Object detection of the bornean orangutan nests using drone and YOLOv5

Rony Teguh[1], I Made Dwijaya Maleh[1], Abertun Sigit Sahay[1], Muhamad Porkab Pratama[2], Okta Simon[2]
[1]Department of Informatics Engineering, Faculty of Engineering, University of Palangka Raya, Palangka Raya, Indonesia
[2]World Wildlife Fund (WWF) Central Kalimantan, Indonesia

Article Info

Article history:

Received Mar 24, 2023
Revised Aug 3, 2023
Accepted Sep 30, 2023

Keywords:

Deep learning
Drone imagery
Ecology
Object detection
Orangutan nest

ABSTRACT

Object detection methods when applied to ecology and conservation can help to identify and monitor endangered species and their habitats. Using drones for this purpose has become increasingly popular due to their ability to cover large areas quickly and efficiently. In this study, we aim to implement object detection using YOLOv5 to detect orangutan nests in forests. To conduct our experiment, we collect drone imagery under different conditions. We propose to use the original YOLOv5 to implement our model. The detection and monitoring of orangutan nests can help conservationists to identify critical habitats, monitor population, and design effective conservation strategies. Additionally, the use of drones can reduce the need for on-the-ground surveys, which can be time-consuming, expensive, and logistically challenging. In our study proposes a model for detecting orangutan nests in forests using a drone and the YOLOv5. Our model predicted 1,970 training images and 414 labeled orangutan nests, with a precision of 0.973, recall 0.949, accuracy mean average precision (mAP)_0.5 is 0.969, and mAP_0.5:0.95 is 0.630. The model finished 217 epochs in 58 hours and had a high object detection accuracy. The model has a 99.9% accuracy in detecting the number of orangutan nests.

Corresponding Author:

Rony Teguh
Department of Informatics Engineering, Faculty of Engineering, University of Palangka Raya
St. Yos Sudarso, Palangka Raya, Central Kalimantan, Indonesia
Email: ronyteguh@it.upr.ac.id

1. INTRODUCTION

The drone technology has recently advanced, and it is possible that they will be used in a variety of areas, including natural disaster management [1], [2], vegetation mapping [3], [4], livestock management [5], [6], animal detection [7], and conservation [8]. Low-cost technology for drone photography can contribute to the visual collection, recognition, identification, and monitoring of data as well as help researchers better identify the use of aerial views taken from the sky with a high-resolution camera [9]–[11]. Scientists have recently begun to deploy remotely piloted drones with autonomous flight capabilities for the purpose of monitoring animals [12]. The drone provides high-resolution geographic data while allowing for quick and frequent identification and monitoring over small to large spatial extents. A variety of studies on shapes, colors, and angles would be easier to locate and recognize. In this paper, we used drone photography to data train an object detection deep learning model to recognize orangutan nests in the open forest and close forest. In this paper, we employ deep learning techniques to detect objects and count nests of orangutans in forests using aerial drone images.

The Borneo orangutans (*Pongo pygmaeus*) build nests on the tops of trees in the forest. The traditional method of measuring orangutan density is to observe and count the population of orangutan nests along ground transects in separate areas of a homogeneous environment [13]. However, the information obtained from ground transects is limited to orangutan activity within a narrow band of habitat, which is constrained by the horizontal distance at which an observer can detect a nest beneath the forest cover unless numerous surveys are conducted across a large area [14], [15]. Small forest fragments and the presence of various land-use types in human-modified landscapes can also result in a complex mosaic of habitats that are challenging to study using ground-transect methods. Aerial surveys allow us to count and identify orangutan nests, which allows us to be sampling the orangutan populations. Combination artificial intelligence technique and computer vision are applied to this work. Finally, our goals were to use deep learning to recognize drone imagery for orangutan nest surveys and to automatically count the number of nests of orangutans. Object detection using drone imagery [16]–[19] and YOLOv5 [20] algorithm vision using open-source artificial intelligence should be further developed and made necessary in wildlife conservation in various natural forest environments.

The YOLOv5 [21]–[23] is based on the YOLO detection architecture and employs a few convolutional neural network optimization techniques [24], including auto learning bounding box anchors, mosaic data augmentation, and the cross-stage partial network. The YOLO model was the first object detector to link the method of predicting bounding boxes with class labels in an end-to-end differentiable network as compared to preceding methods. The three main parts of the YOLOv5 network are the backbone, neck, and head [25]–[27]. Data preprocessing operations are first carried out by the input terminal, including mosaic data augmentation and adaptive image filling. YOLOv5 incorporates adaptive anchor frame calculation on the input to enable it to adjust to various datasets. As a result, when the dataset changes, the initial anchor frame size is automatically set.

A convolutional neural network serves as the framework, gathering and forming image features at various granularities. It primarily employs multiple convolutions and pooling to extract feature maps of various sizes from the input image using a cross-stage partial network (CSP) and spatial pyramid pooling (SPP). The SPP structure enables the feature extraction from different scales for the same feature map and can generate three scale feature maps, which helps improve the detection accuracy. In contrast, the Bottleneck CSP architecture uses less calculation and accelerates inference. To combine and mix image features and pass them on to the prediction, the neck neural network represents several layers. The feature pyramid network (FPN) and the path aggregation network (PAN) feature pyramid structures are applied in the neck network. Strong semantic features are transmitted from the top feature maps to the lower feature maps by the FPN structure. In addition, the PAN structure transfers robust localization features from lower to higher feature maps. Together, these two structures reinforce the quality. This paper discusses how to deploy datasets models for orangutan nests and successfully detect them in situations where distinguishing between different objects like trees, nests, and orangutan individuals. We build the new model and test the orangutan nest model using the original YOLOv5 for detection and counting the spread of orangutan individuals and groups in landscape swamp peat forest.

