# Mangrove Tree Density Detector Using YOLO Based On Darknet Framework

**Ilyas Yudhistira Kurniawan[1], M. Udin Harun Al Rasyid[1] and Sritrusta Sukaridhoto[2]**

[1]*Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, Indonesia*
[2]*Departement of Multimedia Creative, Politeknik Elektronika Negeri Surabaya, Indonesia*

*E-mail address: ilyasofficial1617@gmail.com, udinharun@pens.ac.id, dhoto@pens.ac.id*

**Abstract:** Mangroves are important trees that live in swamp areas. Many mangroves are damaged and need preservation. Monitoring mangroves is important for preservation, but it is difficult access by ground and requires a lot of effort and time. There is currently no system specifically to monitor mangrove forest. Therefore this research created mangrove forest density health system. The system gather data using drone to easily gather data from the forest. Then mangrove tree is being detected using YOLO object detector. Experiment shows that YOLO object detector is able to detect mangrove tree accurately with 95% recall, 88.3% IoU, and 22ms processing time. Then the system calculates the density for mangrove forest health. With this system, environmental surveys and monitoring for mangrove forest can be conducted, and resulting density data can be used for effective preservation action.

**Keywords:** Remote Sensing, Mangrove Health, YOLO, Object Detection, Density Calculation

## 1. INTRODUCTION

Mangroves are trees that live in swamps and brackish water areas. Mangroves located on the boundary of sea and land, and around river estuaries. In recent decades, from 2010 to 2020, the world had lost about 12700 hectares [1]. These mangroves play a protective role against coastal erosion and seawater intrusion into groundwater reservoirs. Prevention of coastal erosion is helpful for preventing the sinking of residential, agricultural, and tourism land. In addition, mangrove acts as carbon storage and the loss of mangrove forests will release carbon into the ecosystem thus accelerate the global warming rate [2]. These crucial roles would be disrupted if the mangrove is damaged, thus we need to create and preserve ideal conditions for the mangrove forests.

For effective preservation, monitoring data is required. Environmental data is used to help people in resource management, disaster precaution, and rehabilitation efforts [3]. Mangrove forest monitoring is so important that the government also puts a lot of resource into it. Especially in Indonesia, one of the countries that have a lot of coastlines, where the mangrove destruction area is increasing [1]. The government issued criteria for ideal mangrove forest density health [4]. The national research agency (BRIN) is also researching on how to monitor mangrove forest. They've created and published guide on mangrove monitoring [5].

They've also created app that helps mangrove monitoring efforts [6].

The problem is that there is no system for monitoring the health of specifically mangrove forest density, using mangrove criteria. Currently mangrove environmental monitoring require significant manpower [5]. To cover large area, it require a lot of effort and time. Especially for mangrove forests where the ground is mud, swamps, and brackish waters make it hard to traverse. Environment monitoring also needs to be done frequently, which will increase the cost and time needed even more [7][8].

Based on mentioned problems, in this research we propose system for mangrove density health system. In the era of emerging technology, there are more and more systems that can aid or replace human tasks for gathering data. These systems can be autonomous and don't require long time and massive effort compared to manual observation. Mainly there are two ways for computer sensors to gather data: in-situ or remote. In-situ means that it uses sensors that touch the object of interest to gather data, for example, IoT system for environment monitoring system [9][10]. Remote, as in remote sensing, means the system uses sensor that don't directly interact with the object of interest [11]. It uses camera and light sensor from afar that is attached to satellite or drone. Remote sensing is used to gather data about an

TABLE I. Related works comparison

| Parameter | Fricker et al. [16] | Itakura et al. [17] | Luo et al. [18] | Donmez et al. [19] | Mubin et al. [20] | Our Work |
|---|---|---|---|---|---|---|
| Multi-object Detection | Yes | Yes | Yes | Yes | Yes | Yes |
| Dense-object Detection | Yes | No | Yes | No | No | Yes |
| Capture Equipment | LiDAR Drone | Ground 360°Cam | RGB Drone | RGB Drone | RGB Sattelite | RGB Drone |
| Mangrove Detection | No | No | No | No | No | Yes |
| Method | CNN | YOLOv2 | Faster R-CNN | CCL Algorithm | CNN LeNet | YOLOv4 |

object from a challenging and hard-to-access location by flying around the object, which is perfect for mangrove density monitoring because of it eliminate problem of hard to access data. This research system uses drone to gather data inside the mangrove forest.

With the development of sophisticated computer vision, technology system can also help process complex processes that humans previously could only do. Object detection is needed to create a system that is able to detect mangrove tree objects similar to human vision. Then the data is processed further to form data to determine the condition of mangrove forest density. A robust object detector uses a neural network algorithm. YOLO, as an object detection neural network, can be used to detect any custom object[12]. YOLO is well known to be capable thus many use it for object detection in their systems[13][14]. Darknet, a framework for neural network, is required in order to develop and use YOLO neural network [15]. Darknet also needed to implement neural networks, especially YOLO, to be integrated with other system components.

The main contribution of this research are :

- Designed and developed system for monitoring mangrove forest health density
- Developed YOLO mangrove tree detector for the system.

The organization of the subsequent section is as follows, related work of tree detection (Section 2), system design (Section 3), experiment and result (Section 4), and finally conclusion (Section 5)

## 2. RELATED WORKS

Table I shows several studies have been conducted related to tree detections. While there is no object detection specifically for mangroves, there are some that similarly detects tree.

