Pest and Plant Disease Identification in Greenhouse Using UAV Images

Tzu-Ming Feng (1), Chao-Hung Lin (1)

[1] Department of Geomatics, National Cheng Kung University, Tainan, Taiwan
Email: dynamo6205@gmail.com ; linhung@mail.ncku.edu.tw ;

KEY WORDS: Deep Learning, Object Detection, UAV

ABSTRACT: Traditional methods on pest and plant disease prevention is visually checking the fields which is labor sensitive and time consuming. In addition, the relative knowledge to identify the pests and the diseases is required during the observation and checking. Therefore, the efficient protection of the crops and the improvement of the yields is still an issue. With the aids of various sensors and developments of artificial intelligent technologies, the crops can be monitored and the pests can be identified automatically. Based on these information, the farmers are able to make corresponding plans for pest prevention, which significantly reduces the labor efforts. Deep learning is a popular technique with various applications, such as object detection. It is possible using deep learning to detect the kinds of pest and calculate the amounts. There are many methods to detect the pest by using deep learning. However, most of the method need to obtain the images manually. The aims of this paper is to observe and detect the pest automatically. In order to get the images automatically, the way to obtain data is flying the Unmanned Aerial Vehicle (UAV) to collect the images from the greenhouse. After collection, deep learning is used to analysis and identify the amounts and the kinds of pests. This method could cost less time and manpower to look into the information of fields.

1. INTRODUCTION

Nowadays, agriculture has gradually declined and the machines are going to replace human to plant and gather vegetables. However, the food and the crops are necessary for everyone to live. When the labor costs decrease but needs increases, the supplier should find the ways to increase crop productivity. One of critical factors causing productivity loss is the control of the diseases and pests. To avoid the economic loss, how to detect the diseases and the pests and do the corresponding solutions is a critical issue to process. Object identification of pest species is important for initiating any kind of pest monitor program (Qiao et al., 2008). There are a lot of methodologies to prevent the pest and monitor the agriculture environment and one of them is the greenhouse. Since the greenhouse is considered as a biophysical system which can automatize the control and plant loops (Ehret et al., 2001). Although the extrinsic factors, such as the weather, typhoons, and others disasters may not have critical influence to the vegetable productivity in the greenhouse, the interior factors will be a serious problem to keep the health of crops. In fact, no automatic methods are generally used to precisely predict and evaluate the status of crops. Most of greenhouses send the staff observe the plants and search for the pests or the diseases periodically. The common methodology to detect the pests in the greenhouses is based on the conventional sticky traps (Pinto-Zevallos and Vänninen, 2013). The numbers and kinds of pests will be counted and identified manually. It means that this work need to rely on human visual judgments (Wise et al., 2007) and it could not be replaced by machine in the past. Apart from this problem, some of pests have similar appearance but the different species and the men who recognize the pests should have corresponding knowledge to avoid the mistakes. Above the problems show the difficulty to let the monitor system automatically, but it is possible nowadays due to the machine learning and the object detection. Based on the visual information (e.g., shape, color, size) taking by the sensors (e.g., camera, radar), the images won't make extra disturbances to the environment (Xia et al., 2014). Using the automatic pest identification system could achieve the ecological intensification, increase the productivity, and minimize the anthropogenic environmental effects in the same time.

In recent years, more and more people use machine learning to identify the pests. However, different pest species and crops have different neural network model to finish the detection. Martin and Thonnat (2007) use a cognitive vision approach that adjusts optimal parameters for segmenting whitefly out of leaves based on adaptive learning techniques. Boissard and Martin (2008) also use cognitive vision approach which include the image segmentation, knowledge-based techniques and feature extraction to detect the location of pest on the leaves. Solis Sánchez et al. (2009) utilize geometric features (e.g., eccentricity, area size) to whitefly scouting by segmenting the insects. Apart from the pests, diseases are also used to recognized by the machine learning. Fuentes and Yoon (2017) use multiple machine learning methodologies (e.g., SSD, VGG-16, R-FCN, ResNet 50-152…) to identify the whitefly and miner.

