

doi: 10.17586/2226-1494-2025-25-1-68-77

УДК 004.93'1

## Метод семантической сегментации данных воздушного лазерного сканирования водоохранных зон Сергей Владимирович Сай<sup>1</sup>, Алексей Владимирович Зинкевич<sup>2</sup>✉

<sup>1,2</sup> Тихоокеанский государственный университет, Хабаровск, 680035, Российская Федерация

<sup>1</sup> [sai1111@rambler.ru](mailto:sai1111@rambler.ru), <https://orcid.org/0000-0002-4506-1857>

<sup>2</sup> [zinkevich1985@mail.ru](mailto:zinkevich1985@mail.ru)✉, <https://orcid.org/0000-0003-3231-8569>

### Аннотация

**Введение.** Выполнена оценка эффективности нейросетевого метода семантической сегментации трехмерных облаков точек, полученных с использованием беспилотного летательного аппарата «Геоскан 401 Лидар». Реализована нейронная сеть, основанная на модели глубокого обучения PointNet++, которая использует метод непосредственной обработки облаков точек. Разработана методика получения и подготовки набора данных с четырьмя классами: земля, растительность, транспортные средства и строительные объекты. Для увеличения точности оценки предложена методика, основанная на аугментации и перераспределении наборов данных. **Метод.** Модель нейронной сети состоит из иерархически построенных блоков, выполняющих выборку, группировку и извлечение признаков. Изменение количества блоков и задание радиуса поиска локальных признаков влияет на точность сегментации и вычислительные затраты. **Основные результаты.** Проведена оценка эффективности метода семантической сегментации трехмерных облаков точек, полученных с помощью беспилотного летательного аппарата. Методика аугментации и перераспределения наборов данных позволила улучшить среднее значение Intersection over Union (IoU) не менее, чем на 35 %. Для полученных данных определен оптимальный радиус в слое группировки, при котором обеспечивается баланс между детализацией и чувствительностью. Установлено, что увеличение количества точек в наборе данных не приводит к существенному росту точности, однако разнообразие используемых наборов данных улучшает эффективность подхода. **Обсуждение.** Разработанный набор данных позволяет повысить эффективность применяемого подхода, в том числе при обучении на иных моделях. Полученные результаты исследования свидетельствуют о перспективности использования предложенных методик и алгоритмов при решении задачи построения цифровых моделей реки Амур и ее основных притоков.

### Ключевые слова

лидарная съемка, воздушное лазерное сканирование, облако точек, семантическая сегментация, нейронная сеть PointNet, гидрологические исследования

### Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 24-11-20024 и Министерства образования и науки Хабаровского края (соглашение № 124С/2024).

**Ссылка для цитирования:** Сай С.В., Зинкевич А.В. Метод семантической сегментации данных воздушного лазерного сканирования водоохранных зон // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 1. С. 68–77. doi: 10.17586/2226-1494-2025-25-1-68-77

## Method of semantic segmentation of airborne laser scanning data of water protection zones

Sergey V. Sai<sup>1</sup>, Alexey V. Zinkevich<sup>2</sup>✉

<sup>1,2</sup> Pacific National University, Khabarovsk, 680035, Russian Federation

<sup>1</sup> [sai1111@rambler.ru](mailto:sai1111@rambler.ru), <https://orcid.org/0000-0002-4506-1857>

<sup>2</sup> [zinkevich1985@mail.ru](mailto:zinkevich1985@mail.ru)✉, <https://orcid.org/0000-0003-3231-8569>

© Сай С.В., Зинкевич А.В., 2025

**Abstract**

This article presents an evaluation of the efficiency of a neural network method for the semantic segmentation of three-dimensional point clouds obtained using the Geoscan 401 Lidar UAV. The proposed implementation of the neural network is based on the PointNet++ deep learning model which directly processes point clouds. A technique has been developed for acquiring and preparing a dataset with four classes: land, vegetation, vehicles, and construction objects. To increase the accuracy of the evaluation, a technique based on augmentation and redistribution of the datasets has been proposed. The neural network model consists of hierarchically constructed blocks that perform sampling, grouping, and feature extraction. Adjusting the number of blocks and setting the search radius for local features affects both the accuracy of segmentation and computational costs. The efficiency of the method for semantic segmentation of three-dimensional point clouds obtained using the Geoscan 401 Lidar UAV has been evaluated. The augmentation and redistribution technique improved the average Intersection over Union (IoU) value by at least 35 %. For the obtained data, the optimal radius in the grouping layer was determined, ensuring a balance between detail and sensitivity. It was found that an increase in the number of points in the dataset does not lead to a significant improvement in accuracy; however, the diversity of the datasets used enhances the method efficiency. The developed dataset increases the effectiveness of the applied approach, including when training other models. The results of this study indicate the potential for using the proposed methods and algorithms in constructing digital models of the Amur River and its main tributaries.

**Keywords**

Lidar survey, airborne laser scanning, point cloud, semantic segmentation, PointNet neural network, hydrological research

**Acknowledgements**

The study was supported by a grant from the Russian Science Foundation No. 24-11-20024 and the Ministry of Education and Science of Khabarovsk Krai (Agreement No. 124C/2024).

**For citation:** Sai S.V., Zinkevich A.V. Method of semantic segmentation of airborne laser scanning data of water protection zones. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 1, pp. 68–77 (in Russian). doi: 10.17586/2226-1494-2025-25-1-68-77

**Введение**

Амур — одна из крупнейших рек мира, имеет десятый по размеру речной бассейн площадью 1,85 млн км<sup>2</sup>, длина р. Амур составляет 2,8 тыс. км. Бассейн р. Амур является одним из самых паводкоопасных районов Российской Федерации, поэтому мониторинг изменений речного русла, поймы, рельефа дна и окружающей местности, которые воздействуют на характер движения воды, является актуальной задачей. Построение эффективной системы прогнозирования можно считать первоочередной задачей. Лучшим решением для построения комплексной прогнозной системы является создание цифровых моделей [1].

