Rice seedlings detection

Donato Francioso
University of Bari "Aldo Moro"
d.francioso7@studenti.uniba.it

Abstract

As the world's population grows, there is an increasing need for prduction of food resources. In particular, rice is one of the most widely produced and consumed crops in the world. It is therefore important to optimize and monitor the harvest to increase the quality and yield of it. In this research, Yolo will be used to survey rice seedlings acquired through high-resolution UAV imagery. These results can then be used for monitoring and control of rice fields.

1. Introduction

As the world's population grows, the demand for food also increases, consequently the use of land aimed for this purpose and the amount of product grown increases. In particular, rice is one of the most demanded crops and produced by the world population [1]. Only in asia there is a production of 677 million tons of rice totaling 89% of total production [2]. Therefore, it is very important to monitor and optimize the fields by helping farmers and companies to optimize harvests, to avoid waste and save on costs. This is helped by precision agriculture that uses modern technologies such as satellites to optimize processes that are normally done manually or not feasible by a human such as monitoring large areas in a very short time. This approach offers many advantages such as minimizing material and resource costs through use of techniques to adjust the amount of fertilizer or water to be used according to the area of use or movement optimization to save diesel fuel. Another advantage is monitoring soil or plant health by reducing the number of pesticides or using the results to analyze yield and crop performance. Monitoring agriculture [3] using satellites has been widely used for years already [4]. Although satellites are autonomous and can cover a very large area, they are limited in both cost and image resolution. In addition, data obtained from satellites must first be processed, so they are not immediate. With the use of drones (UAVs), we can obtain very high resolutions covering an average large area, in addition the cost of a drone is obviously lower than a satellite. Although it needs an operator to fly but the data are immediately accessible and ready for analysis . UAVs have made several applications on combining multispectral data, thermal data, and field information to classify crop species, assess disasters, and monitor plant growth [5-6]. In this research, rice seedlings will be detected on UAV-captured images. These results can be used to cover large areas of rice fields to estimate crop yields and monitor their status, giving the ability to take localized action in a very short time.

2. Related Work

On UAV images for crop monitoring Egli and Hopke [7] developed a lightweight convolutional neural network (CNN) for automated tree species classifying with high-resolution UAV images, Chen et al.[8] applied an object detection network on counting strawberries with ultra-high-resolution UAV images for yield prediction. Yang et al. [9] applied deep-learning to UAV images to estimate rice lodging over a vast area. Li et al. [10] proposed an improved object detection model for high-precision detection of hydroponic lettuce seedlings. Pearse et al. [11] applied a CNN model for tree seedlings detection to map and monitor the regeneration of forests in UAV images. Oh et al. [12] applied the object detection technique to cotton seedling counting in UAV images to analyze the plant density for precision field management. Hsin-Hung Tseng, Ming-Der Yang, R. Saminathan, Yu-Chun Hsu, Chin-Ying Yang and Dong-Hong Wu [13] used machine learning and deep learning models, using the same dataset used in this project, for rice seedling detection. Specifically, they used a support vector machine (SVM), with HOG for feature extraction, and two CNN models, EfficientDet and Faster R-CNN. Compared with the model tested in this project, a much more larger part of the dataset was used for training the models. This work is used to make a comparison with the results obtained from this research.

3. Data process

A public dataset of rice plants [14] was used for this research, by which the Yolov5 model has been trained to de-

tect rice seedlings. Next, a comparison was made with the results of state of art models.

3.1. Dataset description

The dataset consists of orthomosaic image. It is the image stitched from a series of nadir-like view UAV images. The dataset provides 13 images for consecutive growth stages, which were imaged in 2018, 2019 and 2020. For training our model was been used the 2018-08-07 image, in particular were been selected several part of the demo from partition number 1, 3, 5 and 6. That partition have different light and different level of seedling growth. For test has been used the number 4 of the 2018 and some sufields of 2019-03-26 and 2018-08-23.

3.2. Data processing

Roboflow was used for image processing. This tool helps us to easily organize, labielize, and prepare the images to be used. As a first step, images were labelized by selecting rice seedlings, after which a data augmentation was performed with a 90° rotation of the images and with a rotation, of the rice seed boxes, with a rotation between -20° and 20°. From 8 initial training images, a total of 24 images were then generated. The generated dataset provides the labeled images and file with all the position of our rice box.

4. Methodology

YOLO was used for the detection of rice seedlings. YOLO (You Only Look Once) is a CNN model used for Object detection with high performance.

4.1. CNN architecture

CNNs are regularized versions of multilayer perceptrons. Convolutional networks were inspired by biological processes, as the pattern of connectivity between neurons resembles the organization of the animal visual cortex.

