Toward Automatic Urban Forest Inventories with Remote Sensing

Julian Rice, Jonathan Ventura, Dept. of Computer Science & Software Engineering
Camille Pawlak, Matt Ritter, Jenn Yost, Natalie Love, Dept. of Biology
Cameron Gonsalves, G. Andrew Fricker. Dept. of Social Sciences

Tree detector output for a sample urban area in downtown San Luis Obispo, CA. Each red circle indicates a tree detected using our method. The input to our system is 60 cm resolution multispectral imagery from 2018 provided by the US National Agriculture Imagery (NAIP) program.

Introduction

BACKGROUND

Urban forests provide extensive benefits to residents of cities such as controlling microclimate and sequestering carbon.[1] Public street trees are managed by local governments to maximize their benefits. However, public street trees make up only a portion of a city's urban forest and inventories are not located in a central location. The extent of the benefits provided to residents by their urban trees depends on the number of trees in that city, both public and private. When making policy decisions about where the next tree planting is needed most, cities need to be able to account for both the publicly and privately managed regions of urban forest.

OUR RESEARCH

Part of our research has been aggregating and cleaning public street tree inventories into a statewide inventory. This dataset is called the California Urban Forest Inventory (CUFI), and currently has over 7 million points.

Using training data based on the CUFI and manually annotated images, we created a neural network that counts trees using aerial data from the National Agriculture Imagery Program (NAIP). Using this network, we created tree counts for cities within California, and created a spatially explicit point file that has a point for every urban tree in the state. This research allows city managers to create better estimates of the ecosystem services that their urban forests provide and can allow for more robust calculations of the urban forest's ecosystem services and monetary value in the state of California.

