Mangrove Tree Density Detector using YOLO Based on Darknet Framework using RGB Drone Imagery

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@pasca.student.pens.ac.id, udinharun@pens.ac.id, dhoto@pens.ac.id

Abstract: Mangrove preservation is crucial due to their ecological significance impact. Monitoring the health of mangrove forests is essential for preservation strategy, yet it remains challenging and time-intensive, particularly in remote locations. This study aims to create system to automatically assess mangrove density, providing essential data for informed preservation strategies, such as prioritizing reforestation in low density area. Using drones with RGB cameras to capture aerial imagery, enabling remote data collection. The system utilizes the YOLO neural network object detector to automatically detect objects, enabling quantity estimation. Experiment shows that YOLO object detector is able to detect mangrove tree accurately with 95% recall, 88.3% IoU, and 22ms processing time. The system uses 'tiny' model variant to provide more efficient accuracy compared to computation resource, making it suitable for deployment on computer with limited resources. In comparison to standard model that improves the recall by 4%, IoU by 2%, but demands six times more processing time. Then it calculate the covered area using camera transformation formula. Finally, it calculates the density for mangrove forest health, synchronized with GPS location. With the resulting data on mangrove density, evaluations of mangrove forest health become much easier, facilitating effective preservation actions, such as reforestation in area with low mangrove density.

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

1. INTRODUCTION

Mangroves are trees that thrive in swamps and brackish water areas, typically found at the boundary of sea and land, as well as around river estuaries. In recent decades, from 2010 to 2020, the world lost about 12700 hectares [1]. Mangroves play a protective role against coastal erosion and seawater intrusion into groundwater reservoirs. Prevention of coastal erosion helps prevent 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 accelerating the rate of global warming [2]. These crucial roles would be disrupted if the mangrove are damaged, therefore we need to create and preserve ideal conditions for the mangrove forests.

Effective mangrove preservation requires monitoring data. 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 with extensive coastlines, where the area of mangrove destruction is increasing [1]. The most significant mangrove monitoring data is the density data. The government publishes mangrove health assessments with one of the important criteria being density data[4]. The national research agency (BRIN) also researched and published guide on how to conduct mangrove monitoring assessment [5]. They've also created the MonMang2 app, which is used to help measure and collecting mangrove monitoring data [6]. Based on the avalaible density data, conservationists can formulate reforestation programs that prioritize regions with lower mangrove density. Furthermore, the local community will be aware of the potential of mangrove forests for both conservation and tourism purposes [7].

The problem is that currently, there is no system for monitoring the health of mangrove forests, specifically using mangrove density criteria. Currently, mangrove environmental monitoring requires significant manpower to acquire [5]. To cover a large area, a lot of manpower and time are required. 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 [8][9].

TABLE I. Related works comparison

Research	Multi-object Detection	Dense-object Detection	Capture Equipment	Mangrove Detection	Method
Fricker et al. [20]	Yes	Yes	LiDAR Drone	No	CNN
Itakura et al. [21]	Yes	No	Ground 360°Cam	No	YOLOv2
Luo et al. [22]	Yes	Yes	RGB Drone	No	Faster R-CNN
Donmez et al. [23]	Yes	No	RGB Drone	No	CCL Algorithm
Mubin et al. [24]	Yes	No	RGB Sattelite	No	CNN LeNet
Beloiu et al. [25]	Yes	Yes	Aerial Imagery	No	Faster R-CNN
Pu et al. [26]	Yes	Yes	LiDAR Drone	No	LiDAR Analysis
Jemaa et al. [27]	Yes	No	RGB Drone	No	YOLO
Buonocore et al. [29]	No	No	Sattelite and IoT	No	Forest digital twin
Our Work	Yes	Yes	RGB Drone	Yes	YOLOv4

Based on the mentioned problems, in this research, we propose a 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 a long time and massive effort compared to manual observation. There are two main ways for computer sensors to gather data: in-situ and 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 [10][11]. Remote, as in remote sensing, means the system uses sensors that don't directly interact with the object of interest [12]. It uses a camera and light sensor from afar that is attached to a satellite or drone. Remote sensing is used to gather data about an object from a challenging and hard-to-access location by flying around the object, which is perfect for mangrove density monitoring because it eliminates the problem of hard-to-access data. This research system uses drone to gather data inside the mangrove forest.

