

УДК 004.8:528.8

DOI 10.33764/2411-1759-2022-27-3-74-94

Анализ методов и средств искусственного интеллекта для анализа и интерпретации данных активного дистанционного зондирования

А. А. Колесников^{1*}

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск,

Российская Федерация

* e-mail: alexeykw@mail.ru

Аннотация. Данные дистанционного зондирования, как и большинство видов пространственных данных, являются комплексными, динамическими, слабоструктурированными, что затрудняет создание однозначного и универсального процесса их обработки и использования. В то же время развитие аппаратного обеспечения, методов и алгоритмов искусственного интеллекта и машинного обучения привело к тому, что направления информационных технологий находят применение практически во всех областях науки и техники, в том числе и при обработке пространственных данных. В статье сформулированы основные сложности и задачи обработки данных дистанционного зондирования, представлены наиболее распространенные в настоящее время методы и средства их обработки, использующие технологии искусственного интеллекта с целью автоматизации процессов. Рассмотрены возможности использования конкретных алгоритмов и методов искусственного интеллекта для всех этапов обработки данных активного дистанционного зондирования.

Ключевые слова: искусственный интеллект, активное дистанционное зондирование, обработка данных, машинное обучение, облака точек

Введение

Сегодня все большее число задач обработки данных дистанционного зондирования успешно и эффективно решается с помощью машинного и глубокого обучения благодаря возросшей доступности данных и вычислительных ресурсов [1]. Данные дистанционного зондирования ставят новые задачи для методов машинного обучения, поскольку они представляют собой сложноструктурированные данные, обработка которых поднимает некоторые уникальные вопросы, которые преобразуются в нетривиальные задачи:

– данные дистанционного зондирования часто являются мультимодальными, например, от оптических (мульти- и гиперспектральных) и радиолокационных датчиков с синтезированной апертурой (SAR), где геометрия изображения и содержание могут полностью различаться. Совмещение различных типов данных требует объединять эти источники синергетическим образом. Уже на этом этапе, до процессов извлечения информации, методы машинного обучения могут быть использованы для разработки новых архитектур совмещения

изображений одной территории, снятых с разных точек зрения и с другими спектральными характеристиками, при этом как опираясь на существующую трехмерную модель рельефа, так и без нее;

– данные дистанционного зондирования имеют пространственную привязку, то есть они расположены в географическом пространстве. Координаты каждого пикселя однозначно определены, что облегчает поступление информации о пикселях из других источников данных, таких как векторные слои геоинформационных систем, изображения из социальных сетей, содержащих геотеги или значения показателей с определенных датчиков (например, метеостанций), привязанных к определенной точке. С одной стороны, этот факт позволяет значительно упростить проблему слияния данных из различных источников, в то время как с другой стороны, он открывает поле для новых исследований, таких как локализация изображений, изначально не имеющих пространственной привязки, автоматизация проверки достоверности и точности данных на основе их взаимопересечения, реализация приложений дополненной реальности;

– данные ДЗЗ представляют собой геодезические измерения с контролируемым качеством. Это позволяет извлекать все параметры с оценкой точности. Однако, в отличие от подходов, основанных исключительно на данных, роль предварительных знаний об принципах работы датчиков, их возможных ошибках и способах корректировки, качестве промежуточных данных становится еще более важной. Например, для извлечения топографической информации, даже при одном и том же пространственном разрешении, интерферограммы, полученные с использованием однопроходной системы SAR, считаются более важными, чем те, которые получены в режиме повторного прохождения;

– показатели времени становятся практически все более неотъемлемым атрибутом всех значимых исследований. Например, программа Copernicus гарантирует непрерывный сбор данных в течение десятилетий, а спутники Sentinel получают изображение одной и той же точки на Земле каждые шесть дней. Это дает возможность перейти от анализа отдельных изображений к обработке временных рядов. Новые нейросетевые архитектуры должны быть разработаны для оптимального использования временной информации вместе с пространственной и спектральной информацией этих же данных;

– проблемы обработки больших объемов полуструктурированных данных. Например, спутники Sentinel уже получили около 35 петабайт данных, и их сбор продолжается. В крупных исследованиях часто необходимо не только обработать большую часть этих данных, но и отследить зависимости по времени. Таким образом, требуются универсальные алгоритмы, достаточно быстрые и переносимые для применения на любой произвольной территории на всей поверхности Земли. С другой стороны, эти данные хорошо аннотированы и содержат много метаданных. Следовательно, в некоторых случаях большие наборы обучающих данных для алгоритмов машинного обучения могут генерироваться автоматически (полуавтоматически). Для решения большинства из перечисленных выше задач применяются алгоритмы, методы и технологии искусственного интеллекта. По-

скольку разнообразие видов данных дистанционного зондирования и способов их применения очень велико, в данной статье ограничимся анализом средств искусственного интеллекта для обработки и интерпретации данных активного дистанционного зондирования.

Методы и материалы

Перед тем как перейти непосредственно к алгоритмам, методам и технологиям извлечения структур из данных, получаемых методами и инструментами активного дистанционного зондирования, рассмотрим исследование по их предобработке.

В статье [2] представлена архитектура глубокого обучения, которая решает проблему трехмерной семантической сегментации неструктурированных облаков точек. Авторами вводятся дополнительные параметры, которые определяют окрестности точек в начальном мировом пространстве и в пространстве изученных объектов. Окрестности важны, так как они позволяют вычислять локальные или глобальные точечные объекты в зависимости от пространственной протяженности окрестности. Кроме того, включены специальные функции потерь для дальнейшей структуризации пространства изученных точечных объектов на основе расстояний между точками и центроидами их групп.

В статье [3] представлена управляемая на основе структуры данных техника передискретизации облаков точек. Основная идея состоит в том, чтобы изучать многоуровневые объекты на основе опорных точек и расширять набор точек с помощью многоконтурного модуля свертки, неявно в пространстве объектов. Затем модифицированный таким образом объект в виде облака точек разделяется на множество объектов, которые после восстанавливаются для набора точек с повышенной дискретизацией. Нейронная сеть применяется на уровне исправлений с функцией потери соединения, которая способствует тому, чтобы точки с повышенной дискретизацией оставались на базовой поверхности с равномерным распределением. Результаты показывают, что точки с повышенной дискре-

тизацией имеют лучшую однородность и расположены ближе к самой поверхности.

Большинство задач, решаемых на основе данных, получаемых методами активного дистанционного зондирования, относятся к следующим категориям:

- классификация и сегментация SAR данных;
- классификация и сегментация облаков точек;
- регистрация и совмещение облаков точек;
- поиск изменений и прогнозирование перемещений в облаках точек.

Далее подробно рассмотрим каждую из этих категорий.

Классификация и сегментация SAR данных

За последние несколько лет появляется все больше публикаций об исследованиях, связанных с анализом изображений, получаемых с радаров с синтезированной апертурой (англ. Synthetic Aperture Radar, SAR), во многом это связано с открытием доступа к актуальным данным этого типа. Среди этих исследований методы глубокого обучения были в основном применены в типовых задачах автоматизированного распознавания объектов и классификации типов поверхности отдельных территорий.

