

# Approach and Analysis of Yolov4 Algorithm for Rice Diseases Detection at Different Drone Image Acquisition Distances

Fauzan Masykur<sup>1,2</sup>, Kusworo Adi<sup>1</sup>, Oky Dwi Nurhayati<sup>1</sup>

<sup>1</sup> Doctoral Program of Information System, School of Postgraduate Studies Diponegoro University Semarang, Imam Bardjo St No 5, Semarang, Indonesia

<sup>2</sup> Department of Informatics Engineering, Muhammadiyah University of Ponorogo, Indonesia

**Abstract** – Rice production plays an important role in people's lives globally because rice is the most widely consumed staple food for over half of the world's human population. Unfortunately, the rice plants are prone to pests and diseases which may result in a decrease in the rice production. Thus, early and accurate detection of rice diseases is needed. This paper discusses a method for detecting rice diseases called You Only Look Once (YOLO) object detection algorithm version 4. A drone camera was used to acquire the images from four different distances, namely 2 meters, 5 meters, 10 meters and 20 meters. This approach aims to detect the presence of pests to be faster, more accurate and more precise. The precision results at each image capture from the different distances were 46.8%, 48%, 65% and 77.3% with the average loss value of 6.52, 0.54, 1.16 and 2.73

**Keywords** – Drone, YOLOv4, object detection, rice diseases.

## 1. Introduction

In the world, especially in tropical countries, rice plants which are processed into rice are one of the staple foods, including Asia, accounting for 40% of world rice production. Meanwhile, in Indonesia, rice is the majority staple food consumed by the public, with production reaching 31.36 million tons in 2021 [1]. The magnitude of the need for rice requires innovation in maintaining rice production so that it is maintained to meet the basic needs of the community. One of the challenges in maintaining rice production is detecting pests and diseases in rice plants quickly, accurately and precisely [2]. There are various types of diseases that often infect rice plants, namely sheath rot, brownspot, blast disease and bacterial blight [3]. To quickly detect plant diseases, Artificial Intelligence (AI) technology can be used to replace the role of farmers in monitoring plant development. Time and the perspective of each farmer usually determine the results of suspected diseases that attack plants [4]. The existence of AI is able to change the perspective and way of working in the world of agriculture in protecting crops for the better so that crop failures can be avoided. AI with the help of sensors and drones is also able to manage various needs such as only irrigation issues, climate monitoring, nutritional needs so that they are measured according to plant needs [5].

Rice plants that are attacked by disease if handled too late will result in crop failure. The occurrence of fungi begins with the presence of brown spots on the leaves of the rice which is in the vegetative and generative stages [6]. If it is left unchecked or not resolved immediately, the rice grains will not grow, which in turn, decreases the success rate of the harvest. Thus, rice growth from the beginning of the planting period to the harvesting time needs regular checking. Direct observation by farmers definitely takes time and efforts so it is less effective and efficient.

---

DOI: 10.18421/TEM122-39

<https://doi.org/10.18421/TEM122-39>

**Corresponding author:** Fauzan Masykur,  
Doctoral Program of Information System, School of  
Postgraduate Studies Diponegoro University Semarang,  
Imam Bardjo St No 5, Semarang, Indonesia


**Email:** [fauzan@umpo.ac.id](mailto:fauzan@umpo.ac.id)

*Received:* 28 February 2023.

*Revised:* 12 May 2023.

*Accepted:* 21 May 2023.

*Published:* 29 May 2023.

 © 2023 Fauzan Masykur, Kusworo Adi & Oky Dwi Nurhayati; published by UIKTEN. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

The article is published with Open Access at <https://www.temjournal.com/>

Therefore, it requires fast and accurate detection processes to reduce the risk of crop failure.

The development of technology bring up the term smart farming which integrates information and communications technology with traditional agriculture [7]. Technology has penetrated into various fields including agriculture, for example the existence of the Internet of Things [4][8], big data analysis, remote sensing [9], machine learning [10] and unmanned Aerial Vehicle (UAV)[11][12][13]. The use of Unmanned Aerials, also called as drones, in agriculture can reduce observation time and improve efficiency [14]. Drones are able to assist farmers in observing from the air the crop conditions, including irrigation systems, soil varieties, and pests and fungal attacks [15]. Drone imagery has information in the infrared and visual spectral ranges. It can also be extracted to obtain certain information which cannot be known just by looking with the eye.

In this study, the images of rice fields captured by drones, which flew over the fields, were used as datasets in determining the presence or absence of pests. The algorithm is expected to help the object determination processes quicker with more precise results and lower loss rate. Deep learning algorithm with You Only Look Once (YOLO) object detection architecture version 4 was used in this research. Yolo can detect object as quick as 45 frames per second [16]. The drone camera dataset images were taken at various distances with the aim of knowing the distance which could best detecting the object. Thus, the goal of this research is to determine the best image acquisition distance of the drone cameras for detecting and diagnosing pests and diseases on the rice plant leaves.