RGB drone imagery has been used for other vegetation health-related purposes, such as for foliage assessment and pest detection [13][14]. Drones using RGB cameras have the advantage of being more readily available and cheaper compared to LiDAR cameras.

With the development of sophisticated computer vision, technology systems can also help process complex processes that humans previously could only do. Object detection is needed to create a system that can 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[15]. YOLO is popular, well known to be a capable object detector, and easily compatible for integration thus, many use it for object detection in their systems[16][17]. Moreover, the model is very efficient in the sense that it has low computation time with comparable accuracy[18]. Darknet, a framework for neural network, is required to develop and use YOLO neural network [19]. Darknet also

needed to implement neural networks, especially YOLO, to be integrated with other system components.

The main contributions of this research are :

- Designed and developed a system for monitoring mangrove forest health density

- Developed YOLO mangrove tree detector for the system.

- Creating an accessible system by utilizing affordable and commonly available RGB drones, along with low-spec server requirements.

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 detect tree.

Fricker et al. [20] proposed using a convolutional neural network (CNN) to detect individual conifer trees. The convolutional neural network (CNN) model is used on multispectral LiDAR imagery data. The data is obtained from a LiDAR camera on drone. This means that the data includes the distance of each pixel and thus the 3D shape of the object. While it provides valuable data for spatial analysis, it requires significant cost. Detection is used on mixed-conifer trees also to detect specific species of the tree. From the imagery, the forest looks sparse compared to mangrove forest.

Itakura et al. [21] proposed using convolutional neural network YOLO to detect individual tree. The imagery is obtained from trees on their campus using 360° camera carried by a human on the ground instead of a 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. [22] 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 the 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 calculating their density. However, the imagery of the forest appears to be sparse in trees and doesn't evaluate the value of forest health density standard. The forest ecosystem of former coal mine also differs in many ways to mangrove forest.

Donmez et al. [23] 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 a 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 trees.

Mubin et al. [24] 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 plantations, 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.

Beloiu et al. [25] proposed a method utilizing Faster R-CNN for the detection and identification of a mix of coniferous and deciduous tree species within forested areas. Their approach uses openly available aerial RGB imagery, which offers greater detail compared to satellite data. However, the absence of drone usage restricts its applicability to specific sites.

Pu et al. [26] proposed a technique for tree detection using 3D point cloud data gathered from LiDAR Drone. Their method combines various LiDAR analysis techniques, including Mean Shift, Canopy Height Model, Density Scan, and k-Nearest Neighbor classifier. Despite the potential, the accessibility and cost-effectiveness of LiDAR drones are still limited.

Jemaa et al. [27] proposed a method for detecting and assessing the health of orchard apple trees. Their approach combines multispectral near-infrared imagery with drone RGB imagery. It uses a neural network called YOLO to detect the trees. The health assessment for apple trees involves distinct standards and criteria compared to those used for mangroves. This method is specifically used for apple tree plantations rather than dense mangrove forests. Moreover, it requires additional near-infrared data in addition to using RGB drone imagery.

Study somewhat related to the digital twin where objects in the real world are scanned to get virtual data that can then be analyzed. Similar to the research conducted by Moztarzadeh et al. [28] to scan body parts for health analysis. Buonocore et al. [29] proposed a framework for creating a Forest Digital Twin (FDT), which integrates various types of data from satellite imagery and IoT devices. Their approach involves twinning individual trees by integrating real-virtual digital interfaces to capture a wide array of physical, biotic, and environmental variables. In another study conducted by Domenico et al. [30], the Forest Digital Twin (FDT) is utilized to strengthen forest protection actions. The FDT functions by ensuring the traceability, authenticity, and tamper-proof nature of forest-related data. This method, allows the detection of forest damage, allowing for timely intervention and mitigation measures to be implemented.

All of the mentioned studies can 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 require 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 the 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 the 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 is then used to calculate the density of mangrove forests, which are painted over the map.

