

## Chili Plant Classification using Transfer Learning models through Object Detection

Amirul Asyraf Abdul Manan<sup>1</sup>, Mohd Azraai Mohd Razman<sup>1</sup>, Ismail Mohd Khairuddin<sup>1\*</sup> and Muhammad Nur Aiman Shapiee<sup>1</sup>

<sup>1</sup>Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.

**ABSTRACT** – This study presents an application of using a Convolutional Neural Network (CNN) based detector to detect chili and its leaves in the chili plant image. Detecting chili on its plant is essential for the development of robotic vision and monitoring. Thus, helps us supervise the plant growth, furthermore, analyses their productivity and quality. This paper aims to develop a system that can monitor and identify bird's eye chili plants by implementing machine learning. First, the development of methodology for efficient detection of bird's eye chili and its leaf was made. A dataset of a total of 1866 images after augmentation of bird's eye chili and its leaf was used in this experiment. YOLO Darknet was implemented to train the dataset. After a series of experiments were conducted, the model is compared with other transfer learning models like YOLO Tiny, Faster R-CNN, and EfficientDet. The classification performance of these transfer learning models has been calculated and compared with each other. The experimental result shows that the Yolov4 Darknet model achieves mAP of 75.69%, followed by EfficientDet at 71.85% for augmented dataset.

### ARTICLE HISTORY

Received: 12<sup>th</sup> Nov 2020

Revised: 1<sup>st</sup> Dec 2020

Accepted: 17<sup>th</sup> Dec 2020

### KEYWORDS

*Object detection*

*Precision agriculture*

*Transfer learning*

*Crop monitoring*

*Chili plant*

## INTRODUCTION

Agricultural is a potential commodity that has high economic value and a high potential for development. From the supply or production side, Malaysia land area with its Agro-climate diversity enables the development of various crops. Based on the Department of Statistics Malaysia, the agriculture sector contributed 7.1% to the Gross Domestic Product (GDP) in 2019, equal to RM 101.5 billion. Pepper production is one of several selected crops that show the increment of production from 2018 to 2019 [1]. In addition, the production of chillies has slightly escalated from 32.3 thousand tonnes in 2018 to 33.9 thousand tonnes in 2019. In chili production, farmers are the ones who are manually sorting the chili into their categories. Furthermore, this task also involves humans as decision makers to identify the quality of the chili, resulting in manually identifying a lot of time-consuming and labour-intensive tasks. Besides, due to the human visual limitations, product variety is also obtained, and it may also causing inconsistency in the selection process.

The application of object detection in agriculture is to monitor the condition of the crops, especially on a larger production scale. This research seeks to address how object detection can help reduce the time consumption of chili plant identification for crop production and help increase the chili production rate. Therefore, this study proposes to monitor the plant growth rate through image processing. Hence, will ensure that the condition of a plant, which is healthy and non-healthy, can be determined from time to time.

## RELATED WORK

Unmanned Aerial Vehicle (UAV) is equipped with many sensors that can make a notable impact in field-based crop monitoring, providing the information and data from above such as phenotyping, estimation of plant density, or plant quantity proper and systematic order [2]. Many machine learning approaches have been introduced to solve classification problems through image processing and object detection, which employed classifiers such as Support Vector Machine (SVM) and  $k$ -Nearest Neighbor ( $k$ -NN) and not just for crop monitoring [3], [4]. Additionally, machine learning models have improved their accuracy and shown an excellent performance in classifying object-based in a wide range of applications [5]. Besides, there is much more potential for their use in agriculture-focused analysis. Many researchers used machine learning to solve agricultural-related problems such as leaf counting, plant counting, leaf segmentation, and field yield prediction.

On the features extraction aspects, it was studied to use five basic geometric approaches [6]. The detection features they used are the leaf diameter, width, length, perimeter, and area. Another study has proposed using Probabilistic Neural Network (PNN) to classify leaves that used the shape, vein, colour, and texture features in which the author manage to obtain a 93.75% average accuracy [7]. Likewise, research was done to study leaf biometric features and use hidden naive Bayes (HNB) as a classifier [8]. The experiment data shows an accuracy of 97% by using HNB performed with the Flavia dataset.

