Разработка нейросетевой модели для семантической сегментации облаков точек

О.Д. Лешин ¹, Д.С. Григорьев ¹

¹ Томский политехнический университет. пр. Ленина 30, Томск, 634050, Россия

Аннотация

Обработка облаков точек при помощи нейросетевых моделей является относительно новым подходом. Первые работы в этом направлении датируются 2017 годом. Облака точек представляют из себя набор вершин в трехмерном пространстве и характеризуют внешнюю поверхность объекта. В виду особенностей таких данных, типовые полносвязные сверточные сети не могут быть применены. Связано это с тем, что в таких данных должно учитываться не только цветовая составляющая, но и вся геометрическая форма объекта. В данной работе представляют собой данные дистанционного зондирования земли. Для разработки данной модели использовалась уже существующая модель, в основе которой лежит графовое представление облаков точек, а для улучшения результата использовались слои дилатационной свертки. Обучение и тестирование модели проводилось на наборе данных Hessigheim 3D, который был получен при помощи сканирования пригородной местности. В результате численных экспериментов по разработке модели и предобработке набора данных были получены приемлемые результаты по метрикам F1 и overall accuracy.

Ключевые слова

Семантическая сегментация, нейронная сеть, облако точек, дилатационная свертка.

Development of a Neural Network Model for Semantic Segmentation of Point Clouds

O.D. Leshin¹, D.S. Grigoriev¹

¹ Tomsk polytechnic university, ave. Lenin 30, Tomsk, 634050, Russia

Abstract

Processing point clouds using neural network models is a relatively new approach. The first significant works in this direction date back to 2017. Point clouds are a set of vertices in threedimensional space and characterize the external surface of an object. Due to the peculiarities of such data, typical fully connected convolutional networks cannot be applied. This is due to the fact that such data should take into account not only the color component but also the entire geometric shape of the object. This paper presents a neural network model for semantic segmentation of point clouds, which are remote sensing data of the earth. The proposed model is a modification of the original DGCNN model, which is based on the graph representation of point clouds, and atrous convolution layers were used to improve the result. The training and testing of the model was carried out on the Hessigheim 3D dataset, which was obtained by scanning a suburban area. As a result of numerical experiments on model development and data set preprocessing, acceptable results were obtained for the F1 and overall accuracy metrics.

Keywords

Semantic segmentation, neural network, point cloud, dilated convolution.

ORCID: 0000-0003-4647-3148 (О.Д. Лешин); 0000-0001-8810-6691 (Д.С. Григорьев)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

ГрафиКон 2022: 32-я Международная конференция по компьютерной графике и машинному зрению, 19-22 сентября 2022 г., Рязанский государственный радиотехнический университет им. В.Ф. Уткина, Рязань, Россия ЕМАІL: odl3@tpu.ru (О.Д. Лешин); trygx@tpu.ru (Д.С. Григорьев)

1. Введение

Данные дистанционного зондирования земли уже давно используются в различных алгоритмах для получения полезной информации. Такие данные представляют собой изображения, полученные при помощи различной съемочной аппаратуры. Однако изображения не могут передать полную информацию об осматриваемом объекте, так как для этого необходимо сделать несколько снимков, чтобы оценить полностью объект. Для получения более полной информации можно использовать различные 3D сканеры [1]. Данная технология тоже не является новой и ее развитие делает устройства меньше и доступнее. Так же, сканеры умеют определять цвет точки во время сканирования, следовательно, объект получается максимально реалистичным. После сканирования полученное облако точек может быть конвертировано в 3D модель для дальнейшей обработки. При помощи такого сканера можно получить модель определенного объекта или помещения. Не редко такие сканеры используются при строительстве и можно отсканировать целое здание. В некоторых случаях можно отсканировать целый город. Однако для обработки такого количества данных может потребоваться большое количество времени. Как решение можно использовать нейронные сети, для выделения полезной информации из таких данных.

Нейросетевая обработка облаков точек - задача относительно новая. Одной из первых работ в данной области является модель PointNet [2]. Модель PointNet демонстрировала хорошие результаты в задачах семантической сегментации и классификации объектов. В этой работе модель обрабатывала точки отдельно друг от друга, переводя их в пространства большей размерности. После чего все пространства точек объединялись. Эта работа получила продолжение в виде модели PointNet++[3]. Для обработки облака точек использовались перекрывающие области. То есть брались точки из всего облака, и вокруг этих точек строились области с заданным радиусом, после чего из данных областей выделялись характерные признаки. Данная модель легла в основу многих других работ, например, JSNet [4] или ASIS [5]. Такой подход к представлению облака точек показывает хорошие результаты, однако существуют и другие способы, к примеру, представление точек в виде графа.