Автоматизированное распознавание объектов (англ. Automatic Target Recognition, ATR) является важным способом использования SAR во многих областях деятельности [4]. Стандартная последовательность действий состоит из трех этапов: обнаружение объектов, распознавание их контуров и итоговая классификация. Каждый этап имеет тенденцию выполнять намного более сложную и усовершенствованную обработку, чем его предшественник, и основывается на результатах предыдущего этапа обработки. Однако все три этапа можно рассматривать как проблему классификации, и по этой причине глубокое обучение нашло свое применение в рассматриваемой задаче [5]. Одними из первых применяют сверточные нейронные сети (англ. Convolutional Neural Network, CNN) при автоматизированном распознавании образов и тестируют свое решение на стандарт-

ном наборе данных ATR MSTAR [6]. Основной проблемой рассматриваемой задачи является отсутствие достаточного количества обучающих образцов по сравнению с оптическими изображениями. Это может привести к серьезному переобучению и, следовательно, значительно ограничивает возможности обобщения получаемых моделей. Увеличение обучающей выборки является самым правильным способом противодействия переобучению, но это не всегда возможно без значительных затрат. Авторы [7] предлагают дополнительно прореживать все полносвязанные слои из обычных CNN, которые отвечают за большинство обучаемых параметров и на основе этого принципа формируют архитектуру AConvNets. Проводится ряд экспериментов, чтобы проверить способность обобщения данной архитектуры, и она оказалась достаточно устойчивой к тем объектам, которые вызывали ошибки классификации при типовой архитектуре CNN. В этом случае может быть использовано прореживание полносвязанных слоев (или дропаут), которые изначально были разработаны как обучаемый классификатор, поскольку ограниченное число целевых классов можно рассматривать как шаблоны типов объектов, которые извлекаются из данных. Итоговое качество демонстрируется более высоким на этом же наборе данных ATR MSTAR. Авторы [8] по результатам своих исследований сделали общий вывод о том, что увеличение данных является необходимым и наиболее важным шагом для автоматизированного распознавания объектов с использованием CNN. Предлагаются различные стратегии расширения исходных данных, за счет вращения [9], масштабирования [10], интерполяции и других методов генерации данных на основе существующих (аугментации данных) [11]. Авторы [12] используют глубокие сети доверия (англ. Deep Belief Network, DBN) в работе, где стековые ограниченные машины Больцмана (англ. Restricted Boltzmann Machine, RBM) применяются для извлечения признаков, а затем результаты подаются в обучаемый классификатор на основе иерархической системы распознавания (англ. Hierarchical Recognition System, HRS). В статье [13] предлагается исполь-

зовать CNN, чтобы сначала извлечь характерные векторы, а затем передать их в SVM для классификации. CNN обучен с полносвязным слоем, но используются только результаты с предшествующего слоя. Применяется метод систематического дополнения данных, который включает в себя упругие искажения и аффинные преобразования. Он предназначен для имитации типичных ошибок визуализации, таких как изменение угла съемки. Авторы [14] применяют CNN к проблеме детектирования айсбергов/кораблей в открытом море и тестируют свой подход на изображениях TerraSAR-X StripMap. В статье [15] применяют определенный тип глубоких нейронных сетей и совмещают ее с особой архитектурой магистральных сетей (HighWay Network) для распознавания изображений SAR при извлечении контуров судов, в статье [16] применяется CNN для обнаружения объектов на заднем плане в изображениях SAR. Чтобы решить проблему нехватки обучающих образцов, эти ученые использовали программное обеспечение для моделирования данных для обучения. Авторы [17] следуют этой идее и представляют глубокую генеративную нейронную сеть (англ. Generative-Adversarial Network, GAN) для синтеза новых образцов SAR на основе существующих. Генеративно-сверточная нейронная сеть сначала обучается генерировать моделируемое изображение SAR из заданного набора обучающих данных. Во время этого процесса в промежуточных слоях создается пространство признаков. Затем CNN обучается отображать входное изображение SAR в пространство признаков. Цель состоит в том, чтобы разработать расширенную систему автоматизированного распознавания объектов, которая способна интерпретировать ранее невидимую цель в контексте всех известных целей.

Как одна из подзадач обработки SAR данных выделяется в ряде работ, направленных на классификацию поверхности рельефа. Для этой задачи также используется SAR, в частности поляриметрический SAR (PolSAR). Результаты обработки являются еще одним важным ресурсом в радиолокационном дистанционном зондировании. Это очень похоже на задачу сегментации изображения в компью-

терном зрении, но ориентировано на особые типы объектов рельефа местности как естественного, так и искусственного происхождения. Традиционные подходы в наибольшей степени основаны на попарно-поляриметрических целевых параметрах декомпозиции [18]. Дополнительно рассматриваются пространственные закономерности, которые возникают в изображениях SAR высокого разрешения [19]. Глубокое обучение предоставляет такой инструмент для автоматической классификации объектов, которые представляют пространственные закономерности, а также поляриметрические характеристики. В абсолютном большинстве исследований используется по крайней мере один тип неконтролируемых генеративных графических моделей, таких как глубокие сети доверия (DBN), автоэнкодеры (англ. Sparse Autoencoder, SAE) или ограниченные машины Больцмана (RBM). Авторы [20] используют многоуровневое обучение детектированию объектов рельефа для классификации PolSAR, где SAE используется для извлечения полезных функций из изображения PolSAR канала. В статье [21] предлагается глубокий сверточный автоэнкодер (DCAE) для автоматического извлечения признаков и проведения классификации. Он состоит из созданного вручную первого слоя преобразования, который содержит ядра, такие как матрица сигналов уровня серого и фильтры Габора, и второй слой преобразования масштаба, созданный вручную, который объединяет коррелированные соседние пиксели. Остальные слои обучения соответствуют стандартной архитектуре автоэнкодера SAE. Этот подход проверен на изображениях TerraSAR-X с одним вариантом поляризации и высоким разрешением. Авторы [22] позже предложили аналогичную структуру, называемую нейронной сетью с глубоким наблюдением и сжатием сетью (DSCNN), для классификации изображений SAR, которая дополнительно включает гистограмму дескрипторов ориентированного градиента (англ. Histogram of Oriented Gradients, HOG) в качестве созданных вручную ядер. Обучаемые слои автоэнкодера применяют контролируемые штрафы, которые фиксируют соответствующую разницу инфор-

мации между элементами изображения и метками типов объектов, и ограничение, которое повышает локальную инвариантность. Особенность исследования состоит в том, что уменьшение спекл-шумов дает худшую производительность, и авторы делают вывод, что их уменьшение может убирать некоторую полезную информацию.

Авторы публикации [23] проводили исследование по использованию глубоких сетей доверия (DBN) для решения задачи выделения типов городского землепользования и классификации земельного покрова с использованием данных PolSAR. В работе [24] предложено использование SAE в сочетании с данными сверхвысокого разрешения для классификации изображений PolSAR. Несколько уровней слоев автоэнкодера обучаются на суперпиксельной основе. Данные суперпиксели формируются на основе разложенного с помощью матриц Паули псевдоцветного изображения. Выходные данные SAE используются в качестве функции на последнем этапе кластеризации суперпикселей методом ближайшего соседа. Авторы [25] применили сложную

разреженную классификацию изображений «AE to PolSAR», [26] – использовали адаптивные ограниченные машины Больцмана для классификации PolSAR изображений, [27] – предлагают дискриминантные сети доверия (DisDBN) для классификации изображений SAR, в котором особенности дискриминанта изучаются путем объединения обучения ансамбля с сетью глубокого обучения без учителя.

В публикации [28] предлагается сеть с глубокими стеками для классификации изображений PolSAR, которая в основном использует преимущества быстрого вычисления расстояния Вишарта посредством линейной проекции. Предлагаемая сеть направлена на выполнение задачи кластеризации / классификации методом k-средних, где расстояние Вишарта используется в качестве показателя подобия.

Обобщая вышеизложенное, можно на текущий момент предложить схему, представленную на рис. 1, отображающую наиболее подходящие методы искусственного интеллекта на основе нейронных сетей для решения определенных задач обработки SAR данных.



Рис. 1. Методы обработки SAR данных на основе нейронных сетей

Классификация и сегментация облаков точек. В то время как традиционные подходы к трехмерной семантической сегментации обычно опираются только на структурную информацию объектов (то есть геометрии и формы объектов), в последние годы появились методы, объединяющие как визуальные, так и геометрические особенности (например, Structure from Motion (SfM) или Multiview Stereo (MVS) algorithms), которые восстанавливают облака точек из нескольких перекрывающихся изображений [29]. Представ-

ленная авторами работа описывает гибридную методологию для трехмерной семантической сегментации, опираясь как на 2D-, так и на 3D-пространство, и нацелена на исследование того, влияет ли выбор изображения на точность трехмерной семантической сегментации облаков точек. Экспериментальные результаты продемонстрированы на бесплатном онлайн-наборе данных, изображающих городские кварталы вокруг Парижа. Экспериментальная процедура не только подтверждает, что гибридный подход (сочетание гео-