Fricker et al. [16] proposed using convolutional neural network (CNN) to detect individual conifer tree. The convolutional neural network (CNN) model is used on multispectral LiDAR imagery data. The data is obtained from LiDAR camera on drone. This means that the data includes distance of each pixel and thus the 3D shape of the object. While it provide valuable data for spatial analysis, it require significant cost. Detection is used on mixed-conifer tree to also detect specific species of the tree. From the imagery, the forest looks sparse compared to mangrove.

Itakura et al. [17] proposed using convolutional neural network YOLO to detect individual tree. The imagery is obtained from trees on their campus using 360° camera carried by human on the ground instead of flying drone. The result is then combined with 3D reconstruction data to get other tree parameters like diameter and height. Three-dimensional (3D) data can be achieved through the use of the Structure from Motion (SfM) technique, meaning the conversion of a set of two-dimensional (2D) images into 3D data.

Luo et al. [18] proposed using Faster R-CNN, a modified version of Convolutional Neural Network (CNN), to detect individual tree. The imagery data is gathered by the use of drone to capture images of afforestation of the coal mine. Various backbone of convolutional neural networks (CNN) have been used for performance comparison. The system that they propose is capable of quantifying the number of tree and calculate its density. However the imagery of the forest appears to be sparse in trees and doesn't evaluate the value to standard of forest health density. The forest ecosystem of former coal mine also differ in many ways to mangrove.

Donmez et al. [19] proposed using connected component labeling (CCL) algorithm to detect and count individual citrus tree. The data is obtained from drone with multispectral LiDAR camera. The system uses combination of classical computer vision techniques, including grayscale conversion, histogram equalization, and image morphing, in order to extract the edges of the object and thus enable tree object detection. The distribution of citrus tree in agricultural land shows a sparse and homogeneous pattern, which may not be ideal for mangrove tree.
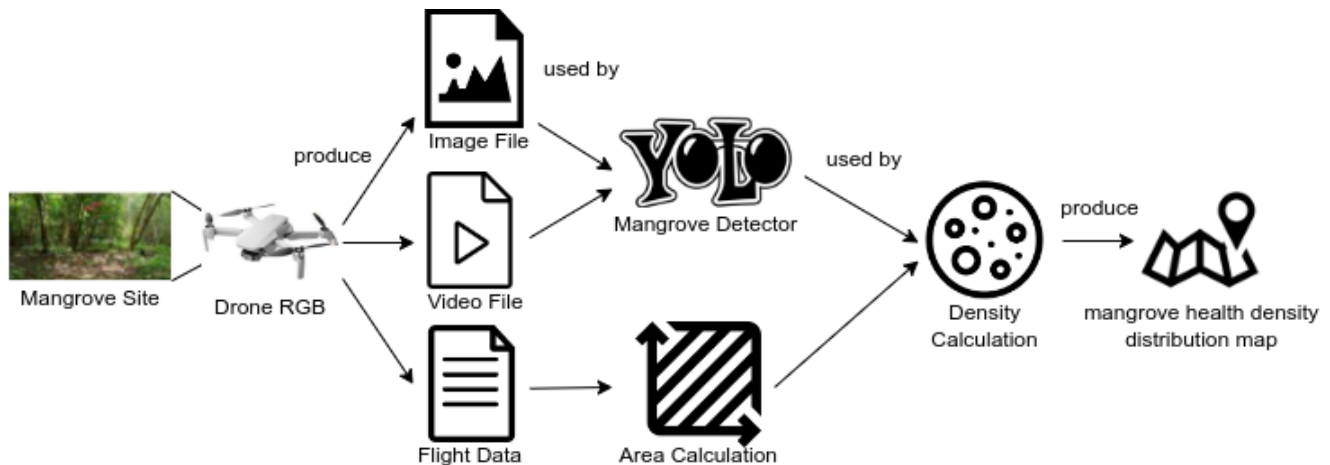
Figure 1. System design to calculate mangrove density health

Mubin et al. [20] proposed using convolutional neural network (CNN) to detect and count individual tree specific to oil palm tree. Two distinct convolutional neural networks (CNN) are developed: one for the detection of young oil palm trees and the other for the detection of mature oil palm trees. By combining resulting data, it can also differentiate between young and mature palm tree. The site used are from palm oil tree plantation, may not be ideal to implement for mangrove tree because the distribution of palm oil tree in agricultural land shows a sparse and homogeneous plantation.

All of the mentioned studies are able to detect individual tree, but none specifically detects mangrove tree. Mangrove forest have different characteristics, such as being more dense and trees located more randomly. Some studies are able to produce tree quantity counters but do not focus on the density data, which is required for producing density-based forest health. some studies required special camera like LiDAR or 360°, which incur more cost. Sattelite can cover a larger area but is limited in resolution and can't capture images from the side. Sattelite data with higher resolution is not open to the public and requires significant cost.

Multi-object detection is required to detect multiple objects inside of one image, which produces quantity for density calculation. Dense object detection is also required because of the nature of mangrove forest which is a dense forest. These became problems for density-based mangrove forest health, which our proposed study solved.

