World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Research Article)

Efficient detection of cacao pod diseases using SSD MobileNetV2 FPN-Lite

Glenn Matthew Garma *

College of Computing Studies, Information and Communication Technology, Isabela State University, Cauayan Campus, Cauayan City, Isabela.

World Journal of Advanced Research and Reviews, 2025, 25(02), 1099-1105

Publication history: Received on 02 December 2024; revised on 01 February 2025; accepted on 04 February 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.25.2.0128

Abstract

Detecting and managing cacao pod diseases is an important task for improving crop yield and quality, especially in regions where agriculture serves as a primary livelihood. In this paper we introduce an object detection model based on the SSD MobileNetV2 FPN-Lite architecture for efficient and accurate detection of cacao pod diseases, focusing on "monilla" and "phytophtora". The model used a feature pyramid network (FPN) to enhance multi-scale detection capabilities to enable the identification of small objects such as early-stage disease symptoms. Evaluation metrics, including mAP, box loss, and classification loss, were used to assess the model's performance. The proposed framework achieved an mAP of 0.83, demonstrating its effectiveness in detecting various cacao pod diseases. With its low computational overhead, the model is optimized for deployment on edge devices, making it a viable solution for real-time disease monitoring in agricultural settings.

Keywords: Convolutional Neural Network; Cacao; Mobilenet; Deep Learning; Artificial Intelligence

1. Introduction

Cacao (Theobroma cacao L.), often referred to as the "food of the gods", is a perennial crop with significant global market potential. Its beans are highly valued and processed into various products, including tablea, cocoa powder, nibs, butter, paste or liquor, and chocolate confectionery. Cacao is developing as an important crop in several countries, particularly in the Philippines, which has competitive advantage for cacao production due to its strategic location, good climatic conditions, and favorable soil. However, several problems cause the country to be unable to meet the current requirements of the growing industry. One of these problems is the prevalence of pests and diseases [1] which destroy as much as 40% of the country's output of cacao [2][3]. Common diseases of cacao in the Philippines are Cacao Black Pod Rot, Cacao Swollen Shoot Virus Disease, and Frosty Pod Rot. Around 80% of cacao growers in the country are small holder farmers [4][5] which for them, managing diseases is a bigger problem compared to larger plantations which has funds and technical expertise. There are many methods being used to address the issues of diseases of Cacao but several experts agrees that early detection and management, such as pruning, is better and to avoid constantly using pesticides, which increases the risk of pesticide resistance [6][7][8]. In recent years, advancements in artificial intelligence (AI) and machine learning have been explored as potential solutions for addressing agricultural challenges. Object detection models, such as the Single Shot Multibox Detector (SSD)[9], have demonstrated capabilities in identifying and classifying plant diseases in research settings. When combined with lightweight architectures like MobileNetV2 [10], these models show promise for deployment on mobile devices, which could make disease detection technology more accessible to farmers, including those in remote areas. This study focuses on developing a mobile application to detect cacao diseases using SSD MobileNetV2 with a Feature Pyramid Network (FPN)[11] Lite backbone, modified to incorporate dilated convolutions. This architecture is designed to enhance the detection of small-scale disease patterns, which are often challenging to identify in real-world field conditions. By leveraging the model's efficiency and accuracy, this research

* Corresponding author: Glenn Matthew Garma

aims to empower cacao farmers with a practical tool for early disease detection and management, ultimately contributing to the sustainability of cacao farming in the Philippines.

2. Material and methods

2.1. Cacao Dataset

Figure 1 Cacao Datasets

The image dataset used in training and validating the object detection model was acquired from the Cocoa Diseases dataset [12][13]. The dataset consists of images of diseases phytophtora and monilla, as well as images of healthy cacao pods. The original dataset consists atotal of 312 images and 1,591 labeled objects. Figure 1 shows a snapshot of the original dataset.