метрических и цветовых данных) может обеспечить более точную семантическую сегментацию, но и демонстрирует важность точки зрения на объект. Таким образом, акцент был сделан на выявлении этого оптимального двумерного вида из нескольких перекрывающихся изображений. Целью данной работы является выполнение семантической сегментации облаков точек городских сцен, использующие как 2D-, так и 3D-объекты. Для выбора оптимального вида сначала все 3D-точки проецируются через матрицы проекции в 2D-пространство. Далее вокруг каждой точки в трехмерном пространстве на плоскости формируется прямоугольный участок с predetermined размерами. Координаты четырех углов, образующих прямоугольный объект, известны и проецируются на разные стороны объекта. Из всех изображений, на которых видна точка, выбираются только те, на которых изображен весь прямоугольник (то есть все четыре угла проецируются внутри изображения). После этого вычисляется площадь проецируемого участка в каждом изображении и получается оптимальное изображение, то есть то, которое обеспечивает максимальную площадь прямоугольника (т. е. изображение с максимальным количеством пикселей, образующим проецируемое изображение – патч). Используя это правило для оптимального выбора изображения, авторами предлагается исключать искаженные и недостаточно подробные виды для извлечения признаков. Для классификации авторами используются случайный лес (англ. Random Forest, RF) и AdaBoost. Итоговая точность (accuracy) составила 60 %. Сочетание двухмерных и трехмерных признаков привело к более точной семантической сегментации трехмерной сцены. Средняя точность равна 50,1 % для базовой процедуры классификации и 60,0 % – для предлагаемой процедуры отсеивания малоинформативных изображений, когда используются случайные леса. Авторы также отмечают, что более сложные методы машинного обучения, такие как условные случайные поля (англ. Conditional Random Fields, CRF), как в [30], могут быть интегрированы в этот конвейер обработки для достижения более высоких показателей точности.

Авторы [31] используют методы искусственного интеллекта для получения актуальных и точных данных о наземной биомассе леса, которые являются необходимым условием для поддержки текущего использования лесов и разработки планов развития лесостроительства и сельского хозяйства. Традиционно инвентаризация лесных массивов проводится путем интеграции спутниковых изображений, в частности оптических данных, с полевыми исследованиями из баз данных инвентаризации лесов. Однако полевые данные ограничены в удаленных и труднодоступных областях. Кроме того, оптический коэффициент отражения обычно насыщается на уровне биомассы высокой плотности и подвержен помехам в виде облаков. Это исследование было направлено на разработку основанного на глубоком обучении процесса для картографирования лесных массивов путем интеграции изображений Landsat 8 и Sentinel-1A с данными LiDAR. Эталонная карта была получена на основе фрагмента данных LiDAR и полевых измерений. Графики LiDAR – стратифицированные случайные образцы лесной биомассы, извлеченные из смоделированных полос LiDAR на эталонной карте, – были приняты в качестве обучающей выборки для традиционных полевых графиков. В дополнение к модели глубокого обучения, а именно к разреженной сети автоэнкодеров (SSAE), были оценены пять различных методов машинного обучения, включая ступенчатые линейные регрессии, методы ближайшего соседа, машину опорных векторов, нейронные сети обратного распространения и случайный лес. Оптические переменные (Landsat 8 OLI), переменные SAR (Sentinel-1A) и их комбинированные переменные были индивидуально обработаны с помощью шести полученных моделей. Результаты показали, что модель SSAE показала наилучшие результаты при оценке биомассы леса. Объединенный набор оптических и микроволновых данных в качестве пояснительных переменных улучшил производительность моделирования по сравнению с данными только оптических или только микроволновых данных, независимо от алгоритмов прогнозирования. Наилучшая точность картирования была получена с помощью модели SSAE с вхо-

дами оптических и микроволновых интегрированных метрик, которые по коэффициенту детерминации составила 0,812, среднеквадратическая ошибка (RMSE) 21,753 Мг/га и относительное RMSE (RMSEr) 14,457 %.

Из-за особенностей отдельных алгоритмов искусственного интеллекта требуются предварительные преобразования исходных облаков точек в трехмерные сетки вокселей. Существует ряд конвертеров для прямого и обратного преобразования, но представление в формате вокселей делает данные излишне объемными. Авторы [32] предлагают новую архитектуру нейронной сети, которая позволяет напрямую использовать облака точек в исходном формате. Сеть, названная PointNet, предоставляет унифицированную архитектуру для классификации объектов, сегментации элементов объектов и семантического анализа сцены. Данная архитектура показывает высокую производительность, точность и стала базовой для более сложных архитектур и ансамблей нейронных сетей.

Авторами статьи [33] представлена архитектура нейронной сети с инвариантной перестановкой для глубокого обучения SO-Net. SO-Net моделирует пространственное распределение объектов в облаке точек путем построения самоорганизующихся карт (англ. Self-Organized Maps, SOM). На основе SOM SO-Net выполняет построение иерархий из отдельных точек и узлов SOM и в конечном итоге представляет входное облако точек набором связанных векторов для каждого объекта. Для входного слоя сети может систематически выполняться формирование дополнительных ограничений путем алгоритма поиска ближайшего соседа. В задачах распознавания, таких как реконструкция облака точек, классификация, сегментация частей объекта и извлечение форм, предлагаемая в статье сеть демонстрирует производительность, которая аналогична или лучше, чем у современных подходов. Кроме того, скорость обучения значительно выше, чем в существующих сетях распознавания облаков точек, благодаря использованию возможностей параллельной обработки и простоте предлагаемой архитектуры.

Еще одной сферой применения технологий искусственного интеллекта является об-

работка облаков точек для сельскохозяйственной робототехники. Одним из основных вариантов использования является текущая оценка состояния посевных культур, например, для борьбы с сорняками. В этом случае ключевым фактором корректности работы является надежное обнаружение и классификация растений и их видов. Автоматическая классификация видов растений по их трехмерным моделям является сложной задачей, потому что некоторые виды визуально выглядят очень похожими и трудно формально с помощью условий и формул описать внешности и различий между растениями в понятной для компьютера форме. Также не всегда возможно надежно выделить структурные элементы растения (листья, ветви, стебель) из массива трехмерных данных лазерного сканирования, необходимые для классификации определенного вида. Один из подходов к решению этой проблемы состоит в том, чтобы определять виды растений, используя набор, состоящий из примеров трехмерных моделей определенных видов и методы искусственного интеллекта. В статье [34] представлен метод классификации видов растений с использованием данных LiDAR-а и алгоритмов машинного обучения. Для этого был разработан набор алгоритмов для выбора наиболее характерных точек, инвариантных к размеру и ориентации всего массива данных. Помимо этих функций, также было выполнено сравнение различных базовых методов машинного обучения с использованием программного обеспечения Weka. Наилучшие результаты (около 98 %) показали логистическая регрессия, машины опорных векторов и нейронные сети. В экспериментах использовались шесть различных видов растений, распространенные в сельском хозяйстве, и около двадцати примеров, в виде трехмерных моделей, для каждого вида.

Как расширение PointNet, PointNet++ может извлекать локальные особенности объектов, что позволяет ему работать лучше, чем PointNet, при обработке данных облака точек. Но на практике часто бывает так, что плотность набора точек обычно меняется в зависимости от местоположения, что делает вычисления с помощью PointNet++ очень тяже-

лыми. Чтобы справиться с этим, авторами [35] предлагается структура сети на основе группирования элементов октодерева для PointNet++, которая называется Octree-Grouping-PointNet++ (OG-PointNet++). Она определяет плотность точек путем построения несбалансированного октодерева для облака точек и группирует точки в соответствии с плотностью точек. Эти группы точек назначаются различным слоям в соответствии с их плотностью, а локальные особенности объекта в каждой группы извлекаются PointNet++. Итоговый результат получается из последнего слоя и используется для классификации и сегментации. Эксперименты авторов показывают его конкурентоспособность во многих задачах обработки облаков точек, таких как классификация объектов и семантическая сегментация.

