

Kenaf plant pest and disease detection using faster regional based convolutional neural network

Alfita Rakhmandasari¹, Wayan Firdaus Mahmudy², Titiek Yulianti³

^{1,2}Faculty of Computer Science, Brawijaya University, Malang, Indonesia

³Indonesian Sweetener and Fibre Crops Research Institute, Malang, Indonesia

Article Info

Article history:

Received Apr 27, 2021

Revised Aug 10, 2021

Accepted Aug 23, 2021

Keywords:

CNN

Faster RCNN

Kenaf plant disease

Kenaf plant pest

Object detection

ABSTRACT

Kenaf plant is a fibre plant whose stem bark is taken to be used as raw material for making geo-textile, particleboard, pulp, fiber drain, fiber board, and paper. The presence of plant pests and diseases that attack causes crop production to decrease. The detection of pests and diseases by farmers may be a challenging task. The detection can be done using artificial intelligence-based method. Convolutional neural networks (CNNs) are one of the most popular neural network architectures and have been successfully implemented for image classification. However, the CNN method is still considered a long time in the process, so this method was developed into namely faster regional based convolution neural network (RCNN). As the selection of the input features largely determines the accuracy of the results, a pre-processing procedure is developed to transform the kenaf plant image into input features of faster RCNN. A computational experiment proves that the faster RCNN has a very short computation time by completing 10000 iterations in 3 hours compared to convolutional neural network (CNN) completing 100 iterations at the same time. Furthermore, Faster RCNN gets 77.50% detection accuracy and bounding box accuracy 96.74% while CNN gets 72.96% detection accuracy at 400 epochs. The results also prove that the selection of input features and its pre-processing procedure could produce a high accuracy of detection.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Wayan Firdaus Mahmudy

Faculty of Computer Science, Brawijaya University

Veteran street, Ketawanggede, Lowokwaru District

Malang City, East Java 65145 Indonesia

Email: wayanfm@ub.ac.id

1. INTRODUCTION

Kenaf plant (*Hibiscus cannabinus* L. Gaud) is a plant that can produce fibre taken from the bark. This plant was originally used as a raw material for burlap sacks to package agricultural product, the wood was used as fuel, and some countries consumed its leaves as vegetables or used as animal feed [1]. In Indonesia, this plant began to be developed commercially and was used as raw material for gunny sacks since 1978/1979 [2]. Currently the use of Kenaf plants is growing, namely as a raw material for the manufacture of geo-textile, particle board, pulp, fibre drain, fibre board, and high quality paper [3]. The emergence of plastic is one of the factors in the decline in the development of this plant, which originally had an area of 26,000 ha to ± 3000 ha. In addition, things that affect the decline in development are low production due to limited land suitable for development in Java and not yet optimally developed on marginal land, high prices for facilities used for production, relatively low fibre prices, high worker wages, and disturbance from pests and disease.

Pest and disease attack on kenaf plants is one of the obstacles that can hinder production. In Indonesia, pests and diseases that are rife in attacking Kenaf plants are sundapteryx, mites, and leaf blight. Symptoms of these pests and diseases appear on the leaves, stems, or roots that are identified by experts or farmers manually. The method of identifying each expert or farmer can be different which causes the accuracy of low disease detection and manual identification also takes longer because farmers have to come to the laboratory to find out the type of disease and to find out which drugs to use. In this study, the authors will identify Kenaf plant diseases based on a picture of the symptoms that appear on the leaves. This identification is to make is easier for researchers and Kenaf plant farmers to identify a disease quickly and precisely.

Several studies have been conducted by using classification methods to detect plant diseases. These methods include back-propagation neural network that achieves 91% accuracy [4] and 92% accuracy [5], and extreme learning machine which gets 66.67% accuracy [6] and 89.19% accuracy [7]. Support vector machine is used which gets 97.6% accuracy [8] and 92.86% accuracy [9]. Fuzzy logic based method is also used such as in [10] which gets 88% accuracy. While the studies report satisfactory results, there is a possibility to increase the accuracy and also reduce the computational time using more powerful methods.

Optimization techniques are also implemented to increase the performance of classification methods such as modified simulated annealing and extreme learning machine which gets 69.7% accuracy [11], particle swarm optimization and extreme learning machine which gets 79.92% accuracy [12], particle swarm optimization and back-propagation neural network that achieves 96.2% accuracy [13], dempster-shafer optimization using genetic algorithm which gets 87.096% accuracy [14], and fuzzy inference systems optimization using Quasi-Newton and genetic algorithms which gets 94% accuracy [15]. One of drawback using hybrid methods is it tend to require higher computational time.