## 2. Literature Review

Deep Learning Algorithm has been used in many different fields from the image classification of aircraft damage [17], object detection [18] to decision making processes [19]. The Yolo Deep Learning algorithm detects class and object positions in the form of a regression with a single convolution neural network [20]. To get the predicted size from a fixed format after convolution on all images, the image be reformatted into 416x416. Then, the image is divided into several grid cells. Each grid cell has the responsibility of detecting the presence of objects in the frame of each bounding box. The grid cell has an accurate prediction level in 5 prediction parameters, namely x, y, w, h and confidence (x, y) which is the target coordinates, and (w, h) which represented width and height of the target object. Confidence is used as a determinant in accepting or rejecting the prediction of the existence of objects within the threshold.

Various studies on object detection show that Yolo's speed in detecting images is 45 frames per second in real time [21]. Meanwhile, research on object detection of apple blossoms conducted by Dihua Wu, et al [22], showed that the apple flower detection using Yolov4 algorithm and the CSPDarknet is 53 frames per second in real time. The results of the tests carried out in the study showed the detection model parameters of the apple flower is 96,74%, model size is reduced by 231,51 MB, inference time is reduced by 39,47% and mAP is 97,31% which is only 0,24% smaller than the previous model [22]. The improvement of Yolov3 model has been carried out by Liu and Zhang [23] with dataset objects in the form of large vehicle images and adaptation to traffic density levels. The K-Means++ algorithm is used to improve the efficiency of anchor box dimension grouping. The experiment resulted in a mAP value of 91.12%. Bandar, et al also applied a dataset of drone imagery results in a plant classification method using a transfer learning approach and CNN (Convolutional Neural Network) [7]. The experiment results show an accuracy detection rate of 92,93%. Meanwhile, in an experiment conducted by Abdul Hafeez [24], drones can be used in the agricultural sector to increase crop yields by monitoring crop conditions and spraying pesticides. Some of the features of the drone are embedded with sensor technology, and used together with spraying areas innovations and the application of AI and remote monitoring.

The calculation of the grid cell probability in the dataset images uses probability calculations by predicting the bounding box as shown in equation 1 below.

$$\frac{\Pr(Class_i | Object) * \Pr(Object) * IOU_{pred}^{truth}}{\Pr(Class_i) * IOU_{pred}^{truth}} = \quad (1)$$

Notes:

$\Pr(Class_i | Object)$  = Probability class-i on object  
 $\Pr(Object)$  = Probability of the object  
 IoU = Intersection of Union

## 3. Proposed Method

This paper proposes the Yolo object detection algorithm version 4 with an aruco marker inserted as a reference point which functions as the coordinate point for each corner to calculate the area of the detected disease infection. The dataset is in the form of drone camera images with various drone height distances, namely 2 meters, 5 meters, 10 meters and 20 meters.

In general, our design chart for detecting objects against plants and calculating the area of the infected area can be presented in figure 1.

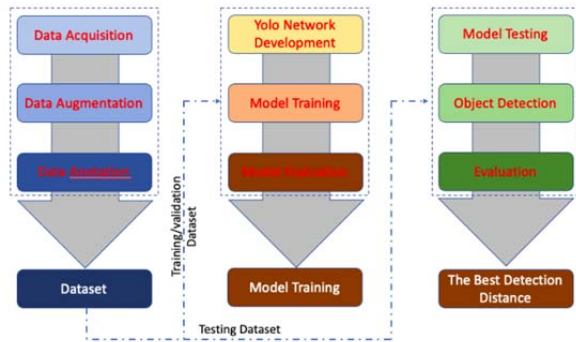


Figure 1. Object Detection Design Chart

### 3.1. Dataset Acquisition

In this step, the datasets in the form of rice plants images were taken and collected using a drone camera. The drone flew over rice farming land at various different height, namely 2 meters, 5 meters, 10 meters and 20 meters. The variation of heights was intended to determine the extent of accuracy and precision in detecting the objects. An illustration of drone image acquisition is presented in Figure 2. The drone collected the images of rice fields at around 08.00 am to 10.00 am, depending on the weather condition such as rain, wind, fog etc. The 3,540 collected images in the datasets were grouped into 2 classes, infected and healthy. Drone camera images of the infected and healthy rice fields (Figure 3) were used as a reference for determining the number of class.



