP for small objects (APS)	NaN
AP for medium objects (APM)	27.50%
AP for large objects (APL)	56.55%

3.2. Per-Species Evaluation

Table 5 are the Average Precision (AP) values for each species detected during the model evaluation.

Table 5. Average Precision (AP) Values For Different Criteria

Species	**AP (%)**
Bali Myna	67.00
Banggai Cardinalfish	49.20
Green Sea Turtle	68.29
Java Green Peafowl	63.39
Komodo Dragon	82.57
Maleo	49.94
Napoleon Wrasse	56.17
Orangutan	39.07
Pygmy Seahorse	42.81
Wallace's Standardwing	30.82

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

3.3. Classification Accuracy and Loss Metrics

Figures 4-6 showing the classification accuracy and the different loss components during the model training process are important for understanding how well the model is learning.

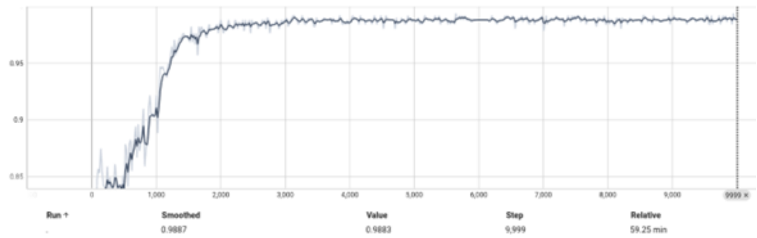
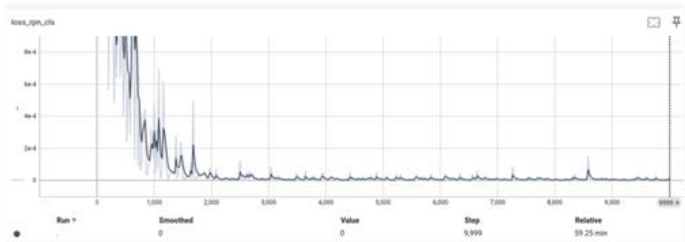
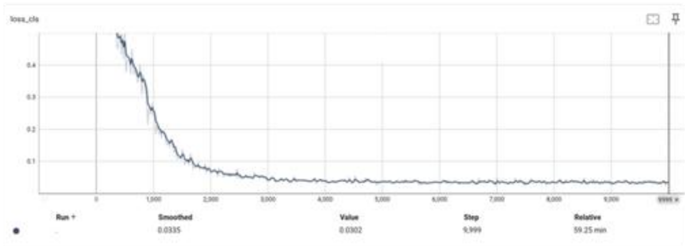


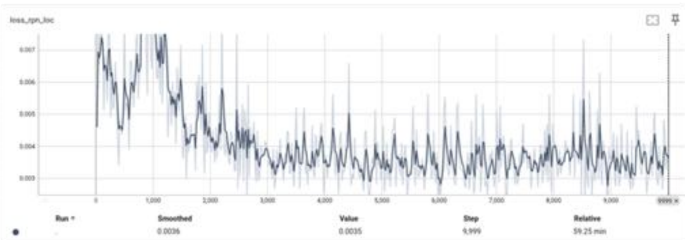
Figure 4. Faster R-CNN classification accuracy during training



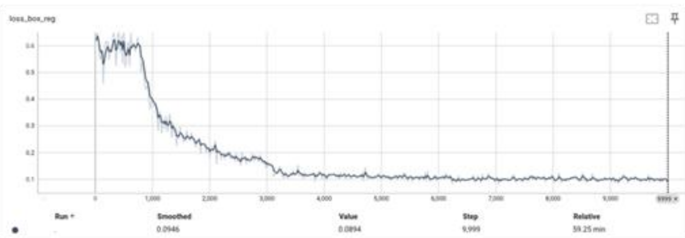
1) RPN classification loss



2) Classification loss



3) RPN location loss



4) Box regression loss

Figure 5. Loss components: 1. RPN classification loss, 2. classification loss, 3. RPN location loss, 4. and box regression loss