There have already been lots of methods to recognize the pests and they have high accuracy. However, how do they get the images and data have one in common. They need to collect the data manually. Boissard and Martin (2008) collect the conventional sticky traps manually and use the scanner to make the dataset. Fuentes and Yoon (2017) collect the dataset by using camera and taking photos in the greenhouse. In this paper, decreasing the cost and time of collecting data is our purpose. To collect dataset conveniently, Unmanned Aerial Vehicle (UAV) is chosen to become our sensor and control it to obtain the conventional sticky trap photos. Deep learning is the methodology to do the object detection and one of them which has the highest accuracy and fastest training speed is called YOLO (Redmon et al., 2016). Flies (Muscomorpha) will be our target and since the photos cover not only the conventional sticky papers but also the crops, the sticky traps will be our first target. After confirm the location of sticky traps, the fly will be our last target to detect.

2. MATERIAL AND METHODS

2.1 Data collection

The study area is the greenhouse at Guantian, Tainan, Taiwan. There are totally 45 greenhouses and the length and width of greenhouse are 40 and 8 meters. The four-row greenhouse are planted with Cabbage and five sticky traps (15cm x 10cm) will be set in one row but the last row won't set (Fig. 1.). It means that there will be fifteen images taken in one greenhouse. The sticky traps were placed 50 cm above the crops and collected after 4 weeks. Before renew the traps, the photos will be taken every week.

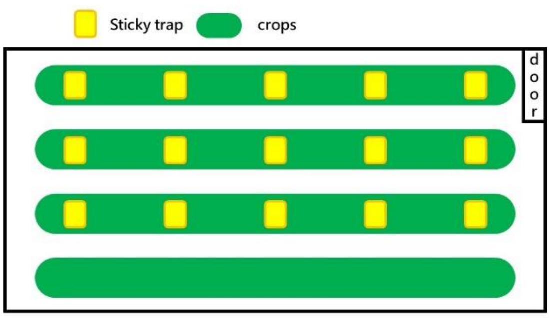

Fig. 1. Greenhouse map showing the locations of sticky traps

The Unmanned Aerial Vehicle (UAV) used in this project is DJI Mavic Pro 2. The UAV flown 1.5 meters above the sticky traps (Fig. 2.) since the wind caused by the UAV is too strong. If the UAV is too close to the crops, it will hurt the crops and violate to our purpose.

Fig. 2. Photo taken from 1.5 meters above the trap

2.2 Object Detection

The common object detection methods include Mask R-CNN, Fast R-CNN, Faster R-CNN and YOLO (You Only Look Once). Every object detection is based on the convolutional neural network (CNN) and the differences between these methods are the architecture of the model. The architecture of CNN model consists of layers which are made up of neural. The neural means the parameters of the layer and the more neural is, the higher accuracy is commonly. However, it will cost longer time to do programing. Usually, the neural could be calculated by the neural of the previous layer, but different kinds of layer have different calculation.

The object detection we use is YOLO. YOLO has 4 versions and the latest one is YOLOv3 (Redmon et al., 2018). Both the pest and trap detection use YOLOv3. The advantage of YOLOv3 is that compared to YOLO and the other version, YOLOv3 improved the accuracy with many tricks and is more capable of detecting small objects. The reason is that YOLOv3 network is fed with input images to predict 3D tensors corresponding to 3 scales, which is (52 x 52), (26 x 26) and (13 x 13) pixels. Based on the scale which we choose, the dimension will be changed:

$$\text{3D tensor Dimension} = \mathrm{N} \times \mathrm{N} \times [\mathrm{box_{number}} \times (\mathrm{box_{cordinate}} + \mathrm{score} + \mathrm{classes})]$$

Where (N x N) is the size of the scale, boxnumber represents the amount of box and each grid predicts 3 boxes, boxcoordinate will be 4 and means x, y, height and width, score represents objectness score and it is defined as 1 and classes means that the class number which you expect.