Figure 2. Data Acquisition Illustration

Drone images collected as many as 3,540 images in the healthy rice field category and the infected rice field category. The cause of the infected rice fields was not specified because no object related to the cause could be clearly detected. Therefore, this paper does not discuss the causes of infection. The rice diseases might be caused by fungi and pests such as leafhoppers, rats, grasshoppers, birds, caterpillars and others.

The drone images only show that the infected areas were in contrast to healthy areas. The infections were detected from 2 weeks old until the plants entered the ovulation period or 60 days old. After the drone imagery has been collected, it is detailed to enter the next step, namely the pre-processing process in the form of labeling according to a predetermined class.

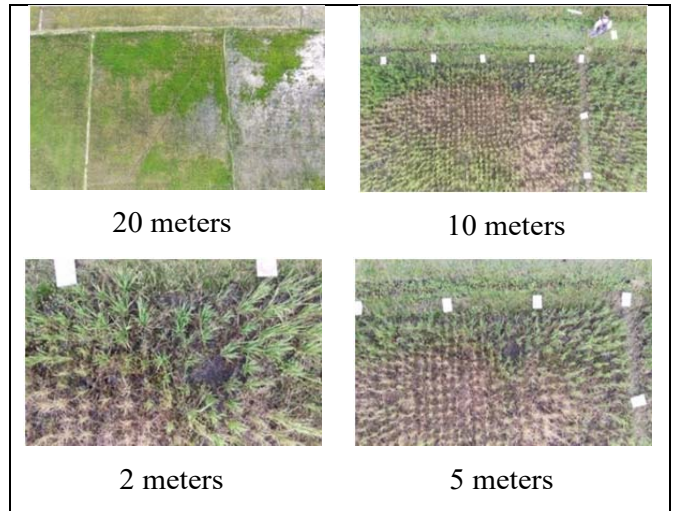


Figure 3. Drone Image of Rice Field

### 3.2. Image labeling

Image labeling is needed as a first step in determining the infected and healthy areas. The labeling process involves all datasets but not all images are labeled as infected or healthy because not all images have infected or healthy areas. Figure 4 presents the image labeling process using labelling software which has 2 classes, namely infected and healthy. Image labeling was conducted to a specific area on an agricultural land. It did not label images of only one plant, as had been used to label object, because the label would be used as a reference to determine the area within the detected area.

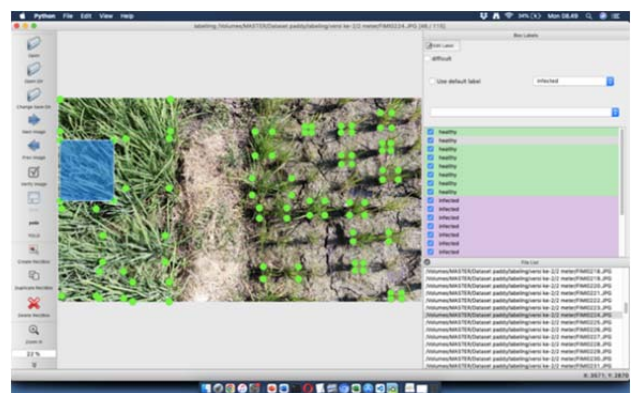


Figure 4. Image Labeling

Dataset images labeling has a limit at each angle which represents the coordinates of the label ( $X_{min}$ ,  $Y_{min}$ ,  $X_{max}$ ,  $Y_{max}$ ) and is used as a measure of model accuracy at the output box and box labels [20].

There are several factors influencing the success of predicting the object detection during image acquisition. Those are light, time of taking the image, natural weather, angle and size of the image.

The datasets used in this study were drone images with dimensions of 3840×2160 and a resolution of 72×72, if you will add data from external sources or taken using a different camera, needs to be done through augmentation processes which consist of scaling, rotating, and resizing. Those processes were depending on the need for equalization of the image. There were two categories of labels for the detected objects which are presented in more details in Table 1.

Table 1. Dataset Labeling

| Class Number | Class Name | Number of Images |
|--------------|------------|------------------|
| 0            | Infected   | 3.540            |
| 1            | Healthy    |                  |

### 3.3. Framework Proposed

Detection of object in the image datasets were conducted using computer vision and image processing technology [25]. Combination of a deep neural network and Yolo object detection algorithm enables the detection of rice plant diseases by making sense of the multiple images [21]. There were three main stages in the research processes, namely dataset construction, training model and evaluation. The Yolov4 foundation structure involves various steps which aims increase the accuracy and efficiency of the processes, mosaic data augmentation, starting from the convolution, batch normalization, cross stage partial connections (CSP), weighted residual connection, cross mini-batch normalization, self adversarial training, mosaic data augmentation, CIoU loss and drop block regularization [26]. Yolov4 being the core of the proposed method has a simple architecture using the CSPDarknet53 backbone. As shown in Figure 5, the input image was resized to 416x416 according to the configuration, which was then inserted into the CSPDarknet53 backbone for the training to achieve the best results.

