

РАЗРАБОТКА НЕЙРОННОЙ СЕТИ ДЛЯ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ОБЛАКОВ ТОЧЕК

*О.Д. Лешин, студент гр. 8ВМ03
Томский политехнический университет
E-mail: odl3@tpu.ru*

Введение

Задача обработки спутниковых снимков земли задача не новая. Существует большое количество алгоритмов для получения полезных данных из такого рода информации. Однако снимки не могут показать полную картину, так как невозможно полностью оценить размер и форму объекта только по одному снимку. В лучшем случае необходимо сфотографировать объект со всех сторон. Существует и другой способ – отсканировать объект и получить его полную модель. Благодаря полученной модели объект можно рассмотреть со всех сторон в подробностях и деталях. Тоже касается и сканирования поверхности земли. При помощи LIDAR [1] сканера можно получить полную сцену помещения, отдельной улицы в городе или даже сам город целиком. После чего можно будет использовать полученные данные для обработки. К примеру, выделение или поиск объектов.

Целью данной работы является реализация нейронной сети для семантической сегментации данных ДЗЗ.

Наборы данных

В данной работе использовался набор данных Hessigheim 3D [2]. В наборе представлены данные местности, в которых содержится информация о координатах точки, ее цвете, коэффициент отражения точки и класс, к которому точка относится. В наборе содержится 11 классов: деревья, крыши, транспортные средства, трава, кустарники, фасады зданий, городское оборудование, почва, печные трубы, вертикальные поверхности, непроницаемые поверхности. На рисунке 1 представлена сцена из данного набора.



Рис. 1. Сцена из набора данных

Описание структуры

Реализованная модель основана на модели DGCNN [3]. Особенность данной модели в том, что она выделяет признаки из облаков точек, представленных в виде графа. Представление облака точек в виде графа представляет собой поиск центральной точки, вокруг которой находится k ближайших соседей. Из полученных данных строится вектор, который имеет вид $(X_c, Y_c, Z_c, X_{ck} - X_c, Y_{ck} - Y_c, Y_{ck} - Z_c)$, где X_c, Y_c, Z_c – это координаты центральной точки, $X_{ck} - X_c, Y_{ck} - Y_c, Y_{ck} - Z_c$ – это ребра графа.

Для улучшения результата было принято решение использовать дилатационную (Atrous) свертку [4]. Дилатационная свертка позволяет увеличить поле зрения, что означает увеличение количества признаков, выделяемых в ходе обработки. К тому же, данный тип свертки аналогичен обычной свертке по вычислительным затратам. Результатом стала архитектура, показанная на рисунке 2.

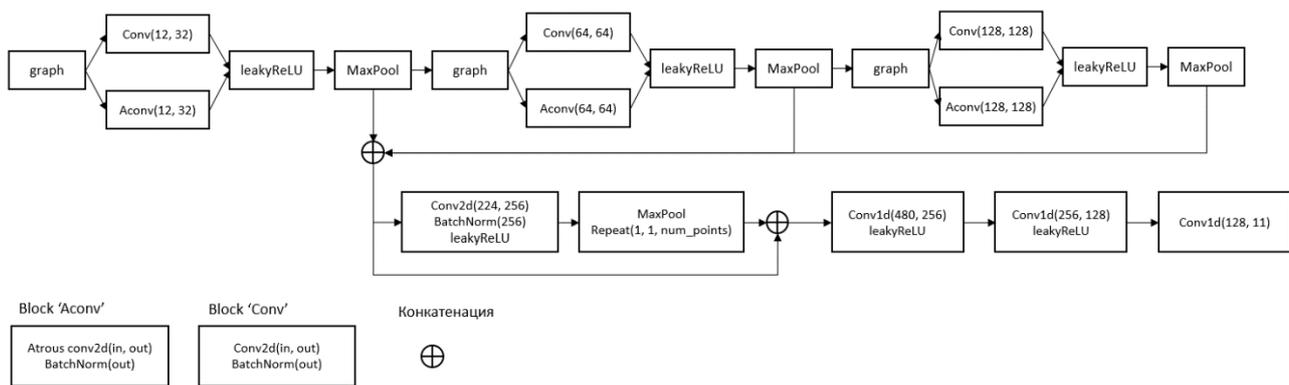


Рис. 2. Архитектура реализованной модели

Тестирование

Для оценки точности модели использовались 2 метрики: accuracy, mean intersection over union. Метрика accuracy – это отношение верно определенных классов ко всем идеальным классам. Метрика mean intersection over union определяется отношением:

$$MIoU = \frac{1}{k} \sum_{i=1}^n \frac{n_{ii}}{t + \sum_{j=1}^k n_{ij} - n_{ii}}. \quad (1)$$

Где k – общее количество классов, n_{ij} $i, j \in 1, \dots, k$, соответственно количество точек принадлежащих классу i , и соотнесенных к классу j . t – общее количество точек, принадлежащих классу i .

По первой метрике модель демонстрирует ~81%, по второй метрике ~64%.

В таблице 1 представлена метрика intersection over union для каждого класса.

Таблица 1. Метрика intersection over union для каждого класса

Класс	Точность
Трава	0.66
Непроницаемая поверхность	0.75
Транспорт	0.0002
Городское оборудование	0.28
Крыша	0.76
Фасад	0.50
Кустарник	0.27
Дерево	0.70
Почва/гравий	0
Вертикальная поверхность	0.28
Печная труба	0

Заключение

В результате проделанной работы была реализована нейронная сеть для семантической сегментации облаков точек с использованием слоев дилатационной свертки. Данная нейронная сеть была обучена и протестирована на наборе Hessigheim 3D и демонстрирует хорошие результаты.

Список использованных источников

- Review on Lidar Technology / Ninad Mehendale, Srushti Neoge
- The Hessigheim 3D (H3D) Benchmark on Semantic Segmentation of High-Resolution 3D Point Clouds and Textured Meshes from UAV LiDAR and Multi-View-Stereo / Michael Kölle, Dominik Laupheimer, Stefan Schmohl, Norbert Haala, Franz Rottensteiner, Jan Dirk Wegner, Hugo Ledoux
- Dynamic Graph CNN for Learning on Point Clouds / Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, Justin M. Solomon
- A Primer on Atrous(Dilated) and Depth-wise Separable Convolutions [Электронный ресурс] – URL: <https://towardsdatascience.com/a-primer-on-atrous-convolutions-and-depth-wise-separable-convolutions-443b106919f5> (дата обращения 20.01.2022).