A. Mangrove Site

The study is conducted in a mangrove forest in Gunung Anyar Mangrove Ecological Park. The site is located in Surabaya City, Indonesia. The site has a river that leads
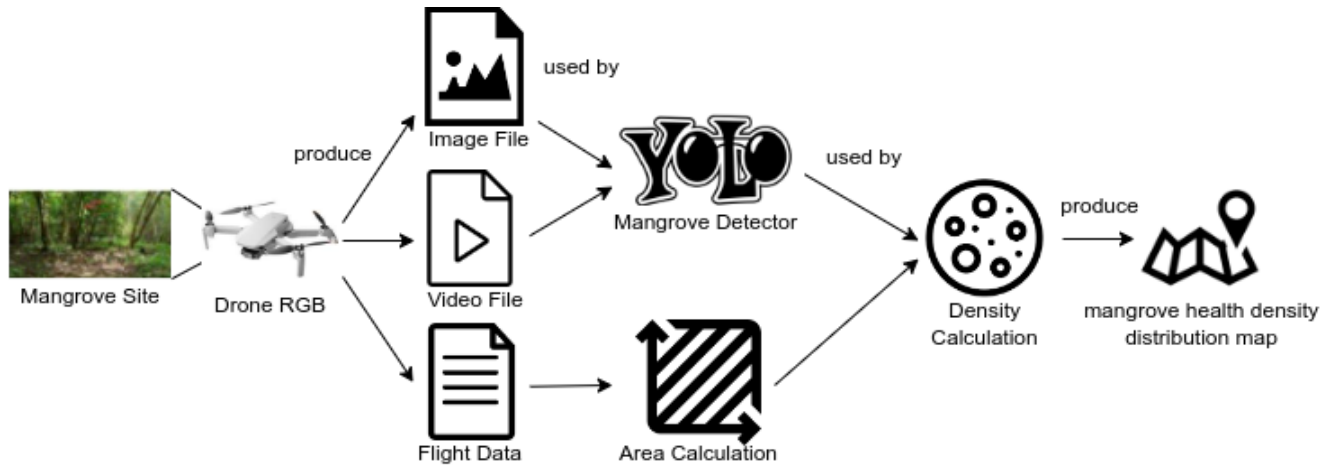
Figure 1. System design to calculate mangrove density health

to the sea, with mangrove forests along the river. There are some designated wooden trails through the mangrove forest. However, mangrove forests along the river require boat ride. Figure 2 shows satellite image of the site, latitude between 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[31]. The site area is about 50.000 square meters.
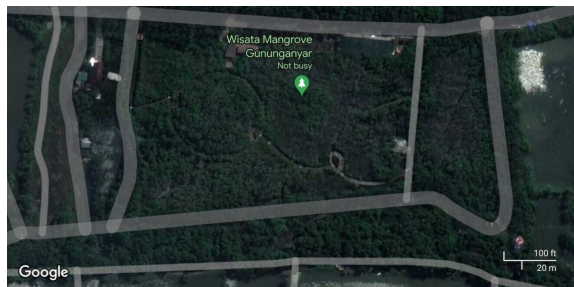

Figure 2. Mangrove site satellite photo

B. Drone RGB

Drone is used for gathering data because it can easily fly rather than traverse on the ground. The drone produces photos, videos, and flight data. Drone are flown into the inside of mangrove forests to penetrate the dense trees. Drone captured imagery from the side of mangrove trees.

Figure 3 shows imagery captured with associated flight data. The bottom of the image shows the recorded flight data, which includes camera specification, GPS value, altitude, and rotational value. Imagery and drone flight data are needed for mangrove density health calculation.

Figure 3. Example of capture with drone flight data

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 a clear visual sight of the drone is obstructed. Image files, video files, and flight data are then used for further calculation of YOLO object detector and area calculation.

To make it easier, the system uses a web interface to communicate with the user. It's easier to use and access web rather than using desktop app or command line interface. To generate a density health map, users must upload both video footage and drone flight data obtained from the drone's SD card.

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 4 shows an example of YOLO's mangrove detection result. The example shows 5 mangrove trees detected. These mangrove quantity data will then be used to calculate the density.

Figure 4. Example scanned result from YOLO mangrove detector

Table II shows examples of YOLO text result containing information of the detected objects. Each line represents precisely one object, its class name, and its relative location in the image. The first value is class name 0, which means it is a 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 that 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

YOLO can accept any image as long as the aspect ratio is the same as images 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, and 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 scans 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 frames 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 is also stored in storage disk as text file. This means the system is capable of handling drone flight files and videos of various lengths and numbers, regardless of RAM capacity.

