Semantic Segmentation of AHN3 Point Clouds with DGCNN

Qian Bai *

Abstract

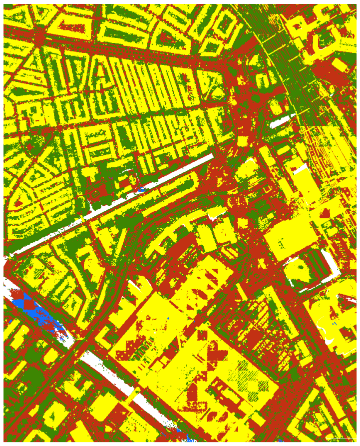
Deep learning approaches designed for 3D point cloud data have made significant progress over the past few years. In this work, we implement DGCNN (Dynamic Graph CNN), which combines PointNet with Graph CNN, and extend its semantic segmentation application from indoor scenes to an aerial point cloud dataset: the Actueel Hoogtebestand Nederland (AHN). Point clouds from the iteration AHN3 are classified into four classes: ground, building, water and others (including vegetation, railways, etc). Moreover, pointwise neural networks usually split the input point cloud into regular blocks before operating on it and process each block independently, which limits the effective range (receptive field) of the network to some extent. Thus, the second aim of this work is to investigate the impact of the effective range on the performance of DGCNN by adjusting two crucial parameters: the block size and the neighborhood size k in k-NN graphs. It turns out that enlarging the block size or k helps to improve the overall accuracy of DGCNN, but cannot ensure better segmentation results from each individual class. With the block size 50 m and $k = 20$, the most balanced F1 scores for all classes and an overall accuracy of 93.28% are achieved. Based on the evaluation for each setting with a certain block size and k, we also manage to further improve the overall accuracy to 93.51% by combining smaller-scale (with block size 30 m) and larger-scale (with block size 50 m) segmentation results, with $k = 20$.

Results

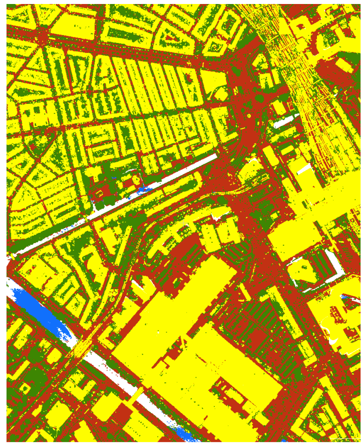
Block size (m)	k	Overall accuracy (%)	Average per class accuracy (%)	Mean IoU (%)
30	20	91.72	81.53	74.94
50	20	**93.28**	89.39	**81.73**
70	20	92.97	**90.90**	79.51
50	15	92.38	88.51	79.98
50	20	93.28	**89.39**	81.73
50	25	**93.30**	89.01	**82.10**
50 & 30	20	**93.51**	**91.60**	82.34
50	15 & 20	93.37	90.48	**82.46**

Table 1: Comparison of the overall accuracy, the average per class accuracy and the mean IoU with different block sizes and k values. "50 & 30" and "15 & 20" in the lowest part indicate results with multi-scale combinations.

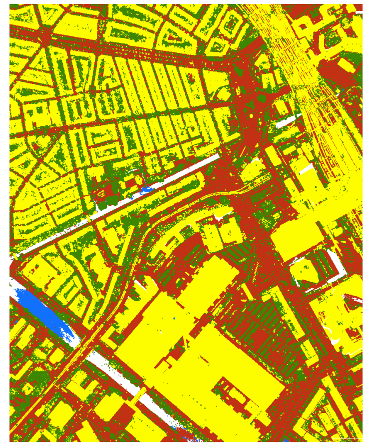
*Dept. of Geoscience and Remote Sensing, Delft University of Technology, The Netherlands