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

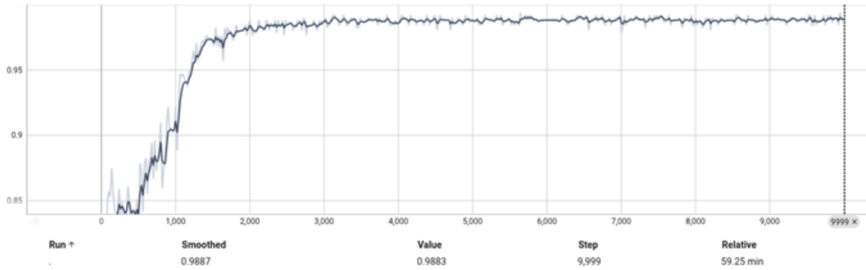


Figure 6. Total loss curve during Faster R-CNN training.

The figure 6 shows the decrease in total loss over time, indicating the model's learning process and optimization as it converges towards a better detection performance

The results of the model's endangered species detection are displayed in the confusion matrix in Figure 7. The columns show the Predicted Labels, and the rows show the True Labels. Whereas off-diagonal values signify incorrect classifications, diagonal values show accurate predictions.

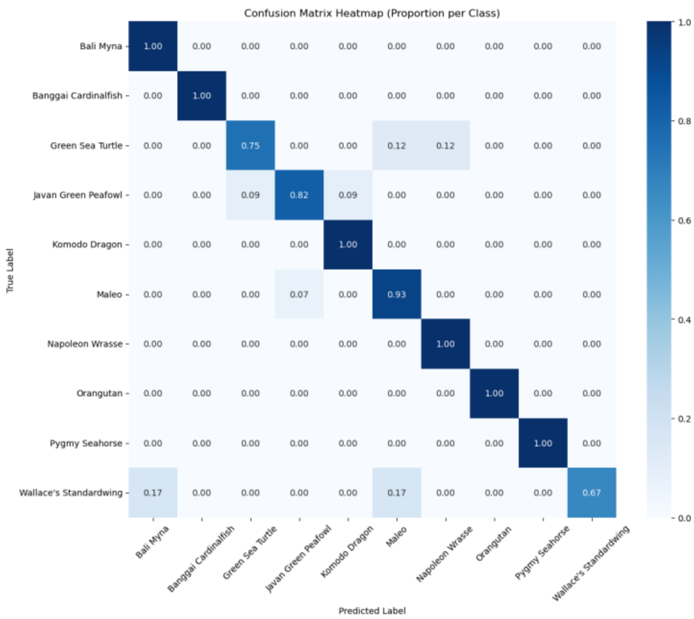


Figure 7. Confusion Matrix for Species Detection

3.4. Training Time and Computational Efficiency

The model's efficiency was also measured by tracking its data processing time in Figure 8-9 and false negative rate in Figure 10. These metrics are critical for understanding the computational efficiency of the model, especially in real-time applications.