Convolutional neural networks (CNNs) as part of deep neural networks are one of the most popular neural network architectures and have been successfully implemented for image classification. For example, the implementation of CNNs gets 73% accuracy in [16], 96.3% accuracy in [17], 96.6% accuracy in [18], and 96.6% accuracy in [19]. By using two architectures namely AlexNet and GoogLeNet, the deep convolutional neural network architecture achieves the top accuracy of 99.35% [20].

From the comparison of these methods, the convolutional neural network method was successfully in detecting with high accuracy. However, the CNN method is still considered a long time in the process, so this method was developed into RCNN by adding a selective search method to find the area of the object. Then the development of RCNN, namely fast region convolutional neural network (Fast RCNN) by changing the RCNN classification from SVM to region of interest pooling (ROI) Pooling for its classification [21]. The two methods were redeveloped become a regional proposal network (RPN) to make computing faster. In research on breast cancer detection cases based on MRI image [22], three methods are compared to the speed of the process and the RCNN method can detect an object within 49 second, Fast RCNN 2.3 second, and faster RCNN 0.2 seconds. In addition to these methods, which pay attention to fast processing times and get efficient result is the SSDMobileNet method. However, in the other research that compared the Faster RCNN and SSDMobileNet methods, Faster RCNN can detect it better by getting 95.57% accuracy [23].

Based on this explanation, we conduct research on the use of faster regional based convolutional neural network (Faster RCNN) method to detect kenaf plant diseases based on symptom images taken on the leave to help the performance of kenaf farmers in detecting symptoms of kenaf plant disease. As the selection of the input features largely determines the accuracy of the classification results, this study contributes by developing a pre-processing procedure to transform the kenaf plant image into input features of faster RCNN. A computational experiment is carried out to prove that the selection of input features and its pre-processing procedure could produce a high accuracy of detection.

2. DATA AND METHOD

The data used in study are obtained by direct observation in the kenaf plantation owned by *PT. ABA Lamongan and Plantation* owned by *Balai Penelitian Tanaman Pemanis dan Serat (BALITTAS)* Karangploso, Malang. Then the collected data are processed in several stages such as preprocessing, training, testing, and classification on the faster RCNN as shown in Figure 1.

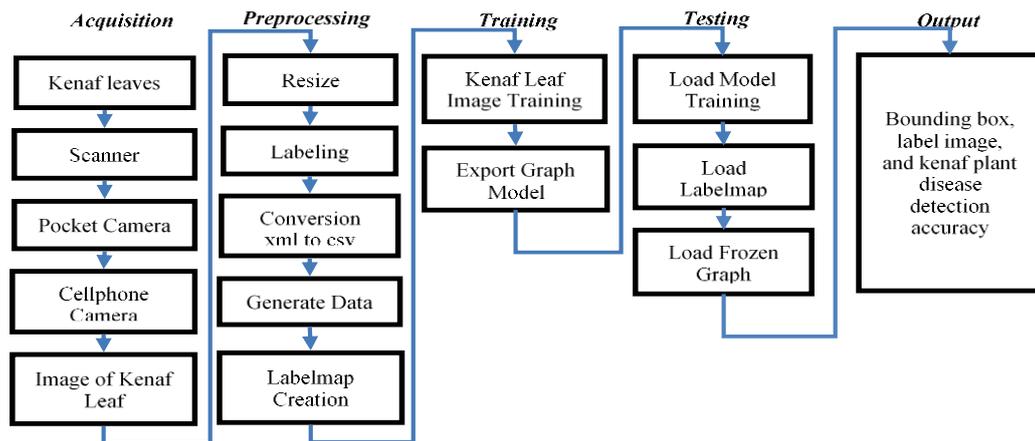


Figure 1. Research diagram

2.1. Data acquisition

The leaves that have been taken during the observation process are then processed to become digital data by taking pictures of the leaves using a scanner, cellphone camera, and pocket camera with white paper on the back of the leaf. The use of various type of cameras is to prove the robustness of the selection of input features and its pre-processing procedure for obtaining the high accuracy of classification. Shooting distance using the camera is about 20-25 cm. The result of data acquisition is shown in Figure 2.

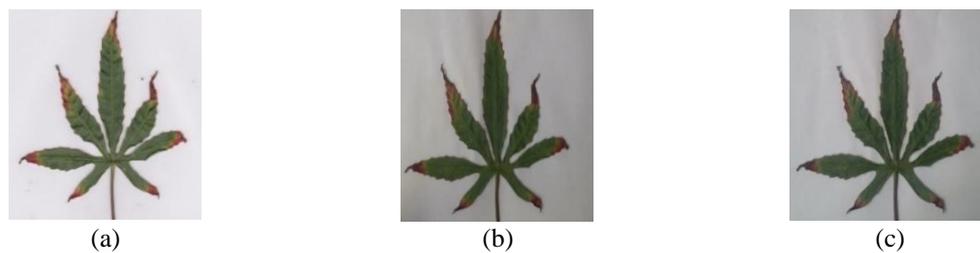


Figure 2. These figures are; (a) scan result; (b) pocket camera result; (c) cellphone camera result

