

ПРИМЕНЕНИЕ ГРАФОВЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ОБЛАКА ТОЧЕК

Будзиаловский П.Б.

Томский политехнический университет, ИШИТР, студент гр. 8В14, e-mail: pbb2@tpu.ru

Научный руководитель: Григорьев Д.С.

Томский политехнический университет, старший преподаватель, e-mail: trygx@tpu.ru

Аннотация

В данной работе представлена реализация модели графовой нейронной сети для решения задачи семантической сегментации облака точек. Проведено сравнение архитектур DGCNN и PointNet на датасете H3D Benchmark. DGCNN показала более устойчивое обобщение, особенно на сложных объектах. Полученные выводы подтверждают перспективность GNN в задачах анализа 3D-сцен и могут быть применимы при разработке систем автономной навигации и картографирования.

Ключевые слова: семантическая сегментация, облако точек, графовые нейронные сети, DGCNN, PointNet.

Введение

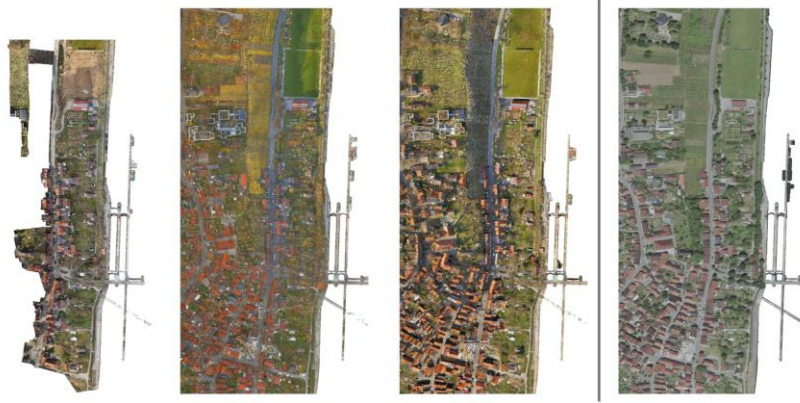
Семантическая сегментация трёхмерного облака точек представляет собой важную задачу в анализе пространственных данных. Её актуальность особенно велика в таких передовых направлениях, как робототехнические системы, автономное управление транспортными средствами, геоинформационные технологии и ряд смежных дисциплин. Современные лидары, устанавливаемые на беспилотных летательных аппаратах, обеспечивают детальное сканирование окружающей среды, формируя высокоточные трёхмерные представления объектов. Однако высокая степень плотности и структурной неоднородности подобных данных обуславливает необходимость в использовании более устойчивых и адаптивных алгоритмов обработки. В связи с этим в последние годы заметно возрос интерес [1] к использованию графовых нейросетевых моделей (GNN), обладающих способностью эффективно обрабатывать разреженные и нерегулярные пространственные структуры [2]. Целью исследования является экспериментальное изучение эффективности графовых нейронных сетей в задаче семантической сегментации трёхмерных облаков точек, а также сравнение их производительности с классическими нейросетевыми архитектурами.

1. В рамках поставленной цели были сформулированы следующие задачи:
2. Реализовать модель на основе графовой нейронной сети DGCNN.
3. Реализовать базовую модель PointNet для сравнения результатов.
4. Провести обучение и тестирование обеих моделей.
5. Сравнить модели по качественным и количественным метрикам, включая общую точность и качество предсказания по классам.

Описание датасета

В рамках данного исследования в качестве исходных данных был использован референсный датасет The Hessigheim 3D (H3D) Benchmark [3, 4]. Этот датасет представляет собой коллекцию трёхмерных облаков точек высокой плотности, дополненных текстурированными полигональными сетками, сформированными с применением лидарных сканеров, установленных на беспилотных летательных аппаратах.

Ключевой особенностью исходных облаков точек является разнообразие: охватываемая детализированные сцены сельской местности, включающие как элементы городской инфраструктуры, так и объекты природного ландшафта. Данные представлены в виде четырёх временных срезов (март 2018 года, ноябрь 2018 года, март 2019 года и март 2016 года). На рисунке 1 представлены доступные эпохи облаков точек.



