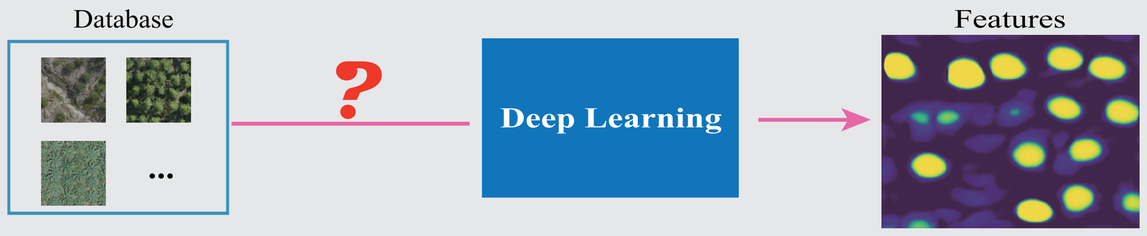
Point in: Counting Trees with Weakly Supervised Segmentation Network

Pinmo Tong, Penghceng Han
Supervised by Prof. Shuhui Bu
Department Of aeronautics, Northwestern Polytechnical University

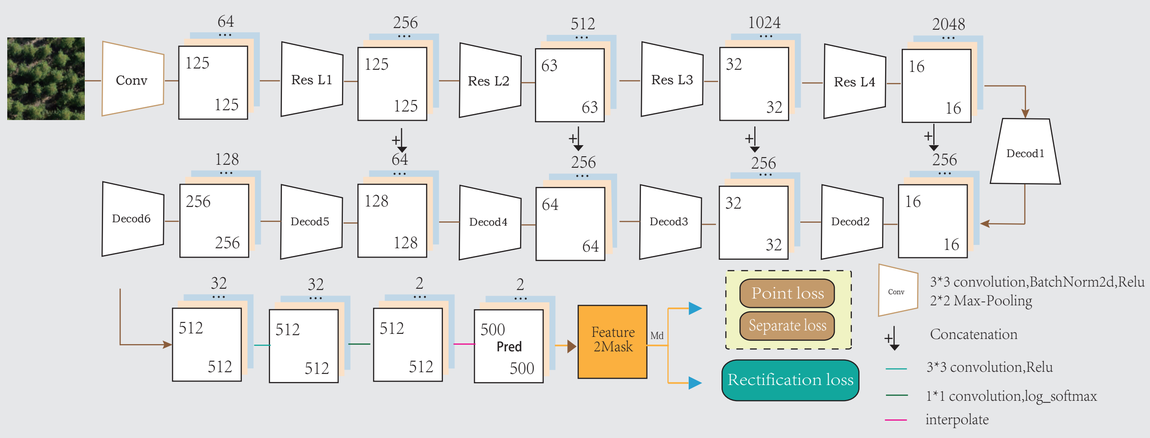
1. Aims and MOTIVATION Of the Paper

Information about the number, location, size and status of economic trees is essential for precision agriculture. Traditional image processing methods require expensive feature engineering and are not end-to-end frameworks. ***Deep learning based feature extraction methods*** demonstrate its great ability in object recognition. However, for large aeral plantations, bounding box or full pixel annotations is time-consuming and easy to get wrong lable data.

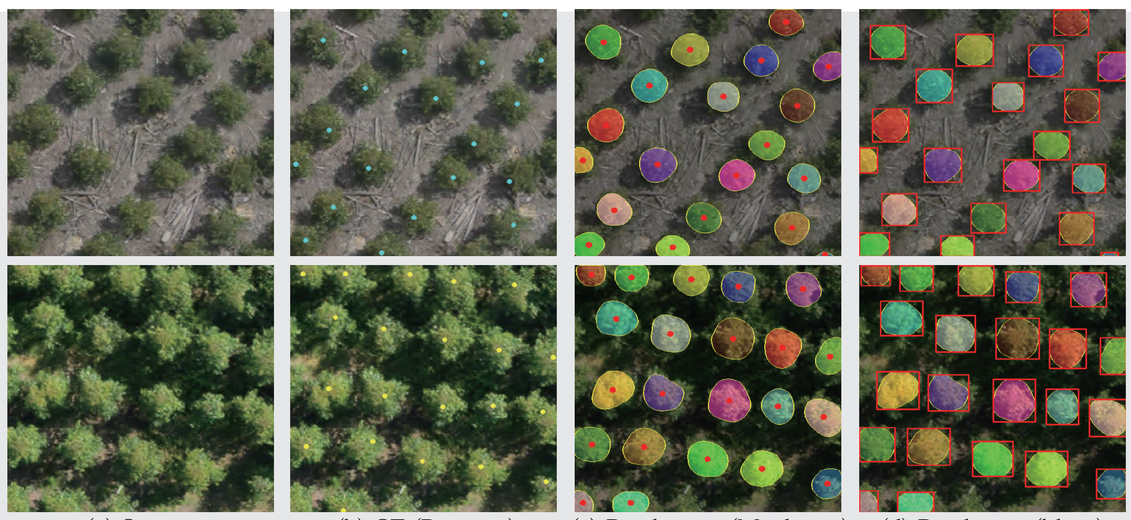
Therefore, the aim of this paper is to obtain the localization and mask of the object at the same time with only ***point supervision***, and make sure the accuracy can be comparable to the full supervised methods.