## 3. SYSTEM DESIGN

Figure 1 shows system design for mangrove health density. Drone RGB captures video, images, and flight data from mangrove site. Then the data is processed by YOLO mangrove detector and area calculation. The resulting data then used to calculate the density of mangrove forests, painted over the map.
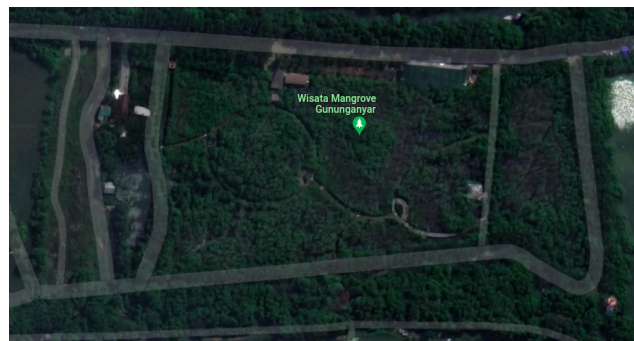


Figure 2. Mangrove site photo



Figure 3. Mangrove site satellite photo

### A. Mangrove Site

The study is conducted in mangrove forest in Gunung Anyar Mangrove Ecological Park. The site is located in Surabaya City, Indonesia. Figure 2 shows Gunung Anyar mangrove site entry gate. The site has a river that leads to the sea, with mangrove forests along the river. There are some designated wooden trails through the mangrove forest. However, mangrove forest along the river require boat ride. Figure 3 shows satellite image of the site, latitude between

Figure 4. Example of capture with drone flight data

7° 19' 50.5452" S and 7° 19' 56.352" S, and longitude between 112° 49' 2.3088" E and 112° 49' 10.3152" E[21]. The site area is about 50.000 meters squared.

### B. Drone RGB

Drone is used for gathering data because it can easily fly rather than traversing on ground. The drone produces photo, video, and flight data. Drone are flown into the inside of mangrove forest to penetrate the dense trees. Drone captured imagery from the side of mangrove tree. Figure 4 shows imagery captured with associated flight data. Bottom of the image shows flight data recorded, which include camera specification, GPS value, drone's altitude, and rotational value. Imagery and drone flight data is needed for mangrove density health calculation.

The drone used to capture the imagery is DJI Mini 2. It has 4K camera and can record the flight data automatically. It can be controlled for a maximum of about 6 kilometers away. The drone possesses the ability to navigate using GPS technology and provide live video feed to the controller, in case when clear visual sight to the drone is obstructed. It is important for the operator to practice caution in order to avoid collision with branches or leaves. If the drone falls into saltwater, then the drone's electronics will be damaged and its hard to retrieve. Image file, video file, and flight data, are then being used for further calculation for YOLO object detector and area calculation.

To make it easier, the system uses web interface to communicate with the user. It's easier to use and access web rather than using desktop app or command line interface. Figure 5 shows web input page to upload Drone's captured videos and flight data. To produce density health map, the user needs to upload video and drone flight data so the server can process the data.

### C. YOLO mangrove detector

YOLO is a convolutional neural network for object detection. It can detect the mangrove object inside any image. Training, choosing model, and optimization are explained more in Section 4. Figure 6 shows an example of YOLO's mangrove detection result. The example shows
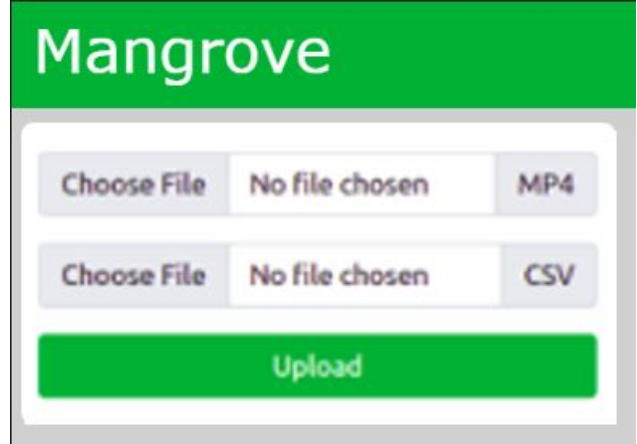


Figure 5. Upload page interface



Figure 6. Example scanned result from YOLO mangrove detector

5 mangrove trees detected. These mangrove quantity data will then be used to calculate the density.

Table II shows example of YOLO text result, containing information of detected object. Each line represents exactly one object, its class name, and its relative location in the image. The first value is class name 0 which means it is mangrove object. The second and third value is the top left relative coordinate of object's position. The fourth and fifth are the horizontal and vertical relative length of object's size. Relative position and length mean that the unit used is relative to the length of the image, between 0 and 1, meaning the position doesn't change even if the image is resized. To produce resulting image, the source image is drawn with boxes in correspond to text data.