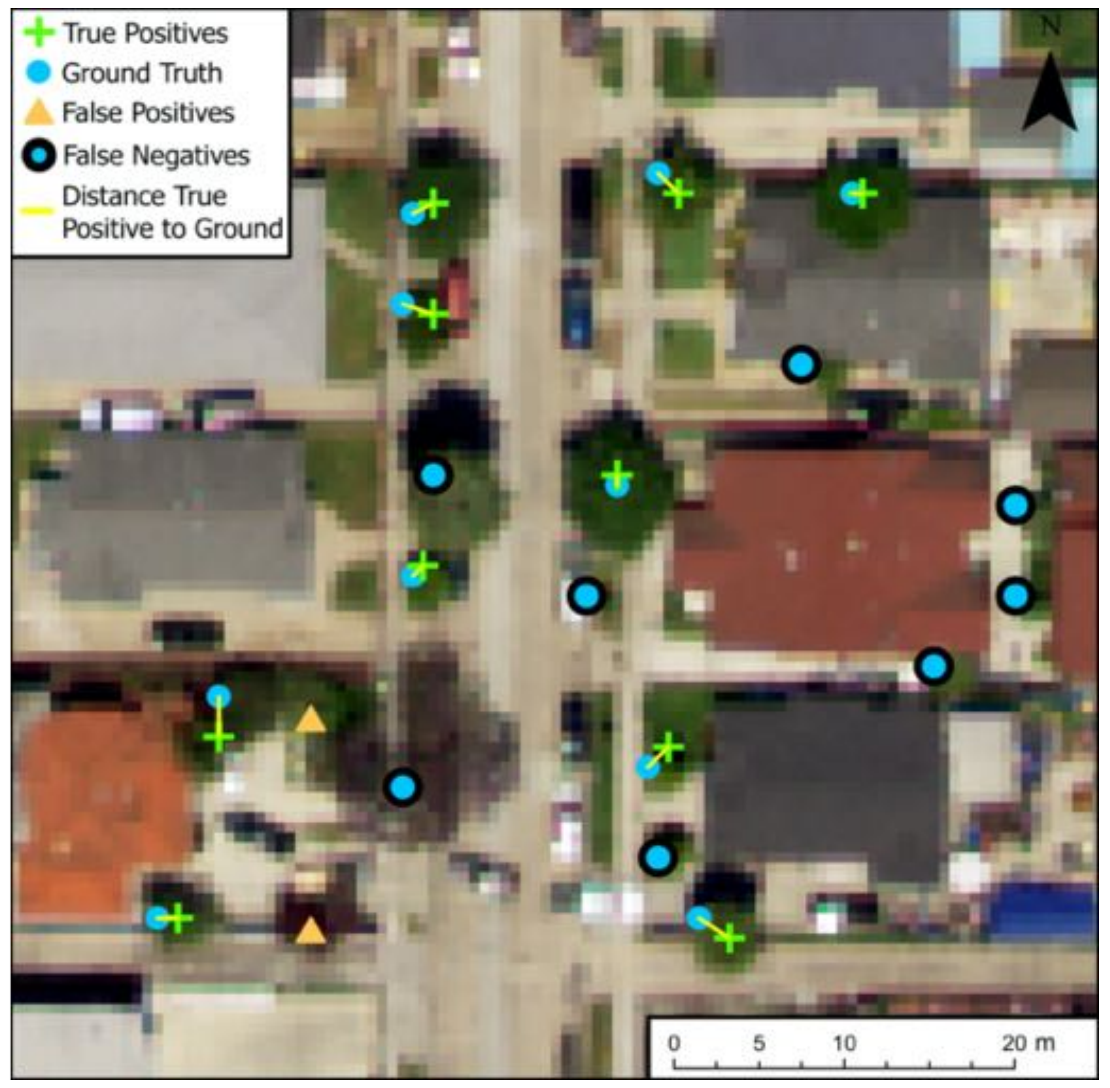
Methods

In our work we adopt a density-based approach to tree counting. The input to the system is a stack of raw and derived rasters. The raster stack is passed through a convolutional neural network which outputs a single-channel density map. Peaks in the density map should correspond to tree locations. We identify tree locations in the density map using local peak finding. Our CNN architecture is based SFANet, a recent state-of-the-art network that is a top performer on several crowd counting benchmarks.[2] SFANet consists of a VGG-16 backbone, a density head, and an attention head.[2,3] We modified the SFANet architecture by adding extra layers to the density and attention heads so that they would output at full-resolution rather than half-resolution. We call our network the High-Resolution SFANet (HR-SFANet).[2] We are using our network to predict tree counts for all urban areas in California.