The dataset for training and validating the object detection model were prepared by first applying important preprocessing step. Auto-orientation ensured that all images were aligned correctly based on their metadata, which is critical for maintaining consistency across the dataset. The images were then resized to a fixed dimension of 640x640 pixels. This ensured uniformity in input size which is a requirement for the object detection model [14]. In addition to preprocessing, augmentation techniques were employed to enhance the diversity and robustness of the dataset. These augmentations simulated real-world conditions, thereby increasing the model's ability to generalize across varied scenarios. Horizontal flipping was applied to generate alternative viewing perspectives of the cacao pods which increased the dataset's variability. Random rotations between -15° and +15° were performed to mimic natural variations in orientation that might occur during image capture. Adjustments to saturation levels, ranging from -25% to +25%, were applied to account for differences in color intensity due to environmental factors. Lastly, exposure adjustments within a range of -10% to +10% were applied to consider the changes in lighting conditions often encountered when capturing images in the field. A summary of the augmentation processes is shown in table 1.

Table 1 Automation Process

Process	Description
Horizontal Flipping	Flip images horizontally
Rotation	Rotate images between -15° and +15°
Saturation Adjustment	Adjust saturation between between -25% and +25%
Exposure Adjustment	Adjust exposure between -10% and +10%

After preprocessing and augmentation, the dataset was expanded and split into three subsets: a training set comprising 753 images, a validation set with 50 images, and a test set containing 15 images. These subsets ensured a balanced approach to training and evaluating the object detection model. Figure 2 shows sample images from the dataset after the augmentation processes

Figure 2 Sample images from the augmented dataset

2.2. MobileNet V2 SSD FPNLite

This study used a Single Shot Multibox Detector (SSD) MobileNetV2 with Feature Pyramid Network-Lite (FPN-Lite) architecture. This model was chosen for its balance between computational efficiency and detection accuracy which is suitable for deployment on mobile and edge devices. The MobileNetV2 backbone network provides a lightweight yet effective feature extractor, while the FPN-Lite enhances the model's ability to detect small objects by aggregating multi-scale feature representations. MobileNetV2 SSD FPN-Lite integrates depthwise separable convolutions in its backbone, significantly reducing the number of parameters compared to standard convolutional networks. The MobileNetV2 backbone introduces inverted residual blocks with linear bottlenecks, ensuring both high efficiency and strong feature extraction capabilities. These blocks improve the model's ability to capture essential features while maintaining a low computational cost. Dilated Convolutions are used in the backbone network to improve the ability to extract features by increasing its receptive field which can enhance the model's sensitivity to fine details[15][16]. The FPN-Lite component is designed to improve small object detection by constructing a top-down pathway that merges high-level semantic information with low-level spatial details. By leveraging multi-scale feature maps, the network can effectively localize and classify objects across varying sizes.

2.3. Model Performance Metrics

To evaluate the performance of the model, the following metrics[17] were utilized, with their corresponding equations adapted from the study[18]:

Precision and Recall: Precision and recall are essential metrics for evaluating the performance of an object detection model. They are used to assess classifier performance in both binary and multiclass classification tasks. Precision evaluates the accuracy of positive predictions, whereas recall measures the model's ability to capture all actual positive instances.

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

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

Mean Average Precision (mAP): The mean of the Average Precision (AP) values calculated for all object classes. It quantifies how well the model balances precision and recall across all classes. AP is represented as:

$$AP = \int_0^1 Precision(r)\, dr$$

Where *r* represents recall, and the integration is over the precision-recall curve. mAP is computed as:

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

Where N is the number of classes.

Box Loss and Classification Loss: Box Loss measures the error in the predicted bounding box coordinates compared to the ground truth while Class Loss quantifies the error in assigning the correct class label to each detected object. The sum of Box Loss and Class Loss represents the total error of the model during training. [19][20]

3. Results and discussion

The model training lasted for 10,200 steps until it was manually stopped due to evident overfitting wherein the Average Precision and Recall of the model on the validation data started to gradually decrease and the Validation Loss started to gradually increase. Figure 3 Shows a graph representing the mAP of the model across all the steps during training. It shows a gradual increase of the mAP which started to plateu on the 5,000th step. The highest mAP recorded during training is 84.62%.