The architecture of YOLOv3 is why it's so special. It uses Darknet-53 and feature pyramid network (FPN) to constitute the architecture (Fig. 3.). Darknet-53 has 53 convolutional layers and ResNet structure. ResNet is used to deal the vanishing gradient issue and let the deep neural network have the shortcut connection. After the Darknet-53, layer connects to the FPN layer and split to 3 scales. FPN is used to make the different scales and combines low-resolution, semantically strong features with high-resolution, semantically weak features via a top-down pathway and lateral connections. Both ResNet and FPN structure improve the accuracy. In addition, there is no pooling layer and fully-connected layer in YOLOv3 and it makes YOLOv3 possible to deal with photos with any sizes.

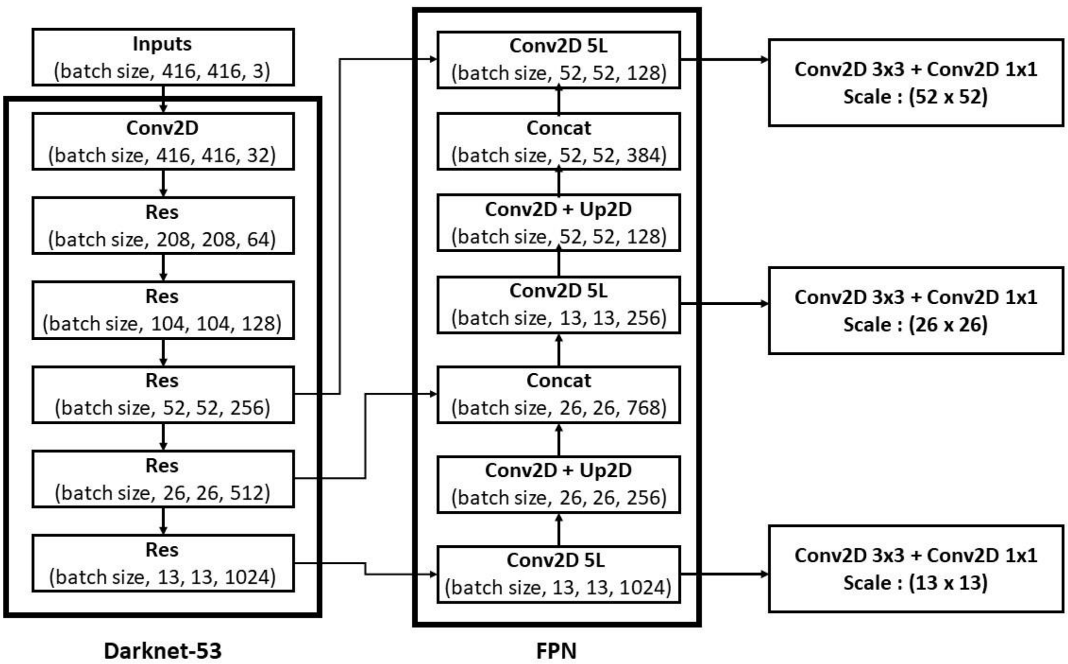

Fig. 3. The architecture of YOLOv3 shows that it consists of Darknet-53, FPN and 3 scales

The disadvantage of YOLOv3 is that size of the input image must be (416 x 416) pixels. Due to that, to resize images is necessary before and after the detection. In this paper, the pest is hard to search after the image resize. To deal with this problem, the original image is clipped based on the location of the sticky trapes. Then, YOLOv3 is used again to identify the pests. According the agriculture experts' suggestion, the fly is the most common pest in study area.