2. METHOD

2.1. Field data collection

The build of dataset used in this study was obtained from aerial photographs taken by a drone camera. The photographs were captured in a forest area and included images of orangutan nests. The camera installed on the drone had a resolution of 12.1 megapixel (MP) and captured images with a size of 6.000×4.000 pixels. The camera was set to take pictures at 2-second intervals, and the drone was equipped with an internal global position system (GPS) to record and marker of the position and altitude information. This study used a straight-line transect survey method [28], [29], which involved analyzing drone images with GPS object location information. Straight-line flight missions have been managed to fly by the drone, covering target forest at an average altitude of 90 and 120 meters above ground level (AGL). Using a deep learning method, sample aerial images obtained during these missions were analyzed to identify the presence of orangutan nests. The analysis was also facilitated by orangutan conservation experts in recognizing the objects in the images. Of the 30 data collection flight, 141 were made to detect individual and group of the orangutan nests. The drones flight mission has three sampling drone survey as shown in Figure 1.

The conventional method for determining orangutan density involves ground transects in different areas of a homogeneous habitat, where orangutan nests are observed. However, this approach is time-consuming, and expensive, and surveys of large geographic areas are not conducted frequently. Therefore, this study aimed to develop an efficient nest object detection system using imagery. The study utilized a collection of aerial images to describe the object of orangutan and nest in Central Kalimantan, Indonesia. In Figure 1, a drone is shown flying straight-line transect missions above a forested area. The drone is used to conduct aerial surveys to detect the existence of orangutan nests and analyze the canopy structure. The study investigated aerial images obtained during these missions to identify the being of orangutan nests.

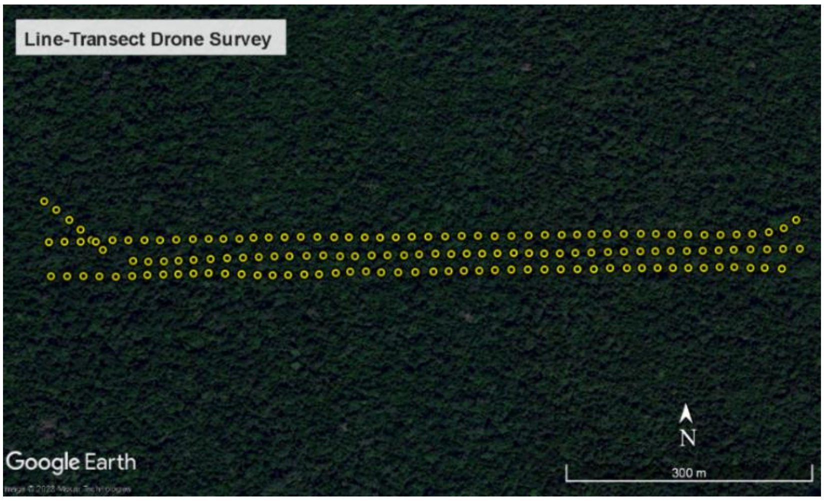

Figure 1. Shows the line transects flight mission using a drone

The dataset for this study were collected in a forest using a Canon Powershoot S100 camera with a resolution of 12.1 megapixel. The images were geotagged using the internal GPS of the camera to marker and record of the position of the object. To reduce ground inaccuracy in GPS locations, photogrammetry software was used to calculate the optimal camera positions. The camera was mounted horizontally in the drone fuselage with its top facing forward, and no filters were used. Photographs were taken every 2-seconds with an exposure time of 1/500 second, an international standardization organization (ISO) speed of 100, and an aperture f-stop of f/2.8 and auto-white balance (AWB) is automatics. In Figure 2 shows a flowchart for object detection and count individual of orangutan nests using YOLOv5 method. The dataset used in this study consists of 1,970 drone images, out of which 141 images were identified by expert orangutan conservation as containing orangutan nests on top of trees. The next phase involves labeling the object of the orangutan nest in the images. The dataset is divided into 70% for 288 train data images containing orangutan nests, 20% for 85 images for data validation, and 10% for 41 images for test data images.