2.2. Preprocessing and training

This stage is carried out after the sample has been in the form of digital data. The results of the leaf image have various sizes, 2548x2664 pixels from scanner, 3264x2448 pixels from cellphone cameras, and 4320x3240 pixels from pocket cameras. An image of these sizes causes the process to run too long, so it takes the resizing process to make the computation process run faster and lighter. The dimensions of the image after resizing are 600x800 pixels. Images that have been reduced in size will be labeled by providing a bounding box for the target object and class name information for each image. There are 3 classes of diseases used, namely leaf blight, sundapteryx, and mites. Each disease class is shown in Figure 3. A same number of images are taken from each device.

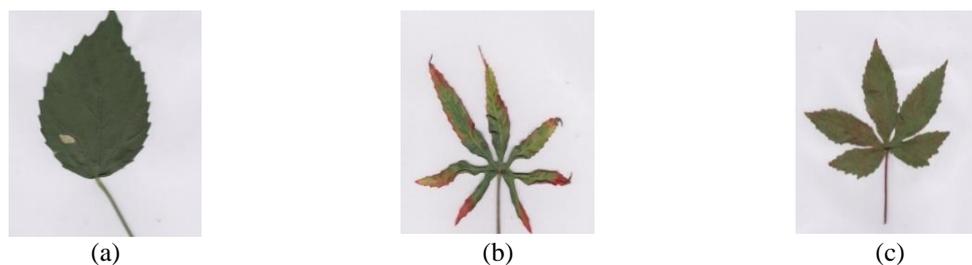


Figure 3. These figures are; (a) leaf blight; (b) sundapteryx; and (c) mites

The training process is run on Tensorflow version 1.15 with the Python programming language on Google colab using the faster RCNN algorithm with the Inception V2 architectural model. Study [24] developed this method by changing the region of selective search proposal in the previous method to region proposal network (RPN) to get the area that allows the object to be detected. The input from the RPN is the feature map resulting from the last convolution layer. The RPN gives the result of several bounding boxes, each of which contains two probability scores for the presence and absence of object at the location. Then these regions are reconstructed using ROI Pooling so that they can be used to classify objects in the proposed region and provide bounding boxes for these objects. The faster RCNN architecture is shown in Figure 4 and the RPN is shown in Figure 5.

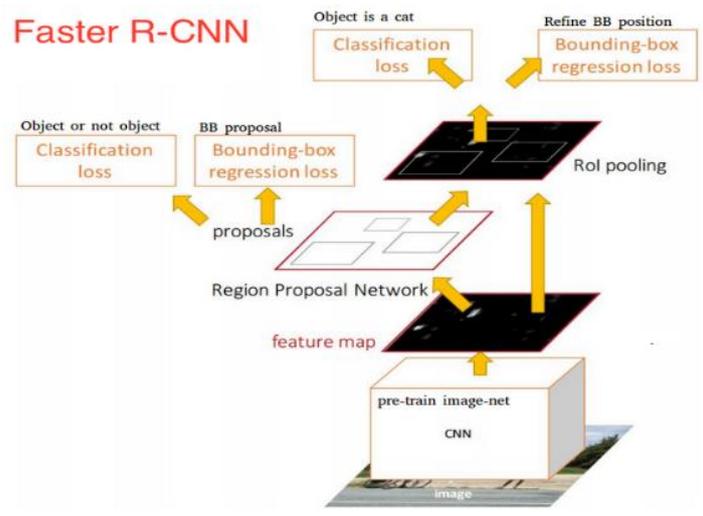


Figure 4. Faster RCNN architecture [25]

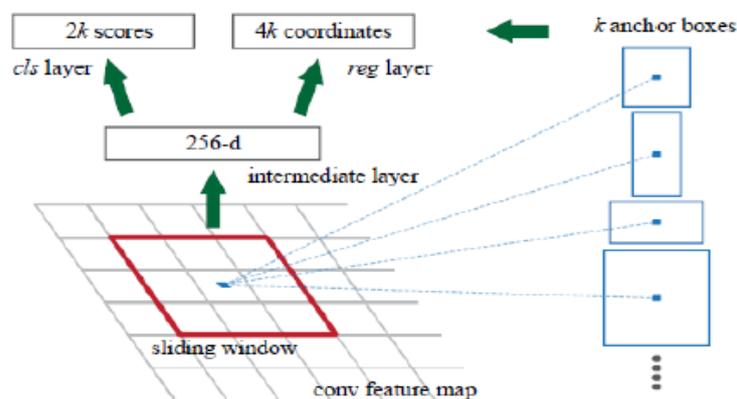


Figure 5. Region proposal network [25]