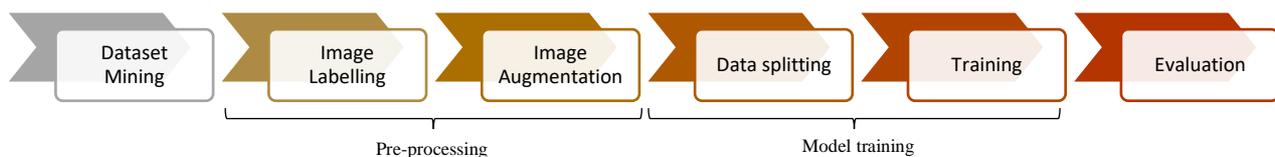
Nowadays, Convolutional Neural Network (CNN) outperform all other traditional methods in computer vision tasks [9], with some of them superior to human ability on ImageNet image classification challenge [10]. Image classification is a task that trying to analyse an entire image as a whole [11]. The image will classify the image by assigning it a specific label obtained by labelling or classifying images into pre-determined categories [12]. AlexNet [13] is one of the convolutional networks that perform image classification. Object detection, however, involves the combination of classification and localisation of various objects in an image. Many object detector convolutional neural networks use a sliding window approach to inspect different regions of an image sequentially [13]. As a result, reducing their speed but have higher accuracy output. Some object detectors such as YOLO take another approach where the model examines the whole image and its context simultaneously, significantly speeding up the process [14].

Deep learning, such as object detection algorithms, is a great way to detect plant health and diseases. A study was conducted to detect diseases from plant leaves that proposed method using four different types of models for the detection, which are DarkNet-19, MobileNet, Inception v2, and ResNet-101. Finally, Darknet-53 achieved a 99.10 % of success rate in the classification [15].

A proposed method of using CNN for the law-protected tree species detection in which they used three different object detection methods, Faster Region-based Convolutional Neural Network (Faster R-CNN), You Only Look Once (YOLO), and RetinaNet [16]. The models were then evaluated, and the acquired around 92% of average precision.

## METHODOLOGY

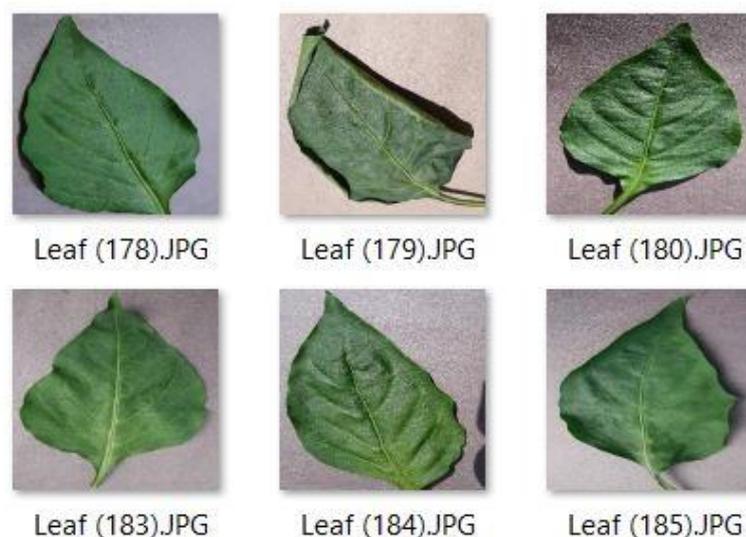
This study aims to answer the research objective to monitor and identify the plants and one or more object detection models trained to detect bird's eye chili. Nonetheless, bird's eye chili classification consists of four steps, namely dataset mining, pre-processing, model training, and evaluation. The processes done is shown in Figure 1 where dataset mining will begin before the pre-processing stage starts. The image labelling will provide the source of images between leaf and chili whereas the image augmentation section is to increase the number images to compare with the initial set of plain composition. The model training will consist of data splitting of the images in turns provide the training of the models. Lastly, the evaluation part will compared the performance of each models and determine the optimum classifier that could detect the chili plant.