Большинство методов картирования видов деревьев основаны на сегментации отдельных деревьев, которые впоследствии классифицируются с использованием набора функций и соответствующего набора условий. Точность классификации для хвойных и лиственных деревьев, применяющих данные воздушного лазерного сканирования, составляет всего около 90 % в случае использования геометрической информации облака точек. Поскольку глубокие нейронные сети (DNN) обладают способностью адаптивно изучать функции из базовых данных, они превзошли классические подходы машинного обучения (ML) в известных наборах эталонных данных, предоставленных сообществом робототехники, компьютерного зрения и дистанционного зондирования. Тем не менее, классификация видов деревьев с использованием процедур глубокого обучения (DL) до настоящего времени представляла незначительный исследовательский интерес. Некоторые исследования были проведены на основе обширного предшествующего поколения изображений или вокселей из трехмерных необработанных данных. Поскольку инновационные DNN напрямую работают с нерегулярными и неупорядоченными трехмерными облаками точек в большом масштабе, целью исследования [36] является использование PointNet++ для семантической маркировки

облаков точек при классификации лиственных и хвойных деревьев. Предлагаемая архитектура включает в себя методику изучения иерархических особенностей массива данных, а также специальные слои, которые способны агрегировать мультимасштабную информацию в соответствии с локальными плотностями точек. Набор данных для экспериментов сформирован из данных лазерного сканирования на территорию Национального парка Баварский лес (366 деревьев/га), включая только ели (хвойные) и буки (лиственные). Этот набор данных был обработан с использованием классического подхода на основе случайного леса (RF) для классификации хвойных (accuracy = 93 %, recall = 80 %) и лиственных деревьев (precision = 82 %, recall = 92 %). Для сравнения результатов PointNet++ прошел обучение и впоследствии оценивался с использованием тех же случайно выбранных наборов данных. Достигнутые результаты классификации по точкам после 163 тренировочных эпох для хвойных деревьев (precision = 90 %, recall = 79 %) и лиственных деревьев (precision = 81 %, recall = 91 %) достаточно высоки, учитывая, что была включена только геометрия без использования цветовой характеристики. Тем не менее, результаты классификации с использованием PointNet++ немного ниже, чем результаты базового метода с использованием RF-классификатора. Причиной этого авторы называют ошибки в обучающих данных и возникающие граничные эффекты. Результаты показывают, что архитектура 3D DNN PointNet++ может быть успешно адаптирована к семантической маркировке больших облаков точек лазерного сканирования для классификации лиственных и хвойных деревьев. Перспективы исследований состоят в интеграции дополнительных параметров для каждой точки облака, таких как интенсивность лазера, нормали к поверхности и мультиспектральные характеристики в качестве исходных параметров данных для DNN. Кроме того, авторами планируется изучить классификацию многочисленных отдельных видов деревьев на основе предварительно сегментированных одиночных деревьев. Авторы [37] в своей работе использовали PointNet++ для неупорядо-

ченных трехмерных данных без первоначальной генерации растровых изображений или вокселей из облаков точек и классифицировали точки облака по отдельности. В эксперименте обучающий набор данных был разделен на кубические блоки с длиной ребра 60 м. Как и стандартный метод PointNet, PointNet++ работал лучше для хвойных деревьев, чем для лиственных деревьев. Одной из причин этого является более стандартизированное представление формы кроны хвойных деревьев. Также авторы столкнулись с эффектами неправильной классификации на краях блоков, так как информация о соседстве между блоками не была предоставлена модели. Одним из многообещающих подходов к решению этой проблемы является использование Superpoint Graphs для семантической маркировки облака точек. Проведенные авторами эксперименты доказывают, что архитектура 3D DNN PointNet++ может быть успешно адаптирована к семантической маркировке больших облаков точек лазерного сканирования для классификации видов деревьев [38].

Обнаружение плоскостей в облаке точек является необходимым условием для решения широкого круга задач.

Принцип минимальной длины описания (MDL) используется для решения нескольких конкурирующих гипотез. В статье [39] представлен новый подход к обнаружению плоскости путем интеграции алгоритмов RANSAC (RANDOM SAMPLE CONSENSUS) и MDL (MINIMUM DESCRIPTION LENGTH). Предлагаемый метод позволяет избежать обнаружения неправильных объектов из-за сложной геометрии трехмерных данных. В статье проверяется работоспособность предложенного метода как на синтетических, так и на реальных данных.

Авторами статьи [40] показано, что прямое расширение существующего алгоритма GAN неприменимо к облакам точек, потому что ограничение, требуемое для дискриминаторов, не определено для заданных данных. Предлагается модификация алгоритма GAN, позволяющая генерировать облака точек (PC-GAN). Используются идеи иерархического байесовского моделирования и неявных генеративных моделей.

Авторы работы [41] представляют полный анализ архитектуры PointNet, чтобы проиллюстрировать, какие функции извлекаются из облаков точек. Также предлагается новая архитектура нейронной сети под названием IPC-Net, чтобы улучшить современные облачные архитектуры точек. Показано, что IPC-Net выделяет больший набор уникальных функций, позволяющих модели производить более точные сегментации по сравнению с архитектурой PointNet.

В статье [42] предлагается PointNetVLAD, представляющая собой комбинацию / модификацию существующих PointNet и NetVLAD, позволяющая осуществлять сквозное обучение и вывод для извлечения глобального дескриптора из заданного трехмерного облака точек. Кроме того, предлагаются новые функции потерь, которые могут обеспечить более дискриминационные и обобщаемые глобальные дескрипторы для решения задачи поиска.

В работе [43] предлагается минимизировать этапы предобработки исходного облака точек путем использования новой архитектуры нейронной сети VoxelNet, которая объединяет извлечение признаков и обучение на основе ограничивающего прямоугольника в одну стадию обучения глубокой нейронной сети. В частности, VoxelNet делит облако точек на одинаково расположенные трехмерные воксели и преобразует группу точек в каждом вокселе в единое представление признаков через введенный слой кодирования признаков вокселей (VFE). Таким образом, облако точек кодируется как описательное объемное представление, которое затем используется для классификации. Эксперименты с набором данных для детектирования автомобилей KITTI показывают, что VoxelNet превосходит современные методы трехмерного обнаружения на основе LiDAR с большим отрывом.

Авторами [44] представляется новый подход к предложению трехмерных объектов, который называется «Сеть предложений генеративных форм» (GSPN), для сегментации экземпляров в данных облака точек.

В работе [45] предлагается сквозной глубокий автокодер для решения неконтролируемых задач обучения на облаках точек.

На стороне кодера применяется расширение на основе графов для продвижения локальных структур поверх PointNet. Предложенная структура декодера теоретически показана как общая архитектура, которая способна реконструировать произвольное облако точек из двумерного отображения облака точек.

Авторы [46] предлагают способ улучшения качества результатов классификации, когда глубокое обучение применяется для фильтрации воздушного облака точек LiDAR. Сравнивается базовый метод, который использует сверточную нейронную сеть (CNN) на основе вокселизации, а затем три метода для улучшения качества результата классификации. Первый метод – это предварительная обработка данных, цель которой – заранее исключить данные, которые явно не находятся на поверхности земли, чтобы эффективно извлечь данные поверхности земли. Предварительная обработка данных может значительно сократить количество целевых точек, и последующая обработка может быть выполнена эффективно. Это также предотвращает ошибочную классификацию шумоподобных точек, плавающих в воздухе, как поверхность земли по сравнению со случаем без предварительной обработки. Второй метод – это изменение сетевой структуры. Чтобы улучшить результат классификации базового метода, изменялась глубина слоя и размер диапазона свертки. Третий метод – это объединение результатов классификации из нескольких сетей. Были интегрированы отдельные результаты из нескольких сетей с различной глубиной слоя и размером свертки, основываясь на базовой обучающей выборке, и исследовалось, улучшились ли результаты. Отмечено, что даже если отдельные результаты были похожи, результаты классификации можно улучшить, объединив результаты.

В статье [47] рассматривается улучшенный подход к обогащению фотограмметрических облаков точек семантической информацией, извлеченной из изображений, чтобы обеспечить большую автоматизацию BIM-моделирования. На основе архитектуры DeepLabv3+ используется семантическая сегментация изображений для извлечения компонентов зданий и объектов интерьеров. Во время фотограмметрической реконструк-

ции сегментированные категории добавляются в облако точек. Любые интерполяции, возникающие во время этого процесса, корректируются автоматически, и точность достигает 51,9 % по способу оценки mIoU. На основе семантической информации дополнительно выравнивается облако точек, корректируется масштаб и извлекается дополнительная информация. Исследование подтверждает, что использование фотограмметрии и глубокого обучения для создания семантически обогащенного облака точек интерьеров дает хорошие результаты. Комбинированное извлечение геометрической и семантической информации дает высокий потенциал для автоматизированной реконструкции модели BIM.

По результатам проведенного анализа автор данной статьи предлагает разделить существующие способы трехмерной сегментации на три класса по типу используемой тематической основы: трехмерные сверточные нейронные сети, нейронные сети, не использующие свертки, и традиционные алгоритмические методы.