The training process begins with the input of image data that has been resized to 600x800. The image data is processed using the Convolutional Neural Network method. Then, the feature map of the convolution layer is mapped to the anchor box which will be labeled based on the predetermined ground truth. The result of the background or foreground labeling are determined based on the Intersection over Union (IoU) value which has threshold of 0.6. IoU is calculated by comparing the areas of the true bounding box (tbb) and predicted bounding box (pbb) [26] based on (1) with the illustration shown in Figure 6.

$$IoU = \frac{area(tbb \cap pbb)}{area(tbb \cup pbb)} > 0.6 \tag{1}$$

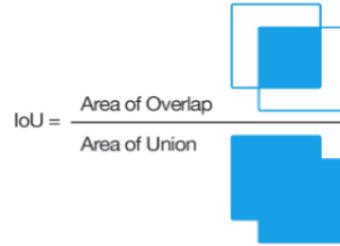


Figure 6. Intersection over union [27]

Anchor box with background labels is not processed by regressors because their true value is low. Regressor Loss Function used is smooth-L1 loss which is located in the upper left corner (x,y) with respect to the ground truth location (t,v) and the logarithm of the depth (w) and height (h) of the box [26] as shown in (2).

$$Lloc(t^v, v) = \sum_{i \in \{x,y,w,h\}}^n smooth_{L1}(t_i^u - v_i) \quad (2)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (3)$$

From this process, a collection of regions with the label foreground is generated. The collection of regions with various sizes from the feature map that has been generated is then processed into ROI Pooling to be reduced to the same size. The results of the ROI Pooling are in the form of a proposal object which is then given a probability value to determine the class and determine the bounding box for each area of the object. The final result is an image with a bounding box containing the object in it and the class name of the object.

2.1. Evaluation metrics

In this study, the evaluation metric used is mean average precision (mAP), precision, and recall. Precision describes the accuracy between data that is predicted to be true positive (TP) by the system with all positive predicted result, namely data that are predicted to be true positive (TP) and predicted false positive (FP). Recall is the success of the system in predicted true positive data from all data that are actually positive, namely data that are predicted to be true positive (TP) and predicted false negative (FN). The equations for precision and recall are shown in (4) and (5). The bounding box issued by the system is a prediction of the boundary of the ground truth and IoU coordinates which are used to set this limit based on the specified threshold. This research model considers all threshold in one metric using mAP. The mAP is a curve that describes precision-recall in one number.

$$precision = \frac{TP}{TP+FP} \quad (4)$$

$$recall = \frac{TP}{TP+FN} \quad (5)$$

In addition, in this study a visual evaluation was also carried out to see the system performance as seen in 20 sample images, each class consisting of 5 images. Visual evaluation results in detection accuracy and bounding box accuracy calculated using (6) and (7) [26].

$$Detection\ accuracy = \left(\frac{TP}{N}\right) \times 100\% \quad (6)$$

$$Bounding\ box\ accuracy = \left(\frac{BB}{TP}\right) \times 100 \quad (7)$$

Detection accuracy is to see the success of the system in detecting a class and the accuracy of the bounding box to see the accuracy of the bounding box in determining the detection location. The accuracy of detection is obtained by comparing the number of diseases detected correctly (TP) and the number of all diseases tested (N). While the accuracy of the bounding box is obtained by comparing the exact location of the bounding box (BB) and the number of diseases detected (TP).

3. RESULTS AND DISCUSSION

The tests are divided into 8 scenarios based on the distribution of the amount of data, different labeling methods, and the distribution of data based on the use of different camera types for data collection. The test scenario is described in Table 1. The data used are 668 data with different numbers for each class. For scenario A and B divided into 532 for training data and 134 for test data. These data are collections of data taken using scanners, cellphone cameras, and pocket digital cameras. The purpose of this test is to see the effect of sharing data with different amounts of data for each class on the results of accuracy. Scenarios C and D contain the same amount of data for each class, namely 58 data for training and 15 data for the test. Scenarios E and F are tests based on data taken using a cellphone camera. For scenario G and H are tests based on data taken using a pocket digital camera. Each scenario uses 2 types of labeling, namely full labeling of 1 leaf and labeling of each leaf finger as shown in Figure 7.