Одной из моделей, использующих такой подход, является модель DGCNN [6]. В данном подходе для всего облака точек или его части, если облако слишком больших размеров, находится центральная точка. Вокруг центральной точки находятся k ближайших соседей, к которым будут строиться ребра графа. Полученный граф будет иметь вид (Xc, Yc, Zc, Xck - Xc, Yck-Yc, Yck-Zc), где Xc, Yc, Zc – это координаты центральной точки, Xck-Xc, Yck-Yc, Yck-Zc – это ребра графа. Данный подход демонстрировал выше результат, чем предыдущий, поэтому было принято решение использовать подход, основанный на графовом представлении облака.

Помимо приведенных выше методов, существуют мультивидовые (multi-view) методы с использованием снимков или панорам. Этот метод реализован в таких моделях, как: DeepPano [7] или MVCNN [8]. В этих работах облако точек конвертируется в 3D модель, после чего полученную модель фотографируют с нескольких сторон или делают один панорамный снимок, постепенно поворачивая модель. Однако такой метод не подходит для решения поставленной задачи, так как облако точек довольно большое и получить правильные снимки с нужных ракурсов становится затруднительно.

2. Описание набора данных

В качестве обучающих данных использовался набор Hessigheim 3D [9]. В наборе представлены данные местности, полученные путем сканирования пригородной территории при помощи LIDAR сканера [10]. В датасете содержится информация о координатах точки, ее цвет, количество отправленных и принятых сигналов от сканера, коэффициент отражающей поверхности и класс, к которому принадлежит точка. В наборе содержится 11 классов:

- Трава (Low Vegetation)
- Непроницаемая поверхность (Impervious Surface)
- Транспорт (Vehicle)

- Городское оборудование (Urban Furniture)
- Крыша (Roof)
- Фасад (Facade)
- Кустарник (Shrub)
- Дерево (Tree)
- Почва/Гравий (Soil/Gravel)
- Вертикальна поверхность (Vertical Surface)
- Печная труба (Chimney)

На рисунке 1 представлена сцена из данного набора.



Рисунок 1 – Сцена из набора Hessigheim 3D

При использовании набора данных была выявлена особенность. Этой особенностью являются отсортированные точки. Точки в наборе данных отсортированы по высоте и классам, к которым они относятся. На рисунке 2 показана данная особенность.



Рисунок 2 – Особенность набора данных

Данные загружаются по их индексам в одном большом массиве. На рисунке показаны индексы трех точек, которые относятся к разным классам. Координаты точек отличаются на десятые доли, но индексы точек разительно отличаются. Для загрузки набор данных делился на отдельные подобласти, так как весь набор целиком обработать не представляется возможным. Исходя из индексов, показанных на рисунке 2, данная подобласть никогда не будет рассматриваться как единое целое, при сколь угодно большом количестве точек в подобласти. Все три класса, будут рассматриваться отдельно друг от друга. В качестве предобработки, был реализован метод, разделяющий облако точек на подобласти, состоящие из ближайших соседей.

3. Описание разработанной модели

Так как подход, основанный на графовом представлении облака точек имел большую точность, было решено взять за основу модель DGCNN. Модель также получила свое развитие в других работах, основанных на усложнении архитектуры, с добавлением слоев, связей или изменении функции ошибки. Для улучшения результата было принято решение использовать дилатационную свертку (Atrous convolutional) [11]. Данный вид свертки помогает увеличить область обзора, путем увеличения пространства между элементами ядра свертки, при этом не требует дополнительных вычислительных ресурсов. Реализованная архитектура модели показана на рисунке 3.



Рисунок 3 – Архитектура разработанной модели

Первым шагом является получение ребер графа. Это происходит при помощи поиска центральной точки области и k ее ближайших соседей. Полученные ребра обрабатываются отдельно блоками «Conv» и «Aconv». Блок «Aconv» содержит в себе слой дилатационной свертки. Результаты обоих блоков поэлементно складываются и отправляются на функцию активации, а в последствии на слой MaxPooling. После данного слоя снова происходит выделение ребер графа. Данный процесс повторяется три раза, а затем все результаты конкатенируются и снова проходят операцию свертки и слой MaxPolling. Полученный результат снова конкатенируется с более ранними и уже проходит финальную обработку.