Трехмерные сверточные нейронные сети автором также предлагается классифицировать на основе критерия, который описывает способ представления исходных данных. Основываясь на этом критерии, первая категория трехмерных сверточных нейронных сетей использует матричное представление, как вариацию вокселей [48], где пустое пространство представляется либо как 0, либо как функция расстояния со знаком [49]. Это представление является простым с точки зрения использования в программном обеспечении и поддерживается всеми основными архитектурами нейронных сетей. Однако, в большинстве трехмерных сцен значительная часть (в среднем от 15 до 35 процентов) трехмерного пространства пуста, но эти ячейки являются составной частью общей матрицы, что приводит к излишнему потреблению памяти и замедленным вычислениям. Чтобы решить эту проблему [50], предложили для представления трехмерного пространства и расчета сверток использовать структуру октодеревя, а также разработали архитектуру нейронной сети OctNet для классификации объектов. Вторая категория сверточных нейронных сетей

использует разреженное трехмерное представление [51]. Разработчики Splat-Net [52] использовали пермутоэдральную решетку на основе прямоугольной сетки в качестве основы представления трехмерных данных при построении архитектуры сверточной нейронной сети, используемой при классификации и семантической сегментации [53]. Третью категорию представляют сверточные нейронные сети на основе псевдонепрерывного представления [54]. В отличие от предыдущих вариантов, они определяют свертки, используя непрерывные ядра в непрерывном пространстве [55]. Однако поиск соседей в непрерывном и нерегулярно заполненном пространстве вычислительно затратен, поскольку требует поиска по многомерному дереву, а не по хеш-таблицам.

Также достаточно популярны архитектуры нейронных сетей, не использующих трехмерные свертки [56]. Авторы [57] предлагают использовать двумерные свертки на поверхности для семантической сегментации объектов. Другой популярный вариант – это использование архитектур на основе PointNet [32]. Варианты PointNet используют набор входных координат в качестве начальных параметров для многослойного перцептрона [37]. Недостатком PointNet является то, что он способен обрабатывать ограниченное количество точек. Разработчики PointCNN [55] устраняют этот недостаток за счет ансамблирующей сети поверх множества базовых сетей, обрабатывающих блоки точек, а также предложили вариант непрерывной трехмерной свертки для нижних уровней PointNet и по результатам экспериментов получили значительное повышение производительности.

Традиционные алгоритмические методы машинного обучения представлены вариантами на основе использования метаточечных графов и деревьев решений. Разработчики Superpoint Graphs [38] предлагают представлять большие облака точек в виде набора взаимосвязанных простых блоков, объединенных метаточками, которые аналогичны суперпиксельным методам при сегментации растровых изображений. Эта структура представлена в виде аннотированного ориентированного графа, названного метаточечным графом. Представление в виде метаточечного графа

имеет преимущество, состоящее в том, что объекты и их отдельные части рассматриваются как единое целое, которые легче классифицировать, в отличие от классификации отдельных точек или вокселей. Кроме того, появляется возможность выделить и подробно описать отношения между соседними объектами, что имеет значение для классификации на основе контекста (автомобили, как правило, находятся над дорогами, потолки окружены стенами и т. д.). Также плюсом такого подхода является то, что размер графа определяется числом простых структур в облаке, а не общим количеством точек, которое обычно на несколько порядков больше. Другим методом, показывающим хорошие результаты, является ансамблирование алгоритмов деревьев решений и марковских случайных полей, и определение итогового класса на основе получаемых значений вероятностей. В этом случае деревья решений являются классификатором, а для обеспечения пространственной когерентности в классификации между точками используются марковские случайные поля [58].

Наиболее распространенными и простыми для обработки представлениями трехмерных данных являются векторы, матрицы и тензоры, но для многомерных пространств такие плотные представления неэффективны из-за разреженности объектов в трехмерном пространстве. Как одно из возможных решений предлагается сохранять только заполненную часть пространства в виде базовой координаты и связанных с ними элементов. Такой вариант представления облака точек является N -мерным расширением разреженной матрицы и известно как разреженный тензор [59]. Данный способ описания данных имеет ряд преимуществ: допускает однородное представление данных в базовых библиотеках для построения и использования нейронных сетей, поскольку большинство из них поддерживают разреженные тензоры: разреженная свертка очень похожа на стандартную свертку, которая доказала свою эффективность как в сегментации двумерных, так и трехмерных данных; разреженная свертка эффективна с точки зрения скорости вычислений [60]. Данный способ представления данных используется в одной из архитектур, названной авторами [61] Spatio-Temporal CNN.

Эта архитектура объединяет архитектуру 3D-UNet с 1D-AutoEncoder для обработки разновременных данных. Также комплексный вариант обработки динамических трехмерных данных предлагается авторами Minkowski Engine, где рассматриваются многомерные сверточные нейронные сети для четырехмерных пространственно-временных данных, трехмерного видео (видеоряд, дополненный данными с камеры глубины) и семимерного пространства, дополненного данными о времени и цветности. По сравнению с другими подходами, которые комбинируют временные данные с рекуррентной нейронной сетью или условными случайными полями (CRF), предложенная архитектура использует однородное представление и последовательные свертки во всех слоях сети. Вместо элементов рекуррентной нейронной сети используется свертка для данных по времени. Экспериментально обосновано, что сверточные нейронные сети с только обобщенными разреженными свертками могут превзойти двумерные или гибридные методы сегментации. Кроме того, использование подобных обобщений делает разреженные четырехмерные сверточные сети более устойчивыми к шумам и выбросам [62]. Но авторы отмечают, что нужно учитывать непоследовательность предсказаний многомерных пространственно-временных обобщенных разреженных сверточных сетей во всем анализируемом пространстве и времени. Для обеспечения согласованности предлагаются условные случайные поля большой размерности, определенные в 7-мерном трехстороннем пространстве (пространство – время – цвет) со стационарной функцией согласованности.

В статье [63] рассмотрена проблема обработки стереометрических данных с использова-

нием искусственной нейронной сети. Предложен метод ориентации облака точек. Рассмотрена проблема извлечения значимых данных из облака точек. Предложен алгоритм выделения нормализованных данных, не поднимающий проблему восстановления изломов на трехмерных поверхностях, что существенно снижает сложность вычислений.

Статья [64] рассматривает ряд популярных методов распознавания трехмерных объектов с помощью искусственных нейронных сетей. Приведено описание наиболее известных архитектур сверточных нейронных сетей, применяемых при анализе трехмерных сцен – PointNet и VoxNet, и необходимые для их работы преобразования исходных данных. Приведены результаты тестирования описанных в статье архитектур PointNet и VoxNet на наборе данных KITTI, содержащем трехмерные облака точек.

По результатам анализа автором сформирован сводный список алгоритмов машинного обучения (включая различные архитектуры нейронных сетей) и конкретных программных решений, классифицированный по типам исходных данных и местам съемки:

- облака точек на основе съемки внутренних помещений (indoor): PointNet, SnapNet, Multiscale SphericalNeighborhoods, PointWeb, SEGCloud;
- облака точек на основе съемки открытой местности (outdoor): SqueezeSeg, LidarPerception, Multiscale SphericalNeighborhoods, Superpoint Graphs, 3D-FCNN-TI, SPLATNet;
- определение составных частей механизмов, транспортных средств, промышленных объектов и т. п.: PointNet, PointNet++, OctNet, Graph-CNN, SGPN, KPConv, PointSIFT, SPLATNet (рис. 2).

