Extraction of Cover Density Map of Rice Seedling Using Transfer Learning and Machine Learning

A. Rezapour, M. Gholizadeh, A. Nayyer, E. Pazouki

***Abstract*—To meet demand for agriculture products, researchers have recently focused on precision agriculture to increase crop production with less input. Crop detection based on computer vision with unmanned aerial vehicle (UAV)-acquired images plays a vital role in precision agriculture. In recent years, machine learning has been successfully applied in image processing for classification, detection, and segmentation. Accordingly, the aim of this study is to detect rice seedlings in paddy fields using transfer learning from model Faster R-CNN. This study relies on a significant UAV image dataset to build a model to detect tiny rice seedlings. The model was also measured with three additional datasets acquired on different dates to evaluate model applicability with various imaging conditions. The results demonstrate that the adoption of transfer learning allows for rapid establishment of object detection applications with promising performance.**

***Index Terms*— machine learning, deep learning, object detection, Faster R-CNN, transfer learning.**

I. Introduction

Demand for agricultural products urges agriculture sectors to accommodate technology to overcome production challenges [1,2]. Prominently, population growth creates constant pressure on the agricultural system to supply more food to fulfil global demand, which drives farmers to adopt modern technologies (such as precision agriculture) in food crop production [3–6]. Globally, precision agriculture plays an important role in increasing the quality of crop production, sustaining crop production, and making decisions based on analyzing large amounts of data and information about crop status obtained from farms. Moreover, it is used for effective fertilizer management and irrigation as well as for labor reduction [7–9]. With the advantages of technology in capturing high-resolution images, particularly by using unnamed aerial vehicle (UAV), large amount of remote sensing data can easily be obtained for analyzing crop yield in precision agriculture. Through the development of Internet of Things (IoT) and computer vision, sensors, and cameras along with machine learning, deep learning and image processing techniques have been getting increasing attention for capturing information and further processing for smart farming to help maintain the sustainability of agricultural production [10]. Smart farming plays a vital role in the agricultural process based on adjusting various agricultural management measures. It provides suggestions and insights for more efficient and effective agriculture production and to solve the challenges in agriculture systems [11].

Moreover, computer vision plays a key role in extracting useful information from the collected image dataset for management of smart farming tasks [10,14]. In recent agricultural operations, machine learning in computer vision has been applied for various object detection and classification tasks through extracting information from images to significantly promote intelligent agriculture [15–19].

As mentioned above, the developments in IoT provide a good platform to collect a large amount of image data with many objects to make meaningful image analysis [20]. To collect image data in agriculture sectors, UAVs or drones are widely used in precision agriculture and many other fields, such as path planning and design, wildlife rescue, weed classification, harvesting, livestock counting and crop and aquatic products damage assessment [21–26]. UAVs can be used to detect potential issues and then obtain high-resolution images to inspect and apply treatments correspondingly. The combination of UAVs and computer vision helps farmers make correct decision by obtaining information from the images [14]. This study focusses on monitoring sowing area via UAVs for identifying rice seedlings and counting them for decision-making regarding the progress of rice seedlings in paddy fields.

Deep learning, which is one branch of machine learning, in object detection can deal with high-density scenes with complex and small objects in images [27,28]. Object detection in computer vision is widely used for various applications. By training with large amounts of image data, object detection can accurately identify the targeted objects and their spatial locations in the images, classify objects from the specified varieties, such as human, animals, crops, plants, and vehicles, and mask the objects within bounding boxes by well-developed algorithms [29–33]. image. Moreover, the challenging task in computer vision is to detect the small objects in an image that lack appearance information to distinguish them from background and similar categories. The precision requirement is higher for accurately locating small objects. A recent review reported detailed information about the use of convolutional neural networks (CNNs) for small-object detection. Their results, based on popularly existing datasets, showed better performance for detecting small objects in terms of multi-scale feature learning, data augmentation, training strategy, context-based detection, and GAN-based detection methods [34]. Based on this evidence,