TABLE II. YOLO text result in Darknet framework

| 1st | 2nd | 3rd | 4th | 5th |
|-----|------|------|------|------|
| 0 | 0.01 | 0.01 | 0.13 | 0.96 |
| 0 | 0.15 | 0.15 | 0.08 | 0.85 |
| 0 | 0.38 | 0.47 | 0.09 | 0.40 |
| 0 | 0.46 | 0.05 | 0.46 | 0.91 |

4

Figure 7. Illustration 3D simulation for area calculation from camera perspective
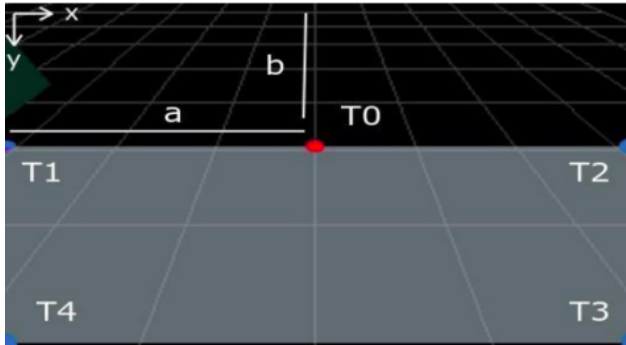


Figure 8. Illustration 3D simulation for area calculation from top-down perspective

YOLO can accept any image as long as the aspect ratio is the same as image that are used for training. This caveat happens because Darknet, a convolutional neural network framework, resizes the image to fit the input into the model. YOLO can handle the resize but if the aspect ratio is different, then the object inside will be stretched which may affect the performance.

To handle video input, the video needs to be dissected into many frames, then the YOLO can be used on the resulting images. Frames have to be linked with drone's flight data to ascertain its origin of capture, camera specification, drone's rotation value, and GPS value. Each frame can have different data as the drone flies and scan different part of the mangrove forests.

Practically, the video is split into its individual frame components, which are then saved as image files in a designated folder on the storage disk. This folder serves as an intermediary storage location for the images that will go through further processing. This storage strategy is used to eliminate memory overflow problem. If individual frame are all stored in computer's random access memory (RAM), it will require a lot of memory, depending on video's length and resolution. Scan result information are also stored in storage disk, as text file. This means the system capable to handle drone flight file and video of various length and numbers, regardless of RAM capacity.

### D. Area calculation

CameraTransform is used for calculating the total area. It can simulate 3D coordinate system from photos given sufficient data, such as camera's sensor size, focal length, elevation, field of view, degree of yaw, degree of roll, degree of pitch, and image size [22]. CameraTransform can convert coordinate inside of an image into 3D coordinate in the real world, meaning the system can calculate an actual area the object resides.

The area selected for calculation is from the bottom of 2D image up to the furthest object detected, shown as a red point. Figure 7 and Figure 8 shows the same 3D simulation but with different perspective. There are four points (*T1*,
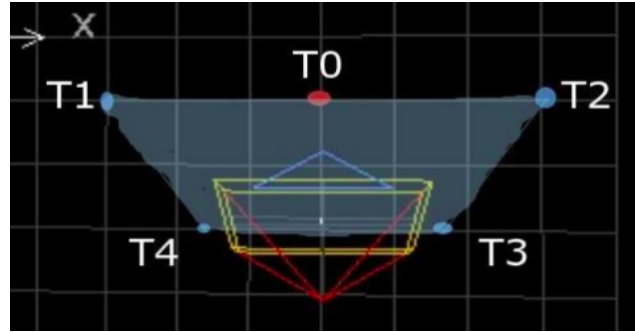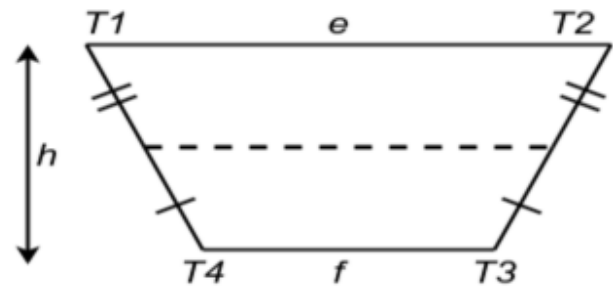


Figure 9. Illustration for trapezoid equation

*T2*, *T3*, and *T4*) to represent points of area boundary. The upper bound of the selected area is not the top part of the image because projection points to the sky and produces invalid 3D coordinate. Also, the furthest object's coordinate namely *T0* is selected because dense mangrove forest might block imagery of other mangrove trees behind them, thus only calculate the area until the furthest visible object. *T0* has *a* and *b* to represent its coordinate. *T0* is only used for reference. The boundary's 2D coordinate is then produced according to Table III. Then the 2D coordinate is projected as 3D coordinate onto the ground plane by CameraTransform.

Figure 9 shows illustration for trapezoid equation (1). Variable *e*, *f* means the length of top line, means the length of bottom line, and *h* means the height of trapezoid.

$$area = \frac{1}{2}(e + f)h \qquad (1)$$

TABLE III. 2D coordinate of area boundary on camera image

| Point | X | Y |
|-------|---|---|
| T0 | a | b |
| T1 | 0 | b |
| T2 | image width | b |
| T3 | image width | image height |
| T4 | 0 | image height |

5

TABLE IV. Mangrove health density standard [4][5]

| Category | Density (*n* tree / *ha*) | Density (*n* tree / $m^2$) |
|---|---|---|
| Bad | < 1000 | < 1 |
| Moderate | ≥ 1000 - < 1500 | ≥ 1 - < 1.5 |
| Good | ≥ 1500 | ≥ 1.5 |



Figure 10. Mangrove health density map



Figure 11. Detail page interface



Figure 12. Preview of mangrove object detection dataset

In 3D space the y-axis is ignored for calculation because the points are projected on ground plane, that height value is zero. Practically, CameraTransform is a Python library. In order to be utilized, it needs to be initialized with the drone's camera information, such as focal length, sensor size, image size, elevation, tilt degree, and roll degree. Next, system invokes the function `spaceFromImage([<x>,<y>])` in order to obtain the 3D space coordinate from 2D image coordinate.