*Рис. 1. Доступные эпохи набора данных НЗД.
Слева направо: март 2018 года, ноябрь 2018, март 2019 и март 2016*

Разработчики датасета выделяют 11 ключевых категорий объектов: низкая растительность, асфальтированные поверхности, транспортные средства, городская инфраструктура (уличные фонари, лавочки, дорожные знаки), крыша, фасад (стены зданий), кустарники, деревья, почва/гравий, вертикальные поверхности (заборы, стены), дымоходы.

В качестве исходных данных для обучения и тестирования была выбрана сцена, относящаяся к эпохе март 2016 года, включающая в себя 2200213 точек, распределенных по классам следующим образом: низкая растительность – 897 950, асфальтированные поверхности – 470 303, транспортные средства – 13 985, городская инфраструктура – 30 189, крыша – 377 328, фасад – 28 858, кустарники – 102 112, деревья – 222 200, почва/гравий – 40 221, вертикальные поверхности – 15 310, дымоходы – 1 757. Основным критерием при выборе послужил сравнительно меньший объём данной подвыборки по сравнению с остальными эпохами. Это позволило существенно сократить время предварительной обработки и обучения моделей, что, в свою очередь, сделало возможным проведение серии сравнительных экспериментов в условиях ограниченных вычислительных ресурсов. Кроме того, несмотря на ограниченный объём, сцена сохраняет репрезентативность: в ней представлены все ключевые классы объектов, предусмотренные аннотацией, что обеспечивает корректность и обоснованность анализа.

Архитектуры моделей

Графовые нейронные сети (Graph Neural Networks, GNN) представляют собой мощный инструмент анализа данных, лишённых регулярной структуры, что делает их особенно востребованными для обработки облаков точек. В отличие от классических свёрточных нейросетей (CNN), ориентированных на упорядоченные входные данные, такие как двумерные изображения или трёхмерные воксельные сетки, GNN эффективно моделируют сложные взаимосвязи между элементами разреженных пространственных представлений.

В данной работе применяется архитектура DGCNN (Dynamic Graph Convolutional Neural Network), являющаяся одной из наиболее распространённых GNN-моделей для анализа облаков точек. Её ключевая особенность – динамическое построение графа в ходе обучения, что позволяет адаптироваться к локальной структуре данных [5]. Основные этапы работы данной сети заключаются в построении графа на основе входного облака точек с использованием алгоритма k-NN [6] и динамическом обновлении графа на каждом слое свёрточной сети, позволяя модели лучше учитывать локальный контекст. EdgeConv – ключевая операция, которая обновляет признаки вершин, агрегируя информацию от соседей (1).

$$x_i' = \sum_{j \in N(i)} h_{\theta}(x_i || x_j - x_i), \quad (1)$$

где x_i – признаки центральной точки, x_j – признаки соседей из множества $N(i)$, h_{θ} – обучаемая нейросеть (свёрточные слои или MLP).

На рисунке 3 представлена архитектура исследуемой модели DGCNN.

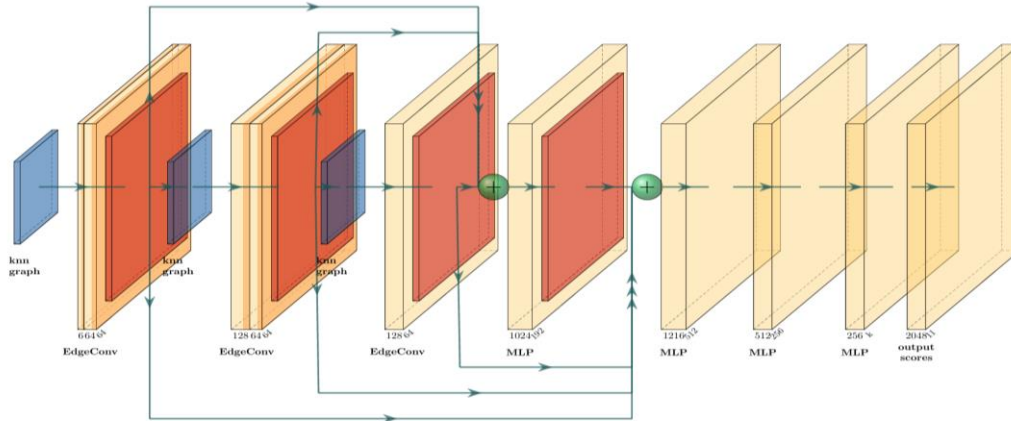


Рис. 2. Архитектура DGCNN

Для сравнения результатов работы графовой нейронной сети использовалась PointNet [7] – классическая нейросетевая архитектура, разработанная для прямой обработки облаков точек с помощью глубокого обучения. В отличие от традиционных подходов, базирующихся на предварительной вокселизации или проекциях на двумерные плоскости, PointNet осуществляет обработку необработанных облаков точек без необходимости их приведения к регулярным структурам.