Table 1. Test scenario

Scenario	Data	Labeling method
A	Using 80% for training data and 20% for testing data	One leaf full data labeling
B	Using 80% for training data and 20% for testing data	Labeling of data for each finger
C	Train and test data for each class are same, 58 data for training and 15 data for testing	One leaf full data labeling
D	Train and test data for each class are same, 58 data for training and 15 data for testing	Labeling of data for each finger
E	Train and test data based on the type of cellphone camera	One leaf full data labeling
F	Train and test data based on the type of cellphone camera	Labeling of data for each finger
G	Train and test data based on the type of pocket digital camera	One leaf full data labeling
H	Train and test data based on the type of pocket digital camera	Labeling of data for each finger

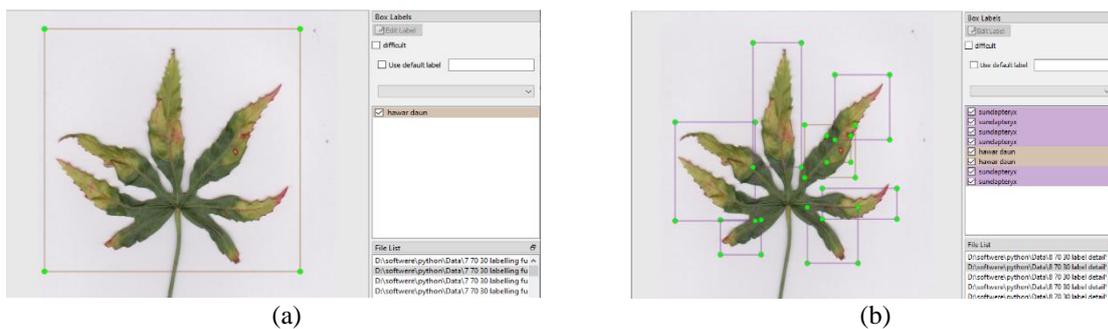


Figure 7. There figures are; (a) full labeling and (b) labeling each leaf finger

From several test scenarios that have been carried out, the values of loss, mAP, recall, detection accuracy and bounding box accuracy are obtained after being run as many as 10000 iterations. The loss value is used to find out how much error a model has. The smaller loss value indicated that the resulting model is getting better. The loss value in Figure 8(a) shows that all scenarios decrease with increasing iterations. However, the test scenario with full labeling has a lower loss value than the test scenario with labeling of each leaf finger.

Mean average precision (mAP) is a value obtained from the comparison between ground truth bounding box and detected bounding box by the system. If the resulting value is higher, the accuracy is also getting better. Figure 8(b) shows that the resulting mAP value increases with increasing iterations. This shows that the model produced by the system can learn to be better. The test scenario with full labeling has a higher mAP value than the test scenario with labeling each leaf finger, but the labeling of each leaf finger moves more stably between iterations.

Average recall (AR) is the average recall in all images, all classes, and all IoU thresholds with a maximum of 1 detection. In Figure 8(c), the AR graph shows the increase in value with each incremental iteration. For data from test scenario with full labeling, AR values are higher than test scenario with labeling each leaf finger. However, the movement of the value is more stable for the labeling of each leaf finger.

The visual evaluation that is carried out resulted in detection accuracy and bounding box accuracy as shown in Table 2. The bounding box is considered correct when all four points match the ground truth. The difference in the results of the bounding box is shown in Figure 9. Figure 9(a) shows a bounding box with full labeling that has the correct position, Figure 9(b) shows the bounding box of the finger labeling indicating the correct position, and (c) shows several bounding boxes generated in 1 prediction. For that case, only 1 is considered true positive and the others are considered false positive.

Table 2. Accuracy results of visual evaluation

Scenario	Detection Accuracy	Bounding Box Accuracy
A	64.16%	83.75%
B	63.295%	96.74%
C	77.502%	72.86%
D	52.27%	94.51%
E	15%	33.35%
F	18.51%	53.33%
G	10%	0%
H	13.19%	39.23%