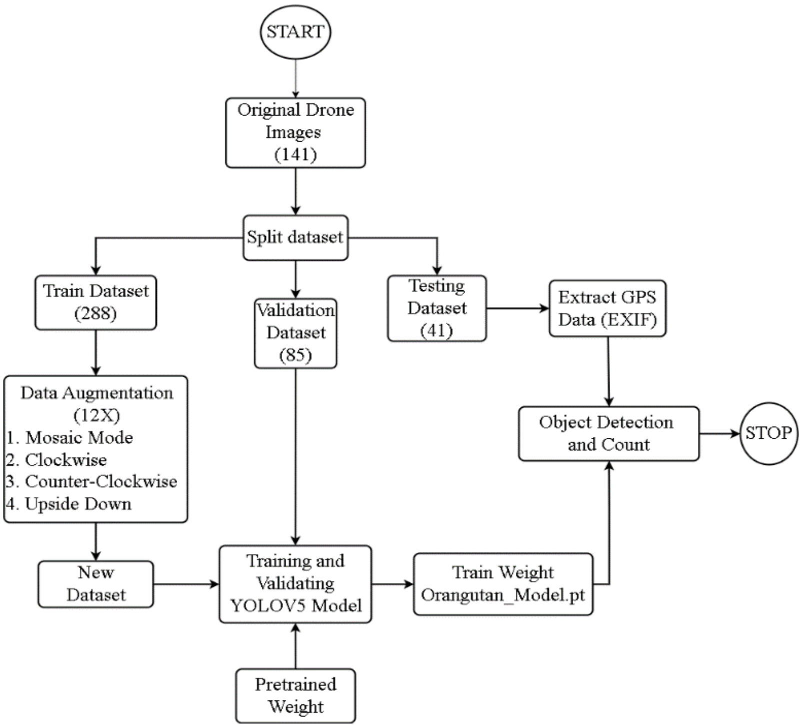

Figure 2. Shows the workflow object detection and count of orangutan nests

2.2. Nests identification and labelling

A total of 1,970 images were collected during line transect drone surveys, and each one had to be manually inspected for orangutan nests. Making a working method for this process took up the majority of the first few months of analysis time, which also included a lot of time spent trying out various approaches and working out any kinks. Using quantum geographic information system (QGIS) software, we will compare our findings to traditional nest counts once we have developed a method for identifying and tagging nests in the imagery. We overlay a grid on a drone image using the GNU image manipulation program (GIMP) open-source software, scroll through it in search of orangutan nests, and then repeat the process. Our study uses drone images and deep learning methods for counting and detecting nests of orangutans has never been done previously by reseacher.

In the line-transects survey, the drone dataset consisted of 1,970 images, and potential orangutan nesting areas were identified in each of these photos. During the labeling process, assistance was received from orangutan conservation experts for the identification of the object of the nests and the individual of the orangutan. We are identification of objects with 141 photos using data patterns discovered about orangutan nests positions on tree branches and the top. To add a label to an image, an object label annotation is given, along with the bounding box and the class name of the object. The object label box can only determine the position if the location of the orangutan nest in the image is known. In the Figure 3 illustrates the process of labeling objects in sample drone images using the open-source software LabelImg [17]. On each object in the image, a box is drawn to show the class label area and separate it from other objects such as trees and leaves.

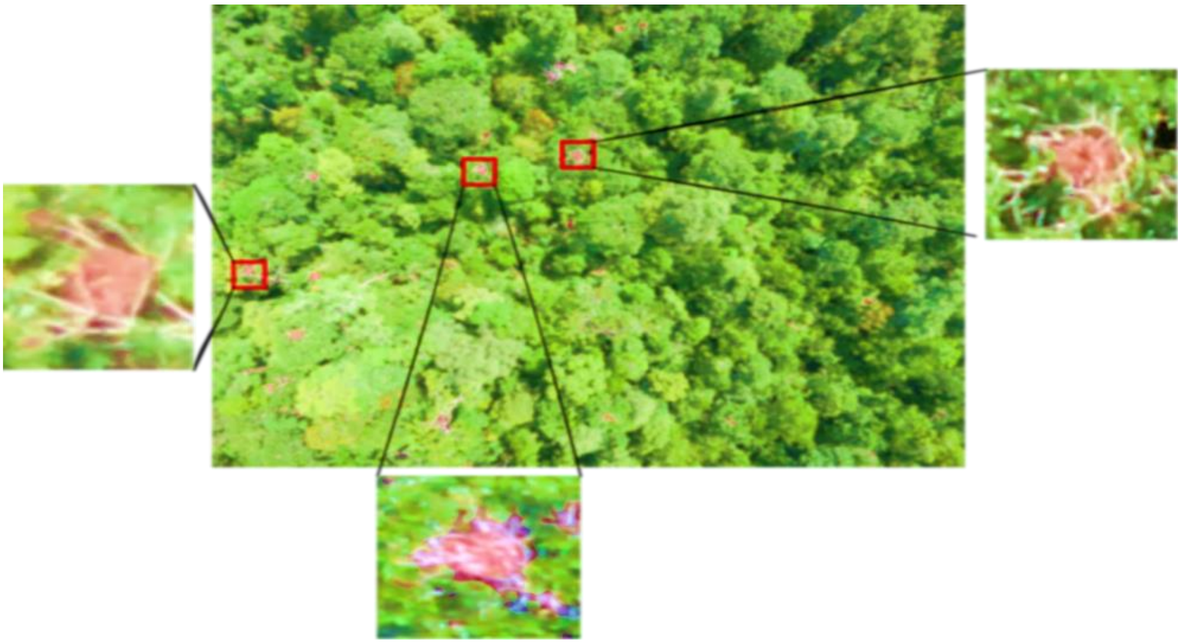

Figure 3. Show the image identification and labeling for orangutan nests

2.3. Data augmentation

In this paper, we use data augmentation as a technique for generating new data samples by making characteristic changes to photos of orangutan nests. To increase the quantity of picture data, using the augmentation technique is a data processing that adds modified photos of current data or newly created synthetic data from existing data. To detection of small objects in this study problem in which attempt to correctly identify small objects in an image. The object of the nest and the orangutan is a small object and need to be able to correctly identify objects even if each individual object is small compared to the photo size. Some especially useful augmentations for small object detection in our model include brightness, blur, noise, mosaic, upside-down, and exposure.

