

AN END-TO-END FRAMEWORK USING 3D CNN AND SVM

Full automation in mobile Lidar data classification



A deep learning architecture called MMCN opens up the possibility for the fully automated classification of highly dense 3D point cloud data acquired from a mobile Lidar system. This offers interesting opportunities in applications such as high-density maps, autonomous navigation and highway monitoring.

The classification of highly dense 3D point cloud data acquired from a mobile Lidar system (MLS) is essential in applications such as high-density maps, autonomous navigation and highway monitoring. MLS data classification algorithms available in the current literature use several parameters or thresholds. The right selection of these parameters is critical for the success of these algorithms under different conditions. This article provides an insight

into a deep learning architecture called multi-faceted multi-object convolutional neural network (MMCN). The MMCN, along with the support vector machine (SVM), provides an end-to-end framework for automatic classification of an input 3D point cloud without manual tuning of any parameters, thus providing a possibility for full automation.

Architecture

The proposed MMCN algorithm uses multiple facets and multiple objects of an MLS segment (Kumar et al., 2020). An object is a subpart of a segment. Multiple objects of the same segment are generated by creating several subparts automatically. The multiple facets or orientations of each object are created by rotating the object about the x, y and z axes. Different objects and facets thus incorporate information from various portions of the same segment. Further, the input and convolutional layers of the MMCN are separate for each facet, whereas the last convolutional layer of each facet merges at the common fully connected layer. The last fully connected layer is the output layer and uses the softmax function. Therefore, the MMCN is able to account for varied and distinct information about a segment. During the training of the MMCN, the above methodology is applied to manually labelled samples. During classification, segments are created with the help of automatically segmented groups of points (see below for more details); these segments do not need to belong to a single class and can have different levels of noise, background clutter and occlusion as is evident in MLS data. Through training, the algorithm learns local structures and variations within a segment along with noise and clutter. Figure 1 shows the architecture of the MMCN for a segmented group of points.



Classification

The trained MMCN is combined with the SVM to provide an end-to-end framework for completely automatic extraction of multiple segments and classification of MLS point clouds. The developed framework takes an MLS point cloud of any size as input and assigns the class label to each data point. Spherical segments around all the points in the MLS dataset are made for a given radius. These segments are passed through the trained MMCN. For the given input segment, the softmax layer of the MMCN generates a probability vector of belonging to different classes. The same probability vector is assigned to all the points within a segment. It may happen that a point can

occur in multiple spherical segments of different radii. In this case, the probability vectors of all such segments are totalled and divided by the number of such segments. This gives the final probability vector of the point for the given radius. Similarly, by varying the radius, multiple sets of probability vectors are calculated for all the points in the MLS dataset. The resulting multiple probability vectors using multiple radii for each point are concatenated and used as a feature vector for training an SVM. The trained SVM then gives the final decision for the classification of the point. Varying the radius around a point forms various shaped segments that incorporate different local neighbourhoods. This provides under-segmentation as well as over-segmentation information to the framework, which is able to learn the variations associated with changing surroundings.

Figure 2: Classification results.

Results

In the developed framework, only the MMCN and SVM need to be trained. Therefore, it is independent of any MLS parameter tuning (multiple radii values can be more or less depending on the available hardware, without requiring any manual tuning). In this framework, only XYZ information about points is required. Further, the network has been developed and trained to learn from segments consisting of noise, occlusion, background clutter and mixed with data points from other classes. The MMCN and SVM have been tested on a Paris-Lille 3D dataset (Roynard et al. 2018) containing more than 142 million data points. The maximum total accuracy and kappa are 96.5% and 93.8%, respectively. More details of the MMCN and SVM can be found in Kumar et al. 2020. Figure 2 and Figure 3 represent the classification results using the MMCN and SVM, and corresponding ground truth, respectively. In the present study, a lower number of convolutional layers, feature maps, neurons and training samples were used due to hardware constraints. That has resulted in some misclassifications.

?			
Fig	jure 3:	Ground	truth.

Conclusion

Classification of highly dense 3D MLS point clouds is a challenging task. During practical implementation on the ground with voluminous real-time (or near real-time) data, manual tuning of MLS parameters as is required in traditional classification approaches is not desirable. Instead, a classification algorithm is desired that can work across different terrains and datasets, with few or no changes in the algorithm. The proposed approach can work on any terrain datasets comprising noise, background clutter and occlusion with no tuning. Hence, this provides a wider practical solution for MLS classification. The next step of development is to implement the proposed approach for industry-level data processing.

Further reading

Kumar, B., Pandey, G., Lohani, B., Misra, S.C., 2020. A framework for automatic classification of mobile Lidar data using multiple regions and 3D CNN architecture. *International Journal of Remote Sensing*. doi: 10.1080/01431161.2020.1734252.

Roynard, X., Deschaud, J.E., Goulette, F., 2018. Paris-Lille-3D: a large and high-quality ground truth urban point cloud dataset for automatic segmentation and classification. arXiv:1712.00032. Available online: https://arxiv.org/abs/1712.00032 (accessed 9 January 2020).

https://www.gim-international.com/content/article/full-automation-in-mobile-lidar-data-classification