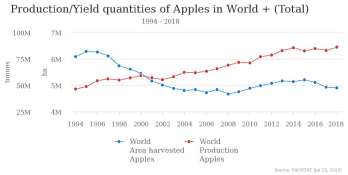
Applying Faster R-CNN and Mask R-CNN on the MinneApple Fruit Detection Challenge

Callum Clark

Introduction

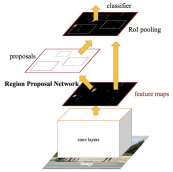
In this thesis we examine the problem of apple detection and localization as an Object Detection problem, applied to the challenging real-world dataset MinneApple.

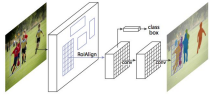
Total global production of apples increased from 46.47 million MT in 1994 to 86.14 million MT tons in 2018 (FAOSTAT)
Use of **technology** in the agricultural food chain had led to greater crop yields from smaller amounts of land



Detection and localization of fruits can improve **yield estimates** as well as various automated picking [4][5] and pruning[6] systems

We train and compare **Faster R-CNN** and **Mask R-CNN** models with various **ResNet backbones** to compare to current state of the art results on the **MinneApple** dataset. A collection of images of apple trees taken from the University of Minnesota's Horticultural Research Center (HRC).





MinneApple Dataset

Previous datasets used either circles or boxes as ground truth labels.
MinneApple gives ground truth labels as **polygons**, allowing for **semantic segmentation**

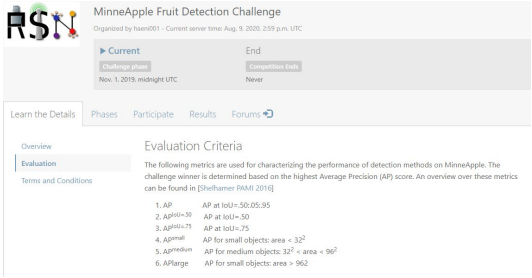
Advantages of MinneApple Dataset

- Highest numbers of **annotations per image**
- High res images of **full apple trees**
- **Large variety** of species
- **Different levels** of illumination
- Many **different sections** of the HRC's orchards

Methods

- For both Faster R-CNN & Mask R-CNN methods we use various **ResNet backbones**
- Residual Networks allow us to **train deeper networks** without the vanishing gradient problem
- We used a **50 layer** and a **101 layer ResNet backbone**
- Also a **101 layer ResNeXt backbone**
- These backbones were pretrained on the ImageNet classification tasks

Evaluation & CodaLab Competition



Experiments

We trained models with different ResNet backbones to compare the results.

Method	Backbone	AP
Faster R-CNN	ResNet-50	0.398
	ResNet-101	0.425
	ResNeXt-101	0.436
Mask R-CNN	ResNet-50	0.386
	ResNet-101	0.394
	ResNeXt-101	**0.441**

Results

Ours against MinneApple paper

Author	Method	AP@ IoU[.50:.05:0.95]	AP@ IoU=.50	AP@ IoU=.75	AP$small$	AP$medium$	AP$large$
Hani[8]	Faster R-CNN	0.438	0.775	**0.455**	0.297	0.578	**0.871**
Hani[8]	Mask R-CNN	0.433	0.763	0.449	0.295	0.571	0.809
Kaka[39]	NA	0.436	0.770	0.453	0.285	[**0.592**]	0.872
Ours	Faster R-CNN	0.436	0.791	0.436	0.291	0.590	0.848
Ours	Mask R-CNN	**0.441**	**0.801**	0.440	**0.300**	0.589	0.861

CodaLab Leaderboard

Results										
#	User	Entries	Date of Last Entry	Team Name	AP at IoU=.50:.05:.95 ▲	AP at IoU=.50 ▲	AP at IoU=.75 ▲	AP for small objects ▲	AP for medium objects ▲	AP for large objects ▲
1	CallumClark	122	07/29/20		0.441 (1)	0.801 (1)	0.440 (3)	0.300 (1)	0.589 (2)	0.861 (3)
2	kuka	228	07/13/20		0.436 (2)	0.770 (3)	0.453 (1)	0.285 (3)	0.592 (1)	0.872 (1)
3	varun_shekar	2	04/05/20		0.426 (3)	0.747 (5)	0.443 (2)	0.273 (6)	0.583 (3)	0.787 (6)
4	wuhaogxu	4	08/31/20		0.422 (4)	0.753 (4)	0.438 (4)	0.274 (4)	0.579 (4)	0.853 (5)
5	leonardolima	2	05/29/20		0.419 (5)	0.744 (6)	0.431 (5)	0.274 (5)	0.571 (5)	0.864 (2)
6	cuge1995	8	09/22/20		0.415 (6)	0.773 (2)	0.412 (6)	0.287 (2)	0.553 (6)	0.272 (7)
7	selimgonen	34	07/20/20		0.160 (7)	0.291 (7)	0.154 (7)	0.034 (7)	0.313 (7)	0.856 (4)

Discussion

Methods

- Benchmark states that Faster R-CNN is the best performer
- Interesting insight is that **Mask R-CNN outperforms Faster R-CNN** (could be due to the Mask R-CNN using **semantic segmentation** and may have learnt to deal with clusters of apples better than the Faster R-CNN which only uses bounding Boxes)

Dataset

- MinneApple still a **small sample of total species** and different conditions around the world significant difference between apples grown in **different continents**, as well as **different countries**
- **New varieties** constantly being bred
- Future datasets need to include these variations to train more general models

Conclusion

Detectron2's **Mask R-CNN** with a **ResNeXt-101 backbone** achieves state of the art accuracy on the MinneApple Dataset

It is currently in first place in the challenge (as of 5/11/2020)

BNAIC 2020 BENELEARN