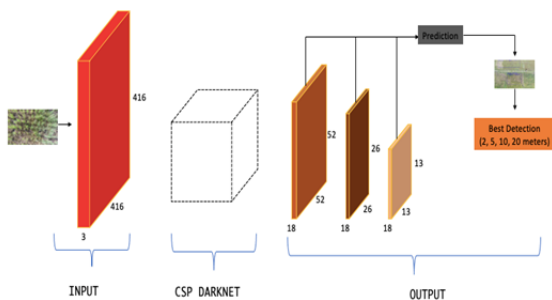


Figure 5. Best Distance Detection in The Yolov4 Architecture

Yolov4 employs end to end operation to immediately predict the bounding box and to accurately classify the objects based on their classes. Several stages in the Yolov4 architecture are image input, feature extraction, bounding box prediction and final detection (Figure 5). In more detail, the stages are as follows [18]:

**Image Input.** The image, which is in the standard YOLO size of 416 x 416, is divided into several s x s grids as the initial formation of a bounding box in predictions.

**Feature extraction.** This stage includes searching for the characteristics of the object class as a differentiator between objects using CSP Darknet. This stage is included in terms of classification and prediction.

**Bounding box prediction.** Each s x s grid cell becomes a delimiter for the bounding box and helps in determining the level of confidence in each box. The constraints which appear have the responsibility of predicting the presence or absence of objects by determining the coordinate points of each corner and midpoint (tx, ty, tw, th and t0).

**The final stage of detection.** This final stage uses the NMS (non-suppression maximum) algorithm to determine the box prediction based on the similarity between boxes. The pixel values that are not optimal will be eliminated, meaning that the highest pixel value will be left while lower pixel values will be replaced with zero so that the frame or only one boundary box is formed.

### 4. Experimental Result and Analysis

This Yolov4 experiment involved dataset images from drone camera with a total dataset of 3.540 images which were classified into two classes, infected and healthy class. Measurement of performance of the detected object was conducted using performance confusion matrix analysis because it is able to help provide a comprehensive assessment of model performance. The confusion matrix includes True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

Accuracy is used as a measurement of performance analysis using the results of the confusion matrix. Confusion matrix as a simple model measurement that is often used in the formation of classification models as shown in equation 2.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (2)$$

Precision is the degree of reliability of the model when it gives "positive" results. When a data is classified as "positive" compared to how precise the model is that the actual label is also "positive".



Determination of the precision value using equation 3.

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

Recall is defined as the degree of reliability of the model to correctly detect data labeled "positive". Recall is also used to define the proportion of the amount of data labeled as "positive" out of all data that is labeled "positive", as shown in equation 4.

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

Mean Average Precision (mAP) is used to evaluate the object detection model taking into account errors, False Positive (FP), and False Negative (FN). Calculation of the mAP is conducted using equation 5.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{5}$$

F1-Score is a combination of recall with precision in a single matrix harmonically. Equation 6 is used to determine the average precision with recall.

$$F1\ Score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \tag{6}$$

Loss Function is used to summarize the loss performance of object detection models with the aim of understanding the final model results. The model shows good performance if the loss value decreases during the model training process.

In this experiment, Yolo object detection version 4 was used with a dataset obtained from drone image acquisition at a distance of 2 meters, 5 meters, 10 meters and 20 meters. Before the training process, the image dataset goes through the labeling stage using Yolo-Mark and Labellmg with 1 class (infected) and 2 classes (infected and healthy).

Labeling with 2 classes is aimed at knowing the accuracy and precision as well as the loss of the model. The training model used several parameters displayed in table 2.

Table 2. Parameters of YOLOv4 Model

| Parameters  | 1 class     | 2 Classes    |
|-------------|-------------|--------------|
| Width       | 416         | 416          |
| Height      | 416         | 416          |
| Batch       | 64          | 64           |
| Max Batch   | 2.000       | 4.000        |
| Subdivision | 16          | 16           |
| steps       | 1.600,1.800 | 3.200, 3.600 |

The results of the training on a dataset labeled 1 class (infected) show the best mean Average Precision (mAP) and the smallest loss was found at a distance of 5 meters. Figure 6 displays a graph showing the results of the model training with datasets taken from different distances.