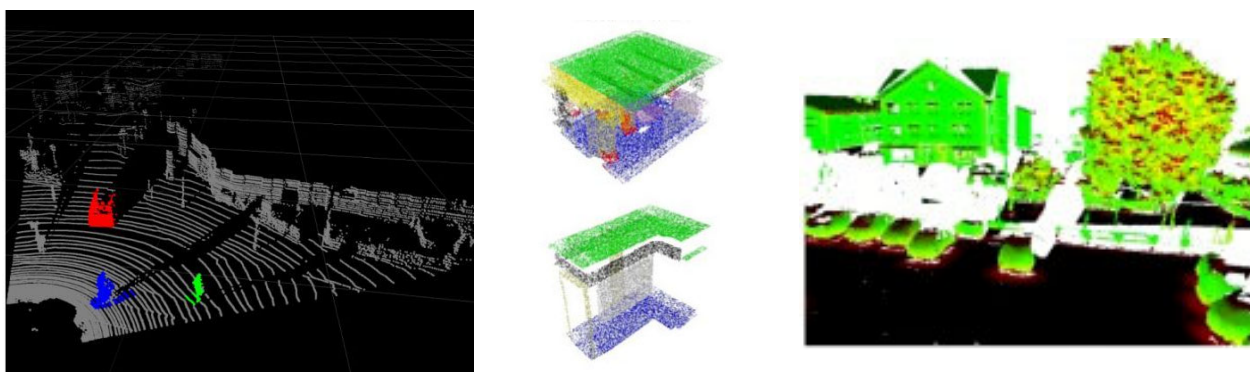


Рис. 2. Примеры обработки облаков точек

Авторами исследования [65] предложен алгоритм семантической сегментации трехмерных сцен, основанный на обработке последовательности изображений на основе сверточных нейронных сетей, с последующей проекцией результатов семантической сегментации на реконструированную трехмерную модель наблюдаемой сцены, представленную облаком 3D-точек. Представлены результаты экспериментального исследования на тестовом наборе ScanNet v2. Достигнутая точность алгоритма с использованием метрики IoU составила 0,5148. Авторы исследования объясняют это как возможными ограничениями предложенного решения, так и качеством разметки обучающего набора данных Pascal VOC и тестового набора данных ScanNet v2, а также возможными ошибками в данных о положении камер.

Регистрация и совмещение облаков точек. Регистрация облака точек – это задача выравнивания трехмерных сканов одной и той же среды, снятых с разных точек. Когда семантическая информация доступна для точек, она может использоваться как приоритет при поиске соответствий для улучшения качества регистрации.

Для задач классификации и сегментации подход PointNet и его последующие варианты / расширения считаются наиболее актуальными. На сегодняшний день применение PointNet для задачи регистрации облака точек остается малоисследованным. В статье [66] утверждается, что сама PointNet может рассматриваться как обучаемая функция «визуализации». Как следствие, классические алгоритмы машинного зрения для выравнивания изображений и облаков точек могут быть использованы для решения этой проблемы, а именно алгоритм Lucas & Kanade (LK). Авторами предлагается модифицировать алгоритм LK, чтобы приспособить функцию формирования изображений PointNet и развернуть PointNet и алгоритм LK в одну обучаемую рекуррентную глубокую нейронную сеть. Описывается архитектура и сравнивается ее производительность с современными в нескольких распространенных сценариях регистрации облаков точек. Архитектура PointNetLK предлагает несколько новых свойств,

в том числе: обобщение по категориям форм и вычислительную эффективность.

Семантическое преобразование нормальных распределений (SE-NDT) – это новый алгоритм регистрации, который уменьшает сложность задачи, используя семантическую информацию для разделения облака точек на набор нормальных распределений, которые затем регистрируются отдельно. В статье авторы [67] расширяют возможности регистрации NDT, используя PointNet, глубокую нейронную сеть для сегментации и классификации облаков точек, для изучения и прогнозирования семантических меток для каждой точки. Также авторами представлен итеративный эквивалент алгоритма ближайшей точки (ICP). Качество и производительность SE-NDT оценивается по сравнению с современным уровнем регистрации облака точек на основе общедоступного набора данных классификации Semantic3d.net. Также обученная модель классификации была протестирована на динамических сценах, используя блоки данных из открытого набора данных KITTI. Эксперименты демонстрируют улучшение регистрации с точки зрения надежности, точности и скорости во всем диапазоне первоначальных ошибок регистрации благодаря включению семантической информации.

Мониторинг состояния объектов. Мониторинг изменений в структуре сооружений является необходимым для большинства объектов. Многие алгоритмы обнаружения изменений используют анализ разности высот с использованием временных данных, таких как LiDAR. Тем не менее, значимые изменения зданий не всегда могут быть обнаружены в городских районах с плотной застройкой из-за влияния теней и визуальных коллизий, вызванных близлежащими.

Авторы [68] предлагают метод пересмотра 3D-данных о зданиях путем интеграции изменения текстуры (крыши и стены) и 3D-изменения формы зданий с использованием STARIMAGER / TLS.

Неотъемлемой частью технологии искусственного интеллекта являются методы оценки результатов и интерпретация получаемых математических моделей.

В работе [69] представлены способы интерпретации математических моделей, получаемых с помощью PointNet. Рассматривается корректная визуализация результатов активации точечных функций для того, чтобы изучить вопрос о том, как глобальные объекты представляют разные классы. Авторами предлагается нейронная сеть, производная от PointNet, названная C-PointNet, для генерации карт ответов с вниманием к классу для изуче-

ния того, на основе какой информации в облаке точек принимает решение PointNet. Эксперименты на ModelNet40 демонстрируют эффективность работы для лучшего понимания механизмов PointNet.

На основе анализа рассмотренных источников автором предлагается следующая классификация архитектур нейронных сетей (рис. 3), наиболее подходящих (на данный момент) для определенных задач обработки облаков точек.

Классификация трехмерных сцен	Семантическая сегментация	Регистрация и совмещение облаков точек	Мониторинг состояния
<ul style="list-style-type: none"> • PointMLP • CurveNet • RPNNet • PAConv • FG-Net • Point-BERT • DGCNN • PointCNN • PointNet++ 	<ul style="list-style-type: none"> • RFCR • SCF-Net • KPConv • SPGraph • SEGCloud • PointNet++ 	<ul style="list-style-type: none"> • MS-SVConv • LMVD • GeDi • D3Feat • 3DSmoothnet • 3DMatch 	<ul style="list-style-type: none"> • PoChaDeNN • HGI-CD • SiamGCN

Рис. 3. Методы обработки облаков точек на основе нейронных сетей

Заключение

Обобщая результаты проведенного обзора, можно отметить следующее.

1. Успешное применение технологий искусственного интеллекта в большинстве процессов обработки данных активного дистанционного зондирования доказывает их актуальность наряду с традиционными алгоритмами и методами.

2. Обработка пространственно-временных данных требует принципиально иных подходов, часто ориентированных на совмещение различных алгоритмов и методов обработки статичных трехмерных данных.

3. Для различных методов и алгоритмов обработки используются различные пред-

ставления трехмерных данных, что затрудняет сравнительный анализ решения одной и той же задачи с помощью разных подходов.

4. Подход на основе нейронных сетей позволяет более успешно по сравнению с алгоритмическим подходом решить ряд задач, но при этом получается своеобразная система «черного» ящика, поскольку интерпретация математических моделей нейронных сетей представляется сложной задачей.

5. Требуется значительная формализация выбора наиболее оптимальной архитектуры (или совмещенных архитектур) нейронных сетей с точки зрения исходных данных, решаемой задачи, точности результатов при обработке данных активного дистанционного зондирования.