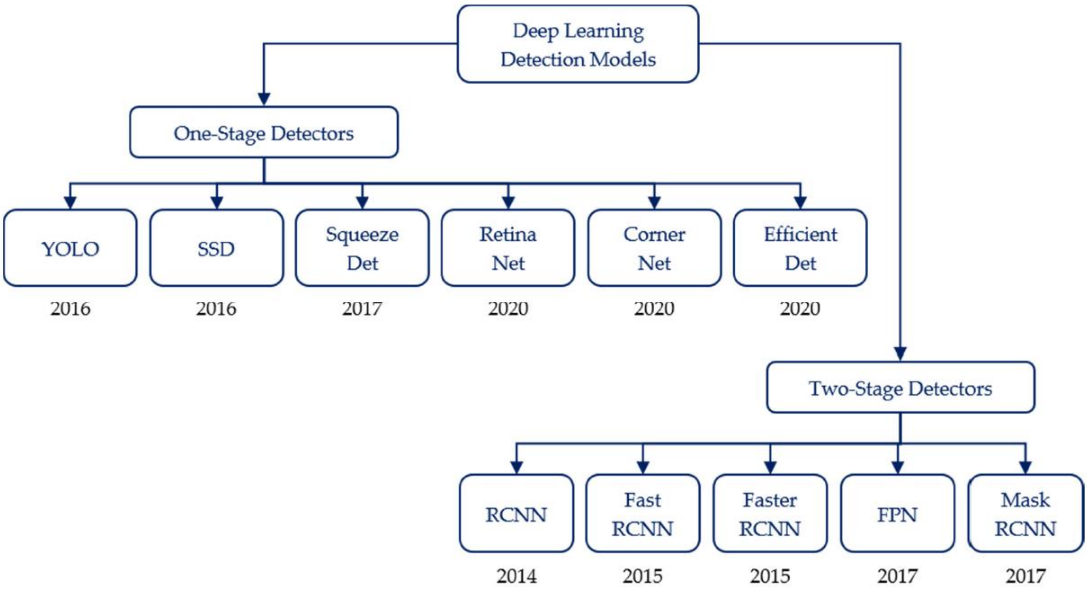

Figure I-1 Development stages of deep learning object detection [30]

this study aims to employ an object-detection model to monitor a single small object in paddy fields using UAVs. This study focusses on the rice seedlings in paddy fields as very tiny objects that can hardly be observed by the human eye to find displacements or missing rice seedlings and count and locate rice seedlings.

Object detection based on machine learning in computer vision has improved enormously in accuracy and speed compared to traditional detection algorithms with feature extraction [35]. It is used for classifying and locating objects in automatic image detection processes based on statistical and geometric features.

The advantage of directly using images in object detection applications is that it allows CNNs to avoid manual feature extraction. In other words, CNNs are one of the most effective algorithms for object detection due to directly extracting features and detecting objects from images. In recent years, impressive improvements have been achieved in CNNs to address the problem of object detection by proposing many algorithms in which the network models are trained by combining local regional perception and feature extraction with a classification process. Tong et al. [34] and Zhang et al. [28] provided detailed reviews about the recent progress of CNN algorithms for objection detection. The development stages of CNN-based object detection models are shown in Figure I-1 [30]. These object detection algorithms based on deep learning are divided into one-stage and two-stage detection algorithms. In one-stage algorithms, features for bounding box regression and class classification are directly extracted by the convolution operations on the output features of backbone networks. Object detection algorithms based on feature map convolution include YOLO [36], SSD [37], SqueezeDet [38], RetinaNet [39], CornerNet [40] and EfficientDet [41]. In two-stage algorithms, regional proposal modules are used to propose targeted object boundary boxes, and features are subsequently extracted from them to predict categories and masking objects. Object detection algorithms based on reginal proposal, such as RCNN [42], Fast R-CNN [43], Faster R-CNN [44], Mask R-CNN [45] and FPN [46], perform better and achieved high mean average precision (mAP). Among them, Mask-RCNN could be used to predict an exact mask within the bounding box of objects to detect single objects in images. This study adopts Faster R-CNN as a two-stage algorithm, to detect rice seedlings in paddy fields due to the advantages of high efficiency, high localization, and high precision for object detection.

This study overcomes a data scarcity problem with lightweight architecture and transfer learning in deep learning for precision agriculture. Besides, this study chooses Faster R-CNN due to its architecture being capable of handling huge variations of feature scales for small-object detection. Overall, this study adopts two-stage object detection architectures to develop tiny-object detection in UAV images to identify rice seedlings for precision agriculture, which has never been done for traditional rice cultivation.

II. Materials and Methods

A. *Data Introduction*

UAVs can be used to help farmers broadly monitor rice growth in the early stage. The field images of rice seedlings were collected by UAVs equipped with cameras and downloaded from an open dataset [47]. Detailed information of camera, UAV, and calibration settings to take images is given in Tables 1 and 2. The study area is located at the Taiwan Agriculture Research Institute, Wu-Feng District, Taichung City, Taiwan, where a long-term field investigation and observation, including UAV imaging and field survey, has been conducted for rice cultivation management experiments. Counting rice seedlings in paddies is one of the keys to calculate density and estimate grain yield. The framework includes four phases: image pre-processing, sub-image generation, object detection with three approaches, and detection result and evaluation.

Table 1 - UAV imaging sensor details [61].

Sensor Description	DJI Phantom 4 Pro
Resolution (H pixel × V pixel)	5472 × 3648
FOV (Hº × Vº)	73.7º × 53.1º
Focal Length (mm)	8.8
Sensor Size (H mm × V mm)	13.2 × 8.8
Pixel Size (µm)	2.41 × 2.41
Image Format	JPG
Dynamic Range	8 bit

Table 2 - Details of flight mission [47].

Parameter	Setting			
Sensor	DJI Phantom 4 Pro			
Acquisition Date	7 August 2018	14 August 2018	12 August 2019	20 August 2019
Time	07:19–07:32	07:03–07:13	14:23–14:44	08:16–08:36
Weather	Mostly clear	Mostly cloudy	Mostly cloudy with occasional rain	Partly cloudy
Avg. Temperature (ºC)	28.9	26.8	26.6	27.5
Avg. Press (hPa)	997.7	992.0	994.1	996.4
Flight Height (m)	21.4	20.8	18.6	19.1
Spatial Resolution (mm/pixel)	5.24	5.09	4.62	4.78
Forward Overlap (%)	80	80	85	85
Side Overlap (%)	75	75	80	80
Collected Images	349	299	615	596

The first phase is orthorectifying and mosaicking images captured by UAVs. The second phase is generating sub-images from orthorectified mosaic images due to the GPU's memory limitation. The third phase shows two-stage object detection architecture—Faster-RCNN. This approach generates detection results with classification and localization predictions that are evaluated with ground truth in the fourth phase. Each rice seedling in sub-images is manually annotated by agricultural experts. A training dataset is used to obtain the best model weights for rice seedling detection and counting.