Для разработки цифровых моделей русла, береговой линии и водоохранной зоны р. Амур и ее основных притоков требуется существенное количество пространственных данных. Поскольку гидрологические исследования часто выполняются в условиях сложного рельефа местности, заболоченных почв, непроходимой густой древесно-кустарниковой растительности, ограждений и других препятствий, то работы наземными методами могут быть невыполнимы. Данные дистанционного зондирования Земли уже давно используются для получения полезной информации. При этом космические снимки представляют информацию среднего и относительно высокого пространственного разрешения, что неблагоприятно сказывается на качестве результатов моделирования. В этих случаях хорошо работают аэрофототопографические методы, позволяющие получить данные сверхвысокого пространственного разрешения. Однако изображения не могут передать полную информацию об осматриваемом объекте, так как для этого необходимо сделать несколько снимков, чтобы оценить объект полностью. Это особенно

актуально на пологих участках поймы с небольшими перепадами высот.

В настоящее время существуют современные методы выполнения топографических работ с беспилотных летательных аппаратов (БПЛА) с применением лидарной съемки. Одним из преимуществ такой съемки является возможность классифицировать полученные точки (земля, растительность, строение и пр.). Также лидар формирует плотное облако точек даже в условиях густой сплошной залесенности и, таким образом, позволяет четко передать фактический рельеф местности. Все это способствует развитию методов воздушной съемки территории с БПЛА, в том числе, проведение мониторинга дна водных объектов и донных наносов с использованием батиметрического лидара, что является предметом исследований в настоящей работе.

Построение модели невозможно без понимания сцены, что является важным шагом в анализе данных облаков точек, полученных от лидара. При решении задачи классификации каждой точке в трехмерном облаке должны быть назначены правильные классы. Эта проблема известна как семантическая сегментация и хорошо изучена в приложениях компьютерного зрения, таких как автономные автомобили, где доступны разнообразные общедоступные наборы данных, например SemanticKITTI [2]. Однако семантическая сегментация недостаточно хорошо изучена для данных сканирования лидара на базе БПЛА. Наиболее заметным отличием является вид сцены и объектов. Данные, которые получены с БПЛА, также более разрежены.

Необходимость подготовки собственных наборов данных становится особенно актуальной в этом контексте. Учитывая уникальные характеристики сканируемых сцен, создание специализированных наборов данных, соответствующих конкретным условиям и

классам, позволяет повысить достоверность и точность модели. В отличие от общедоступных наборов данных, собственные наборы позволяют более точно отражать специфику задач, связанных с семантической сегментацией, таких как различие растительности, транспортных средств и строительных объектов.

Использование нейронных сетей может значительно облегчить обработку большого объема полученной информации, определяя известные классы и автоматически обучаясь на конкретных наборах данных. Однако разработанные универсальные архитектуры нейронных сетей должны быть адаптированы для работы с разными наборами данных и задачами. Важно учитывать особенности каждого нового набора данных, чтобы избежать потери точности.

В настоящей работе предлагается оценка нейросетевого метода для семантической сегментации трехмерных облаков точек, полученных путем воздушного лазерного сканирования с применением БПЛА «Геоскан 401 Лидар». Разработана методика получения и подготовки набора данных с четырьмя классами: земля, растительность, транспортные средства и строительные объекты. Предложена реализация нейронной сети на основе используемой модели глубокого обучения PointNet++ и приведена оценка точности для ряда экспериментов. Для увеличения точности оценки предложена методика, основанная на методах аугментации и перераспределения наборов данных. Цель исследования — разработка методик и алгоритмов для задачи семантической сегментации данных воздушного лазерного сканирования при построении цифровых моделей р. Амур и ее основных притоков.

### Получение данных воздушного лазерного сканирования

Экспериментальные данные получены с помощью лидара Velodyne VLP-32C, установленного на БПЛА «Геоскан 401 Лидар»<sup>1</sup>. Датчик установлен под БПЛА, чтобы иметь возможность измерять поверхность земли сверху. Из-за крепления датчика только около 180° горизонтального поля зрения датчика представляют полезные данные. Полученная информация является набором точек в трехмерном пространстве, представляющих неупорядоченный, неструктурированный массив.

Итоговый набор данных состоит из файлов, в которых находится информация о части водоохранной зоны р. Амур и ее притоках. Каждый файл содержит данные о различной площади сканирования. Всего было получено 18,1 млн точек на площади размером 500 × 700 м. Данные включают примеры открытых территорий, деревьев и кустарников, построек, а также транспортных средств. Высота полета задавалась в пределах 100 м от уровня земли. Растительность и земля появляются во всех частях съемки, в то время как другие объекты встречаются редко. На рис. 1 представлена часть облака

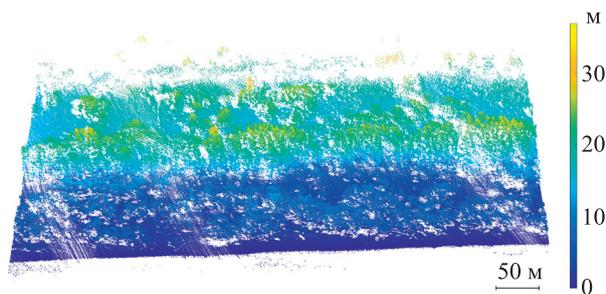


Рис. 1. Часть облака точек, окрашенная в зависимости от значения высоты.

В нижней части рисунка находится граница с рекой  
Fig. 1. Part of the point cloud colored depending on the altitude value.