Работа выполнена в рамках государственного задания Минобрнауки России (тема «Разработка теории и технологических решений контроля состояния защитных сооружений при перекачке нефтепродуктов методами активного дистанционного зондирования», № 0807-2020-0002).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Zhu X. X., Tuia D., Mou L., Xia G. S., Zhang L., Xu F., Fraundorfer F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources // *IEEE Geoscience and Remote Sensing Magazine*. – 2017. – P. 8–36.
2. Engelmann F., Kontogianni T., Schult J., Leibe B. Know what your neighbors do: 3D semantic segmentation of point clouds // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. – 2019. – Vol. 11131 LNCS. – P. 395–409.
3. Yu L., Li X., Fu C. W., Cohen-Or D., Heng P. A. PU-Net: Point Cloud Upsampling Network // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 2790–2799.
4. Dudgeon D. E., Lacoss R. T., Moreira A. An overview of automatic target recognition // *The Lincoln Laboratory Journal*. – 1993. – Vol. 6. – P. 3–10.
5. Chen S., Wang H. SAR target recognition based on deep learning // *International Conference on Data Science and Advanced Analytics*. – 2014. – P. 541–547
6. Keydel E. R., Lee S. W., Moore J. T. MSTAR extended operating conditions: a tutorial // *SPIE 2757, Algorithms for Synthetic Aperture Radar Imagery III*. – 1996.
7. Chen S., Wang H., Xu F., Jin Y. Q. Target classification using the deep convolutional networks for SAR images // *IEEE Transactions on Geoscience and Remote Sensing*. – Vol. 54, No. 8. – 2016. – P. 4806–4817.
8. Morgan D. Deep convolutional neural networks for ATR from SAR imagery // *SPIE 9475, Algorithms for Synthetic Aperture Radar Imagery XXII*. – 2015.
9. Ding J., Chen B., Liu H., Huang M. Convolutional neural network with data augmentation for SAR target recognition // *IEEE Geoscience and Remote Sensing Letters*. – Vol. 13, No. 3. – 2016. – P. 364–368.
10. Du K., Deng Y., Wang R., Zhao T., Li N. SAR ATR based on displacement- and rotation-insensitive CNN // *Remote Sensing Letters*. – 2016. – Vol. 7, No. 9. – P. 895–904.
11. Wilmanski M., Kreucher C., Lauer J. Modern approaches in deep learning for SAR ATR // *SPIE 9843, Algorithms for Synthetic Aperture Radar Imagery XXIII*. – 2016.
12. Cui Z., Cao Z., Yang J., Ren H. Hierarchical recognition system for target recognition from sparse representations // *Mathematical Problems in Engineering*. – 2016. – Vol. 2015, No. 527095.
13. Wagner S. A. SAR ATR by a combination of convolutional neural network and support vector machines // *IEEE Transactions on Geoscience and Remote Sensing*. – Vol. 52, No. 6. – 2016. – P. 2861–2872.
14. Bentes C., Frost A., Velotto D., Tings B. Ship-iceberg discrimination with convolutional neural networks in high resolution SAR images // *European Conference on Synthetic Aperture Radar (EUSAR)*. – 2016.
15. Schwegmann C., Kleynhans W., Salmon B., Mdakane L., Meyer R. Very deep learning for ship discrimination in Synthetic Aperture Radar imagery // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. – 2016.
16. Ødegaard N., Knapskog A.O., Cochin C., Louvigne J.C. Classification of ships using real and simulated data in a convolutional neural network // *IEEE Radar Conference (RadarConf)*. – 2016.
17. Song Q., Xu F., Jin Y.Q. Deep SAR image generative neural network and auto-construction of target feature space // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. – 2017.
18. Jin Y. Q., Xu F. Polarimetric scattering and SAR information retrieval // *Wiley-IEEE*. – 2013.
19. Xu F., Jin Y.Q., Moreira A. A preliminary study on SAR advanced information retrieval and scene reconstruction // *IEEE Geoscience and Remote Sensing Letters*. – 2016. – Vol. 13, No. 10. – P. 1443–1447.
20. Hie H., Wang S., Lie K., Lin S., Hou B. Multilayer feature learning for polarimetric synthetic radar data classification // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. – 2014.
21. Geng J., Fan J., Wang H., Ma X., Li B., Chen F. High-resolution SAR image classification via deep convolutional autoencoders // *IEEE Geoscience and Remote Sensing Letters*. – 2015. – Vol. 12, No. 11. – P. 2351–2355.
22. Geng J., Wang H., Fan J., Ma X. Deep supervised and contractive neural network for SAR image classification // *IEEE Transactions on Geoscience and Remote Sensing*. – 2017. – Vol. 55, No. 4. – P. 2442–2459.
23. Lv Q., Dou Y., Niu X., Xu J., Xu J., Xia F. Urban land use and land cover classification using remotely sensed SAR data through deep belief networks // *Journal of Sensors*. – 2015. – Vol. 2015, No. 538063.
24. Hou B., Kou H., Jiao L. Classification of polarimetric SAR images using multi-layer autoencoders and superpixels // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2016. – Vol. 9, No. 7. – P. 3072–3081.

25. Zhang L., Ma W., Zhang D. Stacked sparse autoencoder in PolSAR data classification using local spatial information // *IEEE Geoscience and Remote Sensing Letters*. – 2016. – Vol. 13, No. 9. – P. 1359–1363.
26. Qin F., Guo J., Sun W. Object-oriented ensemble classification for polarimetric SAR imagery using restricted Boltzmann machines // *Remote Sensing Letters*. – 2017. – Vol. 8, No. 3. – P. 204–213.
27. Zhao Z., Jiao L., Zhao J., Gu J., Zhao J. Discriminant deep belief network for high-resolution SAR image classification // *Pattern Recognition*. – Vol. 61. – 2017. – P. 686–701.
28. Zhang L., Lu D., Moon W. M. PolSAR Image Classification based on QCEA-optimized BP Neural Network // *CGU - CSSS*. – 2014.
29. Adam A., Grammatikopoulos L., Karras E., Protopapadakis E., Karantzas K. A semantic 3D point cloud segmentation approach based on optimal view selection for // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 6th International Workshop LowCost 3D – Sensors, Algorithms, Applications*. Strasbourg. – 2019. – Vol. XLII-2/W17.
30. Riemenschneider H., Bdis-Szomor A., Weissenberg J., Gool L. Learning where to classify in multi-view semantic segmentation // *In Proceedings European Conference on Computer Vision*. – 2014.
31. Shao Z., Zhang L., Wang L. Stacked Sparse Autoencoder Modeling Using the Synergy of Airborne LiDAR and Satellite Optical and SAR Data to Map Forest Above-Ground Biomass // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2017. – Vol. 10, No. 12. – P. 5569–5582.
32. Qi C. R., Su H., Mo K., Guibas L. J. PointNet: Deep learning on point sets for 3D classification and segmentation // *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. – 2017. – P. 77–85.
33. Li J., Chen B. M., Lee G. SO-Net: Self-Organizing Network for Point Cloud Analysis // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 9397–9406.
34. Weiss U., Biber P., Laible S., Bohlmann K., Zell A. Plant species classification using a 3D LIDAR sensor and machine learning // *9th International Conference on Machine Learning and Applications, ICMLA 2010*. – 2010. – P. 339–345.
35. Yao X., Guo J., Hu J., Cao Q. Using deep learning in semantic classification for point cloud data // *IEEE Access*. – 2019.
36. Briechle S., Krzystek P., Vosselman G. Semantic labeling of ALS point clouds for tree species mapping using the deep neural network pointnet++ // *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives 42(2/W13)*. – 2019. – Vol. 42(2/W13). – P. 951–955.
37. Qi C. R., Yi L., Su H., Guibas L. J. PointNet++: Deep hierarchical feature learning on point sets in a metric space // *Proceedings Advances in Neural Information Processing Systems*. – 2017.
38. Landrieu L., Simonovsky M. Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. – Salt Lake City. – 2018.
39. Yang M., Förstner W. Plane Detection in Point Cloud Data // *Proceedings of the 2nd International Conference on Machine Control Guidance*. Bonn. – 2010. – Vol. 1. – P. 95–104.
40. Li C., Zaheer M., Zhang Y., Póczos B., Salakhutdinov R. R. Point cloud GAN // *Deep Generative Models for Highly Structured Data, DGS@ICLR 2019 Workshop*. – 2019.
41. Marulanda F. G., Libin P., Verstraeten T., Nowé A. IPC-Net: 3D point-cloud segmentation using deep inter-point convolutional layers // *International Conference on Tools with Artificial Intelligence, ICTAI*. – 2018. – P. 293–301.
42. Uy M. A., Lee G. PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 4470–4479.
43. Zhou Y., Tuzel O. VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 4490–4499.
44. Yi L., Zhao W., Wang H., Sung M., Guibas L. J. GSPN: Generative shape proposal network for 3D instance segmentation in point cloud // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2019. – P. 3942–3951.
45. Yang Y., Feng C., Shen Y., Tian D. FoldingNet: Point Cloud Auto-Encoder via Deep Grid Deformation // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 206215.