B. CNN Model

Dataset collection is the essential part in object detection. This study used a total of four UAV images of rice paddy fields to train the model for CNN-based rice seedling detection. Each rice seedling in every image was manually labeled using labelImg, an opensource graphical image annotation tool in a pixel-basis with a single separate category (i.e., rice seedlings) from background. To have a sufficient dataset with several rice seedlings in each image, each image was split into several sub-images with each side 512 pixels. A total of 297 sub-images were generated from the four field images, and a training–test split ratio of 80–20 was applied to generate 273 sub-images for training and 60 sub-images for testing. In addition, three separate test datasets acquired on 14 August 2018, 12 August 2019, and 20 August 2019 with 72, 100 and 100 images, respectively, were also included to evaluate the model's applicability to various imaging conditions. Annotating rice seedlings in every sub-image is time-consuming (needing a huge number of person-hours), thus a semi-auto preprocess was adopted for rapid annotation. Datasets can be used to determine the accuracy of rice seedlings detection and counting. The expected study results are looking for raw counts and illustrating the spatial distribution of rice seedlings. Figure II-1 shows annotated images of rice seedlings.

C. Faster R-CNN Model Training

Faster R-CNN is a two-stage object detection network that inherits the robustness of the R-CNN family with a precise detecting capability. To speed up detection, a regional proposal network (RPN) is proposed to replace the selective search algorithm, which has poor GPU utilization. RPN is a fully convolutional network that generates both box regression and box-classification features with a set of predefined anchors as the regression target. After the RPN generates proposals, both proposals from RPN and features from the backbone are passed into the Fast R-CNN detector. Because the design of the anchors already considers multi-scale and multi-ratio patterns, the architecture can feasibly be trained on single-scale images. The visualized architecture of Faster R-CNN ResNet-101 is shown in Figure II-3. A transfer learning strategy is also employed to rapidly establish the detection application. In this study, the ResNet-101 [48] backbone with input image size 640 × 640 pretrained weight is chosen for training the Faster R-CNN model. The changes of parameters in the configuration are the number of classes, training step, batch size and the max number of detection box, which are 1, 25,000, 8 and 200 respectively.

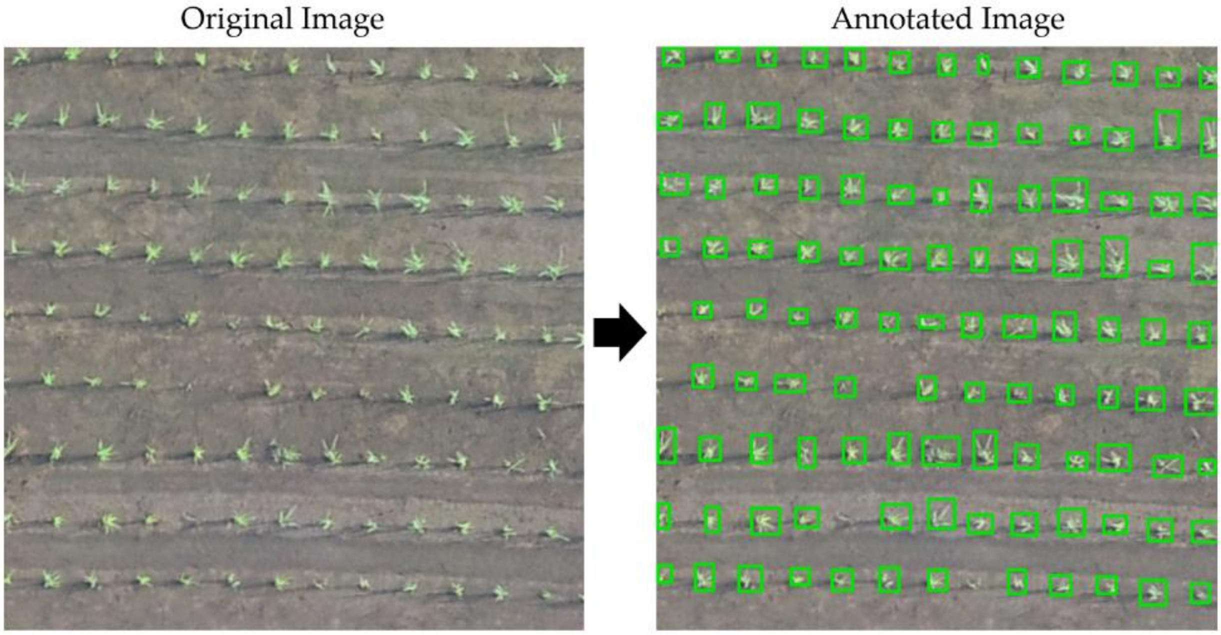

Figure II-1 Rice seedlings annotated with bounding boxes in green on sub-images.