*E. Density Calculation*

One of metrics used to measure mangrove health is based on density [4][5]. Density is calculated by dividing the quantity of mangrove trees by the total area, as given in equation (2). Quantity is obtained from the number of mangrove trees inside of an image, from the result of YOLO mangrove detector.

$$density = \frac{quantity}{area} \qquad (2)$$

Table IV shows density standard for comparing density value produced. Health value is divided into three category: bad, moderate, and good. The area units is converted from hectares (*ha*) to square meters (*m²*) to match it with units used by area calculation. 1 hectare is equal to 10000 square meters. Each density value has GPS position value and can be collected to form mangrove health density map.

*F. Mangrove forest density health map*

Figure 10 shows a map with density data painted as coloured dots. Red dot means bad, yellow dot means moderate, while green dot means good density health. Density map is served as an interactive map onto the website. It provides interactive control such as zoom and move the map. This feature is helpful for users to inspect and pinpoint the location of density data, to help preservation action. Figure 11 shows the details of each density data. It shows detailed information such as GPS location value, quantity, area, density, and scanned image.

## 4. EXPERIMENT AND RESULT

Experiment were conducted to improve YOLO mangrove tree detection performance, in turn, improved the system's performance. YOLO is Convolutional Neural Network (CNN) meaning to customize the detection into mangrove object and improve the detection, training is needed. There are three variations that are expected to improve performance: YOLO model, dataset augmentation, and imagery pixel resolution. The system is also used to monitor the chosen site, Gunung Anyar Mangrove Ecological Park site.

*A. Dataset*

The dataset is composed of photos of mangrove forests with mangrove tree object locations in the photos. Figure 12 shows some examples of photos with its annotation. The dataset consists of 100 photos and has 877 mangrove tree objects. Images and videos is shot at 4k resolution

Figure 13. Darkmark annotation program

(3840×2160 pixels). The dataset divided into 80% training and 20% test, randomly selected.

The dataset is formed from a collection of photos from drone video recordings. Drone video footage is broken down into a collection of photos. Each photo is given information on the location and size of each mangrove tree object. Figure 13 shows Darkmark program, to help annotate and assemble the dataset properly. Darknet requires the same format as the output, as per the YOLO object detection system design. The Darkmark program helps to add object information in the photo, using the mouse, rather than manually by typing. Figure 13 shows a photo with its annotations visualized with a red box. Data is gathered using drone because they make it easier to capture data and gather metadata automatically. Drone is flown with the altitude of 1 to 3 meter.

### B. YOLO Experiment Design

To evaluate the YOLO performance, three parameters were selected:

- Recall, an evaluation of the number of detected mangrove trees when compared to the actual number of mangrove trees [23]. The greater the recall value, the better. The recall performance parameter can evaluate the performance of detecting the number of mangrove trees, so that the quantity results be more accurate.

- IoU, Intersection Over Union, evaluates the accuracy of the location and size of detected mangrove trees when compared to the actual location and size of mangrove trees [23]. The greater the value, the better, so area calculation be more accurate.

- Computation time, evaluate the time required to process the data. The faster the time, the better.

The following preparation needs to be done before conducting the experiment. Photo data needs to be taken from the mangrove forest and annotated. The YOLO model is modified to have the output class to be a single type of object, namely mangrove tree. Darknet installation is done by compiling from the GitHub repository that has been provided.

Inference can be done with any GPU such as NVIDIA MX130. As for training, it is recommended to use the best GPU so that it can be completed quickly because it can take a long time. Experiments will compare computation time during inference. The training parameters are changed according to the recommendations, the iteration is set to 6000 iterations and the learning rate is set to 0.0013. The dataset is divided into 80% for training and 20% for testing. Testing is done to test YOLO based on the experimental parameters.

There are three categories variations that are expected to improve performance: YOLO model variation, dataset augmentation, and resolution variation. Each variation is tested in the experiments to find the best combination for the performance of YOLO detection results. The following is the variation design for the four experiments:

### 1) Experiment design of model variation

YOLO is a convolutional neural network. The structure and configuration of a CNN is called a model. The selected YOLO model is YOLOv4. The YOLOv4Tiny model is one of the modifications of the YOLOv4 model, has the name tiny or small because of its smaller structure with 29 convolution neural layers than the YOLOv4 model which has 127 convolution neural layers [24]. Both model variations will be compared according to predetermined parameters, to choose which one is the best. Thus, two model variations are selected, namely the YOLOv4 model and the YOLOv4Tiny model.

### 2) Experiment design of dataset augmentation

Dataset is the data used for training the model. The performance of neural network models, such as YOLO, is highly dependent on the dataset used. Augmentation is a way or technique of adding artificial data, with the hope to improve detection and generalization. Figure 14 shows augmentation by modifying colors such as darkening or brightening the image. Color augmentation is used to improve generalization to detect the mangrove tree more on the shape rather than specific color. Mangrove tree's leaf and trunk can have different shade of green and brown, depends on camera specification and lightning condition.