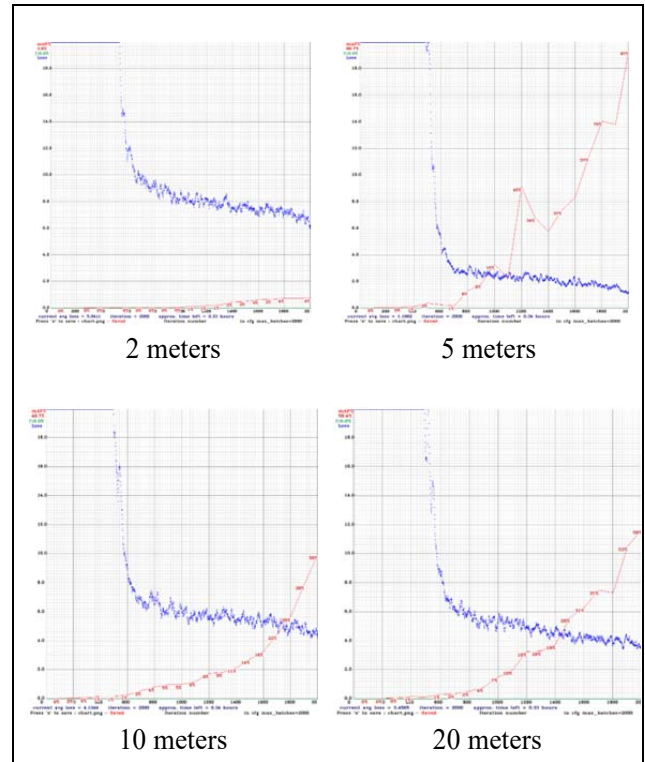


Figure 6. Graph of Training Model Results With 1 Class Label

Image acquisition distance affects the results of mAP and loss. Figure 6 shows that the best mAP value is 97% at an image acquisition distance of 5 meters and the worst value is 4% at an image acquisition distance of 2 meters.

The process of image acquisition and labeling affects the precision results because the acquisition at a distance of 2 meters produces images that are almost the same between infected plant images and healthy plant images so that feature extraction produces almost the same pixel values. In addition, the loss value at the acquisition distance of 2 meters has the greatest value indicating that the model has difficulty differentiating between infected plant images and healthy plant images. All of the experiments in 1 class showed that the built model started to work effectively at the 500th iteration marked by a decrease in the loss value and an increase in the mAP value. Determination of iterations based on 2.000 x number of classes (2.000).

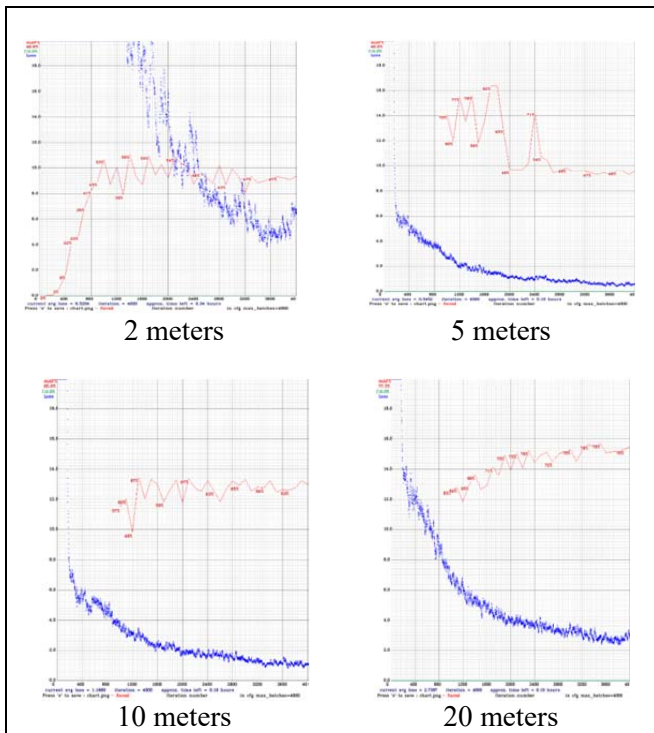


Figure 7. Graph of Training Model Results With 2 Classes Label

The results of the training model with two classes (infected and healthy) show that the best mAP is 77.3% captured from 20 meters distance and the worst value is 46.8% captured from 2 meters. Meanwhile, the largest loss occurs in the acquisition image of 2 meters and the smallest loss occurs in the acquisition image of 5 meters (Figure 7). The results of this experiment show that image acquisition from a distance of 2 meters is the worst result among the various distances of image acquisition datasets in either 1 class or 2 classes. Figures 8 (a) and (b) show graphs of the results of model training with various image acquisition distances.