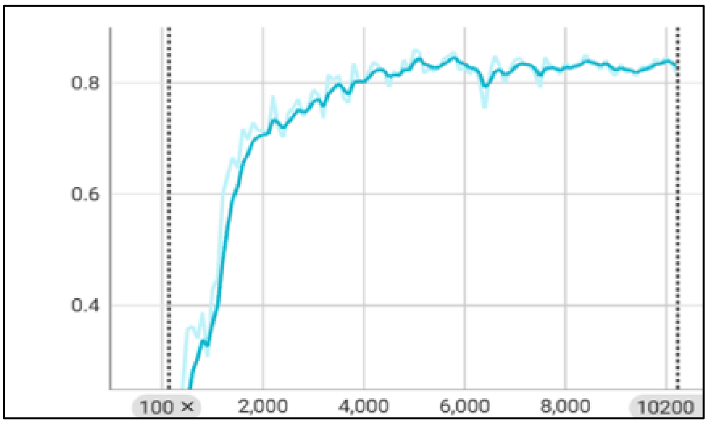

Figure 3 mAP of the Model across 10200 Steps

The Box loss of the model as shown in Figure 4 consistently decreased throughout the training process, reaching the lowest value of 0.0035. This indicates that the model showed steady improvement in localizing the objects withing the bounding boxes.

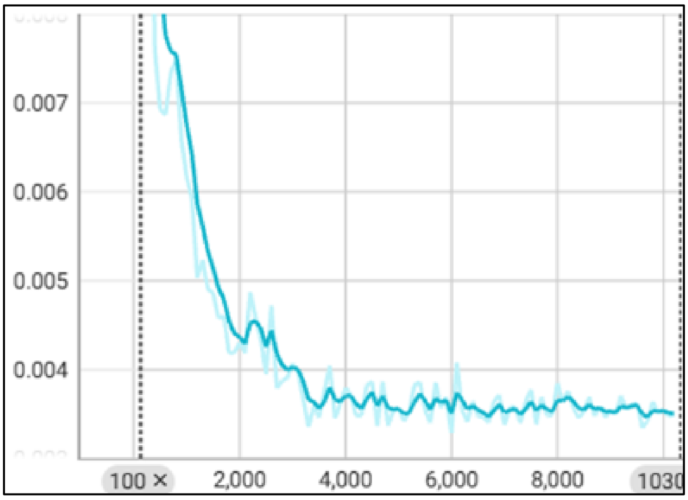

Figure 4 Box Loss of the model throughout training

The Class Loss as shown in Figure 5 started at a higher value and gradually decreased, with the highest value at 0.3256 value until it started to increase again. This suggests that the model was able to learn the object categories to a reasonable extent, but misclassifications still occurred at a moderate rate especially on the near final steps of the training process.