In this study, we are recognizing and analysis the small objects of the nest and individual orangutans with 141 photos. in Figure 4(a) data augmentation generates new images from original dataset for image transformation and rotates clockwise, counterclockwise, and upside down. The original image to transformed into a new image by rotate an image 90 degrees or 180 degrees. In Figure 4(b) every original image to transformed into new data by rotating an image to 45 and -45 degrees. In Figure 4(c), at the time pictures are taken by a drone camera, the results are out of focus. We generate images to append data in blurry conditions. The results of images that are less sharp, and objects that are vague or unclear. Technically, a blurry photo or image is an image with a pixel density with a low resolution. When the images are bright or dark because of changes in sunlight as the camera captures the object. In Figure 4(d) shows how, each original image is transformed into new data by gamma exposing an image to make it brighter or darker to 31% bright and -31% dark. Noise is an obnoxious color spot/grain in a photograph caused by the limits of the electronic sensor system in a digital camera. The presence of spots/grains because of this noise effect is said to be particularly distracting to the view. When shooting at night in auto mode on a digital camera, the camera's response will often force the ISO value higher. However, the higher ISO value will result in greater noise in the photograph. In Figure 4(e) shows how, each original image injects random noise into an image as 8%.

In this paper, we use of Roboflow [18] for image data augmentation to modify image data as custom datasets. The process of modifying a photograph to create variations of orangutan nest photos by employing augmentation techniques including as mosaicking, scaling, and clockwise, counterclockwise, and upside-down processing. We use this technique when a drone is capturing images in a variety of different conditions such as weather, temperature, lighting, and motion. The results produce that by enhancing the original data from the training dataset by 12 times, we have 3,430 new images generated from 288 sample original drone images, as shown in Figure 4.

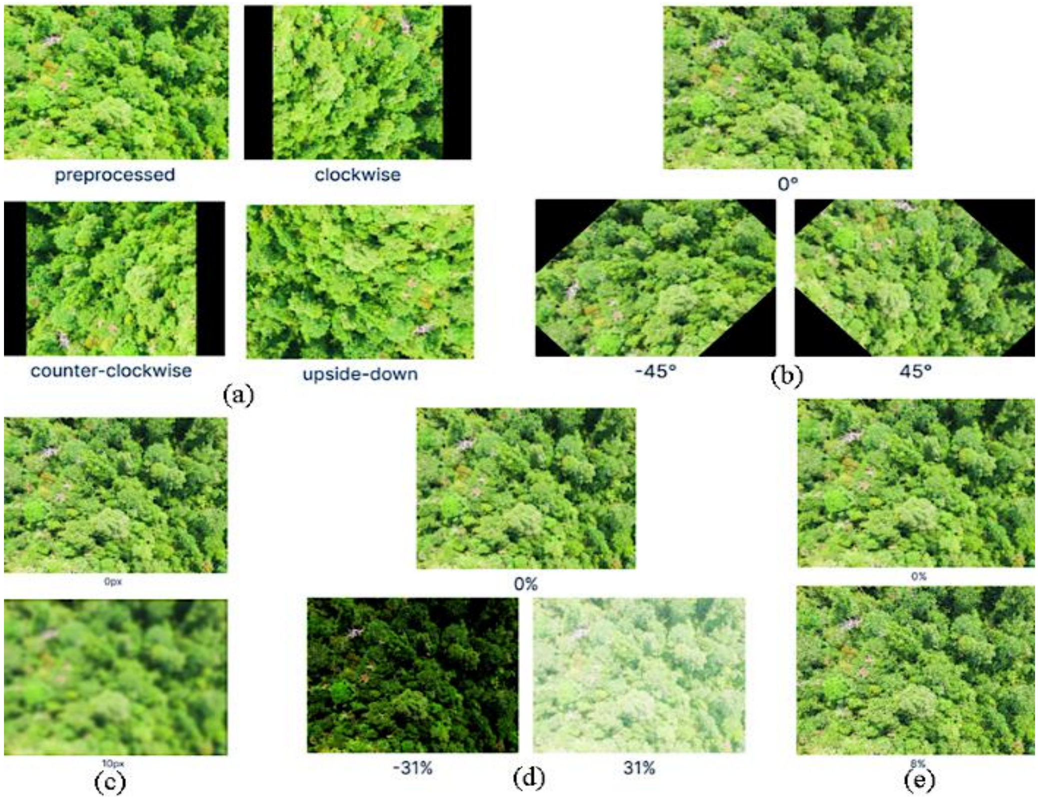

Figure 4. Shows of the image data augmentation using image transformations: (a) clockwise, counterclockwise, upside-down; (b) rotation; (c) blur; (d) exposure; and (e) noise

2.4. Dataset training

The Roboflow was used in this study for model dataset training and augmentation processing. The dataset contained 1,970 photos, 414 of which were shot by a drone in the forest and showed orangutan nests. The data was separated into three subsets: 70% for training (288 photos), 20% for validation (85 images), and 10% for testing (41 images). For the dataset training, mosaic, clockwise, counterclockwise, and upside-down approaches were used. In Table 1, the model takes 16 minutes to complete 1 epoch on a dataset with 3,430 train images and 85 validation images, and 58 hours for 217 epochs to produce precision 0.973, recall 0.949 the mean average precision (mAP) is 0.5 0.969 mAP 0.5:0.95 0.630.