The classical architecture of the CNN consists of :

- **Convolutional Layers:** through windows called kernels applies filters to extract features from the portion of the image under consideration returning a feature map.
- **Pooling layers:** Reduces dimensionality for each feature map leaving only the most important ones.
- **Fully connected layers:** The Fully Connected layer, also called Dense layer, uses the features coming from the convolutional and pooling layers for classifying the input image into different classes.

In figure 1 an image of cnn architecture.

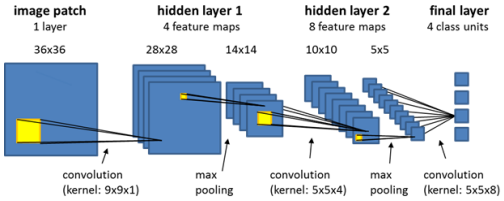

Figure 1. CNN architecture.

4.2. YOLO architecture

YOLO is an object detection algorithm that divides images into a grid system. Each grid cell is responsible for detecting the objects within it. The main idea of it is to produce a collection of bounding boxes of fixed size and, for each of them, calculate the probability that it contains an instance of an object class. It requires very few computational resources.

Yolov5, which is one of the most widely used versions, was used in this study. This version is also available on github.

In figure 2 is showed Yolov5 architecture.

First part consist in Backbone, it takes as input the image and extracts the feature map on which the rest of the network is based. YOLOv5 uses Focus structure with CSPdarknet53 as a backbone. This focus is introduced in this version and is the most important difference between YOLOv4. Second part is Neck, a series of layers which combine the extracted features. The final one is the Head part which uses the features from the neck to predict boxes coordinates (center + width, height) and to classify objects (confidence score).

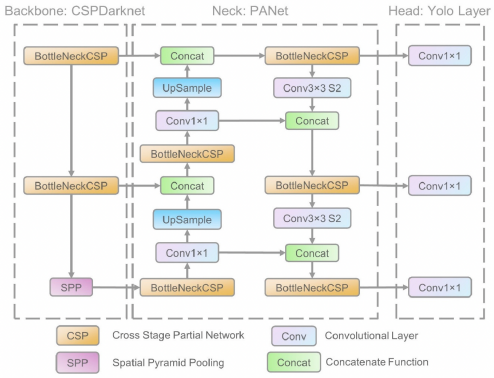

Figure 2. Yolov5 architecture.

4.3. Method application

For rice seed detection, a model was trained from the beginning. Training was then initially done on the dataset generated with Roboflow. Specifically, the model was trained

for 250 epochs with a batch size of 32. The initial weights were chosen from one configuration of Yolov5, the small one, which is lighter than the others . For hyperparameter selection, the three default Yolo configurations were tested, in addition, the internal "evolve" method was used, which generates a minimum of 300 combinations tested for 10 epochs to select the best hyperparameter combination. The method uses a genetic algorithm for optimization which is an algorithm based on the models of biological evolution. This is because in the case of many hyperparameters classical optimization methods have problems with a high dimensional search space and if the correlations among the dimensions in unknown. The best results were obtained with one of the predefined yolo settings. The main values can be seen in Table 1.

Parameter	value
Initial learning rate	0.01
Momentum	0.937
Weight decay	0.0005
IoU threshold	0.2

Table 1. Relevant Hyperparameters.

After this evaluation the model training is performed. The training was completed in about half an hour.

After this step, a simple validation step was performed, and finally a detection could be carried out on multiple images.

4.4. Model Performing

For evaluate the training has been used precision, recall, mAP 0.5, mAP 0.5:0.95 and F1 score.

- **Precision:** is the fraction of relevant instances among the retrieved instances.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

- **Recall:** is the fraction of relevant instances that were retrieved.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

- **F1 score:** is the weighted harmonic mean of precision and recall of a classifier, taking α=1.

Before describe the last two metrics first we need to introduce the Intersection over Union (IoU) for object detection.

IoU is a number that quantifies the degree of overlap between two boxes. It evaluates the overlap of Ground Truth and Prediction region. In figure 3 an example.

$$IoU = \frac{AreaOfOverlap}{AreaOfUnion} \tag{3}$$

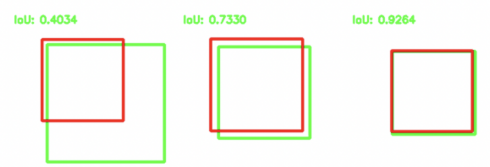

Figure 3. IoU example.

- **mAP 0.5:** is the mean Average Precision with $IoU > 0.5$.
- **mAP o.5:0.95:** is the mean Average Precision at different IoU thresholds, from 0.5 to 0.95.