**Figure 1.** Flow of machine learning process

### Dataset mining

A total of 324 healthy and unhealthy capsicum leaf images as shown in Figure 2 and 298 images of bird' eye chili as in Figure 3 have been collected from capturing images of bird' eye chili from various angles and open-source for model training [17].



**Figure 2.** Sample from PlantVillage Dataset



**Figure 3.** Bird' eye chili images sample

### Pre-processing

The images are then labelled by drawing a bounding box around the objects found in the image. This process was used to allow the object detection algorithm to learn to identify the trained objects. Next, the images were augmented by rotating them at 90-degree angles and upside down and flipping them horizontally. Thus, the total of the images dataset is 1,866 images.

### Model Training

In this section, a python code called training data splitter was used to split the dataset [18]. The dataset was randomly split into three sets: training set, verification and a test set with a distribution of 70, 20 and 10 percent of the images, respectively. A validation set was used for the model's evaluation in the selected iteration of the neural networks.

You Only Look Once, more commonly known as YOLO, is an algorithm used to detect objects of various sizes. YOLO is an object detection architecture composed of a single neural network [14]. The approach YOLO utilised is a single convolutional network that can simultaneously anticipate numerous bounding boxes and makes the model look once at a picture to foresee what objects are shown and where they are. Hence, it makes YOLO fast in detecting objects. Fast YOLO [14] or Tiny YOLO is an alternate implementation of YOLO V4 where it is less accurate but has faster training rate. The Faster Region-based Convolutional Network (Faster R-CNN) is an image processing algorithm that was derived from the Fast R-CNN model [19]. The Fast R-CNN is developed over previous concepts of R-CNN but with more efficiency added to identify objects. The model list from EfficientDet-D0 until EfficientDet-D7. It was claimed that their most basic model structure of EfficientDet-D0 can match YOLOv3 accuracy while being faster [20]. Furthermore, a study shows that EfficientDet-D2 achieves an mAP of 43.0% on the COCO test set, which would be 1% higher than Faster R-CNN [21].

The model is trained on the training set in which employed five constant parameters used for the model. Table 1 describes all the parameters.

**Table 1.** Parameter information

Name	Description
Learning rate	0.01
Momentum	0.9
Decay	0.0005
Batch size	32 images
Total batch	6000

### Performance evaluation

The model is evaluated on the test set and the evaluation metrics called mean Mean Average Precision (mAP), the method of evaluating the accuracy of the object detection model [14][22][23]. Hence, the metrics will be used in this study to evaluate the model's detection performance. Equation 1 shows the formula for Intersection over union (IoU).

$$\text{Intersection over union} = \frac{\text{area of overlap}}{\text{area of union}} \quad (1)$$

If the detection or IoU is more than 0.5, the object detection is classified as True Positive (TP), but if the IoU is less than 0.5, it is classified as False Positive (FP) as it is a false detection. Then, False Negative (FN) is when the model failed to detect the ground truth although the ground truth is in the image.

### EXPERIMENTAL RESULTS

The result of the study regarding the performance of YOLO V4 Darknet, YOLO V4 Tiny, Faster R-CNN, and EfficientDet will be described. An analysis was done to evaluate the models' performance according to the image bounding boxes area in the ground truth to confirm the models' ability and performance in detecting the classes. Figure 4 illustrates the prediction made by YOLOv4 Darknet.