At the bottom of the figure is the border with the river

точек, окрашенного в зависимости от высоты. Эта визуализация позволяет наглядно увидеть топографические особенности местности: более высокие точки окрашены в теплые оттенки, такие как красный и желтый, в то время как более низкие участки отображаются в холодных тонах — от синего до зеленого. В нижней части рисунка отчетливо прослеживается граница с рекой, которая образует естественную линию раздела. Однако из-за сильного отражения света от поверхности воды для используемого лидара — это отражение может создавать ложные сигналы, что затрудняет выявление контуров реки и ее сегментацию.

Полученный набор точек представляет трехмерную сцену или объект. Решение задач классификации и сегментации позволяет извлечь из него пространственную структуру.

### Методы обработки данных

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. Ввиду особенностей таких данных, типовые полносвязные сверточные сети не могут быть применены.

В настоящее время методы классификации облаков точек на основе глубокого обучения можно разделить на четыре типа в зависимости от их сетевой структуры: многопроекционные методы, методы на основе вокселизации, методы на основе графовой свертки и методы на основе обработки точек.

Методы классификации облака точек, основанные на множественных проекциях, аналогичны методам классификации двумерных изображений. Объект рассматривается с разных сторон, формируя ряд проецируемых изображений вокруг трехмерной модели облака точек. Затем применяются методы обработки двумерных изображений, а классы, полученные в результате сегментации в двумерном виде, проецируются обратно на трехмерное облако точек [3–6].

Методы на основе вокселизации преобразуют облака точек в пространственно-одномерную сетку вок-

<sup>1</sup> Геоскан 401 Лидар [Электронный ресурс]. URL: <https://www.geoscan.ru/ru/products/geoscan401/lidar/> (дата обращения: 11.10.2024).

селей. Для выполнения задач классификации обычно используются трехмерные сверточные нейронные сети. Существующие методы, основанные на вокселизации, эффективны для неупорядоченных и неструктурированных облаков точек, однако определение размера вокселя представляет сложную задачу [7–10].

Методы на основе графовой свертки объединяют операции свертки с представлением структуры графа. Эти алгоритмы эффективны для захвата и обработки геометрической структуры облаков точек, однако при обработке разреженных графов модель сети в процессе обучения часто переобучается [11–13].

Методы обработки облаков точек напрямую принимают облака точек в качестве входных данных. Они используют сети облаков точек для извлечения признаков, сохраняя геометрическую информацию входных данных без высоких требований к оборудованию. Широко используемыми моделями в этой категории являются PointNet и PointNet++. Модель PointNet демонстрировала хорошие результаты в задачах семантической сегментации и классификации объектов. В [14] 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++ [15]. Для обработки облака точек использовались перекрывающиеся области, т. е. брались точки из всего облака, и вокруг этих точек строились области с заданным радиусом, после чего из данных областей выделялись характерные признаки.

Из-за потери пространственной информации в многопроекторных методах, высоких вычислительных требований к оборудованию в воксельных методах, сложностей в глобальном обучении признакам в методах на основе графовой свертки, а также особенностей используемых данных, полученных в результате лидарной съемки, в настоящей работе выбрана за основу модель PointNet++.

Исходными данными для обучения нейронной сети является наличие классифицированного облака точек. В связи с чем подготовка размеченного набора данных является важным этапом до начала работы с нейронной сетью, без которого невозможно получить оценку результатов.

### Подготовка классифицированного набора данных

Исходные данные содержат информацию о координатах точек. Данные о классе, к которому принадлежит точка, отсутствуют. Для формирования классифицированного набора данных была проведена предварительная обработка. Предложенная методика включает в себя следующие этапы.

1. Объединение данных из отдельных файлов в общее облако точек. Это связано с обработкой данных от сенсоров БПЛА и сохранением их в файле в виде последовательных сканирований с перекрытием соседних.
2. Разбиение на блоки размером  $50 \times 50$  м и отбрасывание блоков с количеством точек менее 1000.
3. Фильтрация ложных выбросов, наличие которых связано с отражающими поверхностями и сложной геометрией сцены. Используется алгоритм оценки среднего расстояния от текущей точки до 5 соседних точек, после чего точка считается выбросом, если это расстояние превышает указанное пороговое значение.
4. Перевод координат точек из глобальной системы координат в локальную. Алгоритм вычисляет разность между текущими координатами точки и значением центраида облака. При этом высотные отметки, для визуального удобства, считаются от нуля.
5. Ручное разделение точек и объектов по классам.
6. Объединение данных, полученных в разных классах, и перевод точек из локальной системы координат в глобальную.

После проведенной классификации получен набор данных, представляющий наибольший интерес в рамках поставленной задачи. Координаты точек данных приведены в фиксированной системе координат WGS 84. Данные были разделены на четыре класса: земля, растительность, транспортные средства и строительные объекты. Класс «Земля» содержит все типы земли и низкой растительности, которые не могут быть отдельно аннотированы как растительность. Примеры материалов в классе «Земля» включают гравий, траву, песок, камень и почву. В данных класса «Транспортные средства» присутствуют автомобили, грузовики, а также плавательные средства. Класс «Строительные объекты» содержит данные о любых рукотворных объектах, таких как здания, заборы и прочие постройки. Для каждого класса рассчитан его вес относительно общего количества точек по выражению:

$$w_i = \sqrt{\frac{n}{n_i}},$$

где  $w_i$  — вес  $i$ -го класса;  $n_i$  — количество точек  $i$ -го класса;  $n$  — общее количество точек. Информация по полученному набору данных приведена в табл. 1.

На рис. 2 приведен пример облака точек, представленный на рис. 1, но с проведенной классификацией. Наглядно видна разница по дисбалансу классов.