Основная идея модели заключается в применении симметричных функций (например, max pooling), которые делают её инвариантной к перестановке точек. Архитектура включает несколько слоёв многослойных перцептронов (MLP), которые извлекают признаки для каждой точки, а затем выполняет глобальную агрегацию, формируя вектор, описывающий всё облако точек. В задаче семантической сегментации PointNet дополнительно использует локальные признаки, чтобы учитывать пространственные взаимосвязи между точками. На рисунке 3 представлена архитектура классической модели PointNet.

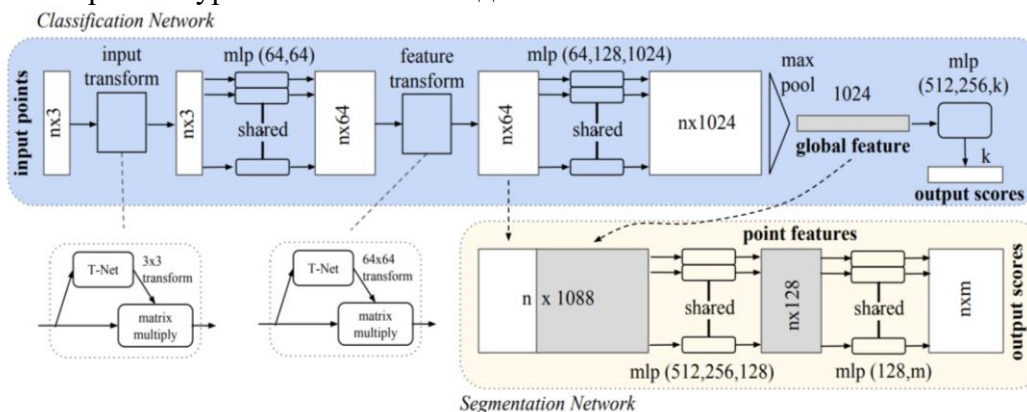


Рис. 3. Архитектура PointNet.

Результаты тестирования

Обучение, валидация и исследование всех моделей производилось с использованием фреймворка PyTorch на облачном сервисе Google Collab. Во всех экспериментах для обеспечения корректного сравнения моделей использовались одинаковые гиперпараметры. В частности, размер батча 16 384 точек, количество эпох 10, в качестве функции потерь использовалась кросс-энтропия, подходящая для текущей задачи.

Для оценки качества сегментации использовались метрики F1-score для каждого отдельного класса и общий F1-score (mean F1), а также Overall Accuracy. Метрики F1-score и OA традиционно используются в задачах классификации и сегментации, так как позволяют

объективно оценить баланс между полнотой и точностью предсказаний [8]. Результаты обучения нейронных сетей по выбранным метрикам представлены в таблицах 1 и 2.

Таблица 1. Обобщённые результаты обучения моделей

Модель	Метрики	
	Overall Accuracy	mF1-score
DGCNN	0.731	0.414
PointNet	0.831	0.481

Таблица 2. Сравнение качества сегментации по классам на обучающей выборке

Класс	F1-score	
	DGCNN	PointNet
низкая растительность	0.760	0.875
асфальтированная поверхность	0.668	0.824
транспортные средства	0.033	0.469
городская инфраструктура	0.145	0.174
крыша	0.804	0.900
фасад	0.419	0.066
кустарники	0.373	0.435
деревья	0.752	0.875
почва/гравий	0.014	0.493
вертикальные поверхности	0.199	0.133
дымоходы	0.010	0.027

Для оценки производительности моделей было проведено тестирование на отложенной тестовой выборке. В таблицах 3 и 4 представлены итоговые результаты, включающие общую точность (OA), средний F1-score (mean F1) и F1-score по классам.

Таблица 3. Обобщённые результаты тестирования моделей

Модель	Метрики	
	Overall Accuracy	mF1-score
DGCNN	0.701	0.404
PointNet	0.167	0.090

Таблица 4. Сравнение качества сегментации по классам на тестовой выборке.