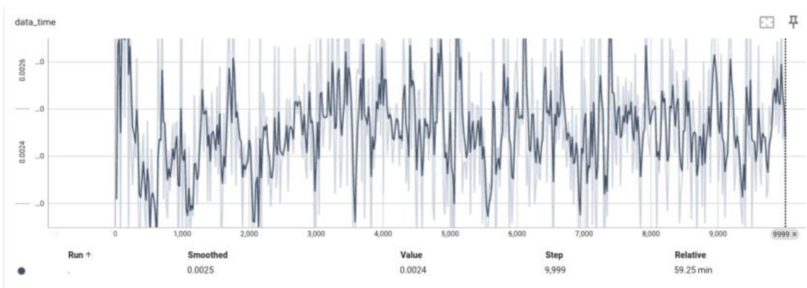


Figure 8. Data processing time during Faster R-CNN training

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

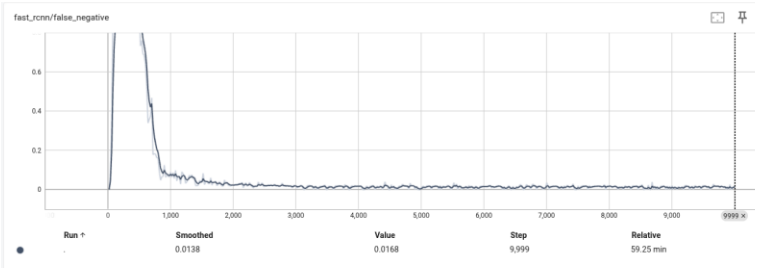


Figure 9. Estimated time remaining during training

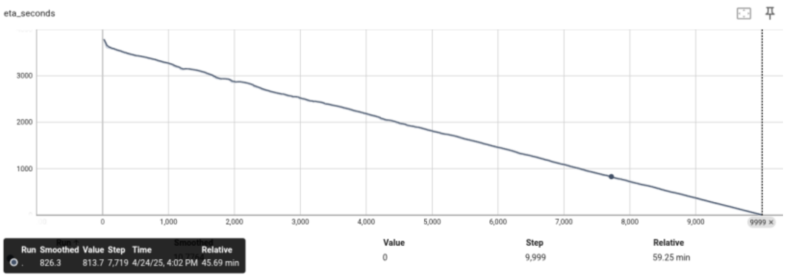


Figure 10. False negative rate during Faster R-CNN evaluation

4. DISCUSSIONS

The experimental results demonstrate that the Faster R-CNN model with a ResNet-101 backbone and Feature Pyramid Network (FPN) is effective in detecting endangered Indonesian species across multiple taxonomic classes. The overall mean Average Precision (mAP) of 54.93% indicates a moderate level of detection accuracy given the diversity of species and the limited size of the custom dataset. The model achieved its highest performance detecting the Komodo Dragon (AP = 82.57%) and Green Sea Turtle (AP = 68.29%), species characterized by relatively large size and distinct morphological features. These findings align with previous research [7,14] showing that Faster R-CNN excels at identifying objects with clear visual characteristics and substantial pixel representation in the image.

Conversely, the model's performance for smaller or more cryptic species, such as Wallace's Standardwing (AP = 30.82%) and Orangutan (AP = 39.07%), was considerably lower. This reflects known challenges in object detection when dealing with small-scale or camouflaged targets in complex natural backgrounds, consistent with observations from studies like Wang et al. (2024) [9]. The relatively poor Average Precision for medium-sized objects (APM = 27.50%) further emphasizes difficulties in detecting species that appear at intermediate scales, which may be due to occlusions, pose variations, or visual similarity to the background environment.

Training curves showing the gradual decrease in total loss and improvements in classification accuracy indicate that transfer learning from COCO-pretrained weights was beneficial, enabling the model to acquire valuable feature representations in spite of the limited data. This is consistent with the literature advocating transfer learning as an effective strategy for ecological image analysis [12].

The confusion matrix reveals some misclassifications between visually similar species within the same taxonomic classes, underscoring the need for more discriminative features or the integration of multimodal data, such as environmental or temporal context, to enhance differentiation. Incorporating additional spatial-temporal information or sensor modalities could be promising directions, as suggested by [10] and [11].

The Faster R-CNN framework provides a balance between accuracy and computational demand when compared to other state-of-the-art object detectors like YOLO and Mask R-CNN. However, future research could investigate ensemble methods or attention mechanisms to improve detection precision, especially for small and camouflaged species [23]. This research supports the advancement of computer vision in environmental informatics, enabling automated, scalable, and context-aware species detection systems.