Наличие дисбаланса классов является проблемой при обучении нейронной сети, что может негативно сказаться на точности ее работы. Для увеличения точ-

Таблица 1. Статистика набора данных по классам точек

Table 1. Dataset statistics by point classes

Класс	Земля	Растительность	Транспортные средства	Строительные объекты
Количество точек	8 017 567	9 867 772	5504	148 518
Вес класса	1,11	1	42,34	8,15

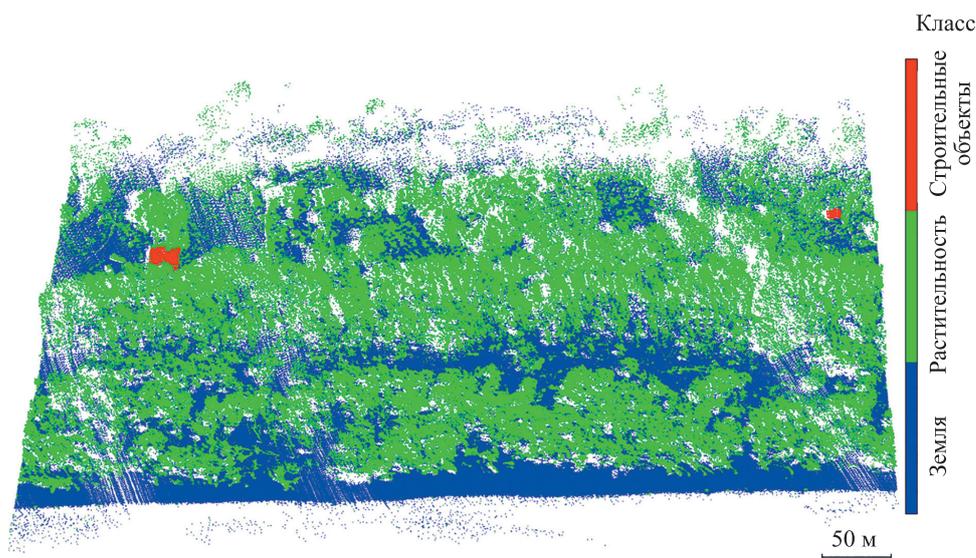


Рис. 2. Часть облака точек, окрашенная по классификации  
 Fig. 2. Part of the point cloud colored by classification

ности обучения были применены методы расширения данных для класса «Транспортные средства» и перераспределение наборов данных для классов «Земля» и «Растительность». Для расширения данных предложено не генерировать синтетические данные, а использовать существующие данные из облака, содержащего точки класса «Транспортные средства». Методика обработки представлена следующими шагами.

1. Отдельно выделены несколько различных облаков точек, содержащих один класс «Транспортные средства».
2. На выделенном наборе проведена аугментация, включающая случайные аффинные преобразования, такие как изменение масштаба, смещение и вращение.
3. Облако, не имеющее изначально класса «Транспортные средства» подвергалось прореживанию. Для этого из них удалялись случайным образом выбранные 15 % точек по классам «Земля» и «Растительность».
4. Полученный после аугментации набор точек инжектировался в прореженное облако и проверялось условие нахождения объекта в границах исходного облака. При несоответствии условия процесс аугментации повторялся.
5. Добавлялся новый объект и проверялась коллизия с ранее добавленным объектом. Процесс повторялся до устранения коллизий. При этом количество инжектированных объектов в среднем составило от двух до четырех.

Полученный в результате расширенный набор данных составил 41,67 млн точек. Информация по расширенному набору данных приведена в табл. 2.

### Структура нейронной сети

Используемая в работе нейронная сеть основана на модели глубокого обучения PointNet++, которая применяет многослойный перцептрон для извлечения признаков из облаков точек и слой максимального объединения (MaxPooling) для извлечения глобальных признаков, необходимых для классификации. Исходная модель также имеет иерархическую структуру для обучения по признакам наборов точек, полученных с использованием алгоритма выборки дальних точек (Farthest Point Sampling, FPS). Таким образом, модель подходит для решения задач сегментации как плотных, так и разреженных облаков точек, что дает решение проблемы неоднородной плотности облаков точек в полученном наборе данных. Упрощенная структура модели PointNet++ приведена на рис. 3.

Модель PointNet++ обрабатывает входные данные облака точек последовательно с помощью типовых блоков SA (Set Abstraction) и FP (Feature Propagation). Блок SA является основной частью подсети классификации PointNet++ и состоит из трех частей: выборка, группировка и извлечение признаков (рис. 4). Выборка с помощью алгоритма FPS уменьшает начальное количество точек от  $N$  до  $N_1$ . Затем следует этап группировки и для каждой выбранной точки происходит

Таблица 2. Статистика расширенного набора данных по классам точек  
 Table 2. Statistics of the extended dataset by point classes

Класс	Земля	Растительность	Транспортные средства	Строительные объекты
Количество точек	17 883 477	23 042 886	326 199	408 729
Вес класса	1,13	1	8,41	7,81



Рис. 3. Упрощенная структура модели PointNet++

Fig. 3. Simplified structure of the PointNet++ model

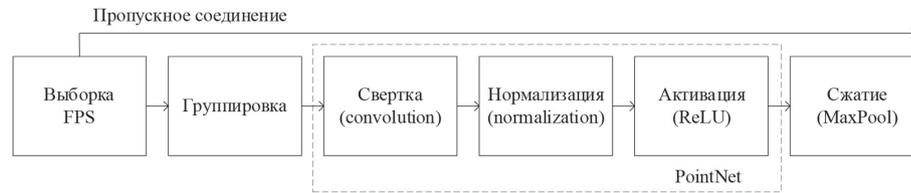


Рис. 4. Структура SA блока

Fig. 4. SA block structure

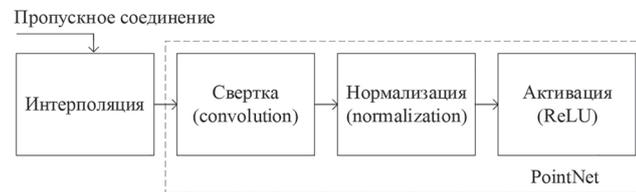


Рис. 5. Структура FP блока

Fig. 5. FP block structure

поиск ближайших  $k$  точек в заданном радиусе, тем самым из входной матрицы  $\mathbf{N} \times \mathbf{D}$  формируется матрица  $\mathbf{N}_1 \times \mathbf{k} \times \mathbf{D}$ . Последним этапом, с помощью сети PointNet, извлекаются глобальные признаки в области выборки, в результате чего получается  $\mathbf{N}_1 \times \mathbf{D}_1$  локальных признаков, где  $D$  и  $D_1$  представляют собой разные размерности признаков.