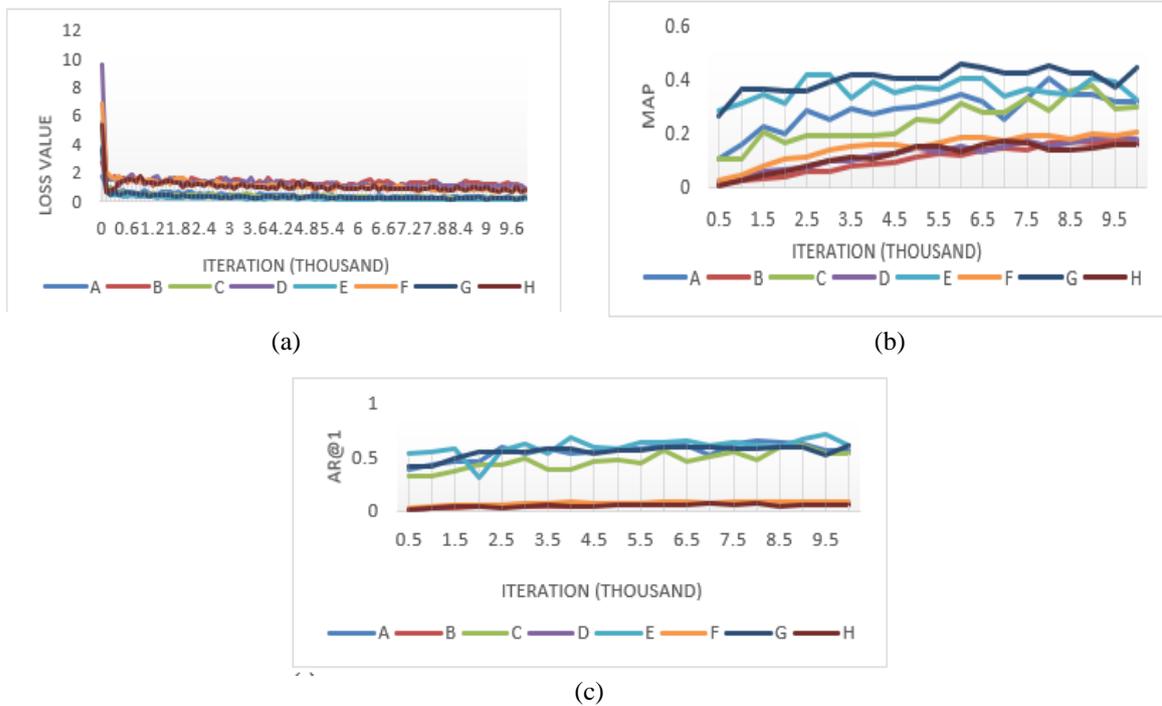


Figure 8. These figures are; (a) comparison of value loss; (b) mAP; (c) average recall

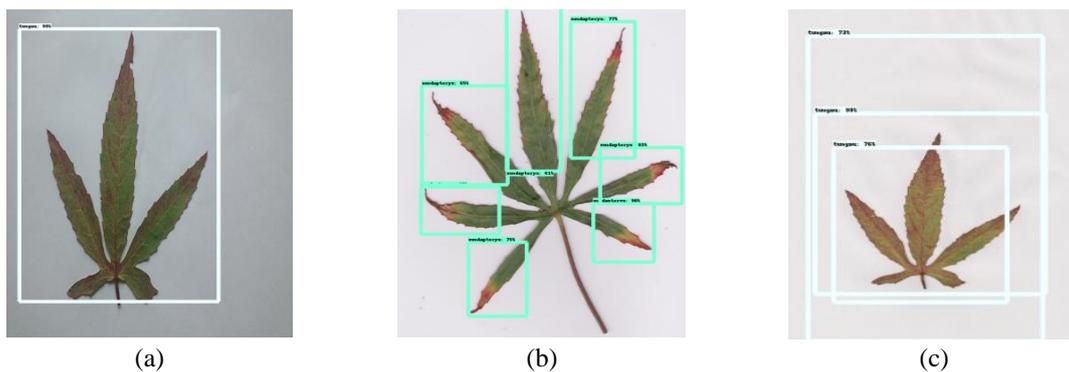


Figure 9. Bounding box; (a) full labeling that has the correct position; (b) finger labeling indicating the correct position; and (c) several bounding boxes generated in 1 prediction

From the results of scenario testing, it is known that the results of the full labeling scenario get better results. On the other hand, the data that was tested based on different types of cameras used to capture the data gave poor results. This is because the amount of data used for training is very small, namely 110 data for all classes.

The best test scenario is the use of the same data sharing for each class, which use 58 data for training and 15 data for testing with full labeling tested into the CNN method with batch 32 and epoch 200 using Google Colaboratory. Because the computation process is too long and the Google Colaboratory usage limit of ± 12 hours had passed since the first process, only 200 epochs are used. Then, 20 leaf images are tested and there are 5 images for each class. The test results obtained a detection accuracy of 45% through visual evaluation. This data has also been used in Fajri, Mahmudy and Yulianti's research [16] using the convolutional neural network method with a training process that is run on a local system of 400 epochs and the time taken is more than 7 hours. When tested on 40 data for each class, the system can predict 122 data correctly. So that it gets an accuracy of 72.96%.

Based on the visual evaluation of the proposed scenario, it can be concluded that the best kenaf plant detection accuracy of 77.50% comes from the scenario of testing data with the same amount of data in each class and the labeling used is full labeling. While the best bounding box accuracy is 96.74% from the test scenario with the use of data divided into 80% for training and 20% for testing and the labeling used is the labeling of each leaf finger. The effect of differences in the amount of data for each class used during training can affect the accuracy of the results. From the test scenario between the use of the same amount of data and the different amount of data each class gets an accuracy of 64.88% and 63.72%.