Figure 15 shows an augmentation by flip horizontal. Regardless of the tilt, to the left or to the right, a mangrove tree is still a mangrove tree. Vertical flip augmentation is not applied because there is no such thing as an upside-down tree. Experiments were conducted to compare the model training results between the augmented dataset and the non-augmented dataset.
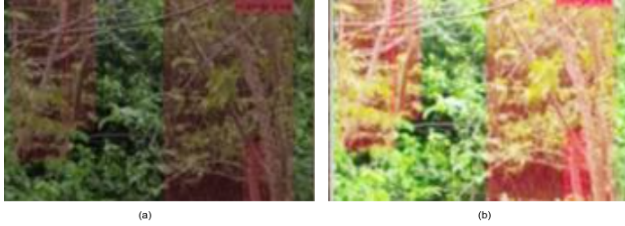
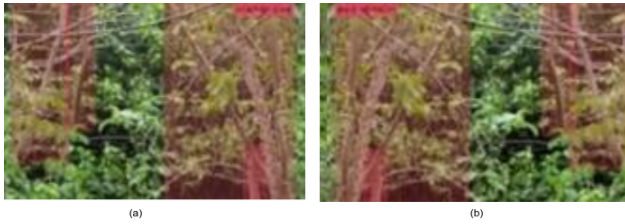Figure 14. Dataset color augmentation. (a) Darkened (b) Lightened



Figure 15. Dataset horizontal flip augmentation. (a) Original (b) Flipped

### 3) Experiment design of photo resolution

Photo resolution is the size of the image in pixels. The difference in photo resolution is suspected to affect the accuracy of detection because a large resolution can hold more imagery information. If humans see a photo with a larger size, they are better able to detect smaller objects in the photo. Three resolutions were selected, which are 512x288 pixels, 1024x576 pixels, and 2048x1152 pixels. The three resolutions were chosen because they meet the darknet resolution requirements, which are multiples of 32. The photo data was resized into three variants of the photo size.

### C. YOLO Experiment Result

The following are the results of the YOLO experiments conducted:

### 1) Experiment result of model variation

Two YOLO models were selected, namely YOLOv4 and YOLOv4Tiny. Both have been trained using the mangrove forest dataset. Table V shows that there are visible differences in experiment result, YOLOv4 has a higher recall difference of 4%, but requires a longer computation time with 135ms. YOLOv4Tiny has a lower recall but requires a lower computation time of 22ms.

### 2) Experiment result of dataset augmentation

The dataset variations chosen are augmented dataset and non-augmented dataset. Both datasets have been used for

TABLE V. Experiment result of YOLOv4 model variation

| Model | Recall | IoU | Time |
|-------|--------|-----|------|
| YOLOv4 | **98%** | **90%** | 135ms |
| **YOLOv4Tiny** | 94% | 88% | **22ms** |

TABLE VI. Experiment result of dataset augmentation variation for YOLOv4 model

| Dataset | Recall | IoU | Time |
|---------|--------|-----|------|
| Augmented | 98% | 90% | 135ms |
| **Non-augmented** | 98% | **93%** | 135ms |

TABLE VII. Experiment result of dataset augmentation variation for YOLOv4 Tiny model

| Dataset | Recall | IoU | Time |
|---------|--------|-----|------|
| Augmented | 98% | 80.5% | 22ms |
| **Non-augmented** | 98% | **88.3%** | 22ms |

training YOLOv4 and YOLOv4Tiny models. So, a total of four variations have been conducted in this experiment:

- YOLOv4 model trained with Augmented Dataset
- YOLOv4 model trained with Non-Augmented Dataset
- YOLOv4Tiny model trains with Augmented Dataset
- YOLOv4Tiny model trained with Non-Augmented Dataset

Table VI and Table VII shows that there is a difference in experiment result especially in the recall parameter. Non-augmented has a higher IoU than augmented. Augmentation might not improve the performance of YOLO; it may even decrease the IoU performance.

### 3) Experiment result of photo resolution

Different photo resolutions are expected to affect the performance of the YOLO object detector. The selected resolution variations are 512x288 pixels, 1024x576 pixels, and 2048x1152 pixels. Each photo resolution variant was trained and tested on the YOLOv4 model and YOLOv4Tiny model. Photo resolution experiments were conducted with a total of six variants, namely:

- YOLOv4 model with resolution of 512x288 pixels
- YOLOv4 model with resolution of 1024x576 pixels
- YOLOv4 model with resolution of 2048x1152 pixels
- YOLOv4Tiny model with resolution of 512x288 pixels
- YOLOv4Tiny model with resolution of 1024x576 pixels
- YOLOv4Tiny model with resolution of 2048x1152 pixels

8

TABLE VIII. Experiment result of photo resolution variation for YOLOv4 model

| Resolution | Recall | IoU | Time |
|---|---|---|---|
| **288px** | **99%** | **93.03%** | 135ms |
| 576px | 98% | 92.95% | 135ms |
| 1152px | 98% | 92.40% | 135ms |

TABLE IX. Experiment result of photo resolution variation for YOLOv4 Tiny model

| Resolution | Recall | IoU | Time |
|---|---|---|---|
| **288px** | **95%** | **88.28%** | 22ms |
| 576px | **95%** | 87.53% | 22ms |
| 1152px | 94% | 79.81% | 22ms |

Table VIII and table IX shows that there is a difference in experiment result, especially in the IoU parameter. The smallest resolution has the best IoU than other resolutions

### D. Mangrove Forest Density Health System

Table X shows sample of processed data. It shows result of detected tree, area calculation, density calculation, and health category. The sample shows that there are some areas with good, moderate, and bad density health. Often if more trees detected, the density health becomes better. However, it also depends on the size of covered area. The larger the area is, the trees will become more scattered. Therefore causes the density value to worsen.