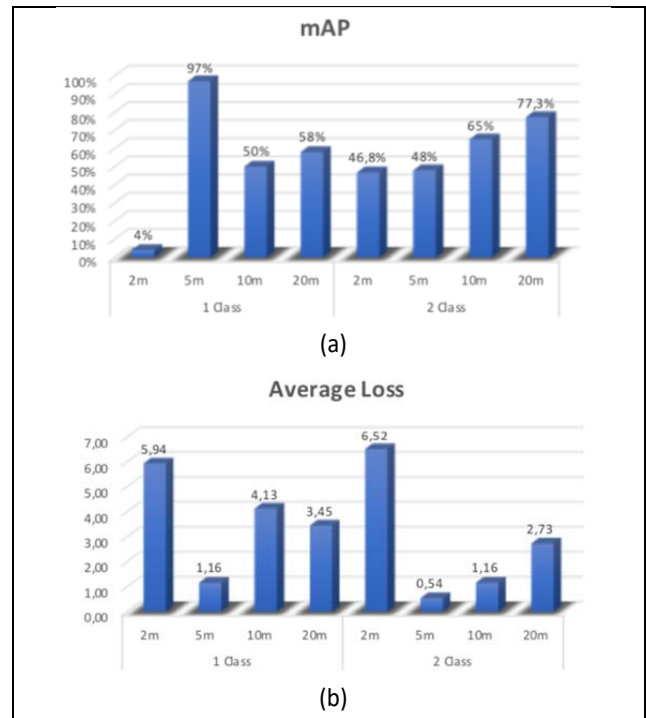


Figure 8. Training Model Results Resume (a) Map (b) Average Loss

The performance of the training results model with a variety of images showed that there are differences in the values because image acquisition is influenced by several things, for example, acquisition distance, light intensity during image acquisition, camera specifications, acquisition time, weather conditions and labeling [27]. The final results of training the model with various image acquisition distances are shown in Table 3. Labeling 2 classes (infected and healthy) makes the best results compared to labeling 1 class. Average Intersection of Union on the type of labeling shows the best value on the type of labeling 2 classes. Meanwhile, the fastest time for detection is 1 second in the labeling using 2 classes.

Table 3. Overall Results of the Training Model Using Yolov4 at Different Image Acquisition Distance

| Label (Class) | Acquisition (meters) | TP  | FP  | FN  | F1-SCORE | Recall | Average IOU (%) | Total Detection Time (Seconds) |
|---------------|----------------------|-----|-----|-----|----------|--------|-----------------|--------------------------------|
| 1             | 2                    | 12  | 97  | 287 | 0,06     | 0,04   | 6,61            | 6                              |
|               | 5                    | 5   | 1   | 0   | 0,91     | 1,00   | 64,75           | 1                              |
|               | 10                   | 29  | 18  | 44  | 0,48     | 0,40   | 39,08           | 5                              |
|               | 20                   | 59  | 24  | 47  | 0,62     | 0,56   | 48,86           | 6                              |
| 2             | 2                    | 200 | 161 | 110 | 0,60     | 0,65   | 37,05           | 1                              |
|               | 5                    | 12  | 7   | 3   | 0,71     | 0,80   | 52,61           | 1                              |
|               | 10                   | 26  | 11  | 9   | 0,72     | 0,74   | 59,67           | 1                              |
|               | 20                   | 57  | 26  | 13  | 0,75     | 0,81   | 54,35           | 1                              |

The results of object detection include the value of the percentage level of detection. For example, Figure 9 shows the object detection values for the healthy and infected labels. The detection value gives rise to a different prediction value because the training process depends on the feature extraction value for each object. In some detection results, a value of 100% is indicated, which means that the detected object has a true 100% accuracy rate, while some other detections show a value of 0.62 or even a value of 0.31. The difference in the value of each detected object indicates the difference in the special characteristics of the object. Figure 9 shows the results of detection at an acquisition distance of 20 meters with the detection results of 8 lands being detected as infected areas and 3 lands being detected as healthy areas.



Figure 9. Object Detection Values For the Healthy and Infected Labels

### 5. Comparison

Research performance is measured by making comparisons with previous research, namely research conducted by Salem, et al [27]. In this study using drones or Unmanned Aerial Vehicles (UAV) as equipment in aerial image acquisition. The height of the drone in image acquisition is set with 3 different distance variations, namely 10 meters, 20 meters and 30 meters and the detectors used are Aggregate Channel Feature (ACF), You Only Look Once (Yolo) MobileNet and Yolo Resnet50. The object detected is human detection as a first step to detect human presence during search and rescue at several locations of disaster victims as a forensic first step. The results showed that the height of the image acquisition distance resulted in different detections with the Yolo MobileNet detector model achieving a significant increase compared to the ACF or Yolo Resnet50 detector model. Complete results of human object detection with various variations of image acquisition distances are presented in full in Table 4. It should be noted that the Yolo detector has a high sensitivity which is influenced by several things, namely the height of the camera, camera resolution, light intensity and the angle at which the image is taken.