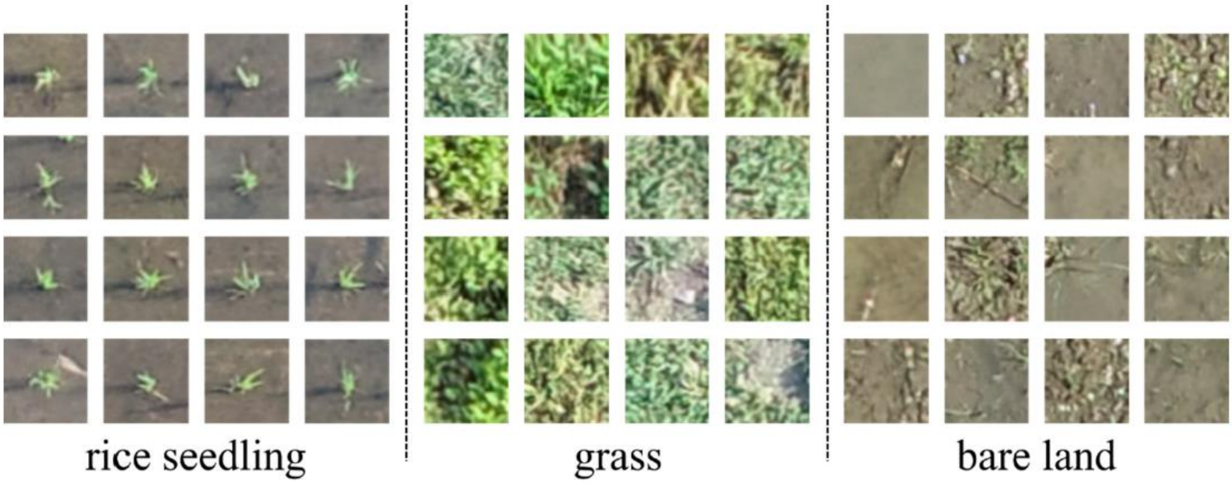

Figure II-2 Examples of annotated images for model training.

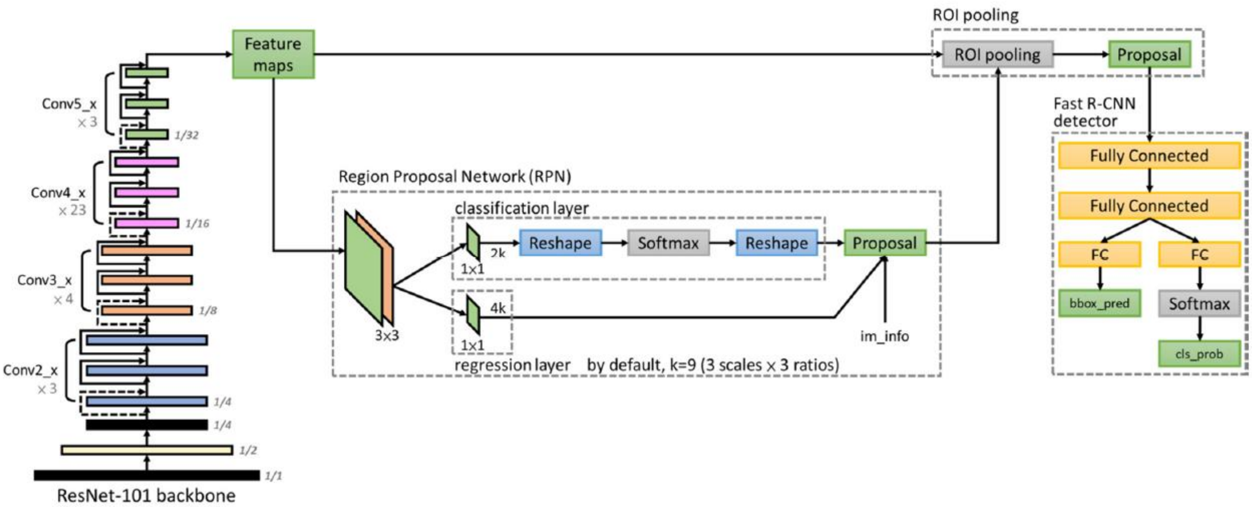

Figure II-3 Architecture of Faster R-CNN ResNet-101.

III. Results and Discussion

The container runs on the hardware specification of one NVIDIA Tesla V100 GPU, four cores of Intel Xeon Gold 6154 CPU, and 90 GB of host memory. The containerized image for the experiments is tensorflow-21.06-tf2 with Tensor- Flow version 2.5 to have the latest function support. Starting from the model building, experiment applies the same scripts from the officially released object detection example of TensorFlow. The only two changes are the configuration files and the pretrained weights. Detailed documentation of usage can be accessed from TensorFlow on GitHub. Detection model is initialized with the COCO 2017 pretrained weights, which provides prior knowledge of object features to give a better and faster model convergence. Figure III-1 shows the visualized detection results, and Figure III-2 shows the visualized AP curve. Faster R-CNN performs 1.000 mAP and 0.996 mIoU on the training data and 0.888 mAP and 0.637 mIoU on the test data. The metrics between training and testing show a big gap that could be caused by overfitting during training.

To simulate the inference for real-time scenarios, images are loaded one by one from the disk for all three models. The CNN-based detection model highly utilizes the GPU to extract features and detect objects parallelly.