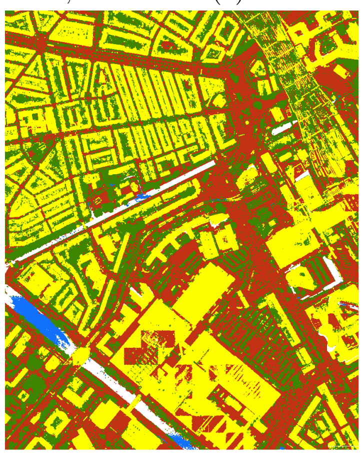
(a) block size = 30 m, $k = 20$



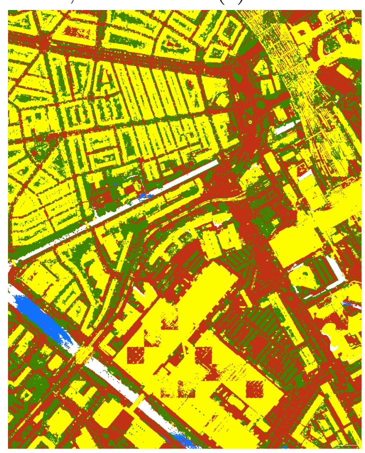
(b) block size = 50 m, $k = 20$



(c) block size = 70 m, $k = 20$



(d) block size = 50 m, $k = 15$



(e) block size = 50 m, $k = 25$

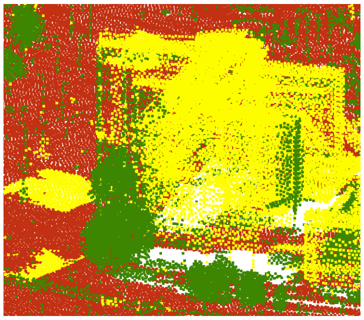
(f) block size = 30 & 50 m, $k = 20$

(g) block size = 50 m, $k = 15$ & 20

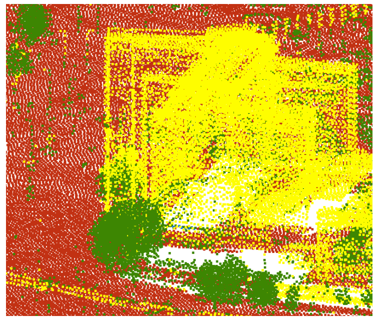
(h) Ground Truth

Figure 1: Segmentation results of a tile in the test area, with ground points in brown, building points in yellow, water points in blue and points from others in green.

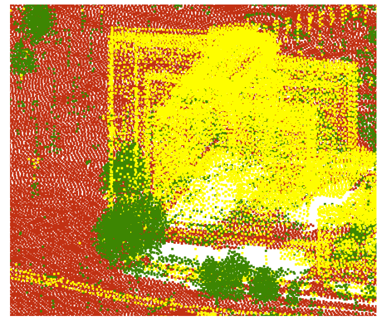
Table 1 summarizes the quantitative results of point cloud semantic segmentation over the test area. Figure 1 and Figure 2 illustrate the segmentation results with different settings over one tile of the test area. Compared to large scales, smaller block size or k can cause some confusion to the classification of building points. The middle of the large roof can be detected as ground points (see Figure 1a), and points on building facades can be labeled as others (trees), as shown in Figure 2a and Figure 2d.



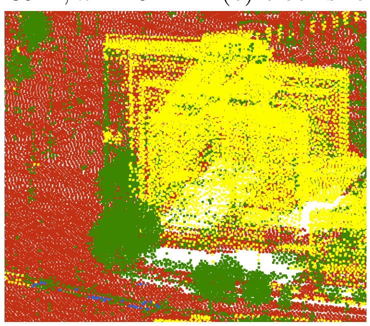
(a) block size = 30 m, $k = 20$



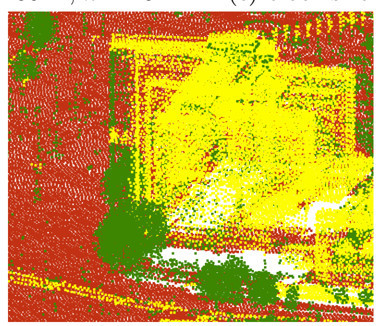
(b) block size = 50 m, $k = 20$



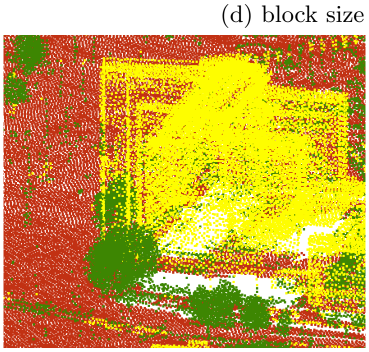
(c) block size = 70 m, $k = 20$



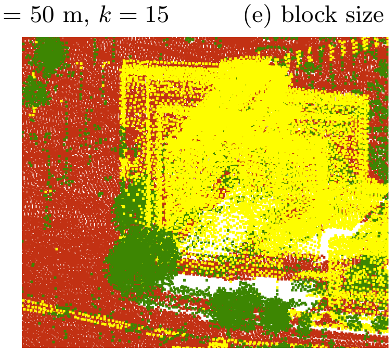
(d) block size = 50 m, $k = 15$



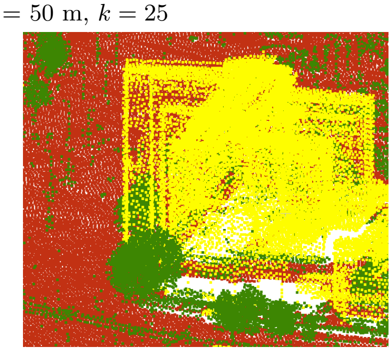
(e) block size = 50 m, $k = 25$



(f) block size = 30 & 50 m, $k = 20$



(g) block size = 50 m, $k = 15$ & 20



(h) Ground Truth

Figure 2: A point cloud subset in 3D view

With small scales, there exist some "block effect" in the test results, which means the edges of some blocks can be clearly seen in the visualization of the segmentation. Most of these problems can be solved when we increase the block size or the neighborhood size k. Beyond that, smaller scales also show advantages in the segmentation of areas, where objects from different classes are highly mixed (see the top right corner of Figure 2a, which indicates the central station in Utrecht, and Figure 1d). When we combine results from two single-scale settings, drawbacks from both scales can be mitigated by taking more confident predictions, as indicated in Figure 2f.