4.1. Performance Based on Object Size

The model performed well in detecting larger objects such as the Komodo Dragon and Green Sea Turtle, with relatively high AP values of 82.57% and 68.29%, respectively. This shows that the model detects larger objects with distinct visual features more easily. However, the model struggled with smaller objects, such as the Orangutan and Pygmy Seahorse, resulting in lower AP values. The model's limitations in handling complex backgrounds and lower contrast are probably the cause of this difficulty in detecting smaller objects.

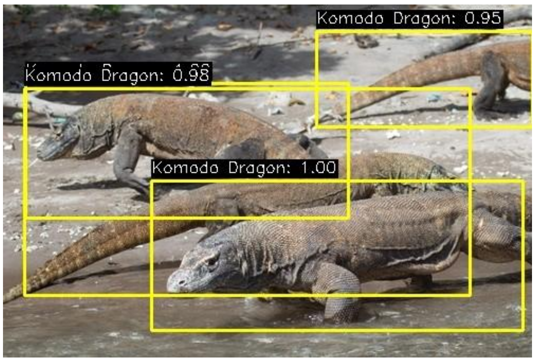


Figure 11. Example of successful detection on a large species (Komodo Dragon)

4.2. Challenges with Complex Backgrounds

In addition to object size, background complexity also plays a significant role in the model’s performance. Species like Wallace’s Standardwing, which blend with their natural environment (e.g., dense forests), were detected with lower accuracy (AP = 30.82%). The model struggles to differentiate objects that have similar textures or colors to the surrounding background, especially in natural habitats with dense vegetation.

Figure 12. Example of detection in a complex background (Wallace’s Standardwing)

4.3. Variation in Performance Across Species

The model’s performance varied significantly across different species. Species with distinct and recognizable physical features, such as the Komodo Dragon and Javan Green Peafowl, showed higher AP values. Conversely, species with smaller sizes and subtler forms, such as Orangutans and Pygmy Seahorses, resulted in lower AP values. The morphology of these species impacts the mod-el’s ability to detect and distinguish them accurately.

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

Figure 13. Example of detection on a smaller species (Pygmy Seahorse) with lower accuracy

4.4. Challenges in Detection and Classification

The model's performance varied across species, with those having distinct features, like the Komodo Dragon and Javan Green Peafowl, showing higher AP values. Smaller species, such as Orangutans and Pygmy Seahorses, had lower AP values due to their subtle features. However, even species with clear physical traits faced detection challenges.

Specifically, the Javan Green Peafowl struggled due to its tendency to blend into complex backgrounds, particularly in dense vegetation, leading to missed detections. Similarly, the Komodo Dragon showed misclassifications in images with varying poses or backgrounds. These issues highlight the model's sensitivity to environmental factors, such as lighting, background complexity, and subject angles.

To address these challenges, further improvements in the dataset, including more varied environmental conditions and poses, and improving the ability of the model to manage complex backgrounds, are crucial for better detection accuracy.

Figure 14. Example where the Banggai Cardinalfish is correctly detected, but nearby coral reefs and other fish are incorrectly classified as Javan Green Peafowl and Komodo Dragon, showing the model's misclassification of the surrounding objects.

4.5. Improving Model Performance

Findings from the evaluation show that the model works effectively for detecting endangered species, particularly those with distinct physical characteristics and larger sizes. However, challenges remain in accurately detecting smaller species and those camouflaged within complex backgrounds. A number of strategies are suggested in order to overcome these constraints and improve the model's capability. First, expanding the dataset to include more images of species with lower average precision (AP), incorporating variations in pose, background, and lighting conditions, can improve the model's generalization ability, as supported by studies [18] and [19]. Second, applying diverse data augmentation techniques can further bolster the model's robustness, especially for small and underrepresented classes, as demonstrated by research in [20] and [21]. Finally, fine-tuning the model through transfer learning

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

from more advanced architectures like Mask R-CNN and YOLOv8 can lead to improved detection accuracy, especially in difficult situations with small or cluttered objects, a finding corroborated by studies [22] and [23].