D. Area calculation

CameraTransform is used to calculate 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 [32]. CameraTransform can convert coordinates inside of an image into 3D coordinates in the real world, meaning the system can calculate the actual area in which the object resides.

Figure 5 and Figure 6 show the same 3D simulation but with a different perspective. The area selected for calculation is from the bottom of 2D image up to the furthest object detected, shown as a red point. There are four points (*T1*, *T2*, *T3*, and *T4*) to represent points of area boundary. The selected area's upper bound is not the image's top part because the projection points to the sky and produces invalid 3D coordinates. Also, the furthest object's coordinate, namely *T0* is selected because dense mangrove forests might block imagery of other mangrove trees behind them, thus only calculating the area until the furthest visible object.
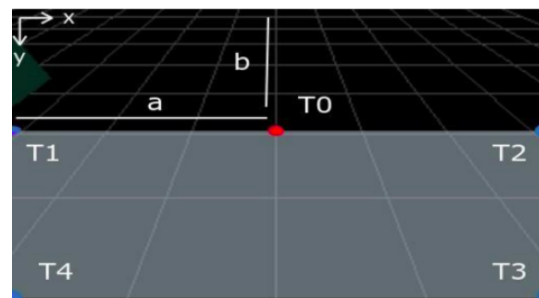

Figure 5. Illustration 3D simulation for area calculation from camera perspective
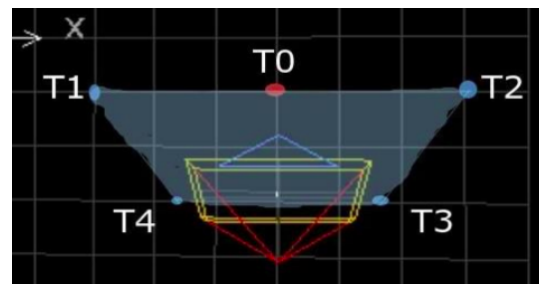

Figure 6. Illustration 3D simulation for area calculation from top-down perspective

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.

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

Figure 7 shows illustration for trapezoid equation (1). Variable *e*, *f* means the length of the top line, means the length of the bottom line, and *h* means the height of the trapezoid.
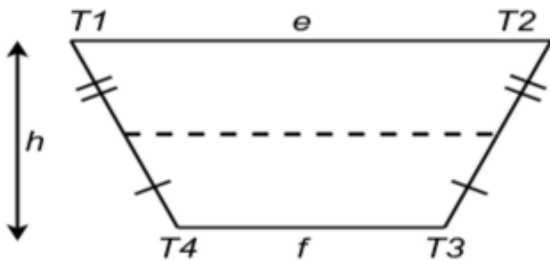
$$area = \frac{1}{2}(e + f)\,h \qquad (1)$$

Figure 7. Illustration for trapezoid equation

In 3D space the y-axis is ignored for calculation because the points are projected on the ground plane, and that height value is zero. Practically, CameraTransform is a Python library. 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>])` to obtain the 3D space coordinate from 2D image coordinate.

E. Density Calculation

One of the 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 the density standard for comparing the density value produced. Health value is divided into three categories: bad, moderate, and good. The area units is converted from hectares (*ha*) to square meters (*m*^{2}) 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 a mangrove health density map.

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

F. Mangrove forest density health map

Figure 8 shows a map with density data painted as colored dots. Red dot means bad, yellow dot means moderate, while green dot means good density health. The density map is served as an interactive map on 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 8. Mangrove health density map

Figure 9 shows the details of each density data. It shows detailed information such as GPS location value, quantity, area, density, and scanned image.
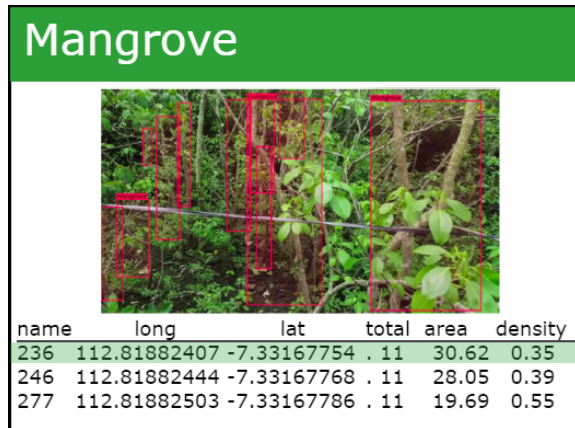
Figure 9. Detail page interface

Figure 10. Darkmark annotation program

4. EXPERIMENT AND RESULT

Experiments were conducted to improve YOLO mangrove tree detection performance, which, 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. Experiment Setup

The following describes the experimental setup needed to explore the best configuration and model for detecting mangrove trees. The dataset explains the data used, metrics describe the evaluation parameters utilized, and the experiment design outlines the parameters and variations being investigated.