Good mangrove density is most often data taken in the midst of mangrove forest. Drone are flown carefully into the forest where there are only mangrove trees around. Regardless of direction, the drone will only see dense mangrove trees. Meanwhile if the data is captured from the edge of the forest or somewhere more open, the density value will be smaller because there are less mangrove trees there relative to coverage area.

TABLE X. Sample of mangrove density health data in Gunung Anyar Ecological Park

| Tree Total | Area ($m^2$) | Density ($n$ tree / $m^2$) | Category |
|---|---|---|---|
| 5 | 2.03 | 2.46 | Good |
| 3 | 1.374 | 2.18 | Good |
| 4 | 3.58 | 1.11 | Moderate |
| 3 | 16.49 | 0.18 | Bad |
| 1 | 3.15 | 0.31 | Bad |
| 4 | 4.15 | 0.96 | Bad |
| 4 | 1.88 | 2.12 | Good |
| 9 | 1.83 | 4.91 | Good |
| 7 | 2.224 | 3.15 | Good |

### E. Discussion

YOLOv4Tiny is the best choice because even though it has a lower recall, it still has a value of 94% which is quite high, and it does not require much time especially if many photos need to be processed. The two variations do not have significant IoU differences with 1% difference. This indicates a trade-off between recall value and computation time.

For experiment of augmentation dataset, there are no improvements by using augmentation. The non-augmented dataset also decreases IoU performance on both models. The best choice is to use the non-augmented dataset. The difference in photo resolution does not affect the recall value but affects the IoU value. The greater the photo resolution, the smaller the IoU value. This is likely due to the error calculation being more precise at a larger resolution because it has more pixels.

Because of these reasons, the YOLOv4Tiny variation was chosen because it has a time performance that is six times faster than YOLOv4. This is important because if the system needs to process a lot of data, it takes a long time. If the system use a better GPU, the processing will be very fast, computing time will not be an issue and can use the YOLOv4 model.

The system that has been made has been proven to be able to analyze data density. It has been tested at the mangrove site. It can easily scans with drones and no longer requires going deep into the mangrove forest.

## 5. CONCLUSION

Monitoring mangrove became very important for preservation effort. However to manually observe and survey require considerable human labor, effort, and time. Currently there is no environment monitoring system specifically for mangrove health. To eliminate those problem, this research developed system to gather and analyze mangrove forest density health. The system uses a combination of drone, YOLO, and CameraTransform. Drone is being used to easily gathering data inside mangrove forest. YOLO is being used to detect mangrove tree. YOLO mangrove object detector demonstrates high accuracy with 95% recall, 88.3% IoU, and 22ms processing time. CameraTransform is being used to calculate the area covered by mangrove trees. By using these combination diperoleh mangrove health density data.

Resulting mangrove health data can be used to improve the effectiveness of future preservation actions, such as to find which areas are particularly bad and need to be prioritized. For future work, we suggest to analyze other mangrove forest health parameters, such as leaf cover, tree height, and diameter. In addition, system should also be deployed to more mangrove sites.

## REFERENCES

[1] Food and Agriculture Organization of the United Nations, "Global forest resources assessment 2020 main report," Food and Agriculture Organization of the United Nations (FAO), Tech. Rep., 2020. [Online]. Available: https://www.fao.org/3/ca9825en/ca9825en.pdf

[2] G. Acharya, "Life at the margins: The social, economic and ecological importance of mangroves," *Madera Bosques*, vol. 8, pp. 53–60, Sep. 2016.

[3] G. M. Lovett *et al.*, "Who needs environmental monitoring?" *Frontiers in Ecology and the Environment*, vol. 5, no. 5, pp. 253–260, 2007.

[4] Environment Ministry of Indonesia, "*Keputusan Menteri Negara Lingkungan Hidup Nomor 201 Tahun 2004 Tentang Kriteria Baku dan Pedoman Penentuan Kerusakan Mangrove*," 2004.

[5] I. W. E. Dharmawan, Suyarso, Y. I. Ulumuddin, B. Prayudha, and Pramudji, *Panduan Monitoring Struktur Komunikasi Mangrove di Indonesia*. PT Media Sains Nasional, 7 2020.

[6] BRIN's Public Relations, "*MonMang, Aplikasi Canggih untuk Monitor Hutan Mangrove*," Indonesia's National Research and Innovation Agency (BRIN), 8 2023. [Online]. Available: https://www.brin.go.id/news/110996/monmang-aplikasi-canggih-untuk-monitor-hutan-mangrove

[7] L. Tang and G. Shao, "Drone remote sensing for forestry research and practices," *Journal of Forestry Research*, vol. 26, no. 4, pp. 791–797, Jun. 2015. [Online]. Available: https://doi.org/10.1007/s11676-015-0088-y

[8] T. A. Gardner *et al.*, "The cost-effectiveness of biodiversity surveys in tropical forests," *Ecology Letters*, vol. 11, no. 2, pp. 139–150, Nov. 2007. [Online]. Available: https://doi.org/10.1111/j.1461-0248.2007.01133.x

[9] E. Prayitno, N. Fahmi, M. U. H. Al Rasyid, and A. Sudarsono, "An implentation of iot for environmental monitoring and its analysis using k-nn algorithm," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, p. 1811, 12 2021.