In this study, deep learning method with two stages was employed to enhance localization accuracy. Wu et al., 2019 [19], used fully convolutional architecture to count rice seedlings in 40 UAV images, resulting in high correlation to the ground truth count (R2 = 0.94), but false positive counting was not considered and adjusted for. Moreover, the size of detected objects was not detected, so it was unable to estimate localization accuracy. In addition to rice seedling counting, the proposed methods in this study can detect the position and size of the rice seedlings.

Furthermore, this study applied transfer learning to reduce the need for a great amount training data and training time (less than 1 h for 500 epochs). Overall, Faster R-CNN models show the capability of real-time inference, as they performed around 20 fps detection. However, the computational cost of the CNN-based model is overstated, because normally the images are read from the camera cache through the bus instead of the disk. Further, the CNN models will be

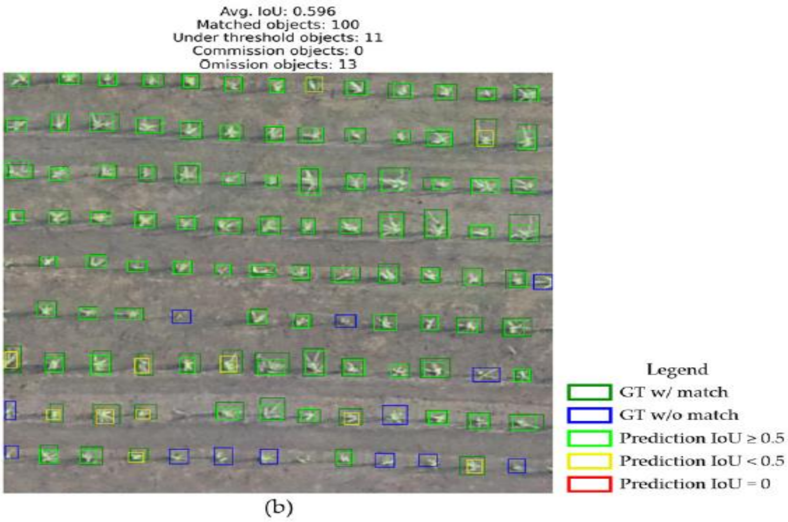

Figure III-2 Example of visualized detection results of Faster R-CNN.

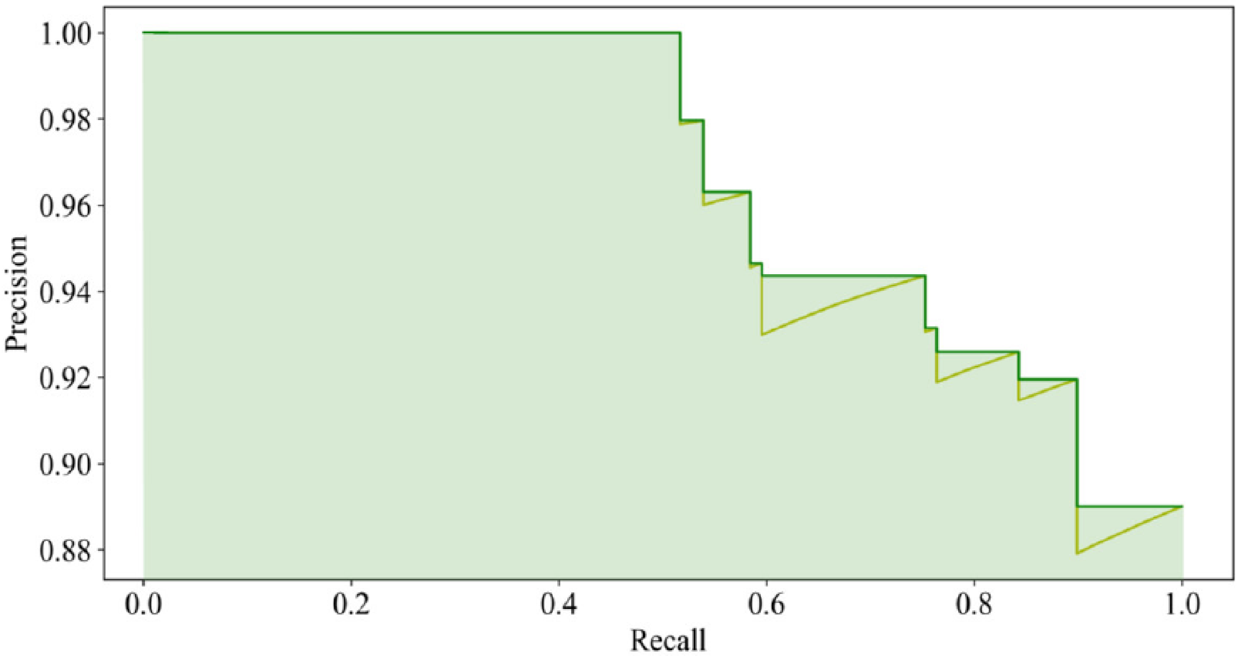

Figure III-1 AP curves of the detection results of Faster

Table 3 - Model performance and computational cost

Training		Test		Computational Cost				
mAP	mIoU	mAP	mIoU	Preprocess	Inference	Visualization	Total	Total
1.000	0.996	0.888	0.637	0.005s	0.042 s	0.003 s	0.050	20.000 fps

optimized to a faster and lighter runtime package to satisfy various deployment environments.