5. CONCLUSION

This study presents the effectiveness of the Faster R-CNN model with a ResNet-101 backbone for detecting endangered species, particularly those with larger sizes and distinct physical features. The evaluation results show that the model performs well in detecting larger species such as the Komodo Dragon and Green Sea Turtle, achieving high average precision (AP) scores. However, the model faces challenges in detecting smaller species and those blending into complex backgrounds, as evidenced by the lower AP values for species like Orangutans and Wallace's Standardwing.

A number of techniques, such as dataset augmentation, dataset enhancement, and model fine-tuning through transfer learning from more sophisticated models, are suggested to address these issues and enhance the model's performance. It is anticipated that these enhancements will boost the model's ability to identify smaller and more camouflaged species, ultimately making it a more reliable instrument for conservation and biodiversity monitoring.

Future research should concentrate on utilizing more varied images to the dataset, exploring more advanced object detection architectures, and integrating additional environmental data to further enhance model reliability and performance in real-world applications. With these advancements, the model can contribute to more effective monitoring and protection of endangered species in their natural habitats. The approach presented herein can be integrated into future intelligent biodiversity monitoring platforms.

REFERENCES

[1] I. Ardiantiono, I. M. R. Pinondang, D. S. Chandradewi, G. Semiadi, F. Pattiselanno, J. Supriatna, J. S. Tasirin, N. L. Winarni, M. Voigt, J. W. Bull, T. Humle, N. J. Deere, and M. J. Struebig, "Insights from 20 years of mammal population research in Indonesia," Oryx, vol. 58, no. 4, pp. 485–492, Jul. 2024. DOI: 10.1017/S0030605323001539.

[2] Y. Purwanto, E. Sukara, P. S. Ajiningrum, and D. Priatna, "Cultural diversity and biodiversity as foundation of sustainable development," Indonesian Journal of Applied Environmental Science and Technology (InJAST), vol. 2, no. 2, pp. 90–97, 2021. [Online]. Available: https://journal.unpak.ac.id/index.php/InJAST/article/view/1976

[3] A. Eryan, "Review on Illegal Wildlife Trade Provisions in Indonesia: Cost-Benefit Analysis and Law Enforcement," Indonesian Journal of International Law, vol. 21, no. 5, Art. 4, 2024. DOI: 10.17304/ijil.vol21.5.1877. [Online]. Available: https://scholarhub.ui.ac.id/ijil/vol21/iss5/4

[4] D. Velasco-Montero, J. Fernández-Berni, R. Carmona-Galán, A. Sanglas, and F. Palomares, "Reliable and efficient integration of AI into camera traps for smart wildlife monitoring based on continual learning," Ecological Informatics, vol. 83, p. 102815, 2024. DOI: 10.1016/j.ecoinf.2024.102815.

[5] L. Jeantet and E. Dufourq, "Improving deep learning acoustic classifiers with contextual information for wildlife monitoring," Ecological Informatics, vol. 77, p. 102256, 2023, doi: 10.1016/j.ecoinf.2023.102256.

[6] F. Gao, T. Liu, H. Wang, H. Shi, C. Yuan, S. Song, B. HaSi, and X. Wu, "Spatial and temporal variation patterns of summer grazing trajectories of Sunit sheep," Ecological Informatics, vol. 78, p. 102322, 2023, doi: 10.1016/j.ecoinf.2023.102322.