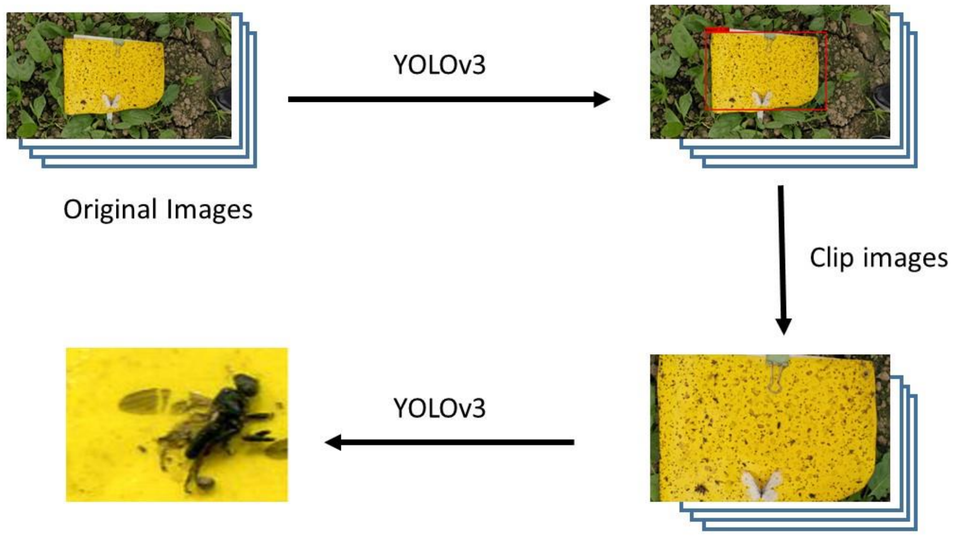

Fig. 4. Detect the area of stick traps and clip the images. Then, the images do the object detection again to get the location of pests

3. RESULTS

Object detection were conducted in two steps: YOLOv3 for sticky traps and for flies.

After remove some images which are blurry, the dataset totally has 909 images. 700 images will be regard as the training data and the others are the testing data. Because the sticky traps can be identified obviously, the training data won't be too much. The Accuracy of the object detection has two: One is IoU and the other is mAp. Since the mAp will be considered as a standard accuracy when the classes is more than one. Intersection over union (IoU) is an evaluation metric used to measure the accuracy of an object detector on a particular dataset.

$$\mathrm{IoU} = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

IoU is a value between 0 and 1. The value which close to 1 means the accuracy of model is high and the value which is close to 0 means the accuracy is low. Fig.5 shows the result of first step and use IoU to be the standard accuracy index. The mean IoU is 0.945263. This means that the sticky trap detection is useful.

Fig. 5. The results of the trap detection

Next, the new dataset will be created after clipping. The amount of dataset won't be changed, but the sizes of dataset are smaller. In this time, the dataset is divided into 800 training data and 109 testing data randomly. After the processing, the results will be like Fig. 6. The IoUs of every test images are showed in Fig. 7. The mean of IOU is 0.72893 and the Fig. 7 show that the IoU values distribute evenly between the 0.5 and 1. The reason cause the low accuracy is the pest overlapping. Sometimes the flies pile and the shape will hard to identify.

Fig. 6. The results of the fly detection. The red area shows the predict location of flies.