Table 1. The performance of model dataset train

Parameter	Best Epoch
Epoch	217
Precision	0.973
Recall	0.949
mAP_0.5	0.969
mAP_0.5:0.95	0.630

2.5. Object detection using YOLOv5

For object detection, we used YOLOv5 original network, which was trained on the pascal visual object classes (VOC) dataset. We chose Yolo because of its speed, but other alternative networks with higher precision can be utilized. YOLOv5 was used to detect every object in each image. The result is a set of bounding boxes (bboxes) around the discovered items. Several bboxes can be associated with the same item. One may relate to the animal's head, while another with other body parts. We utilized a typical non-max suppression technique to merge the overlapping bboxes for the same object. A real-time one-stage object detector YOLOv5, which has a quicker inference time and higher detection accuracy, is perfectly suited to our needs. YOLO has developed into one of the most efficient methods for recognizing objects in both Microsoft COCO datasets and Pascal VOC due to its developer's dedication to enhancing the technology (visual object classes). The benchmark YOLOv5l, the expanded YOLOv5x, and the streamlined preset models YOLOv5s and YOLOv5m make up YOLOv5. The main difference between both sorts of networks is the number of feature-extraction modules and convolution kernels present at various network nodes, with a consequent reduction in model sizes and parameter counts. It employs a single forward CNN to determine the object's class and position using standard CNN technology. In Figure 5 displays the detecting process from the input image to the results. This study main goal is to put a YOLOv5-based framework into practice. It is a brand-new Convolutional Neural Network (CNN) that exhibits higher object detection precision. In this method, the entire image is evaluated using a single neural network, after which it is divided into pieces and probabilities and bounding boxes are predicted for each component. These bounding boxes are weighted according to the estimated likelihood.

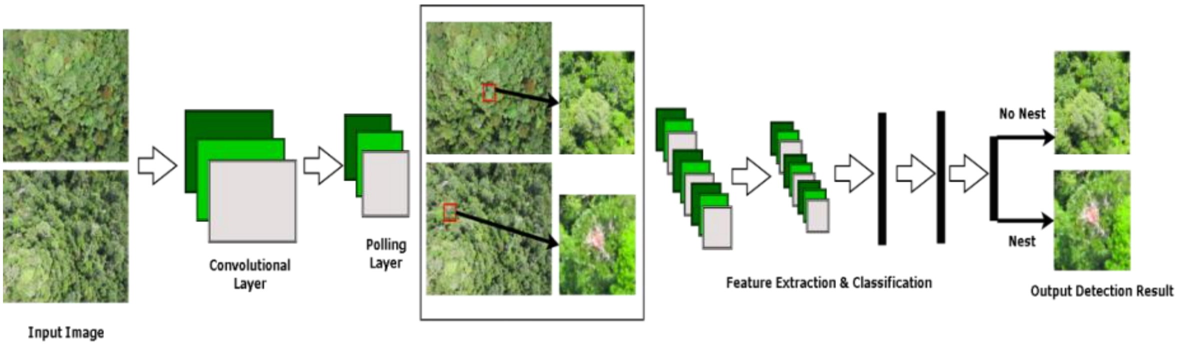

Figure 5. The typical operations of a convolutional neural network CNN are utilized throughout, from the input image to the detection results

3. RESULTS AND DISCUSSION

3.1. Orangutan nests detection

We utilized a collective dataset from drone imagery with labeling. This data was collected in April until September 2019 in the forests of Central Kalimantan, Indonesia. To test our model's ability to identify orangutan nests, we chose 24 photos from our datasets with and without orangutan nests. The dataset was trained using 3,430 photos of orangutan nests and tested with 41 image of orangutan nests. Our research suggests that using deep-learning object identification and imaging drones may identify orangutan in a forest. In Figure 6 shows how, we tested drone data to see if it could identify orangutan nest objects in the forest. An evaluation of image detection performance will be performed during the system testing stage to analyze the level of accuracy and trained weights that can be used to identify orangutan nests. The bounding boxes are drawn to encircle an object and represent the likelihood that it is an orangutan nest if one is found. We compare two photos taken by the camera under various lighting conditions. Our model is capable of swiftly identifying prospective nests and counting the number of nests found in an image.

In this paper, the image data photograph to generated image data as new datasets. The process of generated a quality of image to create variations of orangutan nest photos by employing augmentation techniques including in the high and low quality of image. This method is used when a drone is capturing images in a variety of conditions, such as low light and vibration. Initially, image analysis for both nests and trees was done with the help of local field assistance. The color difference and type of structure that distinguished parts of the tree from the surrounding branches were the primary information used to detect nests in the image. The nests discovered on the ground were later identified in additional drone photos based on GPS location. In Figure 6, show the result of object detection using drone and YOLOv5. Figure 6(a) shows that the orangutan nests are not detected. Figure 6(b) shows how our model can recognize individual orangutan nests in the wild. Our approach recognizes and counts orangutan group nests in Figure 6(c).