[10] M. U. H. Al Rasyid, M. H. Mubarrok, and J. A. Nur Hasim, "Implementation of environmental monitoring based on kaa iot platform," *Bulletin of Electrical Engineering and Informatics*, vol. 9, pp. 2578–2587, 12 2020.

[11] J. B. Campbell, R. H. Wynne, and V. A. Thomas, *Introduction to remote sensing, sixth edition*. London, England: Guilford Press, Oct. 2022.

[12] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020.

[13] A. Rasyid *et al.*, "Pothole visual detection using machine learning method integrated with internet of thing video streaming platform," in *2019 International Electronics Symposium (IES)*. IEEE, Sep. 2019.

[14] E. D. Fadhillah, P. C. Bramastagiri, R. Sigit, S. Sukaridhoto, A. Brahmanta, and B. S. B. Dewantara, "Smart odontogram: Dental diagnosis of patients using deep learning," in *2021 International Electronics Symposium (IES)*. IEEE, Sep. 2021.

[15] A. Bochkovskiy, "Darknet," 2021. [Online]. Available: https://github.com/AlexeyAB/darknet

[16] G. A. Fricker, J. D. Ventura, J. A. Wolf, M. P. North, F. W. Davis, and J. Franklin, "A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery," *Remote Sensing*, vol. 11, no. 19, p. 2326, Oct. 2019. [Online]. Available: https://doi.org/10.3390/rs11192326

[17] K. Itakura and F. Hosoi, "Automatic tree detection from three-dimensional images reconstructed from 360° spherical camera using YOLO v2," *Remote Sensing*, vol. 12, no. 6, p. 988, Mar. 2020. [Online]. Available: https://doi.org/10.3390/rs12060988

[18] M. Luo *et al.*, "Individual tree detection in coal mine afforestation area based on improved faster RCNN in UAV RGB images," *Remote Sensing*, vol. 14, no. 21, p. 5545, Nov. 2022. [Online]. Available: https://doi.org/10.3390/rs14215545

[19] C. Donmez, O. Villi, S. Berberoglu, and A. Cilek, "Computer vision-based citrus tree detection in a cultivated environment using UAV imagery," *Computers and Electronics in Agriculture*, vol. 187, p. 106273, Aug. 2021. [Online]. Available: https://doi.org/10.1016/j.compag.2021.106273

[20] N. A. Mubin, E. Nadarajoo, H. Z. M. Shafri, and A. Hamedianfar, "Young and mature oil palm tree detection and counting using convolutional neural network deep learning method," *International Journal of Remote Sensing*, vol. 40, no. 19, pp. 7500–7515, Jan. 2019. [Online]. Available: https://doi.org/10.1080/01431161.2019.1569282

[21] "Google maps gunung anyar mangrove ecological site," Google Maps, accessed: October 5, 2023. [Online]. Available: https://maps.app.goo.gl/RJP6CoKdCnRmV7Sa6

[22] R. C. Gerum *et al.*, "CameraTransform: A python package for perspective corrections and image mapping," *SoftwareX*, vol. 10, p. 100333, Jul. 2019. [Online]. Available: https://doi.org/10.1016/j.softx.2019.100333

[23] D. Shah, "Mean average precision (map) explained: Everything you need to know," Mar 2022. [Online]. Available: https://www.v7labs.com/blog/mean-average-precision

[24] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-YOLOv4: Scaling cross stage partial network," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jun. 2021.

10

**Ilyas Yudhistira Kurniawan** received the Bachelor of Applied Computer Engineering degree from Departement of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya (PENS), Indonesia, in 2022. He is currently pursuing master degree in master's degree in applied informatics and computer engineering from Politeknik Elektronika Negeri Surabaya (PENS), Indonesia. His research interests include neural networks, internet of things (IoT), web development, and their system integration.

**M. Udin Harun Al Rasyid** received the B.Sc. degree from the Department of Informatics Engineering, Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 2004, and the Ph.D. degree in computer and communication network program from the College of Electrical Engineering and Computer Science (CECS), National Taiwan University of Science and Technology (NTUST), Taiwan, in 2012. He is currently an Associate Professor with the Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya (PENS), Indonesia. He heads the research group of EEPIS Wireless Sensor Networks (EWSN). His research interests include wireless sensor networks (WSNs), wireless body area networks (WBANs), the Internet of Things (IoT), and Web technology.

**Sritrusta Sukaridhoto** received the B.E. degree in electrical engineering, computer science program from Sepuluh Nopember Institute of Technology, Indonesia, in 2002 and the Ph.D. degree in Communication Networks Engineering from Okayama University, Japan, in 2013. He joined at Politeknik Elektronika Negeri Surabaya (PENS), Indonesia, as a lecturer in 2002. He stayed at Tohoku University, Japan, in 2004, as a visiting researcher. From 2017, He becomes Head of Human Centric Multimedia Lab, received several research grants, and also has several collaborations with government and industries. He is a technology enthusiast, his research interests include computer networks, human-centric, immersive multimedia technology, and Industrial Internet of Things. He has received several academic awards, best paper awards, and IEEE Young Researcher Award in 2009. He is a member of IEEE.