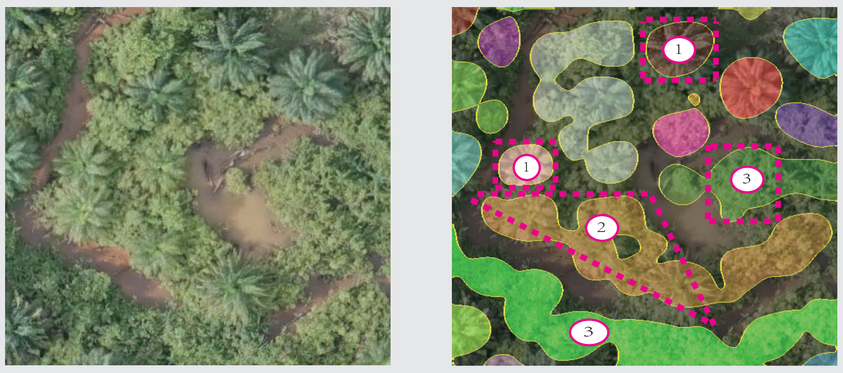
2. The base Algorithm



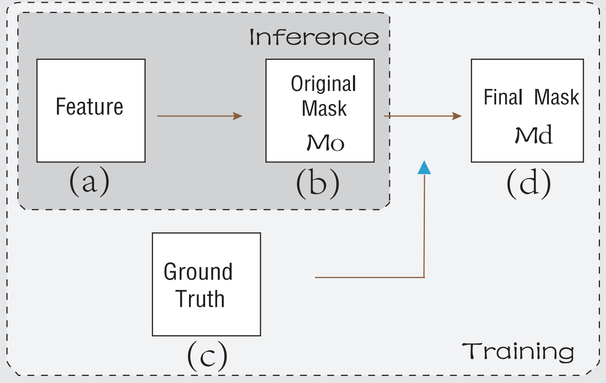
Our project address:
https://github.com/tongpinmo/WTCNet

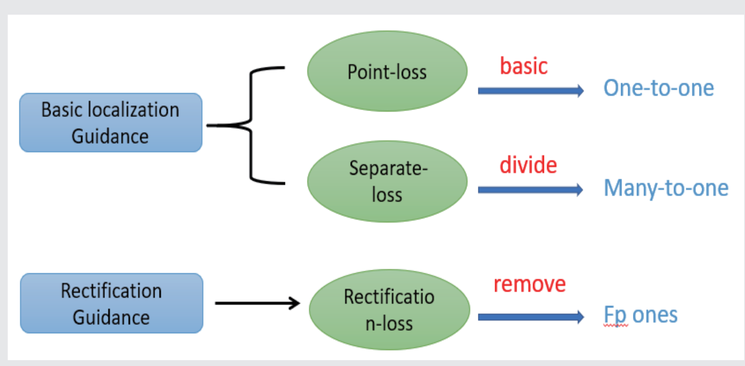


(a) Source (b) GT (Point in) (c) Prediction (Mask out) (d) Prediction(bbox)



(a) Original image (b) Different Masks





Complete multiple tasks

The annotation of the framework is only point supervision, and the output consists of counting, position, mask.

Challenges of the task

There will be occlusion, overlap and scale challenges due to the complicated species and background.