[7] A. Magdy, M. S. Moustafa, H. M. Ebied, and M. F. Tolba, "Lightweight Faster R-CNN for object detection in optical remote sensing images," Scientific Reports, vol. 15, no. 1, p. 16163, May 2025. doi: 10.1038/s41598-025-99242-y.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," arXiv preprint arXiv:1506.01497, 2015, doi:

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

10.48550/arXiv.1506.01497.
[9] X. Wang, Y. Zhang, and Z. Li, "Deep learning-based object detection for endangered species in satellite imagery," Applied Sciences, vol. 14, no. 11, p. 4443, 2024. [Online]. Available: https://doi.org/10.3390/app14114443.
[10] Y. Zhang, D. Liu, and Q. He, "Remote sensing and machine learning for monitoring biodiversity," Remote Sensing, vol. 16, no. 8, p. 1350, 2023. [Online]. Available: https://doi.org/10.3390/rs16081350.
[11] M. Liu, Y. Chen, and X. Zhao, "Spatial-temporal features for object detection in complex environments," International Journal of Applied Geospatial Research, vol. 28, no. 2, pp. 103–110, 2024. [Online]. Available: https://doi.org/10.1016/j.jag.2024.103732.
[12] Y. Hu, J. Wang, and Z. Zhang, "Biodiversity monitoring using high-resolution aerial imagery and deep learning models," Ecological Informatics, vol. 23, pp. 102–115, 2023. [Online]. Available: https://doi.org/10.1016/j.ecoinf.2023.102383.
[13] R. Jones, S. Smith, and L. Thomas, "Agricultural impact on endangered species distribution and the integration of environmental data," Agronomy, vol. 14, no. 10, p. 2194, 2024. [Online]. Available: https://doi.org/10.3390/agronomy14102194.
[14] H. Li, W. Zhang, and S. Liu, "Advanced object detection for endangered species in remote sensing data using Faster R-CNN," Entropy, vol. 24, no. 3, p. 353, 2023. [Online]. Available: https://doi.org/10.3390/e24030353.
[15] iNaturalist, "Endangered Species Observations for Place in Indonesia," accessed May 27, 2025. [Online]. Available: https://www.inaturalist.org/observations?place_id=6966&view=species.
[16] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll´ar, and C. L. Zitnick, “Microsoft COCO: Common objects in context,” in European Conference on Computer Vision (ECCV), Springer, 2014, pp. 740–755.
[17] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, “Detectron2,” Facebook AI Research, 2019. [Online]. Available: https://github.com/facebookresearch/detectron2.
[18] E. Karakullukcu, "Leveraging convolutional neural networks for image-based classification of feature matrix data," Expert Systems with Applications, vol. 281, p. 127625, 2025, doi: 10.1016/j.eswa.2025.127625.
[19] R. Shinoda and K. Shiohara, "PetFace: A Large-Scale Dataset and Benchmark for Animal Identification," in Computer Vision – ECCV 2024, Springer Nature Switzerland, Cham, 2025, pp. 19–36, doi: 10.1007/978-3-031-72649-1_2.
[20] Y. Fang, Y. Liu, Y. Liu, and Q. Yang, "Data Augmentation for Object Detection via Controllable Diffusion Models," in Proc. IEEE Winter Conf. on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2024. [Online]. Available: https://openaccess.thecvf.com/content/WACV2024/papers/
[21] J. Wu, Y. Chao, and J. Liu, "Enhancing Multi-modal Object Detection with Data Augmentation, Focal Loss, and Model Ensembling," in Pattern Recognition. ICPR 2024 International Workshops and Challenges, vol. 15617, Cham: Springer, 2025. doi: 10.1007/978-3-031-88217-3_21. [Online]. Available: https://doi.org/10.1007/978-3-031-88217-3_21
[22] H. Fujita, M. Itagaki, K. Ichikawa, Y. K. Hooi, K. Kawahara and A. Sarlan, "Fine-tuned Surface Object Detection Applying Pre-trained Mask R-CNN Models," 2020 International Conference on Computational Intelligence (ICCI), Bandar Seri Iskandar, Malaysia, 2020, pp. 17-22, doi: 10.1109/ICCI51257.2020.9247666.
[23] I. Iqra and K. J. Giri, "SO-YOLOv8: A novel deep learning-based approach for small object detection with YOLO beyond COCO," Expert Systems with Applications, vol. 280, p. 127447, 2025, doi: 10.1016/j.eswa.2025.127447. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417425010693.
[24] M. J. Mubarok, E. M. Yuniarno, Y. O. Sihombing and M. H. Purnomo, "Semantic Segmentation of Sentinel-2 Satellite Image for Rice Growth Phase Classification Using Deep Learning," 2024 IEEE International Conference on Imaging Systems and Techniques (IST), Tokyo, Japan, 2024, pp. 1-6, doi: 10.1109/IST63414.2024.10759168.
[25] M. Willi, R. T. Pitman, A. W. Cardoso, C. Locke, A. Swanson, A. Boyer, M. Veldthuis, and L. Fortson, "Identifying animal species in camera trap images using deep learning and citizen