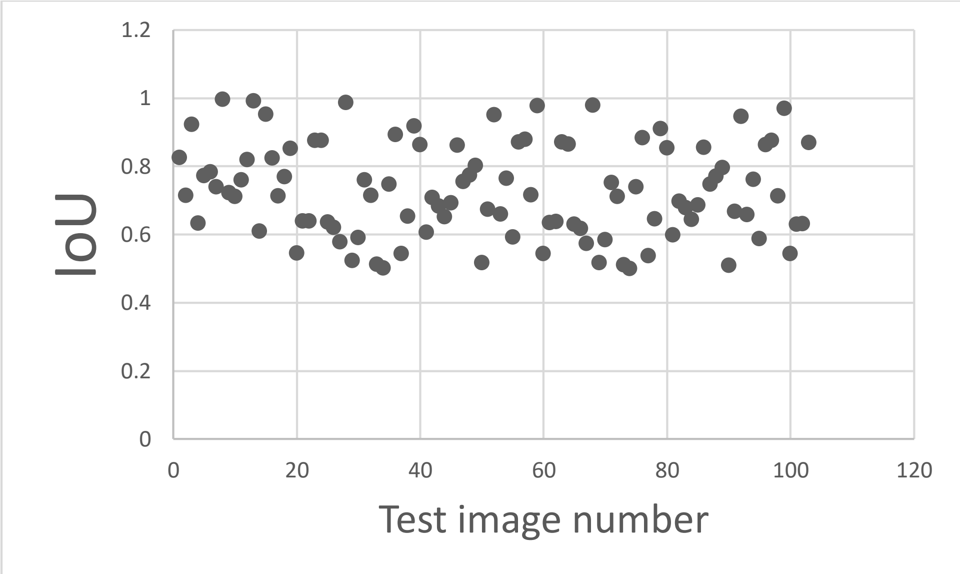

Fig. 7. The IoU values of number 1 ~109 images

4. CONCLUSION

By using the Unmanned Aerial Vehicle to be the sensor collecting the image, this method is useful to obtain the high resolution image. Different from others paper, though the data is more convenient to acquire, the target (e.g., sticky traps) in the image maybe not like what we expect. Therefore, the preprocessing is necessary to gain the expected data. In addition, the way to fly the UAV is critical also. Based on the crop species, the height of UAV should be set to avoid the damage to the crops. The results show that the accuracy of sticky traps detection is enough to do the next step.

However, the pest object detection didn't have the great accuracy. The reason may have two: one is the environment of the greenhouse. The lightness, the angle between traps and ground and the weather may be the factors to effect the results. The rain will make the traps be wet and the color will transform to dark; the other reason is the pest overlapping. The flies may pile each other and it will increase the difficulty to detect the pest.

We prove that the UAV is feasible to be the sensor but the object detection could be improved in the future. Our purpose will be improving accuracy in the future.

5. REFERENCES

Cho, J., Choi, J., Qiao, M., Ji, C.W., Kim, H.Y., Uhm, K.B., Chon, T.S., 2007. Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis. Int. J.Math. Comput. Simul. 346 (246), 244

Ehret, D. L., Lau, A., Bittman, S., Lin, W., Shelford, T., 2001. Automated monitoring of greenhouse crops. Agronomie: Agric. Environ. 21 (4), 403–414.

Xia, C., Chon, T.S., Ren, Z., Lee, J.M., 2014. Automatic identification and counting of small size pests in greenhouse conditions with low computational cost

Boissard, P., Martin, V., Moisan, S., 2008. A cognitive vision approach to early pest detection in greenhouse crops. Comput. Electron. Agric. 62 (2), 81–93

Martin, V., Thonnat, M., 2007. A cognitive vision approach to image segmentation.Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence Washington, DC, USA, pp. 265–294.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement.

Qiao, M., Lim, J., Ji, C.W., Chung, B.K., Kim, H.Y., Uhm, K.B., Myung, C.S., Cho, J., Chon, T.S., 2008. Density estimation of Bemisia tabaci (Hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. J. Asia Pac. Entomol. 11 (1), 25–29.

Gaston, K.J., O'Neill, M.A., 2004. Automated species identification: why not? Philos. Trans.R. Soc. Lond. B Biol. Sci. 359 (1444), 655–667.

Koumpouros, Y., Mahaman, B.D., Maliappis, M., Passam, H.C., Sideridis, A.B., Zorkadis, V.,2004. Image processing for distance diagnosis in pest management. Comput. Electron. Agric. 44 (2), 121–131