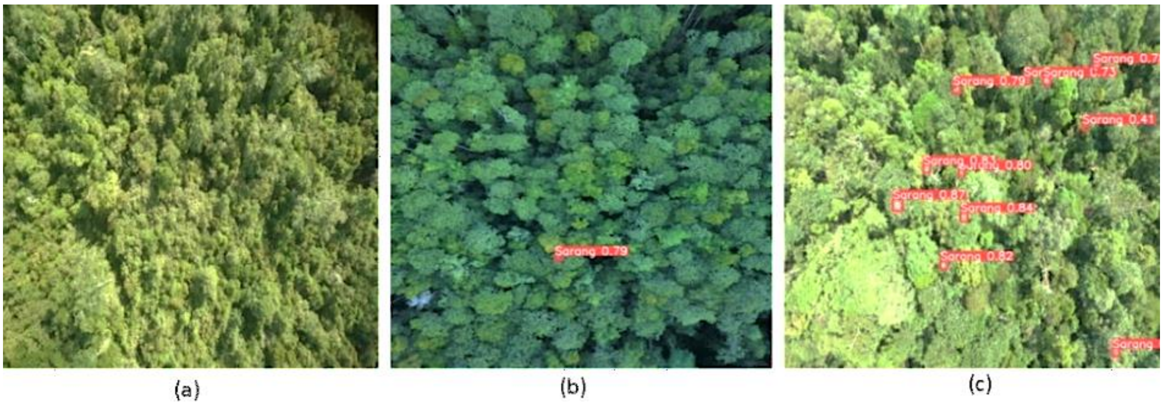

Figure 6. Show the result of object detection for (a) non- nests, (b) individual nests, and (c) group nests

3.2. Orthophoto of drone survey

The information from a drone survey line transects used to determine the presence of orangutan nests is shown in Figure 7. Our system uses the sample drone image data to find two nest objects in the forest. Orangutan nests found between the branches of forest trees can be identified and counted using this system. In the forest, we investigated the potential use of drones to locate orangutan nests and evaluated the effect of picture resolution, seasonality, tree type, nests height, and color on nests detectability. It's only that it can be challenging to pinpoint the location in a single image with multiple nesting items and only one GPS metadata.

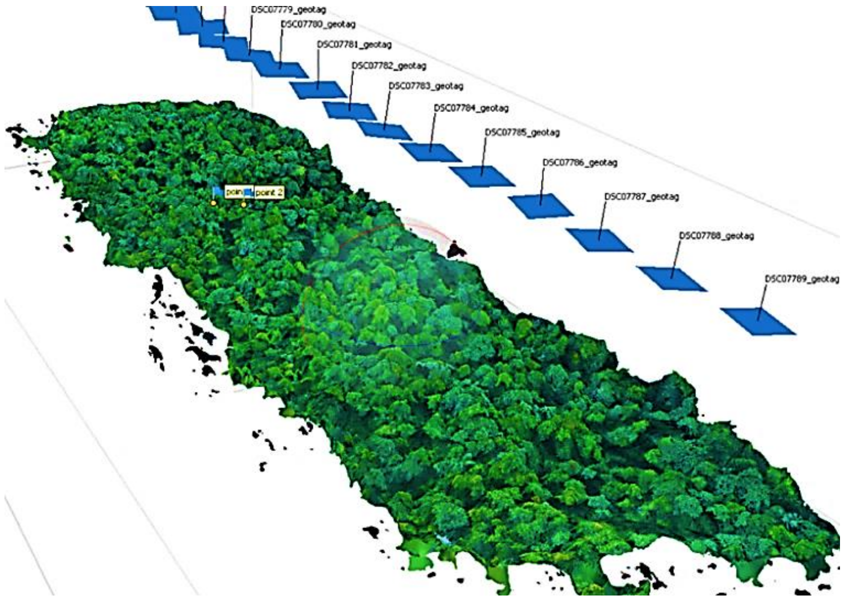

Figure 7. Line transects survey and drone detection of orangutan nest object in orthophoto

3.3. Distribution of orangutan nest

In this study, we used drone data for each line transect plot suspected of containing information on the presence of orangutan nests. We recognize 24 images with the potential for nests from all of these. The results of our model were then compared to the ability of local experts to recognize a nest object around of tree. In Table 2, we used YOLOv5 to detect objects in the forest and identified the findings as nests. The experiments in Table 2 to test our model from the best training weight model discovered that the model had an average accuracy of 99.9% in detecting the number of orangutan nests and that YOLOv5 can recognize orangutan nests in photos with canopy structure in the forest.