Feature to mask strategy

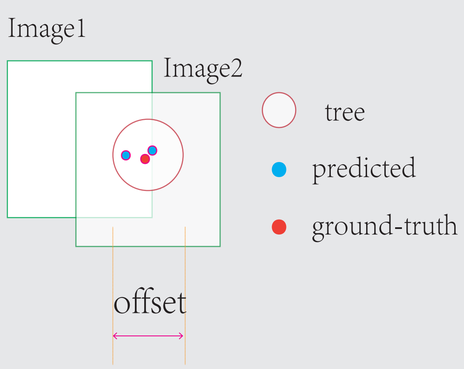
Since there will be three main kind of mask situations which shown in (b), the Feature2Mask strategy is proposed to deal with it. In the training process, the final mask is getted with output of Feature Extractor Module and ground truth. In the inference process, the mask is only obtained from Feature.

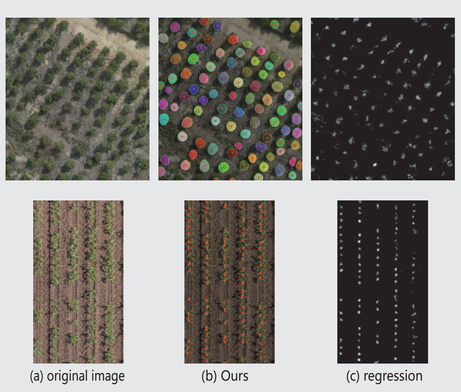
Loss function

The loss function consists of two parts: Basic localization guidance and rectification guidance. The loss is designed to encourage the network to output a single mask for per object.

3. Experiments

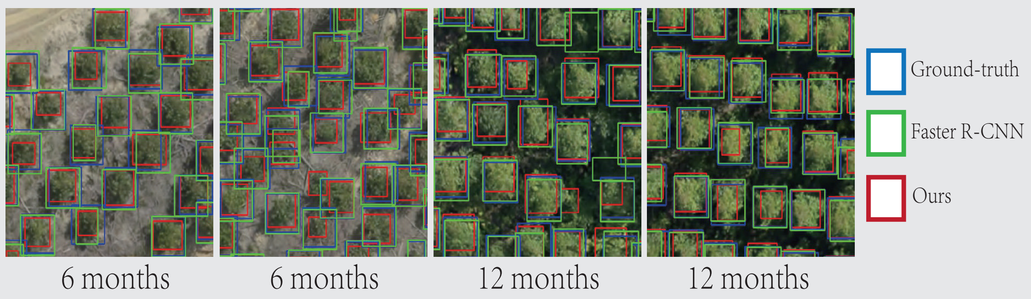
Optimal Parameters Selection and Counting Visualization





(a) original image (b) Ours (c) regression

Experiments on Mask Performance



6 months 6 months 12 months 12 months

Recognition

Table 1. The quantitative results of parameters setting over the Acacia dataset.

THE QUANTITATIVE RESULTS OF PARAMETERS SETTING OVER THE ACACIA DATASET.

Parameters		Acacia-6			Acacia-12		
d_α	d_β	TPR	Prec	F_1	TPR	Prec	F_1
60	30	0.963	0.977	0.970	0.924	0.918	0.921
60	20	0.895	0.908	0.901	0.761	0.756	0.758
40	**40**	**0.975**	**0.991**	**0.983**	**0.967**	**0.950**	**0.958**
40	30	0.968	0.977	0.972	0.928	0.911	0.920
40	20	0.904	0.912	0.908	0.765	0.751	0.758
30	30	0.968	0.974	0.971	0.929	0.900	0.915
30	20	0.905	0.910	0.908	0.768	0.745	0.756

Table 2. The counting performance compared with different methods over datasets.

THE COUNTING PERFORMANCE COMPARED WITH DIFFERENT METHODS OVER DATASETS.

Dataset	Acacia-6		Acacia-12		Oil Palm		Sorghum Plant	
Measures	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
MCNN [12]	12.32	52.06	21.48	100.15	4.48	5.53	11.07	15.03
HA-CCN [23]	4.12	18.42	6.07	30.94	3.67	4.81	4.01	9.59
CAN [24]	3.35	12.06	5.61	21.58	2.49	4.12	2.11	3.62
Ribera [20]	-	-	-	-	-	-	**1.9**	**2.7**
Ours	**2.309**	**3.455**	**4.621**	**6.025**	**2.270**	**3.345**	2.33	3.43

4. Discussion & Conclusions

Contributions:

Propose a novel weakly supervised network architecture to estimate the number, location, and mask of trees in an image, without any notion of bounding boxes or region proposals.

Introduce a feature2mask strategy and a novel loss function that encourages the network to output a single mask for per tree.

Evaluate our method on proposed datasets. The results show the superiority of our method in three tasks: counting, localization, and mask performance.

Future Work:

Explore new features of the proposed method and improve the efficiency to deal with complicated shapes.