В результате для входного облака точек  $A = \{a, x\}$  на выходе SA блока получается выходное облако  $B = \{b, y\}$ , где  $a \in R^{N \times 3}$  и  $x \in R^{N \times D}$  — координаты и класс входных точек, а  $b \in R^{N_1 \times 3}$  и  $y \in R^{N_1 \times D_1}$  — координаты и класс выходных точек. Данные первого SA блока подаются на вход следующему SA блоку, повторяя этот процесс. И хотя количество центральных точек уменьшается, поле охвата увеличивается, что позволяет захватывать больше информации о локальных признаках точек. Стоит отметить, что каждый следующий SA блок имеет в два раза больший радиус поиска, чем предыдущий.

Блок FP интерполирует точки, а затем объединяет их с вычисленным классом из SA блока (рис. 5). Далее с помощью PointNet вычисляются классы точек более высокого уровня. Таким образом, перенос информации о классах в исходное облако (этап семантической сегментации) реализован подобно U-net сети.

Последний слой в модели PointNet++ выполняет линейную свертку для получения оценок классов для каждой точки. Категория с наибольшей оценкой является результатом классификации данной точки.

Результатирующая структура сети может содержать различное количество SA блоков, что влияет на точ-

ность сегментации и вычислительные затраты. В экспериментальной части работы исследованы варианты сети, включающие от трех до 5 SA блоков, а также различные радиусы поиска в слое группировки.

### Метрики оценки и параметры обучения сети

Количество входных точек сети может свободно задаваться, однако для лучшего использования вычислительных ресурсов (Graphics Processing Unit) и процесса пакетного обучения в сеть подается подготовленное облако  $A$  с количеством точек равным 8192. Для этого к входным данным применяются алгоритмы, включающие изменение размера (равномерное прореживание облака) и минимально-максимальную нормализацию.

Полученная ранее статистика весов классов показала дисбаланс, который при обучении нейронной сети может привести к смещенным предсказаниям и низкому качеству распознавания. Чтобы минимизировать дисбаланс классов использована функция потерь взвешенной кросс-энтропии:

$$WeightedCrossEntropy = - \sum_{t=1}^C w_t p_t \log_e(y_t),$$

где  $C$  — количество классов;  $w_t$  — вес для класса  $t$ ;  $p_t$  — истинная метка для класса  $t$ ;  $y_t$  — предсказанная вероятность метки для класса  $t$ .

В качестве алгоритма оптимизации используется оптимизатор Adam с числом epoch равным 20. Обучение

проводится с применением метода градиентного спуска. Скорость обучения установлена на 0,0005. Размер пакета (batch) равен 16.

Используемыми метриками оценки выбраны: пересечение по объединению (Intersection over union, IoU), среднее значение IoU для всех категорий (mIoU) и общая точность (overall accuracy, OA):

$$IoU = \frac{TP_i}{TP_i + FP_i + FN_i}; \quad mIoU = \frac{\sum_i^N IoU_i}{N};$$

$$OA = \frac{TP + TN}{TP + TN + FP + FN},$$

где TP (True Positives) — истинно положительные; TN (True Negatives) — истинно отрицательные; FP (False Positives) — ложноположительные; FN (False Negatives) — ложноотрицательные;  $N$  — количество классов.

Более высокое значение IoU указывает на то, что сеть предсказывает более точные границы сегментации. Метрика OA измеряет точность классификации модели, рассчитываемую как доля правильно классифицированных точек и общее количество точек.

### Методика эксперимента и результаты

Набор данных разделен случайным образом на обучающий и тестовый. На тестовый набор отводится 10 % данных.

В первой части эксперимента проведена оценка точности нейронной сети, обученной на схожем классифицированном наборе данных лидарной съемки с БПЛА — Dayton Annotated Laser Earth Scan (DALES), который содержит более 150 млн точек [16]. Выполнено сравнение точности при подаче на вход нейронной сети набора данных DALES и созданного набора дан-

ных. Различие классов в созданном наборе данных и наборе данных DALES решалось объединением классов DALES в более общие классы. В частности, классы «Легковые автомобили» и «Грузовики» объединены в более общий класс «Транспортное средство». Оставшиеся классы, за исключением «Земля» и «Растительность», объединены в класс «Строительные объекты» (табл. 3).

Полученные результаты показывают, что класс «Земля» сегментируется с высокой точностью в обоих наборах данных. В то же время уникальные характеристики сцен и объектов созданного набора данных не позволяют достичь высокой достоверности и точности сегментации. В частности, в классе «Строительные объекты» созданного набора данных, метрика IoU показывает низкую точность границ сегментации. Также можно отметить, что увеличение количества точек в наборе данных не дает существенного роста точности. Достижение более высоких результатов возможно как за счет обучения на созданном наборе данных, так и за счет перераспределения параметров в блоках поиска глобальных и локальных признаков.