In table 2 are showed the training results.

Metric	Value
Precision	0.913
Recall	0.841
mAP 0.5	0.890
mAP 0.5:0.95	0.325

Table 2. Train metrics.

In figure 4 is showed the precision and the box loss during the training, while in figure 5 and 6 are showed Precision-Recall and F1_score curve.

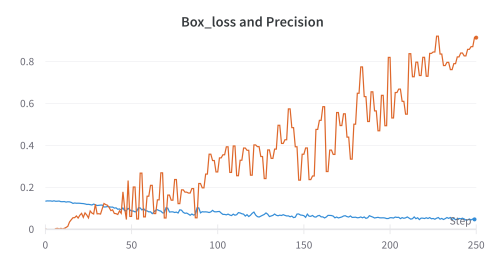

Figure 4. Precision and box_loss.

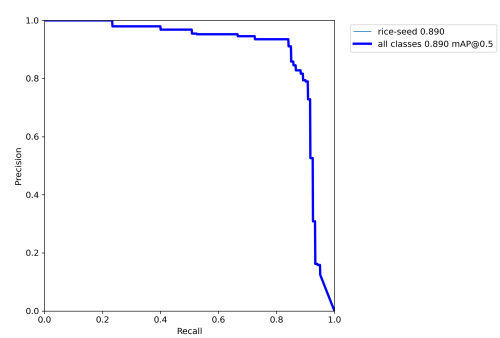

Figure 5. Precision-Recall curve.

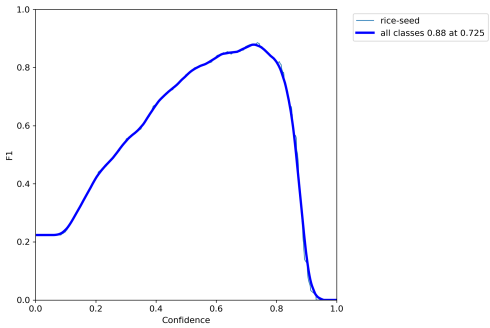

Figure 6. F1_curve.

For validation are tested a small part of the demo number 7 with a total of 500 items. The results are showed in table 3. The obtained values are very similar to the training, only recall and mAP_0.5 is slightly lower.

Metric	Value
Precision	0.939
Recall	0.721
mAP 0.5	0.762
mAP 0.5:0.95	0.308

Table 3. Validation metrics.

Figure 7 shows a survey performed on demo number 4 with a confidence threshold of 0.4. As can be seen only very few seedlings were not detected.

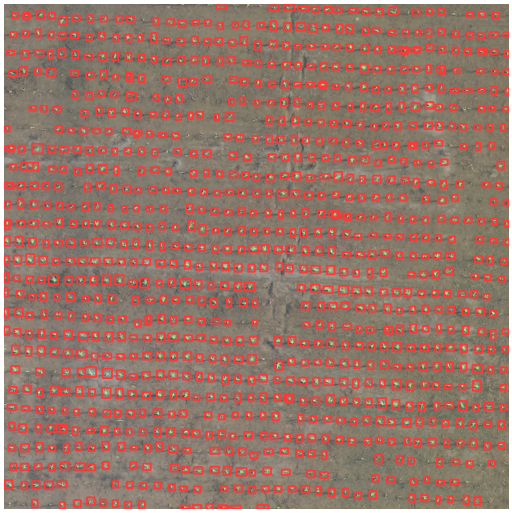

Figure 7. Test on n.4 demo.

The model was also tested on completely different fields and Figure 8 and 9 show the results. In this case it was not able to perfectly detect the whole field but still good results were obtained. In the case of Figure 8 we have grown rice plants so YOLO has difficulty detecting it, this is because it was trained on much smaller seedlings, in fact there are some boxes that are overlapped. There is about a two-week difference between the growth of the training and this test. In Figure 9 there are very few overlapping boxes and the detection worked better. However, it can be seen that the soil is extremely different, it has a coloration very similar to rice seedlings, so it is much more difficult to spot them.

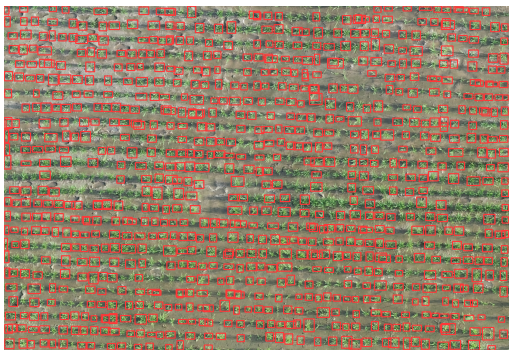

Figure 8. Detect on 2018-08-23 subfield.