Vol. 6, No. 6, December 2025, Page. 5435-5449
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.4793

science," Methods in Ecology and Evolution, vol. 10, no. 1, pp. 80–91, 2019.
[26] BirdLife International, "Leucopsar rothschildi,"The IUCN Red List of Threatened Species, 2021, e.T22710912A183006359. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/22710912/183006359
[27] G. R. Allen and T. J. Donaldson, "Pterapogon kauderni (amended version of 2007 assessment)," The IUCN Red List of Threatened Species, 2025, e.T63572A274966962. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/63572/274966962
[28] J. A. Seminoff, "Chelonia mydas (amended version of 2004 assessment)," The IUCN Red List of Threatened Species, 2023, e.T4615A247654386. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/4615/247654386
[29] BirdLife International, "Pavo muticus," The IUCN Red List of Threatened Species, 2018, e.T22679440A131749282. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/22679440/131749282
[30] T. Jessop, A. Ariefiandy, M. Azmi, C. Ciofi, J. Imansyah, and D. Purwandana, "Varanus komodoensis," The IUCN Red List of Threatened Species, 2021, e.T22884A123633058. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/22884/123633058
[31] BirdLife International, "Macrocephalon maleo," The IUCN Red List of Threatened Species, 2021, e.T22678576A194673255. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/22678576/194673255
[32] B. Russell, "Cheilinus undulatus," The IUCN Red List of Threatened Species, 2004, e.T4592A11023949. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/4592/11023949
[33] M. Ancrenaz, M. Gumal, A. J. Marshall, E. Meijaard, S. A. Wich, and S. Husson, "Pongo pygmaeus (amended version of 2023 assessment)," The IUCN Red List of Threatened Species, 2024, e.T17975A259043172. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/17975/259043172
[34] V. Vaniartha and R. Pollom, "Hippocampus bargibanti," The IUCN Red List of Threatened Species, 2024, e.T10060A250508253. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/10060/250508253
[35] BirdLife International, "Semioptera wallacii," The IUCN Red List of Threatened Species, 2017, e.T22706140A118483106. Accessed: Jun. 17, 2025. [Online]. Available: https://www.iucnredlist.org/species/22706140/118483106
[36] M. A. Tabak et al., "Machine learning to classify animal species in camera trap images: applications in ecology," Methods in Ecology and Evolution, vol. 10, no. 4, pp. 585–590, 2019, doi: 10.1111/2041-210X.13120.
[37] T. Fischer, "Label Studio: Data labeling software for machine learning," Journal of Open Source Software, vol. 6, no. 59, p. 2872, 2021, doi: 10.21105/joss.02872.
[38] K. Lin et al., "Microsoft COCO: Common objects in context," in Proc. European Conference on Computer Vision (ECCV), 2014, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48.
[39] L. Ratner et al., "Snorkel: Rapid training data creation with weak supervision," in Proc. VLDB Endowment, vol. 11, no. 3, pp. 269–282, 2017, doi: 10.14778/3157794.3157809.
[40] J. Krippendorff, Content Analysis: An Introduction to Its Methodology, 3rd ed., Sage Publications, 2012, pp. 221–229.