Table 4. Comparison Result

| Training Data          | Testing Data | Test Precision (AP) |               |      | Average Test AP |               |      |
|------------------------|--------------|---------------------|---------------|------|-----------------|---------------|------|
|                        |              | YOLO MobileNet      | YOLO ResNet50 | ACF  | YOLO MobileNet  | YOLO ResNet50 | ACF  |
| 10 m                   | 10 m         | 0.93                | 0.37          | 0.8  | 0.80            | 0.22          | 0.50 |
|                        | 20 m         | 0.76                | 0.16          | 0.5  |                 |               |      |
|                        | 30 m         | 0.70                | 0.13          | 0.2  |                 |               |      |
| 20 m                   | 10 m         | 0.88                | 0.24          | 0.5  | 0.87            | 0.20          | 0.50 |
|                        | 20 m         | 0.82                | 0.26          | 0.6  |                 |               |      |
|                        | 30 m         | 0.91                | 0.11          | 0.4  |                 |               |      |
| 30 m                   | 10 m         | 0.77                | 0.19          | 0.2  | 0.83            | 0.11          | 0.37 |
|                        | 20 m         | 0.78                | 0.08          | 0.3  |                 |               |      |
|                        | 30 m         | 0.94                | 0.05          | 0.6  |                 |               |      |
| <b>Average Test AP</b> |              | 0.83                | 0.18          | 0.46 |                 |               |      |

Table 4 shows that the Yolo MobileNet precision is higher than the ACF and Yolo Resnet50 detectors. While the average precision (AP) is 0,83 for Yolo MobileNet, 0,18 for Yolo Resnet50 and 0,46 for ACF. Thus, a low false-negative rate relates to higher recall and a low false positive relates to higher precision.

### 6. Conclusion

This paper has discussed the implementation of Yolo v4 object detection experiment with the aim of knowing its accuracy at the various image acquisition

distances. The results of the experiment show that labeling with 2 classes, namely infected and healthy, produces better values than labeling into 1 class. The best mAP value is produced by the image acquisition distance from a distance of 20 meters worth 77.3% and the smallest loss value is produced by the image acquisition distance from a distance of 5 meters with a value of 0.54.

Thus, the final result of this study is that the best distance for detecting pests on rice plant leaves with image acquisition using drones is at a distance of 20 meters.