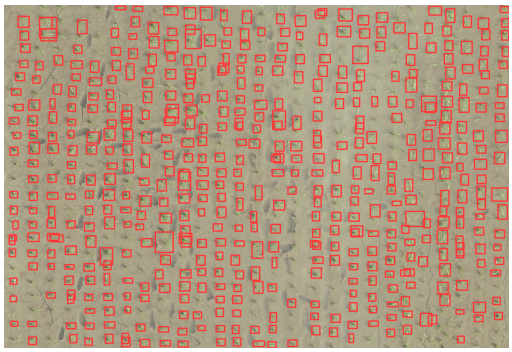

Figure 9. Detect on 2019-03-26 subfield.

5. Model Comparison

The created model was compared with a three models from the state of the art [14]. The first is an SVM machine learning model with HOG for feature extraction. The second and third models are CNN models. In particular the first is EfficientDet, that is a highly scalable one-stage object detection architecture. It inherits the prior-developed model EfficientNet as the backbone for feature extraction. The second model is a Faster R-CNN, it is a two-stage object detection network that inherits the robustness of the R-CNN family with a precise detecting capability.

5.1. Comparison

In table 4 and 5 are showed the mAP metric for all models.

Model	mAP
HOG-SVM	0.700
EfficientDet	0.955
Faster R-CNN	1.000
YOLOv5	0.890

Table 4. mAP for each model for Training phase.

Model	mAP
HOG-SVM	0.702
EfficientDet	0.837
Faster R-CNN	0.888
YOLOv5	0.762

Table 5. mAP for each model for Test phase.

Yolo model is trained and tested on a smaller number of labels then state of art but the results are equally good. We can see that the value on the test is slightly lower than the CNN models but higher than the machine learning method.

6. Conclusion

As could be seen from the previous sections, we get very good results with YOLO despite being trained with a smaller number of training images. These numbers, make YOLO a very good model for rice seedling detection and thus of excellent use in precision farming models. Since it is also very light and fast, it can be used in different machine without having to spend a lot of money in sophisticated architecture.

7. Bibliography

References

[1] Bekele Shiferaw, Boddupalli M. Prasanna, Jonathan Hellin, Marianne Banziger *Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security.*

[2] *Food and agriculture organization of the united nations*

[3] Xiangming Xiao , Stephen Boles , Steve Frolking , Changsheng Li , Jagadeesh Y. Babu , William Salas , Berrien Moore III *Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images*

[4] Sanford w. Downs, Jr. Astrionics Laboratory *Remote sensing in agriculture*. NASA technical memorandum.

[5] Geun-Ho Kwak and No-Wook Park, *Impact of Texture Information on Crop Classification with Machine Learning and UAV Images*, Department of Geoinformatic Engineering, Inha University.

[6] Ming-Der Yang , Kai-Siang Huang , Yi-Hsuan Kuo , Hui Ping Tsai , and Liang-Mao Lin, *Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery*, Department of Civil Engineering, National Chung Hsing University.

[7] Sebastian Egli ,and Martin Hopke *CNN-Based Tree Species Classification Using High Resolution RGB Image Data from Automated UAV Observations*, Laboratory for Climatology and Remote Sensing, Philipps-University of Marburg.

[8] Yang Chen , Won Suk Lee , Hao Gan , Natalia Peres , Clyde Fraisse, Yanchao Zhang and Yong He, *Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages*

[9] Ming-Der Yang, Hsin-Hung Tseng , Yu-Chun Hsu and Hui Ping Tsai, *Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images*.

[10] Zhenbo Lia, Ye Lia, Yongbo Yanga, Ruohao Guoa, Jinqi Yanga, Jun Yue, Yizhe Wang, *A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN*.

[11] Grant D. Pearse, Alan Y.S. Tan, Michael S. Watt, Matthias O. Franz, Jonathan P. Dash, *Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data*.

[12] Sungchan Oh , Anjin Chang, Akash Ashapure, Jinha Jung , Nothabo Dube, Murilo Maeda, Daniel Gonzalez and Juan Landivar, *Plant Counting of Cotton from UAS Imagery Using Deep Learning-Based Object Detection Framework*.

[13] Hsin-Hung Tseng, Ming-Der Yang, R. Saminathan, Yu-Chun Hsu, Chin-Ying Yang and Dong-Hong Wu, *Rice Seedling Detection in UAV Images Using Transfer Learning and Machine Learning*

[14] Ming-Der Yang, Hsin-Hung Tseng, Yu-Chun Hsu, Chin-Ying Yang, Ming-Hsin Lai and Dong-Hong Wu *A UAV Open Dataset of Rice Paddies for Deep Learning Practice*.