The effect of different labeling methods on training data can also affect the results of detection accuracy and bounding box accuracy. The detection accuracy for full labeling is on average 70.83% and for labeling each finger leaf an average of 57.78%. However, the average bounding box accuracy for full labeling is 78.30% and the labeling of each leaf finger is 95.62%. From the accuracy results obtained, it can be concluded that data with full labeling can be detected better. Therefore, every detection given to correctly labeled data also has a bounding box with an accurate position as well. What makes the accuracy of data detection with the label of each leaf finger low can also be due to the calculation that each finger must be detected. Indeed, from some images, not all fingers were detected, but enough to represent the detection of its class.

The distribution of data based on the different types of cameras used to capture the image less visible effect and produce poor accuracy. Because the amount of data captured by each type of camera is very small for training, it can lead to many images that are not detected properly or even many images are not detected at all. Therefore, an additional preprocessing mechanism is needed to produce relatively the same values for input from various types of cameras. When compared to CNN in the best test scenario, faster RCNN gets better results and faster computation time. Training via google colaboratory, faster RCNN can complete 10000 iterations for 3 hours. While CNN completed 100 epochs with a batch size of 32 and got a detection accuracy of 45%. When run through the local system, CNN achieved 72.96% accuracy at epoch 400 for more than 7 hours.

4. CONCLUSION

This study proves that the faster RCNN provides better results than CNN. The faster RCNN gets 77.50% detection accuracy and bounding box accuracy 96.74% while CNN gets 72.96% detection accuracy at 400 epochs. The results also prove that the selection of input features and its pre-processing procedure could produce a high accuracy of detection. Furthermore, the computational experiment proves that the faster RCNN has a very short computation time by completing 10000 iterations in 3 hours compared to convolutional neural network (CNN) completing 100 iterations at the same time.

REFERENCES

- [1] J. M. Dempsey, "Fibre Crops," USA: The University Presses of Florida, Gainesville, 1975.
- [2] E. L. Arumingtyas, "Kenaf: its prospect in Indonesia," *Journal of Biological Researches*, vol. 20, no. 2, pp. 21-26, 2015.
- [3] A. H. N. Huda, M. T. Ramlah, A. Zakiah, G. S. Aisyah and S. E. Suhaila, "Reclamation of Nutrient from Kenaf Retted Wastewater using Newly Designed Ultrafiltration Membrane," in *5th Brunei International Conference on Engineering and Technology*, 2014, pp. 1-6, doi: 10.1049/cp.2014.1063.
- [4] R. Gupta and T. Verma, "Tomato Leaf Disease Detection using Back Propagation Neural Network," *Int. Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 8, pp. 529-538, 2020, doi: 10.35940/ijitee.h6531.069820.
- [5] V. Ramya and M. A. Lydia, "Leaf Disease Detection and Classification using Neural Networks," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 11, pp. 207-210, 2016, doi: 10.11113/jt.v77.6463.
- [6] T. H. Saragih, D. M. N. Fajri, W. F. Mahmudy, A. L. Abadi, and Y. P. Anggodo, "Jatropha curcas disease identification with extreme learning machine," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 12, no. 2, pp. 883-888, 2018, doi: 10.11591/ijeecs.v12.i2.pp883-888.