## References

- [1]. Statistik, B. P. (2022). *Luas Panen dan Produksi Padi di Indonesia 2021*. Badan Pusat Statistik. Retrieved from: <https://www.bps.go.id/publication/2022/07/12/c52d5c3e530c363d0ea4198/luas-panen-dan-produksi-padi-di-indonesia-2021.html> [accessed: 05 January 2023].
- [2]. Masykur, F., Adi, K., & Nurhayati, O. D. (2022). Classification of Paddy Leaf Disease Using MobileNet Model. In *2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED)* (1-4). IEEE.
- [3]. Ramesh, S., & Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. *Information Processing in Agriculture*, 7(2), 249–260. Doi: 10.1016/j.inpa.2019.09.002
- [4]. Chen, W. L., Lin, Y. B., Ng, F. L., Liu, C. Y., & Lin, Y. W. (2020). RiceTalk: Rice Blast Detection Using Internet of Things and Artificial Intelligence Technologies. *IEEE Internet of Things Journal*, 7(2), 1001–1010.
- [5]. Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture*, 4, 58–73.
- [6]. Prasetyo Purwo Nugroho, Sp. P. (2019). *Penyakit Kresak pada Tanaman Padi*. Retrieved from: <https://lisa.id/petani/artikel/5c8b158ad324d09579ad5fde> [accessed: 06 March 2023]
- [7]. Al-Rami, B., Alheeti, K. M. A., Aldosari, W. M., Alshahrani, S. M., & Al-Abrez, S. M. (2022). A New Classification Method for Drone-Based Crops in Smart Farming. *International Journal of Interactive Mobile Technologies*, 16(9), 164–174. Doi: 10.3991/ijim.v16i09.30037
- [8]. Masykur, F., Prasetyo, A., Widaningrum, I., Cobantoro, A. F., & Setyawan, M. B. (2020). Application of Message Queuing Telemetry Transport (MQTT) Protocol in the Internet of Things to Monitor Mushroom Cultivation. *7th International Conference on Information Technology, Computer, and Electrical Engineering, ICITACEE 2020 - Proceedings*, 135–139. Doi: 10.1109/ICITACEE50144.2020.9239118
- [9]. Wahab, I., Hall, O., & Jirstrom, M. (2018). Remote sensing of yields: Application of UAV imagery-derived ndvi for estimating maize vigor and yields in complex farming systems in Sub-Saharan Africa. *Drones*, 2(3), 1–16. Doi:10.3390/drones2030028
- [10]. Jasim, W. N., Abued, S., Almola, S., Alabiech, M. H., & Harfash, E. J. (2022). *Citrus Diseases Recognition by Using CNN Model 2 Literature review*, 46, 85–94.
- [11]. van der Merwe, D., Burchfield, D. R., Witt, T. D., Price, K. P., & Sharda, A. (2020). Drones in agriculture. *Advances in agronomy*, 162, 1-30.
- [12]. Junos, M. H., Mohd Khairuddin, A. S., Thannirmalai, S., & Dahari, M. (2022). Automatic detection of oil palm fruits from UAV images using an improved YOLO model. *Visual Computer*, 38(7), 2341–2355.
- [13]. Inoue, Y., & Yokoyama, M. (2019). Spectral and 3d measurement by drone-based remote sensing of farmland-geo-Information for smart farming. *Journal of the Japan Society for Precision Engineering*, 85, 236–242.
- [14]. Puri, V., Nayyar, A., & Raja, L. (2017). Agriculture drones: A modern breakthrough in precision agriculture. *Journal of Statistics and Management Systems*, 20(4), 507–518.
- [15]. Wang, Q., Liu, Q., & Zhang, S. (2020). Implementation of Drone System in Survey for Tomato Chlorotic Spot Virus. *Journal of Extension*, 58(2), 1–10.
- [16]. Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition*, 6517–6525. Doi: 10.1109/CVPR.2017.690
- [17]. Adi, K., Widodo, C. E., Widodo, A. P., & Margiati, U. S. (2022). Detection of Foreign Object Debris (Fod) Using Convolutional Neural Network (Cnn). *Journal of Theoretical and Applied Information Technology*, 100(1), 184–191.
- [18]. Yin, Y., Li, H., & Fu, W. (2020). Faster-YOLO: An accurate and faster object detection method. *Digital Signal Processing: A Review Journal*, 102, 102756. Doi: 10.1016/j.dsp.2020.102756
- [19]. Nurhayati, O. D., Bachri, O. S., Supriyanto, A., & Hasbullah, M. (2018). Graduation prediction system using artificial neural network. *International Journal of Mechanical Engineering and Technology*, 9(7), 1051–1057.
- [20]. Li, X., Qin, Y., Wang, F., Guo, F., & Yeow, J. T. W. (2020). Pitaya detection in orchards using the MobileNet-YOLO model. *Chinese Control Conference*, 6274–6278. Doi: 10.23919/CCC50068.2020.9189186
- [21]. Morbekar, A., Parihar, A., & Jadhav, R. (2020). Crop disease detection using YOLO. In *2020 international conference for emerging technology (INCET)*, 1-5. IEEE.
- [22]. Wu, D., Lv, S., Jiang, M., & Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Computers and Electronics in Agriculture*, 178, 105742. Doi: 10.1016/j.compag.2020.105742
- [23]. Liu, J., & Zhang, D. (2020). Research on Vehicle Object Detection Algorithm Based on Improved YOLOv3 Algorithm. *Journal of Physics: Conference Series*, 1575(1). Doi: 10.1088/1742-6596/1575/1/012150
- [24]. Hafeez, A., Husain, M. A., Singh, S. P., Chauhan, A., Khan, M. T., Kumar, N., ... Soni, S. K. (2022). Implementation of drone technology for farm monitoring & pesticide spraying: A review. *Information Processing in Agriculture*. Doi: 10.1016/j.inpa.2022.02.002
- [25]. Akanksha, E., & Rao, P. R. K. (2021). A Feature Extraction Approach for Multi-Object Detection Using HoG and LTP. *International Journal of Intelligent Engineering and Systems*, 14(5), 259–268. Doi: 10.22266/ijies2021.1031.24
- [26]. Dhiaegana, R. N. (2020). Penerapan Convolutional Neural Network untuk Deteksi Pedestrian pada Sistem Autonomous Vehicle. *Institut Teknologi Bandung*. Retrieved from: [https://informatika.stei.itb.ac.id/~rinaldi.munir/TA/Makalah\\_TA\\_Renjira.pdf](https://informatika.stei.itb.ac.id/~rinaldi.munir/TA/Makalah_TA_Renjira.pdf) [accessed: 4 March 2023]
- [27]. Salem, M. S. H., Zaman, F. H. K., & Tahir, N. M. (2021). Effectiveness of Human Detection from Aerial Images Taken from Different Heights. *TEM Journal*, 10(2), 522. Doi: 10.18421/TEM102-06