1) Dataset

The dataset is composed of photos of mangrove forests with mangrove tree object locations in the photos. Figure 11 shows some examples of photos with its annotation. The dataset consists of 100 photos and has 877 mangrove tree objects. Images and videos are shot at 4k resolution (3840×2160 pixels). The dataset was 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 10 shows Darkmark program to help annotate and assemble the dataset properly[33]. 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 typing.

Figure 11 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 at an altitude of 1 to 3 meters.

Figure 11. Preview of mangrove object detection dataset

2) Metrics

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 [34]. The greater the recall value, the better. The recall performance parameter can evaluate the performance of detecting the number of mangrove trees to make the quantity results 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 [34]. The greater the value, the better, so area calculation is more accurate.

- Computation time, evaluate the time required to process the data. The faster the time, the better. For computation, a budget laptop is used with specification of an Intel Core i5-8250U 4 core processor 3.4GHz, Nvidia MX130 graphics card, and 12GB RAM

3) YOLO Experiment Design

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 be a single object type, namely a 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, as 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 of 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:

- 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 [18]. 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.

- 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 of improving detection.
  Figure 12 shows augmentation by modifying colors, such as darkening or brightening the image. Color augmentation improves generalization to detect the mangrove tree more on the shape rather than a specific color. Mangrove tree's leaves and trunks can have different shades of green and brown, depending on camera specifications and lighting conditions.
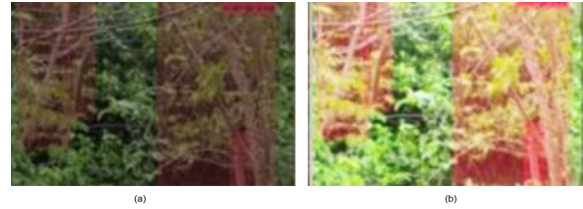

Figure 12. Dataset color augmentation. (a) Darkened (b) Lightened

Figure 13 shows an augmentation by flip horizontal. Regardless of the tilt, to the left or 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 and non-augmented datasets.
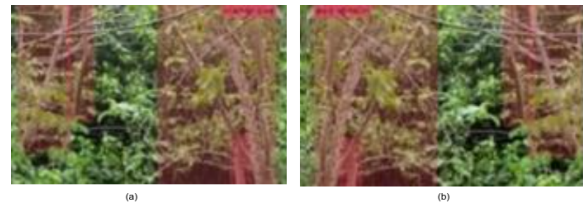

Figure 13. Dataset horizontal flip augmentation. (a) Original (b) Flipped

- 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 can better 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, multiples of 32. The photo data was resized into three variants of the photo size.

B. 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 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.

TABLE V. Experiment result of YOLOv4 model variation

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

2) Experiment result of dataset augmentation

The dataset variations chosen are augmented dataset and non-augmented dataset. Both datasets have been used to train the 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 show a difference in experiment results, 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.

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

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

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.

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

C. Mangrove Forest Density Health System

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

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

Good mangrove density is most often data taken amid mangrove forest. Drone are flown carefully into the forest where only mangrove trees are 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 fewer mangrove trees there relative to the coverage area.

D. Discussion

YOLOv4Tiny is the best choice because even though it has a lower recall, it still has a value of 94% which is relatively 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 the 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 because the error calculation is 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.

The limitation of this research is that the detection is specific to mangrove trees. It is also limited by the capabilities of the drone. The RGB camera on the drone cannot scan the image of mangrove trees if it is occluded. In addition, the image can only be taken from a location the drone can reach, enough space to fly the drone, and within the flying range. Even so, the RGB drone used is more cost-effective and, if necessary, can bought a more powerful drone.