Table 4 - Evaluation of AP and IoU on four datasets.

Date	AP	IoU
7 August 2018	0.888	0.637
14 August 2018	0.981	0.686
12 August 2019	0.986	0.871
20 August 2019	0.739	0.382

Table 5 - Evaluation of model performance on four datasets.

Date	F1- Score	Precision	Recall
7 August 2018	0.783	0.855	0.780
14 August 2018	0.904	0.948	0.747
12 August 2019	0.790	0.972	0.615
20 August 2019	0.480	0.583	0.345

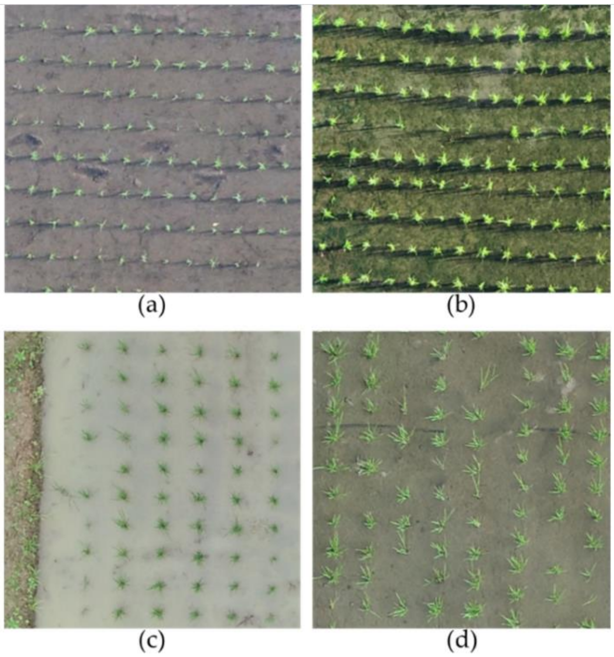

Figure III-3 An example of the test images on four datasets on (a) 7 August 2018, (b) 14 August (c) 12 August 2019 and (d) 20 August

A. Model Evaluation of Different Datasets

The detection model was evaluated with four test datasets, which were different in the planting year, growth day, location, and environmental conditions. The AP and IoU metrics of the model are listed in Table 4.

To discuss this issue, an example of the test images on four different datasets was selected and visualized (Figure III-3), and comparison of the test results is visualized as Figure III-4. Figure III-3 shows the variances of paddy environment, seedling size and illumination. The variance of seedling sizes is due to the different image acquisition date, which can be categorized into three sizes of side length 20, 25 and 30 pixels. The paddy environment can be categorized into four situations with combinations of ponding management and growth of algae.

An example of the detection results with comparisons between the three models is shown in Figure 18. Precision was calculated by the number of predictions for which the IoU is equal or greater than 0.5 (light green boxes) divided by the total number of predictions. Recall was calculated by the number of light green boxes divided by the total number of ground truth boxes.

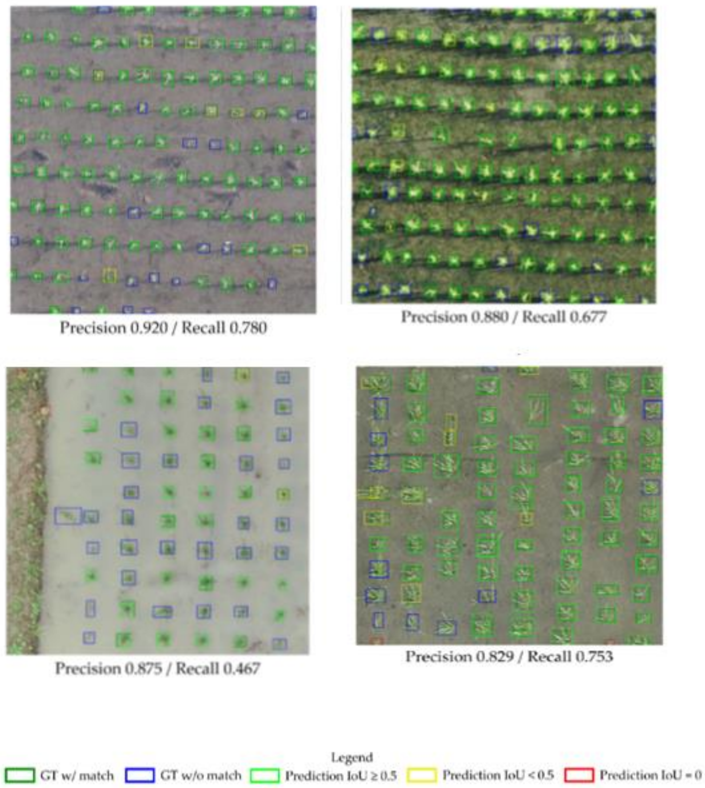

Figure III-4 An example of detection results with precision and recall metrics on four datasets.

IV. Conclusion

Tiny-object detection in UAV images is a challenging task in practical applications. Long computation time and slow speed due to memory consumption are the first-priority causes. Complex background and scenes, high density areas and random textures of fields can also decrease the performance of small-object detection. In this study, small rice seedlings are presented in highly noisy environments that influence the detection using deep learning on UAV imagery.