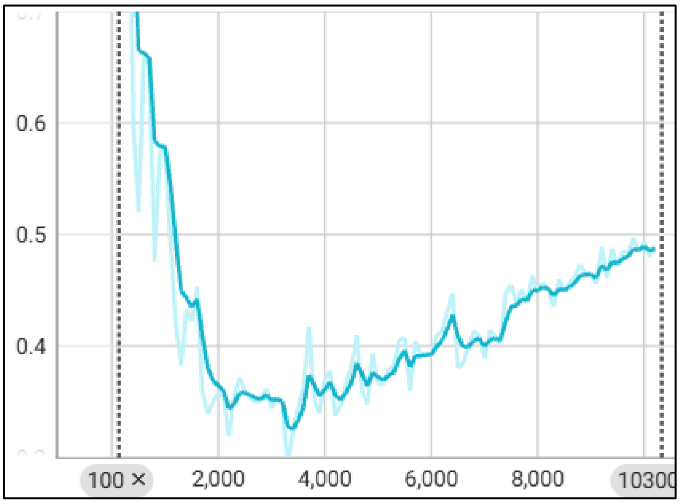

Figure 5 Class Loss of the model throughout training

3.1. Model Testing

The trained model was tested on the test dataset to visualize the prediction bounding boxes of the model. The test dataset consists of the images that were not included in the training data nor in the validation data. Based on the testing as shown in figure 5, the model accurately identify whether a certain cocoa pod is healthy or damaged by either monilla or phytophtora disease.

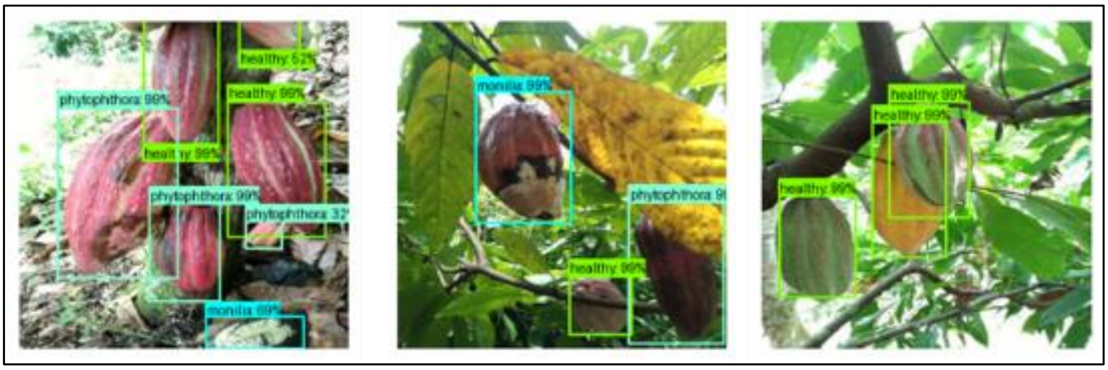

Figure 6 Object Detection on Test Dataset

4. Conclusion

This study demonstrates the effective detection of cacao pod diseases using the SSD MobileNetV2 FPN-Lite architecture. By utilizing a lightweight architecture and leveraging feature pyramid networks, the model achieved high accuracy in localizing and classifying cacao diseases, making it suitable for deployment on mobile and edge devices. The model's mean Average Precision (mAP) achieved highest value of 84.62%, indicating a strong overall detection performance. The box loss consistently decreased, reflecting the model's improvement in bounding box localization, with the lowest value of 0.0035 representing a high level of localization accuracy. On the other hand, the class loss lowest value of 0.3256. This suggests that while the model effectively learned to classify the objects, some misclassifications persisted, especially as the model reached higher training steps and began to overfit.

The final model was tested on unseen dataset, and the object detection results indicated accurate identification of both healthy and diseased cacao pods even on the ones which are further from the camera. The model successfully distinguished between Monilla and Pythophtora, showing its practical applicability in real-world scenarios for early disease detection.

Future work could focus on refining the classification component, such as by experimenting by other techniques to enhance the model's classification accuracy. The Current Dataset could be expanded by collecting more images of cacao pods at different stages of disease progression, as well as from different environmental conditions. The dataset could also be enhanced to detect other common cacao diseases. The trained object detection model in this study is mainly optimized for deployment on mobile and edge devices. This study could be extended by implementing the model on mobile devices and evaluate th scalability and performance of the application. The detection tool could also be integrated with other technology to provide farmers with actionable insights and decision support.