Во второй части эксперимента проведена оценка точности нейронной сети, обученной на созданном наборе данных, при различном количестве SA блоков. Радиус поиска ближайших  $k$  точек на этапе группировки первого SA блока равен 10 % от размера нормализованного блока данных (табл. 4).

В общем случае увеличение количества SA блоков повышает средние значения OA и IoU, позволяя захватывать сложные структуры данных, однако существенно увеличивает вычислительные затраты. В частности, количество обучаемых параметров сети составляет 0,21, 0,89 и 3,5 млн параметров для SA от трех до 5, соответственно. Увеличение количества SA блоков также делает модель более чувствительной к классам, имеющим больший вес, что также сказывается на качестве сегментации мелких объектов. Наличие

Таблица 3. Метрики, полученные в результате сегментации сетью, обученной на наборе данных DALES

Table 3. Metrics obtained as a result of segmentation by a network trained on the DALES dataset

Класс	Набор данных DALES			Созданный набор данных		
	OA	IoU	mIoU	OA	IoU	mIoU
Земля	0,87	0,75	0,59	0,99	0,90	0,37
Растительность	0,82	0,72		0,41	0,41	
Транспортные средства	0,41	0,21		0,02	0,01	
Строительные объекты	0,79	0,69		0,66	0,16	

Таблица 4. Метрики, полученные в результате сегментации сетью, обученной на созданном наборе данных

Table 4. Metrics obtained as a result of segmentation by a network trained on the created dataset

Класс	Три SA блока			Четыре SA блока			5 SA блоков		
	OA	IoU	mIoU	OA	IoU	mIoU	OA	IoU	mIoU
Земля	0,89	0,83	0,31	1	0,86	0,28	0,99	0,91	0,40
Растительность	0,66	0,42		0,26	0,25		0,62	0,57	
Транспортные средства	0	0		0	0		0	0	
Строительные объекты	0,002	0,002		0,05	0,03		0	0	

Таблица 5. Метрики, полученные в результате сегментации сетью, обученной на расширенном наборе данных  
 Table 5. Metrics obtained as a result of segmentation by a network trained on the augmented dataset

Класс	Три SA блока			Четыре SA блока			5 SA блоков		
	OA	IoU	mIoU	OA	IoU	mIoU	OA	IoU	mIoU
Земля	0,99	0,92	0,47	0,99	0,93	0,53	0,99	0,92	0,54
Растительность	0,61	0,58		0,63	0,59		0,62	0,59	
Транспортные средства	0	0		0,68	0,24		0,40	0,14	
Строительные объекты	0,92	0,39		0,51	0,35		0,70	0,50	

Таблица 6. Метрики, полученные в результате сегментации сетью с уменьшенным радиусом поиска в SA блоках  
 Table 6. Metrics obtained as a result of segmentation by a network with a reduced search radius in SA blocks

Класс	Три SA блока			Четыре SA блока			5 SA блоков		
	OA	IoU	mIoU	OA	IoU	mIoU	OA	IoU	mIoU
Земля	0,99	0,93	0,58	0,99	0,92	0,52	0,99	0,92	0,64
Растительность	0,68	0,65		0,63	0,62		0,61	0,58	
Транспортные средства	0,38	0,27		0,48	0,28		0,45	0,26	
Строительные объекты	0,67	0,45		0,47	0,27		0,86	0,81	

в созданном наборе данных сильного дисбаланса классов, несмотря на использование функции потерь взвешенной кросс-энтропии, не позволяет сети правильно определять классы «Транспортные средства» и «Строительные объекты».

В третьей части эксперимента проведена оценка точности нейронной сети, обученной на расширенном наборе данных, который был создан по предложенной методике в разделе «Подготовка классифицированного набора данных» для уменьшения дисбаланса классов, в частности класса «Транспортные средства» (табл. 5).

Полученные результаты показали увеличение точности и качества работы сети с данными, подвергнутыми балансировке классов. Ошибки в определении класса «Транспортные средства» в сети с тремя SA блоками связано с невозможностью захватить достаточное количество признаков для корректной сегментации мелких объектов. Увеличить точность возможно за счет изменения радиуса поиска в слое группировки. Следует учесть, что большой радиус может привести к потерям детализации в локальных структурах, поскольку важно учитывать мелкие детали, которые могут быть потеряны в процессе обобщения. В то же время слишком маленький радиус может привести к избыточной чувствительности и шуму, поскольку группировка может быть основана на нескольких случайных точках.

В четвертой части эксперимента проведена оценка точности нейронной сети, обученной на расширенном наборе данных. Дополнительно был вдвое уменьшен радиус поиска на этапе группировки (табл. 6).

Уменьшение радиуса поиска повышает оценку точности при определении класса «Транспортные средства» для модели с тремя SA блоками, однако теряет больше признаков класса «Строительные объекты». В то же время модель с 5 SA блоками позволяет получить более точные оценки по всем классам. Можно отметить, что оптимальный радиус должен быть определен в зависимости от специфики решаемой задачи и

набора данных. Для используемых данных оптимальным можно считать радиус равный 5 % от размера нормализованного блока данных. Тем не менее, невысокий IoU по классу «Транспортные средства» указывает на необходимость добавлять больше разнообразных объектов этого класса в обучающую выборку, в частности, плавательных средств.

## Заключение

В работе оценена эффективность нейросетевого метода семантической сегментации трехмерных облаков точек, полученных с использованием беспилотного летательного аппарата «Геоскан 401 Лидар». Разработанная методика подготовки набора данных, включающего четыре ключевых класса — земля, растительность, транспортные средства и строительные объекты — позволила создать основу для качественной сегментации.