- [7] A. Bhatia, A. Chug, and A. Prakash Singh, "Application of extreme learning machine in plant disease prediction for highly imbalanced dataset," *Journal of Statistics and Management Systems*, vol. 23, no. 6, pp. 1059-1068, Aug. 2020, doi: 10.1080/09720510.2020.1799504.
- [8] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41-49, 2017, doi: 10.1016/j.inpa.2016.10.005.
- [9] M. Khan, "Detection and Classification of Plant Diseases Using Image Processing and Multiclass Support Vector Machine," *International Journal of Computer Trends and Technology*, vol. 68, no. 4, pp. 5-11, Apr. 2020, doi: 10.14445/22312803/IJCTT-V68I4P102.
- [10] V. Mahajan and N. R. Dhumale, "Leaf Disease Detection Using Fuzzy Logic," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 7, no. 6, pp. 6801-6807, 2018, doi: 10.1109/SPIN.2015.7095350.
- [11] T. H. Saragih, W. F. Mahmudy, A. L. Abadi, and Y. P. Anggodo, "Application of extreme learning machine and modified simulated annealing for jatropha curcas disease identification," *International Journal of Advances in Soft Computing and its Applications*, vol. 10, no. 2, 2018.
- [12] M. W. Alauddin, W. F. Mahmudy, and A. L. Abadi, "Extreme Learning Machine Weight Optimization Using Particle Swarm Optimization to Identify Sugar Cane Disease," *Journal of Information Technology and Computer Science*, vol. 4, no. 2, pp. 127-136, 2019, doi: 10.25126/jitecs.201942116.
- [13] M. Chanda and M. Biswas, "Plant disease identification and classification using Back-Propagation Neural Network with Particle Swarm Optimization," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1029-1036, doi: 10.1109/ICOEI.2019.8862552.
- [14] T. H. Saragih, W. F. Mahmudy, and Y. P. Anggodo, "Optimization of Dempster-Shafer's believe value using genetic algorithm fo identification of plant diseases Jatropha curcas," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 12, no. 1, pp. 61-68, 2018, doi: 10.11591/ijeecs.v12.i1.pp61-68.
- [15] J. B. Velandía, C. E. R. Calderón, and D. D. L. Lara, "Optimization techniques on fuzzy inference systems to detect Xanthomonas campestris disease," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 4, pp. 3510-3518, 2021, doi: 10.11591/ijece.v11i4.pp3510-3518.
- [16] D. M. N. Fajri, W. F. Mahmudy, and T. Yulianti, "Detection of Disease and Pest of Kenaf Plant using Convolutional Neural Network," *Journal of Information Technology and Computer Science*; vol. 6, no. 1, pp. 18-24, 2021.
- [17] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Computational Intelligence and Neuroscience*, vol. 1, no. 1, pp. 1-11, 2016, doi: 10.1155/2016/3289801.
- [18] P. Bharali, C. Bhuyan, and A. Boruah, "Plant Disease Detection by Leaf Image Classification Using Convolutional Neural Network BT-Information, Communication and Computing Technology," *Springer Singapore*, 2019, pp. 194-205.
- [19] M. S. AL-Huseiny, N. K. Abbas, and A. S. Sajit, "Diagnosis of arrhythmia based on ECG analysis using CNN," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 9, no. 3, June 2020, doi: 10.11591/eei.v9i3.2172.
- [20] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016, doi: 10.3389/fpls.2016.01419, [Online]. Available: <https://www.frontiersin.org/article/10.3389/fpls.2016.01419>
- [21] Y. P. Chen, Y. Li, and G. Wang, "An Enhanced Region Proposal Network for object detection using deep learning method," *PLOS ONE*, vol. 13, no. 9, p. e0203897, Sep. 2018, doi: 10.1371/journal.pone.0203897, [Online]. Available: <https://doi.org/10.1371/journal.pone.0203897>
- [22] K. G. Chaudhari, "Comparitive analysis of CNN models to diagnose Breast Cancer," *International Journal of Innovative Research in Science*, vol. 7, no. 10, pp. 8180-8187, 2018, doi: 10.2139/ssrn.3729056.
- [23] S. Ghoury, C. Sungur, and A. Durdu, "Real-Time Diseases Detection of Grape and Grape Leaves using Faster R-CNN and SSD MobileNet Architectures," in *International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES 2019)*, 2019, pp. 39-44.
- [24] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Neural Information Processing Systems (NIPS)*, pp. 1-14, 2015, doi: 10.1109/TPAMI.2016.2577031.
- [25] D. Alamsyah and M. Fachrurrozi, "Faster R-CNN with Inception V2 for Fingertip Detection in Homogenous Background Image," *Journal of Physics: Conference Series, International Conference on Information System, Computer Science and Engineering*, vol. 1196, no. 1, 2019, pp. 1-6, doi: 10.1088/1742-6596/1196/1/012017.
- [26] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 658-666.
- [27] R. Padilla, S. Netto, and E. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *International Conference on Systems, Signals and Image Processing*, Jul. 2020, pp. 1-7.

BIOGRAPHIES OF AUTHORS

Alfita Rakhmandasari, obtained a Bachelor of Computer Science degree in Computer Science, Brawijaya University, Indonesia in 2016. Currently studying to take a master's degree in Master of Computer Science, Brawijaya University, Indonesia. Her research interests are machine learning and image processing.



Wayan Firdaus Mahmudy, obtained a Bachelor of Science degree from the Mathematics Department, Brawijaya University in 1995. His Master in Informatics Engineering degree was obtained from the Sepuluh Nopember Institute of Technology, Surabaya in 1999 while a PhD in Manufacturing Engineering was obtained from the University of South Australia in 2014. He is a Senior Lecturer at Department of Computer Science, Universitas of Brawijaya (UB), Indonesia. His research interests include optimization of combinatorial problems and machine learning.



Titiek Yulianti, obtained a bachelor's degree in agriculture from the Plant Pests and Diseases study program in 1983 from the Bogor Agricultural Institute. She took his Masters in Plant Pathology study program at the University of Melbourne Australia in 1997. She obtained doctoral degree from the University of Western Australia in the Biological Control study program. She is a principal researcher in *Badan Penelitian Tanaman Pemanis dan Serat* (Balitas), Karangploso, Malang, Indonesia in the field of research on pests and plant diseases.