Table 2. Object detection and count results by our model

No	File Image	Local Expert	Our Model	No	File Image	Local Expert	Our Model
1	DSC00084_geotag	1	1	13	DSC08403_geotag	1	1
2	DSC00169_geotag	3	3	14	DSC08477_geotag	1	1
3	DSC00177_geotag	1	1	15	DSC09091_geotag	1	1
4	DSC07644_geotag	1	1	16	DSC09455_geotag	2	1
5	DSC07670_geotag	1	1	17	DSC09458_geotag	1	1
6	DSC07672_geotag	3	1	18	DSC09462_geotag	1	1
7	DSC07783_geotag	1	1	19	DSC09545_geotag	1	1
8	DSC07794_geotag	1	1	20	DSC09663_geotag	1	1
9	DSC07804_geotag	1	1	21	DSC09944_geotag	1	2
10	DSC07819_geotag	1	1	22	DSC09944_geotag	1	2
11	DSC07827_geotag	1	1	23	DSC09968_geotag	1	1
12	DSC07960_geotag	1	3	24	DSC09968_geotag	1	1

4. CONCLUSION

Object detection technology can be used to identify, monitor, and conserve endangered species and their habitats. This study proposes a new dataset model for the count and detection of orangutan nests in forests using a drone and the YOLOv5 algorithm. The algorithm is a state-of-the-art object detection algorithm that has been used in various applications such as detecting and tracking people, vehicles, and animals. Implementing this method can have significant implications for the conservation of orangutans and their habitats, as it can help conservationists identify key habitats, monitor population trends, and design effective conservation strategies. Additionally, the use of drones can reduce the need for on-the-ground surveys, which can be time-consuming, expensive, and logistically challenging. Combination drone imagery and deep learning method are applied successfully to this work. Finally, our goals were to use deep learning to recognize drone imagery for orangutan nests and automatically count the number of nests of orangutans in the canopy cover. The application of multi-class for nests, orangutans, and fruit trees in the forest will continue to be studied.

REFERENCES

[1] S. M. S. Mohd Daud *et al.*, "Applications of drone in disaster management: A scoping review," *Science and Justice*, vol. 62, no. 1, pp. 30–42, Jan. 2022, doi: 10.1016/j.scijus.2021.11.002.

[2] S. Ivanova, A. Prosekov, and A. Kaledin, "A survey on monitoring of wild animals during fires using drones," *Fire*, vol. 5, no. 3, p. 60, Apr. 2022, doi: 10.3390/fire5030060.

[3] S. M. Hamylton *et al.*, "Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches," *International Journal of Applied Earth Observation and Geoinformation*, vol. 89, p. 102085, Jul. 2020, doi: 10.1016/j.jag.2020.102085.

[4] S. Ecke *et al.*, "UAV-based forest health monitoring: a systematic review," *Remote Sensing*, vol. 14, no. 13, p. 3205, Jul. 2022, doi: 10.3390/rs14133205.

[5] M. A. Alanezi, M. S. Shahriar, M. B. Hasan, S. Ahmed, Y. A. Sha'aban, and H. R. E. H. Bouchekara, "Livestock management with unmanned aerial vehicles: a review," *IEEE Access*, vol. 10, pp. 45001–45028, 2022, doi: 10.1109/ACCESS.2022.3168295.

[6] S. J. Hong, Y. Han, S. Y. Kim, A. Y. Lee, and G. Kim, "Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery," *Sensors (Switzerland)*, vol. 19, no. 7, p. 1651, Apr. 2019, doi: 10.3390/s19071651.

[7] P. C. Gray *et al.*, "A convolutional neural network for detecting sea turtles in drone imagery," *Methods in Ecology and Evolution*, vol. 10, no. 3, pp. 345–355, Jan. 2019, doi: 10.1111/2041-210X.13132.

[8] J. J. López and M. Mulero-Pázmány, "Drones for conservation in protected areas: Present and future," *Drones*, vol. 3, no. 1, pp. 1–23, Jan. 2019, doi: 10.3390/drones3010010.

[9] J. Valente, B. Sari, L. Kooistra, H. Kramer, and S. Mücher, "Automated crop plant counting from very high-resolution aerial imagery," *Precision Agriculture*, vol. 21, no. 6, pp. 1366–1384, May 2020, doi: 10.1007/s11119-020-09725-3.

[10] Z. Sun *et al.*, "Pine wilt disease detection in high-resolution UAV images using object-oriented classification," *Journal of Forestry Research*, vol. 33, no. 4, pp. 1377–1389, Nov. 2022, doi: 10.1007/s11676-021-01420-x.

[11] M. F. Albaghdadi and M. E. Manaa, "Unmanned aerial vehicles and machine learning for detecting objects in real time," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 6, pp. 3490–3497, Dec. 2022, doi: 10.11591/eei.v11i6.4185.

