DUBLINCITY: ANNOTATED LIDAR POINT CLOUD AND ITS APPLICATIONS

[1 Results and Applications 2](#_Toc163729904)

[1.1 3D Point Cloud Object Classification 2](#_Toc163729905)

[1.2 Image-based 3D Reconstruction 2](#_Toc163729906)

[1.3 Formulas for calculations: 4](#_Toc163729907)

# Results and Applications

## 3D Point Cloud Object Classification

The performance analysis of the dataset for object classification problem is showcased by using the state-of-the-art models namely, Table 1: PointNet [20], PointNet++ [22] and SO-Net [14]. These three models directly work on unstructured point cloud datasets. They learn the global point cloud features that have been shown to classify forty man-made objects of the ModelNet40 [38] shape classification benchmark. The 3D point cloud dataset is comprised of a variety of outdoor areas (i.e. university campus and city centre) with structures of facades, roads, door, windows and trees as shown in Figure 2. In order to study the classification accuracy on the three CNN-based models, a dataset of 3982 objects of 5 classes (i.e. doors, windows, facades, roofs and trees) is gathered. To evaluate the three models, the dataset is split into a ratio of 80 : 20 for training and testing respectively. While training, for each sample, points on mesh faces are uniformly sampled according to the face area and normalised into a unit sphere (i.e. -1 to +1). Additionally, data augmentation techniques are applied on-the-fly by randomly rotating the object along the up-axis and jittering the position of each point by Gaussian noise with zero mean and 0.02 standard deviation. Each model is trained for 100 epochs. In Table 1, the performance of the three trained models in a different point cloud input setting using the Overall and Average class accuracy (as used in [20, 22]) is shown. It is observed that with an increase in the number of points per objects, the performance of the three models increases. Amongst all the three networks, the So-Net architecture performs the best. This is in consistence with the results in [14]. However, there is still a huge potential in the improvement of the performance scores. This is primarily because dataset is challenging in terms of structural similarity of outdoor objects in the point cloud space namely, facades, door and windows.

Table 1: Overall and Avg. class classification scores using the state-of-the-art models on the dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| #Points | PointNet [20] | | PointNet++ [22] | | So-Nets [14] | |
| Avg. Class | Overall | Avg. Class | Overall | Avg. Class | Overall |
| 512 | 24.17 | 35.17 | 39.47 | 45.56 | 41.89 | 48.74 |
| 1024 | 38.84 | 50.13 | 44.65 | 62.91 | 45.73 | 63.54 |
| 2048 | 46.77 | 59.68 | 49.23 | 63.42 | 49.34 | 64.55 |
| 4096 | 48.77 | 60.68 | 51.23 | 64.42 | 50.34 | 65.55 |

## Image-based 3D Reconstruction

In order to carry out the experiment, the complete reconstruction pipeline is evaluated as per [11]. This is because the ground truth for the camera positions to specifically evaluate the camera poses from an SfM algorithm are not available, only the GPS position is known (Figure 1). The open-source software selected for the reconstruction is COLMAP (SfM [26] and MVS [28]), since it is reported as the most successful in different scenarios in the latest comparisons carried out (see Section 2). Furthermore, it has advantages over other methods [4, 18] because



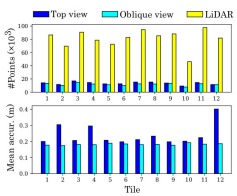


Figure : On the left, the area of the city covered by the tiles in the comparison (in green). On the right, a comparison of the number of points and the mean accuracy per tile

it gives the possibility of handling a large amount of data without running out of memory. In this experiment, COLMAP is applied with the same configuration to each set of images and as a result, two dense point clouds are obtained. The configuration includes, apart from the default parameters, using a single camera per flight path and the vocabulary tree method [27] for feature matching. This was selected because it is the recommend mode for large image collections (several thousands). Moreover, as in COLMAP there is no option implemented to enforce GPS priors during SfM computation, we follow the recommendation of applying the geo-registration after obtaining the sparse point cloud.

As pointed out in [11], the mean distance between the point clouds can be affected by outliers. Hence, they propose to use the following measurements for further study: precision, recall, and F score (3). The precision, P, shows the accuracy of the reconstruction, the recall, R, is related to how complete the reconstruction is, and the F score, F, is a combination of both. They are defined in for a given threshold distance d. In the equations, I is the image-based reconstruction point cloud, G is the ground truth point cloud, | · | is the cardinality, distI→G(d) (1) are the points in I with a distance to G less than d and distG→I(d) is analogous (i.e. distA→B(d) = {a ∈ A | min b∈B ka−bk2 < d}, A and B being point clouds).

## Formulas for calculations:

1. Formula for calculating the percentage of reconstruction accuracy :

|  |  |
| --- | --- |
|  | () |

1. Formula for calculating the percentage of reconstruction completeness :

|  |  |
| --- | --- |
|  | (2) |

1. Formula for calculating overall accuracy :

|  |  |
| --- | --- |
|  | (3) |

where – the set threshold state;

– the reconstruction point cloud;

– the ground point cloud;

— distance from the reconstruction point cloud to the ground point cloud;

— distance from ground point cloud to reconstruction point cloud.