The research could be integrated with the Monmang2 app, an existing mangrove monitoring application published by the National Research Agency BRIN[6]. Currently, mangrove monitoring relies on the Monmang2 app, which requires manual data collection by filling out digital form. Our research assists by collecting density data with synchronized GPS location, thereby reducing the workload on environmental surveyors. In potential future works, our research could automatically gather additional detailed data beyond density, such as trunk circumference, height, canopy, and species, while synchronizing with IoT data, such as tidal height, water pH, pollution levels, and oxygen content. It could potentially be utilized to automate the entire data collection process for a more comprehensive and integrated assessment of mangrove health.

5. CONCLUSION

Monitoring mangroves became very important for preservation efforts. However, manually observing and surveying requires considerable human labor, effort, and time. Currently, there is no environment monitoring system specifically for mangrove health. To eliminate those problems, this research developed a system to gather and analyze mangrove forest density health. The system uses a combination of drone, YOLO, and CameraTransform. Drone is being used to gather data inside the mangrove forests easily. 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 combinations, mangrove health density data is obtained.

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

This study on mangrove density has the potential to be applied to other forests that require data on health density. Tree detection for different forests would need to be customized for different tree types.

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.

[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.

[7] L. Hakim, D. Siswanto, and N. Makagoshi, "Mangrove conservation in east java: the ecotourism development perspectives," *Journal of Tropical Life Science*, vol. 7, no. 3, pp. 277–285, 2017.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

[13] L. Parra, D. Mostaza-Colado, S. Yousfi, J. F. Marin, P. V. Mauri, and J. Lloret, "Drone rgb images as a reliable information source to determine legumes establishment success," *Drones*, vol. 5, no. 3, 2021.

[14] F. Betti Sorbelli, L. Palazzetti, and C. M. Pinotti, "Yolo-based detection of halyomorpha halys in orchards using rgb cameras and drones," *Computers and Electronics in Agriculture*, vol. 213, p. 108228, 2023.

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

[16] 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.

[17] 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.

[18] 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.

[19] A. Bochkovskiy, "Darknet : open source neural network framework," https://github.com/AlexeyAB/darknet, 2021, computer Software.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] M. Beloiu, L. Heinzmann, N. Rehush, A. Gessler, and V. C. Griess, "Individual tree-crown detection and species identification in heterogeneous forests using aerial rgb imagery and deep learning," *Remote Sensing*, vol. 15, no. 5, 2023.

[26] Y. Pu, D. Xu, H. Wang, X. Li, and X. Xu, "A new strategy for individual tree detection and segmentation from leaf-on and leaf-off uav-lidar point clouds based on automatic detection of seed points," *Remote Sensing*, vol. 15, no. 6, 2023.

[27] H. Jemaa, W. Bouachir, B. Leblon, A. LaRocque, A. Haddadi, and N. Bouguila, "Uav-based computer vision system for orchard apple tree detection and health assessment," *Remote Sensing*, vol. 15, no. 14, 2023.

[28] O. Moztarzadeh, M. B. Jamshidi, S. Sargolzaei, F. Keikhaee, A. Jamshidi, S. Shadroo, and L. Hauer, "Metaverse and medical diagnosis: A blockchain-based digital twinning approach based on mobilenetv2 algorithm for cervical vertebral maturation," *Diagnostics*, vol. 13, no. 8, 2023.

[29] L. Buonocore, J. Yates, and R. Valentini, "A proposal for a forest digital twin framework and its perspectives," *Forests*, vol. 13, no. 4, 2022.

[30] G. Di Domenico, L. Cesaretti, E. Pontuale, and V. Di Stefano, "Forests management in the metaverse: the potential of forest digital twin and blockchain for preventing crimes." Salerno, Italy: 4th Conference of Associazione Italiana Società Scientifiche Agrarie (AISSA) Under40, July 12–13 2023.

[31] "Google maps gunung anyar mangrove ecological site," Google Maps, accessed: October 2023.

[32] R. C. Gerum *et al.*, "CameraTransform: A python package for perspective corrections and image mapping," *SoftwareX*, vol. 10, p. 100333, Jul. 2019.

[33] S. Charette, "Darkmark: Image markup and annotation for darknet framework machine learning," https://github.com/stephanecharette/DarkMark/, 2021, computer Software.

[34] D. Shah, "Mean average precision (map) explained: Everything you need to know," Mar 2022.

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.