3D Convolutional Neural Networks for Tree Detection using Automatically Annotated LiDAR data

Ananya Gupta[1,2], Jonathan Byrne[1], David Moloney[1], Simon Watson[2], Hujun Yin[2]
[1]Intel Corporation
[2]The University of Manchester

Introduction

LiDAR data provides a useful means to collect information for a number of tasks such as forest surveying and urban planning (Treepedia, 2015). However, this data needs to be annotated in order for it to be useful. Currently most of this annotation is done manually as is the case in urban green cover surveys, where identifying trees in the cities is a long, laborious process.

This work focuses on making this urban tree annotation process autonomous by the means of using data present in high quality LiDAR scans to automatically label trees. This annotated data can then be used in conjunction with deep learning to identify trees in LiDAR scans without the requisite information.

Methods

In order to identify trees in LiDAR scans, ground points are first identified and filtered using a Progressive Morphological Filter. This filtered scan is then voxelized in a sparse 3D hierarchical data structure, VOLA (Byrne et al., 2017), in order to reduce the input resolution. A 2 bits per voxel approach is used to encode additional information such as colour, intensity and number of returns information.

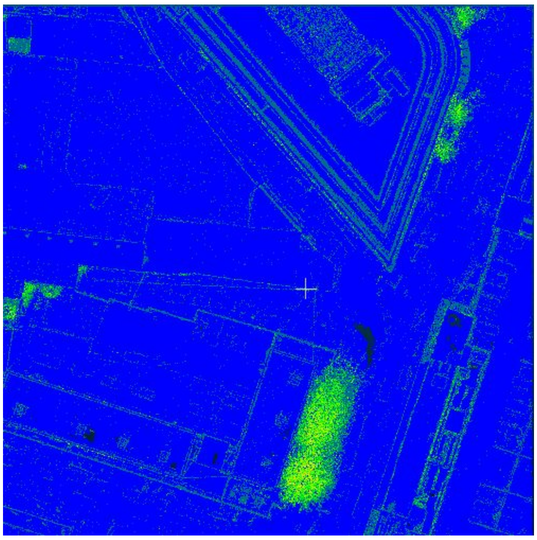

Figure 1: LiDAR scan showing decreasing number of returns: Green>Blue

LiDAR laser pulses can be reflected once (as in the case of a flat surface such as the ground) or multiple times (from edges of buildings, trees etc.). Based on the insight that tree regions have a high number of returns, voxels with a high number of returns are identified and retained, followed by connected component analysis to isolate individual tree canopies. A horizontal bounding box is fitted around the tree canopies and is extended to ground in the vertical dimension in order to capture the tree trunks.

The trees identified in this case are used as positive samples to train a 3D convolutional neural network (CNN) (Maturana et al., 2015) for tree detection. The structure of the network can be seen in Figure 2. A number of non-tree regions are extracted from the LiDAR scans as negative training samples for the network. The training data is augmented by adding noise, rotating the data around the horizontal plane and by jittering it in all 3 dimensions.

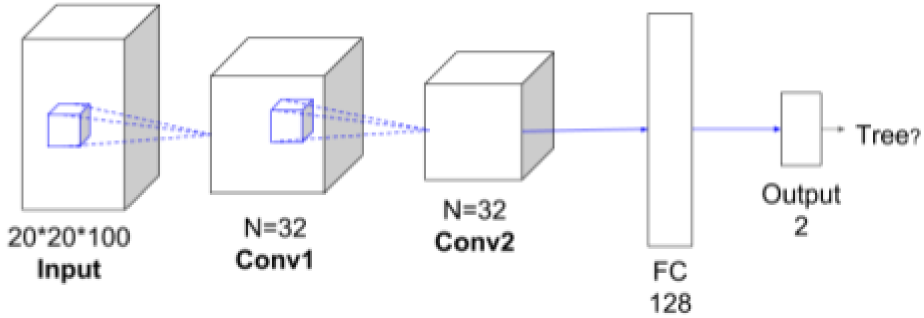

Figure 2: Structure of 3D Convolutional Neural Network

Results

The tree annotation method was tested on a dense airborne LiDAR dataset of Dublin city (Laefer et al., 2015). A subset of the scan was manually annotated for 535 trees and the algorithm was able to correctly identify 469 trees. It returned 56 regions incorrectly identified as trees and had a precision of 0.88 with a recall of 0.89 with 0.88 as the overall accuracy.

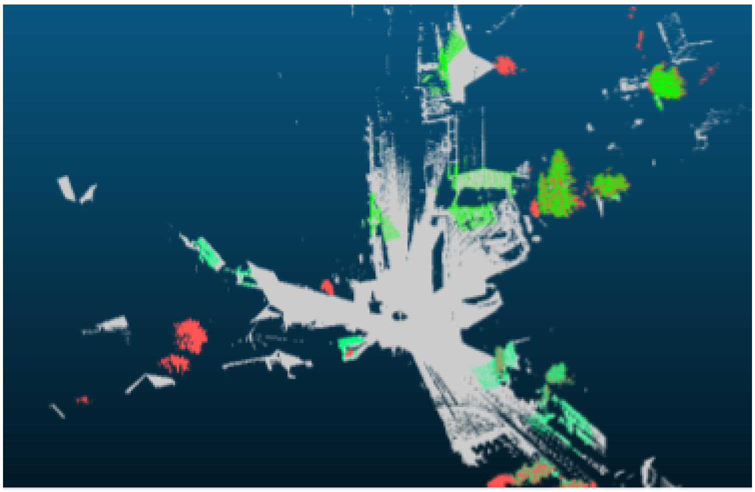

Figure 3: Results of 3D CNN for tree detection; Red: Ground Truth, Green: Trees identified by CNN

The 3D CNN was tested on a publicly available ground based LiDAR dataset which contains both rural and urban scenes (Hackel et al., 2017). The results show that the neural network is able to correctly identify most trees in scan with a few false positives. It misses some of the trees right at the edges of the scan which is possibly due to the missing data in those regions.

Discussion and Conclusion

The methods described in this work provide a way to annotate LiDAR data for trees using the number of returns information. The results show that the algorithm is able to correctly identify most of the trees in an urban setting.

This work also introduces a CNN to detect trees in LiDAR scans which do not contain the number of returns information. The results in this case show that the network is able to identify trees in ground based LiDAR scans which essentially only have depth information (not complete 3D), while being trained on data which is from an aerial LiDAR scan which has full 3D information.

References

Byrne, J., Caulfield, S., Xu, X., Pena, D., Baugh, G., & Moloney, D. (2017). Applications of the VOLA Format for 3D Data Knowledge Discovery . In *International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery* (pp. 1–8).

Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., & Pollefeys, M. (2017). Semantic3D.net: A new Large-scale Point Cloud Classification Benchmark. http://doi.org/10.5194/isprs-annals-IV-1-W1-91-2017

Laefer, D. F., Abuwarda, S., Vo, A.-V., Truong-Hong, L., & Gharibi, H. (2015). 2015 Aerial Laser and Photogrammetry Survey of Dublin City Collection Record. http://doi.org/10.17609/N8MQ0N

Maturana, D., & Scherer, S. (2015). VoxNet: A 3D Convolutional Neural Network for real-time object recognition. *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, 922–928. http://doi.org/10.1109/IROS.2015.7353481

Treepedia. (2015). MIT Senseable City Lab. Retrieved from http://senseable.mit.edu/treepedia