Использование в нейронной сети модели глубокого обучения PointNet++ продемонстрировало высокую точность в сегментации, что подтверждает целесообразность применения данного подхода. Увеличение количества агрегирующих блоков делает модель более чувствительной к классам, имеющим больший вес, что сказывается на качестве сегментации мелких объектов. Для дальнейшего повышения точности оценки отдельных классов предложена методика, основанная на аугментации и перераспределении наборов данных, что позволило значительно улучшить результаты сегментации. В частности, среднее значение Intersection over Union было улучшено, минимум на 35 %.

Дополнительное увеличение точности возможно при учете специфики решаемой задачи и набора данных. Большой радиус группировки приводит к потерям детализации в локальных структурах, в то время как слишком маленький радиус приводит к избыточной чувствительности и шуму, поскольку группировка может быть основана на нескольких случайных точках.

Для решаемой задачи радиус поиска определен экспериментально и составляет 5 % от размера нормализованного блока данных.

На основе результатов экспериментальных оценок получено, что увеличение количества точек в наборе данных не дает существенного роста точности. В то же время наличие большего разнообразия доступных и используемых наборов данных, полученных с помощью

разных лидаров при различных условиях сканирования, должно повысить эффективность применяемого подхода.

Полученные результаты исследования свидетельствуют о перспективности использования предложенных методик и алгоритмов для задач семантической сегментации при разработке цифровых моделей реки Амур.

### Литература

1. Неров И.О., Краснопедев С.М., Бугаец А.Н., Беликов В.В., Глотко А.В., Борисова Н.М., Васильева Е.С., Кролевецкая Ю.В. Опыт создания цифровой модели рельефа для гидродинамических расчетов в бассейне р. Амур // Вестник ДВО РАН. 2021. № 6 (220). С. 45–55. [https://doi.org/10.37102/0869-7698\\_2021\\_220\\_06\\_04](https://doi.org/10.37102/0869-7698_2021_220_06_04)
2. Behley J., Garbade M., Milioto A., Quenzel J., Behnke S., Stachniss C., Gall J., SemanticKITTI: a dataset for semantic scene understanding of LiDAR sequences // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2019. P. 9296–9306. <https://doi.org/10.1109/iccv.2019.00939>
3. Angrish A., Bharadwaj A., Starly B. MVCNN++: Computer-Aided design model shape classification and retrieval using multi-view convolutional neural networks // Journal of Computing and Information Science in Engineering. 2021. V. 21. N 1. P. 011001. <https://doi.org/10.1115/1.4047486>
4. Feng Y., Zhang Z., Zhao X., Ji R., Gao Y. GVCNN: group-view convolutional neural networks for 3D shape recognition // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 264–272. <https://doi.org/10.1109/cvpr.2018.00035>
5. Li M., Cao Y., Wu H. Three-dimensional reconstruction for highly reflective diffuse object based on online measurement // Optics Communications. 2023. V. 533. P. 129276. <https://doi.org/10.1016/j.optcom.2023.129276>
6. Sfikas K., Pratikakis I., Theoharis T. Ensemble of PANORAMA-based convolutional neural networks for 3D model classification and retrieval // Computers and Graphics. 2018. V. 71. P. 208–218. <https://doi.org/10.1016/j.cag.2017.12.001>
7. Riegler G., Ulusoy A.O., Geiger A. OctNet: learning deep 3D representations at high resolutions // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 6620–6629. <https://doi.org/10.1109/cvpr.2017.701>
8. Choy C., Gwak J., Savarese S. 4D spatio-temporal ConvNets: Minkowski convolutional neural networks // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 3070–3079. <https://doi.org/10.1109/cvpr.2019.00319>
9. Hua B.-S., Tran M.-K., Yeung S.-K. Pointwise convolutional neural networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 984–993. <https://doi.org/10.1109/cvpr.2018.00109>
10. Li W., Luo Z., Xiao Z., Chen Y., Wang C., Li J. A GCN-Based method for extracting power lines and pylons from airborne LiDAR data // IEEE Transactions on Geoscience and Remote Sensing. 2022. V. 60. P. 1–14. <https://doi.org/10.1109/TGRS.2021.3076107>
11. Wang Y., Sun Y., Liu Z., Sarma S.E., Bronstein M.M., Solomon, J.M. Dynamic graph CNN for learning on point clouds // ACM Transactions on Graphics. 2019. V. 38. N 5. P. 146. <https://doi.org/10.1145/3326362>
12. Liang Z., Yang M., Deng L., Wang C., Wang B. Hierarchical depthwise graph convolutional neural network for 3D semantic segmentation of point clouds // Proc. of the International Conference on Robotics and Automation (ICRA). 2019. P. 8152–8158. <https://doi.org/10.1109/ICRA.2019.8794052>
13. Zhao Y., Zhou F., Guo B., Liu B. Spatial temporal graph convolution with graph structure self-learning for early MCI detection // Proc. of the IEEE 20th International Symposium on Biomedical Imaging (ISBI). 2023. P. 1–5. <https://doi.org/10.1109/isbi53787.2023.10230807>
14. Qi C.R., Su H., Mo K., Guibas L.J. PointNet: deep learning on point sets for 3D classification and segmentation // Proc. of the IEEE conference on computer vision and pattern recognition (CVPR). 2017. P. 1–9. <https://doi.org/10.1109/cvpr.2017.16>