[12] N. Bonnin, A. C. Van Andel, J. T. Kerby, A. K. Piel, L. Pintea, and S. A. Wich, “Assessment of chimpanzee nest detectability in drone-acquired images,” *Drones*, vol. 2, no. 2, pp. 1–17, Apr. 2018, doi: 10.3390/drones2020017.
[13] S. Milne *et al.*, “Drivers of bornean orangutan distribution across a multiple-use tropical landscape,” *Remote Sensing*, vol. 13, no. 3, pp. 1–16, Jan. 2021, doi: 10.3390/rs13030458.
[14] A. E. Johnson, C. D. Knott, B. Pamungkas, M. Pasaribu, and A. J. Marshall, “A survey of the orangutan (Pongo pygmaeus wurmbii) population in and around Gunung Palung National Park, West Kalimantan, Indonesia based on nest counts,” *Biological Conservation*, vol. 121, no. 4, pp. 495–507, Feb. 2005, doi: 10.1016/j.biocon.2004.06.002.
[15] M. Ancrenaz *et al.*, “Aerial surveys give new estimates for orangutans in Sabah, Malaysia,” *PLoS Biology*, vol. 3, no. 1, p. e3, Dec. 2005, doi: 10.1371/journal.pbio.0030003.
[16] R. Walambe, A. Marathe, and K. Kotecha, “Multiscale object detection from drone imagery using ensemble transfer learning,” *Drones*, vol. 5, no. 3, p. 66, Jul. 2021, doi: 10.3390/drones5030066.
[17] O. Sahin and S. Ozer, “YOLODrone: Improved YOLO architecture for object detection in drone images,” in *2021 44th International Conference on Telecommunications and Signal Processing (TSP)*, Jul. 2021, pp. 361–365, doi: 10.1109/TSP52935.2021.9522653.
[18] C. Chen *et al.*, “YOLO-Based UAV Technology: a review of the research and its applications,” *Drones*, vol. 7, no. 3, p. 190, Mar. 2023, doi: 10.3390/drones7030190.
[19] A. Hanafi, L. Elaachak, and M. Bouhorma, “Machine learning based augmented reality for improved learning application through object detection algorithms,” *International Journal of Electrical and Computer Engineering*, vol. 13, no. 2, pp. 1724–1733, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1724-1733.
[20] L. Xie, Y. Xue, and J. Ye, “UAV aerial photography target detection algorithm based on improved YOLOv5,” *Journal of Physics: Conference Series*, vol. 2284, no. 1, p. 12024, Jun. 2022, doi: 10.1088/1742-6596/2284/1/012024.
[21] F. Jubayer *et al.*, “Detection of mold on the food surface using YOLOv5,” *Current Research in Food Science*, vol. 4, pp. 724–728, 2021, doi: 10.1016/j.crfs.2021.10.003.
[22] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A review of yolo algorithm developments,” *Procedia Computer Science*, vol. 199, pp. 1066–1073, 2021, doi: 10.1016/j.procs.2022.01.135.
[23] S. M. Abas, A. M. Abdulazeez, and D. Q. Zeebaree, “A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 200–213, Jan. 2022, doi: 10.11591/ijeecs.v25.i1.pp200-213.
[24] T. Diwan, G. Anirudh, and J. V. Tembhurne, “Object detection using YOLO: challenges, architectural successors, datasets and applications,” *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, Aug. 2023, doi: 10.1007/s11042-022-13644-y.
[25] P. Mittal, R. Singh, and A. Sharma, “Deep learning-based object detection in low-altitude UAV datasets: A survey,” *Image and Vision Computing*, vol. 104, p. 104046, Dec. 2020, doi: 10.1016/j.imavis.2020.104046.
[26] S. Li, Y. Li, Y. Li, M. Li, and X. Xu, “YOLO-FIRI: Improved YOLOv5 for infrared image object detection,” *IEEE Access*, vol. 9, pp. 141861–141875, 2021, doi: 10.1109/ACCESS.2021.3120870.
[27] M. Boukabous and M. Azizi, “Image and video-based crime prediction using object detection and deep learning,” *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 1630–1638, Jun. 2023, doi: 10.11591/eei.v12i3.5157.
[28] L. Silveira, A. T. A. Jácomo, and J. A. F. Diniz-Filho, “Camera trap, line transect census and track surveys: A comparative evaluation,” *Biological Conservation*, vol. 114, no. 3, pp. 351–355, Dec. 2003, doi: 10.1016/S0006-3207(03)00063-6.
[29] F. Braga-Pereira *et al.*, “Congruence of local ecological knowledge (LEK)-based methods and line-transect surveys in estimating wildlife abundance in tropical forests,” *Methods in Ecology and Evolution*, vol. 13, no. 3, pp. 743–756, Mar. 2022, doi: 10.1111/2041-210X.13773.

BIOGRAPHIES OF AUTHORS

Rony Teguh received the M. Eng. degree in Information Technology from the Gadjah Mada University, Indonesia, in 2010 and 2014, respectively. He is currently pursuing a Ph.D. degree in System Science and Informatics from the Graduate School of Information Science and Technology, Hokkaido University, Japan. His current research interests include Data Science, Wireless Sensor Networks, and Artificial Intelligence. He has some collaborations with other (institutions): BRIN, WWF Indonesia, BRGM, CIFOR, USAID, JICA, Hokkaido University, Melbourne University, and Michigan State University. He can be contacted at email: ronyteguh@it.upr.ac.id.

I Made Dwijaya Maleh received a bachelor’s degree in informatics engineering from the University Palangka Raya, Indonesia. His current research interests include Artificial Intelligence and Computer Vision. He can be contacted at email: madej250@gmail.com.

Abertun Sagit Sahay received the M. Eng. degree in Information Technology from the Gadjah Mada University, Indonesia, in 2010 and 2012, respectively. His current research interests Information Science. He can be contacted at email: abertun@gmail.com.

Muhammad Porkab Pratama Experienced Geographic Information System Officer with a demonstrated history of working in the environmental services industry. Skilled in ArcGIS, Spatial Analysis, Ecology, Microsoft Office, and Geographic Information Systems (GIS). Strong information technology professional with a bachelor's degree focused on Forestry from Universitas Palangka Raya (UPR). He can be contacted at email: mpratama@wwf.id

Okta Simon received a bachelor's degree in Forestry from Palangka Raya University (UPR), Central Kalimantan, Indonesia. His current research interests include conservation and ecology. He can be contacted at email: osimon@wwf.id