This study presents a machine learning model, Faster R-CNN, on UAV images to detect tiny rice seedlings. The datasets are semi-annotated with preprocessing of image processing and manual verification to reduce the cost of labor. This approach is sure to generate usable datasets rapidly.

The class achieved above 85% F1-score in both training and testing. However, explosive data growth is one of the problems that needs to be solved in practical applications.

In this study, the CNN model was transferred from pretrained models to develop a well-generalized model with high detection accuracy and rapidity. The pretrained model was well-trained by splitting four paddy images into 297 sub-images (each image sized 512 ×512 ×3) by annotating each rice seedling in every sub-image. To verify model applicability with various imaging conditions, Faster R-CNN was applied to the rest of the images and three additional datasets acquired on different dates for model testing. We have shown that Faster R-CNN has a performance with mAP of 0.888, 0.981 and 0.986 and mIoU of 0.637, 0.686 and 0.871 on the first three test datasets. Moreover, the CNN-based model had acceptable detection results, with 0.739 mAP and 0.382 mIoU (Faster RCNN) on the fourth dataset, even though huge variances exist between test datasets and training datasets.

Further study will focus on detecting rice seedlings with more variety in imaging conditions, such as illumination, tone, color temperature, blur, and noise. The models can be retrained using these additional images to adapt to more image changes. Further, optimizing model parameters to reduce computational time and increase prediction accuracy is needed to enable models to be deployed in environments with tight resources.

Acknowledgment

We would love to thank our parents and partners for their generous support and love.

References

[1] Lencucha, R.; Pal, N.E.; Appau, A.; Thow, A.-M.; Drope, J. Government policy and agricultural production: A scoping review to inform research and policy on healthy agricultural commodities. Glob. Health 2020, 16, 11.W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Bochtis, D.D.; Sørensen, C.G.C.; Busato, P. Advances in agricultural machinery management: A review. Biosyst. Eng. 2014, 126, 69–81.

[3] Josephson, A.L.; Ricker-Gilbert, J.; Florax, R.J.G.M. How does population density influence agricultural intensification and productivity? Evidence from Ethiopia. Food Policy 2014, 48, 142–152.

[4] Ricker-Gilbert, J.; Jumbe, C.; Chamberlin, J. How does population density influence agricultural intensification and productivity? Evidence from Malawi. Food Policy 2014, 48, 114–128.

[5] Fróna, D.; Szenderák, J.; Harangi-Rákos, M. The challenge of feeding the world. Sustainability 2019, 11, 5816

[6] Le Mouël, C.; Lattre-Gasquet, D.; Mora, O. Land Use and Food Security in 2050: A Narrow Road; Éditions Quae: Versailles, France, 2018.

[7] Zhang, Y. The Role of Precision Agriculture. Resource 2019, 19, 9.

[8] Singh, P.; Pandey, P.C.; Petropoulos, G.P.; Pavlides, A.; Srivastava, P.K.; Koutsias, N.; Deng, K.A.K.; Bao, Y. 8-Hyperspectral remote sensing in precision agriculture: Present status, challenges, and future trends. In Hyperspectral Remote Sensing; Pandey, P.C., Srivastava, P.K., Balzter, H., Bhattacharya, B., Petropoulos, G.P., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 121–146.

[9] Yang, C.Y.; Yang, M.D.; Tseng, W.C.; Hsu, Y.C.; Li, G.S.; Lai, M.H.; Wu, D.H.; Lu, H.Y. Assessment of Rice Developmental Stage Using Time Series UAV Imagery for Variable Irrigation Management. Sensors 2020, 20, 5354.

[10] Tian, H.; Wang, T.; Liu, Y.; Qiao, X.; Li, Y. Computer vision technology in agricultural automation—A review. Inf. Process. Agric. 2020, 7, 1–19.

[11] Saiz-Rubio, V.; Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. Agronomy 2020, 10, 207.

[12] Carolan, M. Publicising Food: Big data, precision agriculture, and co-experimental techniques of addition. Sociol. Rural. 2017, 57, 135–154

[13] López, I.D.; Corrales, J.C. A Smart Farming Approach in Automatic Detection of Favorable Conditions for Planting and Crop Production in the Upper Basin of Cauca River. In Proceedings of the Advances in Information and Communication Technologies for Adapting Agriculture to Climate Change, Popayán, Colombia, 22–24 November 2017; pp. 223–233.

[14] Zhao, W.; Yamada, W.; Li, T.; Digman, M.; Runge, T. Augmenting crop detection for precision agriculture with deep visual transfer learning—A case study of bale detection. Remote Sens. 2021, 13, 23.

[15] Gomes, J.F.S.; Leta, F.R. Applications of computer vision techniques in the agriculture and food industry: A review. Eur. Food Res. Technol. 2012, 235, 989–1000.

[16] Rose, D.C.; Chilvers, J. Agriculture 4.0: Broadening responsible innovation in an era of smart farming. Front. Sustain. Food Syst. 2018, 2, 87

[17] Deng, R.; Jiang, Y.; Tao, M.; Huang, X.; Bangura, K.; Liu, C.; Lin, J.; Qi, L. Deep learning-based automatic detection of productive tillers in rice. Comput. Electron. Agric. 2020, 177, 105703

[18] Vasconez, J.P.; Delpiano, J.; Vougioukas, S.; Auat Cheein, F. Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation. Comput. Electron. Agric. 2020, 173, 105348.