### References

1. Nerov I.O., Krasnopeyev S.M., Bugaets A.N., Belikov V.V., Glotko A.V., Borisova N.M., Vasileva E.S., Krolevetskaya Yu.V. Experience in digital terrain model generation for hydrodynamic modeling in the Amur River basin. *Vestnik of Far Eastern Branch of Russian Academy of Sciences*, 2021, no. 6 (220), pp. 45–55. (in Russian). [https://doi.org/10.37102/0869-7698\\_2021\\_220\\_06\\_04](https://doi.org/10.37102/0869-7698_2021_220_06_04)
2. Behley J., Garbade M., Milioto A., Quenzel J., Behnke S., Stachniss C., Gall J., SemanticKITTI: a dataset for semantic scene understanding of LiDAR sequences. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 9296–9306. <https://doi.org/10.1109/iccv.2019.00939>
3. Angrish A., Bharadwaj A., Starly B. MVCNN++: computer-aided design model shape classification and retrieval using multi-view convolutional neural networks. *Journal of Computing and Information Science in Engineering*, 2021, vol. 21, no. 1, pp. 011001. <https://doi.org/10.1115/1.4047486>
4. Feng Y., Zhang Z., Zhao X., Ji R., Gao Y. GVCNN: Group-View Convolutional Neural Networks for 3D Shape Recognition. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 264–272. <https://doi.org/10.1109/cvpr.2018.00035>
5. Li M., Cao Y., Wu H. Three-dimensional reconstruction for highly reflective diffuse object based on online measurement. *Optics Communications*, 2023, vol. 533, pp. 129276. <https://doi.org/10.1016/j.optcom.2023.129276>
6. Sfikas K., Pratikakis I., Theoharis T. Ensemble of PANORAMA-based convolutional neural networks for 3D model classification and retrieval. *Computers and Graphics*, 2018, vol. 71, pp. 208–218. <https://doi.org/10.1016/j.cag.2017.12.001>
7. Riegler G., Ulusoy A.O., Geiger A. OctNet: learning deep 3D representations at high resolutions. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6620–6629. <https://doi.org/10.1109/cvpr.2017.701>
8. Choy C., Gwak J., Savarese S. 4D spatio-temporal ConvNets: Minkowski convolutional neural networks. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3070–3079. <https://doi.org/10.1109/cvpr.2019.00319>
9. Hua B.-S., Tran M.-K., Yeung S.-K. Pointwise convolutional neural networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 984–993. <https://doi.org/10.1109/cvpr.2018.00109>
10. Li W., Luo Z., Xiao Z., Chen Y., Wang C., Li J. A GCN-Based method for extracting power lines and pylons from airborne LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, vol. 60, pp. 1–14. <https://doi.org/10.1109/TGRS.2021.3076107>
11. Wang Y., Sun Y., Liu Z., Sarma S.E., Bronstein M.M., Solomon, J.M. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 2019, vol. 38, no. 5, pp. 146. <https://doi.org/10.1145/3326362>
12. Liang Z., Yang M., Deng L., Wang C., Wang B. Hierarchical depthwise graph convolutional neural network for 3D semantic segmentation of point clouds. *Proc. of the International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8152–8158. <https://doi.org/10.1109/ICRA.2019.8794052>
13. Zhao Y., Zhou F., Guo B., Liu B. Spatial temporal graph convolution with graph structure self-learning for early MCI detection. *Proc. of the IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, 2023, pp. 1–5. <https://doi.org/10.1109/isbi53787.2023.10230807>
14. Qi C.R., Su H., Mo K., Guibas L.J. PointNet: deep learning on point sets for 3D classification and segmentation. *Proc. of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, pp. 1–9. <https://doi.org/10.1109/cvpr.2017.16>

15. Qi C.R., Li, Y, Su H., Guibas L.J. PointNet++: deep hierarchical feature learning on point sets in a metric space // arXiv. 2017. arXiv:1706.02413v1. <https://doi.org/10.48550/arXiv.1706.02413>
16. Varney N., Asari, V.K., Graehling Q. DALES: a large-scale aerial LiDAR data set for semantic segmentation // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2020. P. 1–10. <https://doi.org/10.1109/cvprw50498.2020.00101>
15. Qi C.R., Li, Y, Su H., Guibas L.J. PointNet++: deep hierarchical feature learning on point sets in a metric space. *arXiv*, 2017, arXiv:1706.02413v1. <https://doi.org/10.48550/arXiv.1706.02413>
16. Varney N., Asari, V.K., Graehling Q. DALES: a large-scale aerial LiDAR data set for semantic segmentation. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2020. pp. 1–10. <https://doi.org/10.1109/cvprw50498.2020.00101>

#### Авторы

**Сай Сергей Владимирович** — доктор технических наук, профессор, профессор, Тихоокеанский государственный университет, Хабаровск, 680035, Российская Федерация, [sc 14042260800](https://orcid.org/0000-0002-4506-1857), <https://orcid.org/0000-0002-4506-1857>, [sai1111@rambler.ru](mailto:sai1111@rambler.ru)

**Зинкевич Алексей Владимирович** — кандидат технических наук, доцент, Тихоокеанский государственный университет, Хабаровск, 680035, Российская Федерация, [sc 57283303300](https://orcid.org/0000-0003-3231-8569), <https://orcid.org/0000-0003-3231-8569>, [zinkevich1985@mail.ru](mailto:zinkevich1985@mail.ru)

*Статья поступила в редакцию 17.10.2024*  
*Одобрена после рецензирования 10.12.2024*  
*Принята к печати 24.01.2025*

#### Authors

**Sergey V. Sai** — D.Sc., Full Professor, Pacific National University, Khabarovsk 680035, Russian Federation, [sc 14042260800](https://orcid.org/0000-0002-4506-1857), <https://orcid.org/0000-0002-4506-1857>, [sai1111@rambler.ru](mailto:sai1111@rambler.ru)

**Alexey V. Zinkevich** — PhD, Associate Professor, Pacific National University, Khabarovsk 680035, Russian Federation, [sc 57283303300](https://orcid.org/0000-0003-3231-8569), <https://orcid.org/0000-0003-3231-8569>, [zinkevich1985@mail.ru](mailto:zinkevich1985@mail.ru)

*Received 17.10.2024*  
*Approved after reviewing 10.12.2024*  
*Accepted 24.01.2025*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»