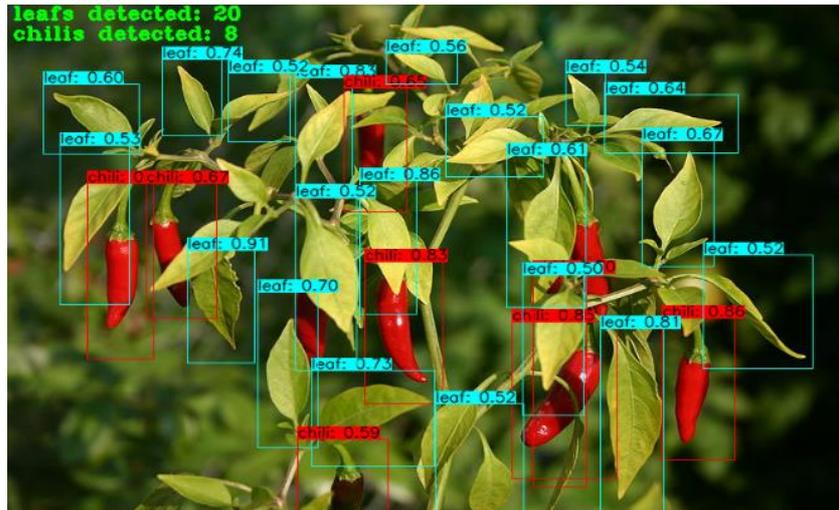


Figure 4. Predictions made by YOLOv4 Darknet

### Accuracy of the models

Experiments were conducted from altering this dataset's composition proved that augmenting the data set yielded nearly a 10% increase in mAP from plain composition to augmented set for all models, as shown in Figure 5. For instance, YOLOv4-Tiny increases from 55.53% to 63.85%. Even though Faster R-CNN has the lowest accuracy at 54.45% for plain composition, it surpassed YOLOv4-Tiny when using augmented data with 67.61%, which may be due to the architecture simplicity of Faster R-CNN [14]. YOLOv4 Darknet model has the highest accuracy that rose from 63.67% to 75.69% from plain to augmented composition, which EfficientDet model came short with its best value of 71.85% for augmentation composition. Overall, the result suggests that as more data gets added, the algorithms become more accurate as augmented images outperform the plain composition for all models

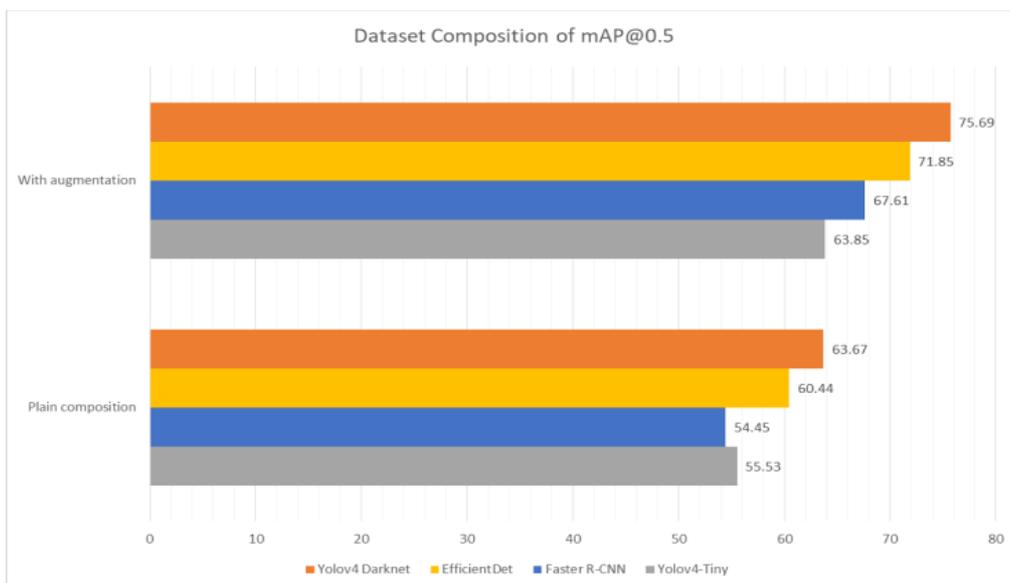


Figure 5. mAP of the dataset after trained on the models

## CONCLUSION

The present study evaluated different object detection models in the classification of bird's eyes chili. It was shown the model can identify the bird's eyes chili and its leaves with a satisfying result. The model is also able to inform the quantity of the objects that are successfully detected. From four algorithms used for the comparisons, YOLOv4 Darknet, YOLOv4 Tiny, Faster R-CNN, and EfficientDet, the experiment result reveals that YOLOv4 Darknet performs well against the other models at the highest mAP of 75.69% when using dataset composition with augmentation.

