Bush Detection for Vision-based UGV Guidance in Blueberry Orchards: Data Set and Methods

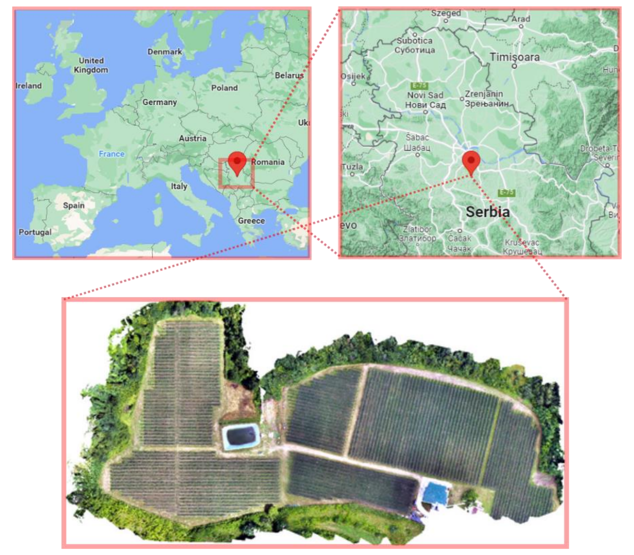
Vladan Filipović, Dimitrije Stefanović, Nina Pajević, Željana Grbović, Nemanja Đurić, Marko Panić
BioSense Institute, Novi Sad, Serbia

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

BioSense INSTITUTE

Motivation

- Advances in blueberry production create the need for bush detection
- Existing applications focus on tree trunk detection for orchard guidance, more complex structures are not yet explored
- Bush detection is important in:
  - agriculture (berry production)
  - forestry (wildfire prevention)



Blueberry orchard location

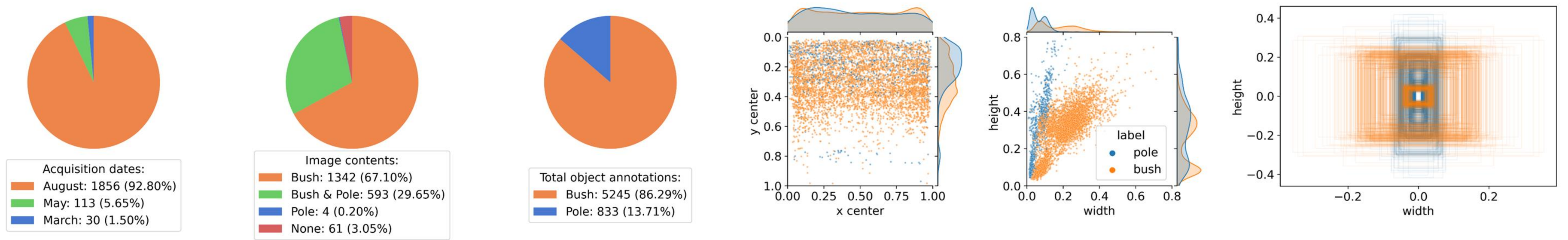
Dataset

- 2000 RGB images captured using a robotic platform and OAK-D devices
- Single orchard in Babe village, Serbia
- Three dates during the growing season
- 20 image sequences for autonomous movement (translation and rotation)
- Two classes of objects: bushes and poles

Robot and OAK-D device

Translation sequence example

Rotation sequence example



Dataset statistics: acquisition dates, object numbers, positions and shapes

Dataset features and variability

Variation in camera angle and height

Variation in bush shape: from young and sparse to mature and dense

Various poles: hale netting poles, bumblebee hive legs, and lamp posts

Variation in contrast and shadows: from light and diffuse to dark and opaque

Various occlusions: different types and quantities of weed and branches

Views from irregular heights, positions, angles; out-of-context objects

Examples of background images with no objects present

Antares FLEXIGROBOTS

Link to the dataset on Zenodo: FlexiGroBots Ground-level Blueberry Orchard Dataset v1 - RGB Bush Detection Dataset

Labeling procedure

- Bushes have undefined borders (unlike cars, cats, and dogs)
- Focus on the base of the plant (inverted triangle, contact with the ground)
- To converge to the standard bush shape, 400 images are labeled by 10 annotators and used to train the initial model, which is later used as a starting point for further labeling (semi-autonomous procedure)
- The whole dataset is then relabeled and corrections are done manually

Difference in annotations between 5 different annotators

Baseline model

- YOLOv5 🚀 model is used in 3 complexity levels: nano, small, and medium
- Train/val/test split is done based on sequences to prevent data leakage
- Augmentation: horizontal flipping, HSV scaling, and mosaic augmentation

Results and conclusion

- Decent model performance in most cases, including shadows, weeds, and occlusions (ex. 1-4); trouble with ambiguity in annotations, far away objects and image rotation (ex. 5-7); good generalization abilities with unseen objects (ex. 8); detection speed suitable for real-time applications.
- **Models achieved promising metrics, setting a good basis for further work**

Resulting detection metrics

Model	All				Bush				Pole				Param count	Latency [ms]
	P	R	mAP 50	mAP 50-95	P	R	mAP 50	mAP 50-95	P	R	mAP 50	mAP 50-95		
YOLOv5n	0.940	0.790	0.859	0.440	0.927	0.840	0.912	0.479	0.954	0.740	0.805	0.401	1.7M	96.4
YOLOv5s	0.892	0.800	0.873	0.472	0.882	0.871	0.909	0.500	0.903	0.730	0.838	0.444	7.0M	179.8
YOLOv5m	0.930	0.797	0.872	0.489	0.935	0.860	0.924	0.510	0.924	0.735	0.820	0.467	20.8M	313.2

Examples of resulting detections