[19] Wu, J.; Yang, G.; Yang, X.; Xu, B.; Han, L.; Zhu, Y. Automatic counting of in situ rice seedlings from UAV images based on a deep fully convolutional neural network. Remote Sens. 2019, 11, 691.

[20] Wolfert, S.; Ge, L.; Verdouw, C.; Bogaardt, M.-J. Big data in smart farming—A review. Agric. Syst. 2017, 153, 69–80.

[21] Yang, M.D.; Boubin, J.G.; Tsai, H.P.; Tseng, H.H.; Hsu, Y.C.; Stewart, C.C. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. Comput. Electron. Agric. 2020, 179, 105817

[22] Ward, S.; Hensler, J.; Alsalam, B.H.; Gonzalez, L. Autonomous UAVs wildlife detection using thermal imaging, predictive navigation and computer vision. In Proceedings of the IEEE Aerospace Conference, IEEE Aerospace Conference, Big Sky, MT, USA, 5–12 March 2016; pp. 1–8.

[23] Yang, M.D.; Huang, K.S.; Kuo, Y.H.; Tsai, H.P.; Lin, L.M. Spatial and spectral hybrid image classification for rice lodging assessment through UAV imagery. Remote Sens. 2017, 9, 583.

[24] Driessen, C.; Heutinck, L. Cows desiring to be milked? Milking robots and the co-evolution of ethics and technology on Dutch dairy farms. Agric. Hum. Values 2014, 32, 3–20.

[25] Li, D.; Sun, X.; Elkhouchlaa, H.; Jia, Y.; Yao, Z.; Lin, P.; Li, J.; Lu, H. Fast detection and location of longan fruits using UAV images. Comput. Electron. Agric. 2021, 190, 106465.

[26] Soares, V.H.A.; Ponti, M.A.; Gonçalves, R.A.; Campello, R.J.G.B. Cattle counting in the wild with geolocated aerial images in large pasture areas. Comput. Electron. Agric. 2021, 189, 106354.

[27] Zhang, Y.; Chu, J.; Leng, L.; Miao, J. Mask-Refined R-CNN: A network for refining object details in instance segmentation. Sensors 2020, 20, 1010.

[28] Zhang, Q.; Liu, Y.; Gong, C.; Chen, Y.; Yu, H. Applications of deep learning for dense scenes analysis in agriculture: A review. Sensors 2020, 20, 1520.

[29] Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. Comput. Electron. Agric. 2018, 147, 70–90.

[30] Murthy, C.B.; Hashmi, M.F.; Bokde, N.D.; Geem, Z.W. Investigations of object detection in images/videos using various deep learning

techniques and embedded platforms—A comprehensive review. Appl. Sci. 2020, 10, 3280.

[31] Zou, X. A review of object detection techniques. In Proceedings of the 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA), Xiangtan, China, 10–11 August 2019; pp. 251–254.

[32] Liu, L.; Ouyang, W.; Wang, X.; Fieguth, P.; Chen, J.; Liu, X.; Pietikäinen, M. Deep learning for generic object detection: A survey. Int. J. Comput. Vis. 2020, 128, 261–318

[33] Etienne, A.; Ahmad, A.; Aggarwal, V.; Saraswat, D. Deep Learning-Based Object Detection System for Identifying Weeds Using UAS Imagery. Remote Sens. 2021, 13, 5182.

[34] Tong, K.; Wu, Y.; Zhou, F. Recent advances in small object detection based on deep learning: A review. Image Vis. Comput. 2020, 97, 103910.

[35] Yang, M.D.; Tseng, H.H.; Hsu, Y.C.; Tsai, H.P. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images. Remote Sens. 2020, 12, 633.

[36] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788

[37] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In European Conference on Computer Vision; Springer: Berlin/Heidelberg, Germany, 2016; pp. 21–37.

[38] Wu, B.; Iandola, F.; Jin, P.H.; Keutzer, K. SqueezeDet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 446–454.

[39] Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. IEEE Trans. Pattern Anal. Mach. Intell. 2020, 42, 318–327.

[40] Law, H.; Deng, J. CornerNet: Detecting Objects as Paired Keypoints. Int. J. Comput. Vis. 2020, 128, 642–656.

[41] Tan, M.; Pang, R.; Le, Q.V. EfficientDet: Scalable and efficient object detection. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 10778–10787.

[42] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

[43] . Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV'15), Washington, DC, USA, 7–13 December 2015; pp. 1440–1448.

[44] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. arXiv 2015, arXiv:1506.01497.

[45] He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV'17), Venice, Italy, 22–29 October 2017; pp. 2961–2969.

[46] Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'17), Kalakaua Ave, HI, USA, 21–26 July 2017; pp. 2117–2125.

[47] Yang, M.D.; Tseng, H.H.; Hsu, Y.C.; Yang, C.Y.; Lai, M.H.; Wu, D.H. A UAV Open Dataset of Rice Paddies for Deep Learning Practice. Remote Sens. 2021, 13, 1358.

[48] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.