References

[1] Cacao – Industry Strategic Science and Technology Plans (ISPs) platform. 2024. Available from: https://ispweb.pcaarrd.dost.gov.ph/isp-commodities/cacao/

[2] International Cocoa Organization. Pests & Diseases. 2023 May 12. Available from: https://www.icco.org/pests-diseases/

[3] BusinessMirror. Pests, diseases destroy up to 40% of food crops–FAO. 2022 Mar 30. Available from: https://businessmirror.com.ph/2022/03/30/pests-diseases-destroy-up-to-40-of-food-crops-fao/

[4] Irader MKP. Agencies address cacao pest management concerns. DOST-PCAARRD Portal. 2016 Apr 13. Available from: https://www.pcaarrd.dost.gov.ph/index.php/quick-information-dispatch-qid-articles/agencies-address-cacao-pest-management-concerns

[5] Bobiles S, Elegado F, Millena C, Merca F. Small-scale cacao (Theobroma cacao L.) fermentation process utilizing cacao pod husk. Food Research. 2022;6(4):236–245. Available from: https://doi.org/10.26656/fr.2017.6(4).502

[6] Solpot TC. Cacao (Theobroma cacao) diseases: threats to the Philippines cocoa industry. Philippine Plant Disease Bulletin. 2020;1(1):2-4.

[7] Niogret J, Sanchez V, Marelli J. Increasing the resilience of cacao to major pest and disease threats in the 21st century. ACIAR Proceedings, No. 149, Proceedings of an Asia-Pacific Regional Cocoa IPM Symposium. 2020.

[8] Babin R. Pest management in organic cacao. In: CAB International eBooks. 2017. p. 502–518. Available from: https://doi.org/10.1079/9781780644998.0502

[9] Hoopen GMT, Krauss U. Biological control of cacao diseases. In: Springer eBooks. 2016. p. 511–566. Available from: https://doi.org/10.1007/978-3-319-24789-2_17

[10] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C, et al. SSD: Single Shot MultiBox Detector. In: Lecture Notes in Computer Science. 2016. p. 21–37. Available from: https://doi.org/10.1007/978-3-319-46448-0_2

[11] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: Inverted residuals and linear bottlenecks. arXiv. 2018. Available from: https://doi.org/10.48550/arxiv.1801.04381

[12] Yang L, Zhang X, Wang L, Zhu M, Li J. Lite-FPN for Keypoint-based monocular 3D object detection. arXiv. 2021. Available from: https://doi.org/10.48550/arxiv.2105.00268

[13] Hernán TSL, Enrique MCF, Edgardo NPP, Feisar-Enrique-Moreno-Corzo. Prototipo de aplicación móvil para la identificación de mazorcas de cacao enfermas haciendo uso de visión por computadora y aprendizaje de máquina. 2020. Available from: http://hdl.handle.net/20.500.12749/13367

[14] Cocoa Diseases - Dataset Ninja. 2024. Dataset Ninja. Available from: https://datasetninja.com/cocoa-diseases

[15] Keras documentation: Object Detection with RetinaNet. Available from: https://keras.io/examples/vision/retinanet/

[16] Yu F, Koltun V. Multi-Scale context aggregation by dilated convolutions. arXiv. 2015. Available from: https://doi.org/10.48550/arxiv.1511.07122

[17] Yen C, Tsao C. Lightweight convolutional neural network for chest X-ray images classification. Sci Rep. 2024;14(1).

[18] Classification: Accuracy, recall, precision, and related metrics. Google for Developers. 2024. Available from: https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall

[19] Vera DB, Oviedo B, Casanova WC, Zambrano-Vega C. Deep Learning-Based Computational Model for Disease Identification in Cocoa Pods (Theobroma cacao L.). arXiv. 2024. Available from: https://doi.org/10.48550/arxiv.2401.01247

[20] Jiang S, Qin H, Zhang B, Zheng J. Optimized loss functions for object detection and application on nighttime vehicle detection. Proc Inst Mech Eng Part D J Auto Eng. 2021;236(7):1568–1578. Available from: https://doi.org/10.1177/09544070211036366