4. Тестирование

Для проведения тестирования было решено использовать метрики F1 и overall accuracy (OA). Метрика OA – это отношение верно определенных классов ко всем идеальным классам. Метрика F1 вычисляется по следующей формуле:

$$F1 = \frac{tp}{tp + \frac{1}{2}(fp + fn)}$$
(1)

где tp – верно определенный класс точки, fp – точка определена как принадлежащая к определенному классу, но таковой не является, fn – точка принадлежит к одному классу, но определена, как принадлежащая к другому. По метрике OA модель показала результат ~ 81%, против 77% у оригинальной модели и 65% у модели PointNet. Результаты по метрике F1 для разработанной модели и сравнение с другими моделями показаны в таблице 1.

| | Таблица 1 – Метрика F1 и сравнение с другими методами | | |
|------------------------------|---|-------|----------|
| Класс | Разработанная модель | DGCNN | PointNet |
| Трава | 0.84 | 0.81 | 0.56 |
| Непроницаемая поверхность | 0.86 | 0.83 | 0.71 |
| Транспорт | 0.32 | 0.47 | 0.16 |
| Городское оборудование | 0.40 | 0.33 | 0.14 |
| Крыша | 0.89 | 0.84 | 0.55 |
| Фасад | 0.57 | 0.54 | 0.001 |
| Кустарник | 0.21 | 0.20 | 0 |
| Дерево | 0.88 | 0.82 | 0.36 |
| Почва/гравий | 0.38 | 0.32 | 0.25 |
| Вертикальная поверхность | 0.73 | 0.70 | 0.45 |
| Печная труба | 0.02 | 0 | 0 |
| Среднее значение | 0.55 | 0.53 | 0.30 |

На рисунке 4 показаны примеры результаты сегментации.



Рисунок 4 – Результаты сегментации а) Облака точек в формате RGB б) Оригинальные метки классов в) Результаты сегментации

Как можно увидеть из таблицы 1 и рисунка 4 большая часть точек распознается верно. Однако самая главная проблема остается с классом «Печная труба». Данный класс представлен наименьшим количеством точек, а также объекты, которые состоят из этих точек, являются небольшими по размеру. И не смотря на то, что при обучении использовались веса классов, для балансирования данных, эта проблема все равно остается актуальной.

5. Заключение

В работе предложена новая архитектура для семантической сегментации облаков точек, основанная на модели DGCNN с использованием слоев дилатационной свертки. Результаты, полученные в ходе численных экспериментов на наборе данных Hessigheim 3D показывают, что разработанная модель демонстрирует приемлемые результаты по метрикам OA и F1 в задаче семантической сегментации данных дистанционного зондирования земли. Также разработанная модель демонстрирует результаты выше, чем оригинальная модель и модель PointNet по метрикам OA и F1.

6. Список источников

- [1] Яков Бондарев. Что такое 3D-сканирование? [Электронный ресурс]: 3D Integration. URL: https://i3d.ru/blog/dlya_mozayki/chto-takoe-3d-skanirovanie/ (дата обращения 21.02.2022).
- [2] PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation / Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) / IEEE. 2017.
- [3] PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space / Charles R. Qi, Li Yi, Hao Su, Leonidas J. Guibas/ Conference on Neural Information Processing Systems (NIPS). 2017.
- [4] Lin Zhao, Wenbing Tao. JSNet: Joint Instance and Semantic Segmentation of 3D Point Clouds // Proceedings of the AAAI Conference on Artificial Intelligence. 2020. № 34(07). C. 12951-12958
- [5] Associatively Segmenting Instances and Semantics in Point Clouds / Xinlong Wang, Shu Liu, Xiaoyong Shen, Chunhua Shen, Jiaya Jia // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) / IEEE. 2019.
- [6] Dynamic Graph CNN for Learning on Point Clouds / Yue Wang [и др.] // ACM Transactions on Graphics. 2019. № 5.
- [7] DeepPano: Deep Panoramic Representation for 3-D Shape Recognition / Baoguang Shi, Song Bai, Zhichao Zhou, Xiang Bai // IEEE Signal Processing Letters. 2015. № 22. C. 2339 – 2343.
- [8] Multi-view Convolutional Neural Networks for 3D Shape Recognition / Hang Su, Subhransu Maji, Evangelos Kalogerakis, Erik Learned-Miller // 2015 IEEE International Conference on Computer Vision (ICCV) / IEEE. 2015.
- [9] 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 [и др.] // ISPRS Open Journal of Photogrammetry and Remote Sensing / ISPRS. 2021.
- [10] Ninad Mehendale. Srushti Neoge. Review on Lidar Technology // SSRN Electronic Journal / SSRN. 2020.
- [11] Aadhithya Sankar. A Primer on Atrous(Dilated) and Depth-wise Separable Convolutions [Электронный ресурс]: Towards Data Science URL: https://towardsdatascience.com/a-primeron-atrous-convolutions-and-depth-wise-separable-convolutions-443b106919f5 (дата обращения 9.03.2022).