Results

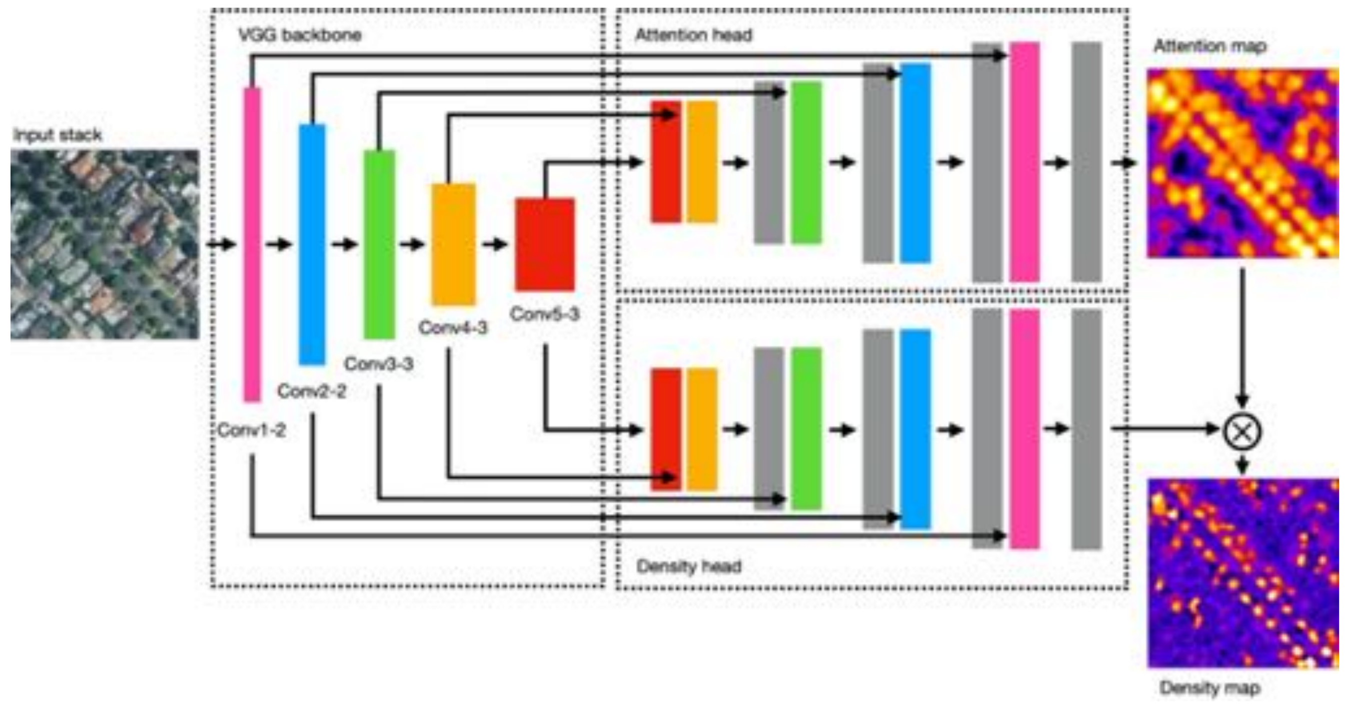
Currently, we have achieved a precision of 0.76, which out-preforms similar research.[4] Our best recall and F-score is 0.73, and our best root mean square error is 3.65. Our overall best model uses the blue, green, red, and near-infrared bands, as well as an NDVI as inputs. Our tree counter estimates that there are over 39.3 million trees within California's urban areas. The CUFI represent over 7 million street trees. The jump from 7 to 39 million trees counted has implications for calculating the value of urban forests for management purposes.

Results from our test data in Claremont, CA. This is an area where the neural network counted trees with high accuracy. Green crosses indicate true positives, blue circles indicate our training data, blue circles with a dark outline indicate false negatives, and orange triangles indicate false positives. We lose accuracy when trees are very dense, shrubs are very large, or there are many shadows in the imagery.



An example where the counter has lower accuracy due to small trees and shadows in Santa Monica, CA. We are currently adding training data to our model to attempt to increase accuracy.

Heat map of the number of predicted trees for all urban areas in California.



The inputs, outputs and structure of our neural network.

Resolution	Channels	Aggregation	Precision	Recall	F-score	RMSE [m]
Full	GR	max	0.71	0.68	0.70	3.86
Full	GRN	max	0.71	0.72	0.72	3.77
Full	GRNV	max	0.73	**0.73**	**0.73**	3.75
Full	BGR	max	0.70	0.71	0.71	3.73
Full	BGRN	max	0.71	**0.73**	0.72	3.68
Full	BGRNV	max	**0.76**	0.70	**0.73**	**3.65**
Full	BGRNV	sum	0.73	0.70	0.71	3.87
Half	BGRNV	max	0.71	**0.73**	0.72	3.85

Results from our model outputs and model comparison. The best scores are bold and underlined. Our complete model (highlighted in gray) is the best performer.

REFERENCES

1 Livesley, S. J., McPherson, E. G., & Calfapietra, C. (2016). The urban forest and ecosystem services: impacts on urban water, heat, and pollution cycles at the tree, street, and city scale. *Journal of environmental quality*, *45*(1), 119-124.

2 Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

3 Zhu, L., Zhao, Z., Lu, C., Lin, Y., Peng, Y., & Yao, T. (2019). Dual path multi-scale fusion networks with attention for crowd counting. *arXiv preprint arXiv:1902.01115*.

4 Zamboni, P., Junior, J. M., Silva, J. D. A., Miyoshi, G. T., Matsubara, E. T., Nogueira, K., & Gonçalves, W. N. (2021). Benchmarking Anchor-Based and Anchor-Free State-of-the-Art Deep Learning Methods for Individual Tree Detection in RGB High-Resolution Images. *Remote Sensing*, *13*(13), 2482.