Класс	F1-score	
	DGCNN	PointNet
низкая растительность	0.641	0.262
асфальтированная поверхность	0.666	0.003
транспортные средства	0.025	0.019
городская инфраструктура	0.114	0.000
крыша	0.816	0.515
фасад	0.441	0.077
кустарники	0.421	0.023
деревья	0.623	0.063
почва/гравий	0.006	0.002
вертикальные поверхности	0.265	0.013
дымоходы	0.003	0.012

Результаты экспериментов показывают, что на тренировочной выборке модель PointNet демонстрирует более высокие показатели точности по сравнению с DGCNN, однако на тестовой выборке её качество сегментации значительно ниже. Такое поведение свидетельствует о переобучении PointNet: модель хорошо запоминает обучающие данные, но испытывает трудности с обобщением на новые сцены. Это обусловлено архитектурными особенностями: в PointNet отсутствует явный учёт локального контекста. Поэтому без добавления дополнительных признаков (цвет точки, её прозрачность и тд.) снижается способность модели адаптироваться к сложным пространственным структурам.

Анализ результатов по классам показывает, что графовые модели особенно хорошо справляются с сегментацией сложных объектов, таких как фасады зданий, крыши и деревья. Это связано с тем, что в таких классах пространственные взаимосвязи между точками имеют ключевое значение, и модели, работающие с графами, лучше улавливают эти структуры. Однако, несмотря на это, DGCNN плохо справляется с распознаванием дымоходов и транспорта, в связи с малым количеством представленных точек данного класса в датасете, и почвы, так как поверхность состоит из множества мелких, хаотично расположенных частиц, между которыми отсутствуют стабильные пространственные взаимосвязи.

Заключение

В данной работе рассмотрено применение графовых нейронных сетей для задачи семантической сегментации облака точек, полученного с лидара, установленного на БПЛА. Было проведено сравнение DGCNN с классической архитектурой PointNet, и результаты показали, что использование графовых сверток позволяет существенно улучшить качество сегментации. Эксперименты подтвердили, что графовые методы более эффективно извлекают локальные и глобальные признаки, что особенно важно для объектов со сложной структурой.

Таким образом, результаты исследования показывают, что графовые нейронные сети являются перспективным направлением для семантической сегментации облаков точек, и их дальнейшее развитие может привести к еще более точным и эффективным методам обработки данных лидара.

Список использованных источников

1. Khemani B., Patil Sh., Kotecha K., Tanwar S. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions [Электронный ресурс]. – URL: journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00876-4 (дата обращения 30.03.2025).

2. Intel Tiber AI Studio – The Essential Guide to GNN (Graph Neural Networks) [Электронный ресурс]. – URL: cnvrg.io/graph-neural-networks/ (дата обращения 30.03.2025).

3. Kölle M., Laupheimer D., Schmohl S., Haala N., Rottensteiner F., Wegner J.D., Ledoux H. The Hessigheim 3D (H3D) Benchmark on Semantic Segmentation of High-Resolution 3D Point Clouds and Textured Meshes from UAV LiDAR and Multi-View-Stereo [Электронный ресурс]. – URL: arxiv.org/pdf/2102.05346 (дата обращения 30.03.2025).

4. Institute for Photogrammetry and Geoinformatics – Hessigheim Benchmark [Электронный ресурс]. – URL: ifpwww.ifp.uni-stuttgart.de/benchmark/hessigheim/default.aspx (дата обращения 30.03.2025).

5. Wang Y., Sun Y., Liu Z., Sarma S.E., Bronstein M.M., Solomon J.M. Dynamic Graph CNN for Learning on Point Clouds [Электронный ресурс]. – URL: arxiv.org/pdf/1801.07829 (дата обращения 30.03.2025).

6. GeekforGeeks – K-Nearest Neighbor (KNN) Algorithm [Электронный ресурс]. – URL: geeksforgeeks.org/k-nearest-neighbours/ (дата обращения 30.03.2025).

7. Qi C.R., Su H., Mo K., Guibas L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation [Электронный ресурс]. – URL: arxiv.org/pdf/1612.00593 (дата обращения 30.03.2025).

8. Sokolova M., Lapalme G. A systematic analysis of performance measures for classification tasks [Электронный ресурс]. – URL: atour.iro.umontreal.ca/rali/sites/default/files/publis/SokolovaLapalme-JIPM09.pdf (дата обращения 30.03.2025).