## ACKNOWLEDGEMENT

The authors would like to acknowledge Universiti Malaysia Pahang for funding this study under the Research Grant (RDU200332).

## REFERENCES

- [1] "Department of Statistics Malaysia Official Portal," *Value of gross output of agriculture sector registered an annual growth rate of 11.1 per cent to RM91.2 billion*, Apr-2019. .
- [2] J. C. O. Koh, M. Hayden, H. Daetwyler, and S. Kant, "Estimation of crop plant density at early mixed growth stages using UAV imagery," *Plant Methods*, vol. 15, no. 1, pp. 1–9, 2019.
- [3] M. N. A. Shapiee, M. A. R. Ibrahim, M. A. M. Razman, M. A. Abdullah, R. M. Musa, and A. P. P. Abdul Majeed, "The Classification of Skateboarding Tricks by Means of the Integration of Transfer Learning and Machine Learning Models," *Lect. Notes Electr. Eng.*, vol. 678, pp. 219–226, 2020.
- [4] S. Puteh, N. F. M. Rodzali, M. A. M. Razman, Z. Z. Ibrahim, and M. N. A. Shapiee, "Features Extraction of Capsicum Frutescens (C.F) NDVI Values using Image Processing," *MEKATRONIKA*, vol. 2, no. 1, pp. 38–46, Jun. 2020.
- [5] A. Karami, M. Crawford, and E. J. Delp, "Automatic Plant Counting and Location Based on a Few-Shot Learning Technique," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5872–5886, 2020.
- [6] S. G. Wu, F. S. Bao, E. Y. Xu, Y. X. Wang, Y. F. Chang, and Q. L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," *ISSPIT 2007 - 2007 IEEE Int. Symp. Signal Process. Inf. Technol.*, pp. 11–16, 2007.
- [7] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Leaf Classification Using Shape, Color, and Texture Features," pp. 225–230, 2013.
- [8] H. F. Eid, A. E. Hassanien, and T. H. Kim, "Leaf Plant Identification System Based on Hidden Naïve Bays Classifier," *Proc. - 2015 4th Int. Conf. Adv. Inf. Technol. Sens. Appl. AITS 2015*, no. August, pp. 76–79, 2016.
- [9] M. Z. Alom *et al.*, "The history began from AlexNet: A comprehensive survey on deep learning approaches," *arXiv*. 2018.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification - IEEE Conference Publication," *Proc. IEEE Int. Conf. Comput. Vis. 2015 Int. Conf. Comput. Vis.*, 2015.
- [11] S. Wang and Z. Su, "Metamorphic Testing for Object Detection Systems," *arXiv*. 2019.
- [12] P. Kamavisdar, S. Saluja, and S. Agrawal, "A Survey on Image Classification Approaches and Techniques," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 2, no. 1, pp. 1005–1009, 2013.
- [13] Alex Krizhevsky, "ImageNet Classification with Deep Convolutional Neural Networks," *ImageNet Classif. with Deep Convolutional Neural Networks*, pp. 1–1432, 2012.
- [14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016.
- [15] B. Anjanadevi, I. Charmila, A. Ns, and R. Anusha, "An Improved Deep Learning Model for Plant Disease Detection," *Int. J. Recent Technol. Eng.*, vol. 8, no. 6, pp. 5389–5392, 2020.
- [16] A. A. dos Santos *et al.*, "Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVS," *Sensors (Switzerland)*, vol. 19, no. 16, pp. 1–11, 2019.
- [17] "PlantVillage Dataset | Kaggle." [Online]. Available: <https://www.kaggle.com/emmarex/plantdisease>. [Accessed: 28-Jul-2021].
- [18] "GitHub - bigrobins/Training-Data-Splitter: A simple python script to split data into train, test, and validation sets."
- [19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [20] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," *arXiv*, pp. 10781–10790, 2019.
- [21] "Detection Evaluation, Common Objects in Context."
- [22] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 6517–6525, 2017.
- [23] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *Comput. Vis. Pattern Recognit.*, vol. 34, no. 5, pp. 241–247, Apr. 2018.