46. Yotsumata T., Sakamoto M., Satoh T. Quality improvement for airborne lidar data filtering based on deep learning method // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2020. – Vol. XLIII-B2-2020. – P. 355–360.
47. Gülch E., Obrock L. S. Automated semantic modelling of building interiors from images and derived point clouds based on deep learning methods // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2020. – Vol. XLIII-B2-2020. – P. 421–426.
48. Tchapmi L. P., Choy C. B., Armeni I., Gwak J. Y., Savarese S. Segcloud: Semantic segmentation of 3d point clouds // International Conference on 3D Vision (3DV). – 2017.
49. Dai A., Chang A. X., Savva M., Halber M., Funkhouser T., Niener M. Scannet: Richly annotated 3D reconstructions of indoor scenes // Proceedings Computer Vision and Pattern Recognition (CVPR). – 2017.
50. Riegler G., Ulusoy A.O., Geiger A. Oct-net: Learning deep 3d representations at high resolutions // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2017.
51. Su H., Jampani V., Sun D., Maji S., Kalogerakis E., Yang M., Kautz J. SPLATNet: Sparse Lattice Networks for Point Cloud Processing // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2018. – P. 2530–2539.
52. Graham B. Sparse 3d convolutional neural networks // British Machine Vision Conference. – 2015.
53. Adams A., Baek J., Davis M.A. Fast high-dimensional filtering using the permutohedral lattice // Proceedings Computer Graphics Forum. – 2010. – Vol. 28. – P. 753–762.
54. Hermosilla P., Ritschel T., Vazquez P.P., Vinacua A., Ropinski T. Monte-Carlo convolution for learning on non-uniformly sampled point clouds // ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2018). – 2018.
55. Li Y., Rui B., Mungchao S., Wei W., Xinhan D., Baoquan C. PointCNN: Convolution On X-Transformed Points // NeurIPS 2018. – 2018.
56. Pan H., Liu S., Liu Y., Tong X. Convolutional neural networks on 3D surfaces using parallel frames. – 2018. – arxiv preprint arXiv: 1808.04952.
57. Tatarchenko M., Park J., Koltun V., Zhou Q. Tangent convolutions for dense prediction in 3D // CVPR. – 2018.
58. Gall Y. L., Thomas H., Goulette F., Deschaud J., Marcotegui B. Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods // 2018 International Conference on 3D Vision (3DV). Verona. – 2018.
59. Choy C. B., Xu D., Gwak J., Chen K., Savarese S. 3D-r2n2: A unified approach for single and multi-view 3D object reconstruction // Proceedings of the European Conference on Computer Vision (ECCV). – 2016.
60. Zeng A., Song S., Niener M., Fisher M., Xiao J., Funkhouser T. 3Dmatch: Learning the matching of local 3D geometry in range scans // CVPR. – 2017.
61. Zhao Y., Li X., Huang H., Zhang W., Zhao S., Makkie M., Zhang M., Li Q., Liu T. 4D Modeling of fMRI Data via Spatio-Temporal Convolutional Neural Networks (ST-CNN) // IEEE Transactions on Cognitive and Developmental Systems. – 2019.
62. Choy C., Gwak J., Savarese S. 4D Spatio-Temporal ConvNet: Minkowski Convolutional Neural Network // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2019.
63. Абросимов М. А., Бровко А. В. Метод нормализации облака точек, подлежащего обработке с помощью искусственной нейронной сети // Информационно-коммуникационные технологии в науке, производстве и образовании ICIT-2017. – Воронеж, 2017. – С. 262–268.
64. Бабаев А. М. Нейросетевые технологии распознавания трехмерных объектов // Международный журнал гуманитарных и естественных наук. – 2019. – Т. 39, No. 12-2. – С. 74–76.
65. Каздорф С. Я., Першина Ж. С. Алгоритм семантической сегментации трехмерных сцен // Cloud of Science. – 2019. – Vol. 6, No. 3. – P. 451–461.
66. Aoki Y., Goforth H., Srivatsan R. A., Lucey S. Pointnetlk: Robust & efficient point cloud registration using pointnet // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2019. – P. 7156–7165.
67. Zaganidis A., Sun L., Duckett T., Cielniak G. Integrating Deep Semantic Segmentation into 3-D Point Cloud Registration // IEEE Robotics and Automation Letters. – 2018. – Vol. 3, No. 4. – P. 2942–2949.
68. Neidhart H., Sester M. Identifying building types and building clusters using 3D-laser scanning and GIS-data // Machine Learning. – 2003.

69. Zhang B., Huang S., Shen W., Wei Z. Explaining the PointNet: What Has Been Learned Inside the PointNet? // IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2019. – P. 71–74.

Об авторах

Алексей Александрович Колесников – доцент, кандидат технических наук, доцент кафедры картографии и геоинформатики.

Получено 14.03.2022

© А. А. Колесников, 2022

Analysis of methods and tools of artificial intelligence for analysis and interpretation of active remote sensing data

A. A. Kolesnikov¹*

¹ Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation

* e-mail: alexeykw@mail.ru

Abstract. Remote sensing data, like most types of spatial data, are complex, dynamic, semi-structured, which makes it difficult to create an unambiguous and universal process for their processing and use. At the same time, the development of hardware, methods and algorithms of artificial intelligence and machine learning has led to the fact that the areas of information technology are used in almost all areas of science and technology, including the processing of spatial data. The article formulates the main difficulties and tasks of processing remote sensing data, presents the most common methods and tools for their processing at present, using artificial intelligence technologies to automate processes. The possibilities of using specific algorithms and methods of artificial intelligence for all stages of processing data from active remote sensing are considered.

Keywords: artificial intelligence, active remote sensing, data processing, machine learning, point clouds

REFERENCES

1. Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine* (pp. 8–36).
2. Engelmann, F., Kontogianni, T., Schult, J., & Leibe, B. (2019). Know what your neighbors do: 3D semantic segmentation of point clouds. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11131 LNCS, 395–409.
3. Yu, L., Li, X., Fu, C.W., Cohen-Or, D., & Heng, P. A. (2018). PU-Net: Point Cloud Upsampling Network. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 2790–2799).
4. Dudgeon, D. E., Lacoss, R. T., & Moreira, A. (1993). An overview of automatic target recognition. *The Lincoln Laboratory Journal*, 6, 3–10.
5. Chen, S., & Wang, H. (2014). SAR target recognition based on deep learning. *International Conference on Data Science and Advanced Analytics* (pp. 541–547).
6. Keydel, E. R., Lee, S. W., & Moore, J. T. (1996). MSTAR extended operating conditions: a tutorial. *SPIE 2757, Algorithms for Synthetic Aperture Radar Imagery III*.
7. Chen, S., Wang, H., Xu, F., & Jin, Y.Q. (2016). Target classification using the deep convolutional networks for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), 4806–4817.
8. Morgan, D. (2015). Deep convolutional neural networks for ATR from SAR imagery. *SPIE 9475, Algorithms for Synthetic Aperture Radar Imagery XXII*.
9. Ding, J., Chen, B., Liu, H., & Huang, M. (2016). Convolutional neural network with data augmentation for SAR target recognition. *IEEE Geoscience and Remote Sensing Letters*, 13(3), 364–368.
10. Du, K., Deng, Y., Wang, R., Zhao, T., & Li, N. (2016). SAR ATR based on displacement- and rotation-insensitive CNN. *Remote Sensing Letters*, 7(9), 895–904.

11. Wilmanski, M., Kreucher, C., & Lauer, J. (2016). Modern approaches in deep learning for SAR ATR. *SPIE 9843, Algorithms for Synthetic Aperture Radar Imagery XXIII*.
12. Cui, Z., Cao, Z., Yang, J., Ren, H. (2016). Hierarchical recognition system for target recognition from sparse representations. *Mathematical Problems in Engineering, 2015*, P. 527095.
13. Wagner, S. A. (2016). SAR ATR by a combination of convolutional neural network and support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 52(6)*, 2861–2872.
14. Bentes, C., Frost, A., Velotto, D., & Tings, B. (2016). Ship-iceberg discrimination with convolutional neural networks in high resolution SAR images. *European Conference on Synthetic Aperture Radar (EUSAR)*.
15. Schwegmann, C., Kleyhans, W., Salmon, B., Mdakane, L., & Meyer, R. (2016). Very deep learning for ship discrimination in Synthetic Aperture Radar imagery. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*.
16. Ødegaard, N., Knapskog, A. O., Cochin, C., & Louvigne, J. C. (2016). Classification of ships using real and simulated data in a convolutional neural network. *IEEE Radar Conference (RadarConference)*.
17. Song, Q., Xu, F., & Jin, Y. Q. (2017). Deep SAR image generative neural network and auto-construction of target feature space. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*.
18. Jin, Y. Q., & Xu, F. (2013). Polarimetric scattering and SAR information retrieval. *Wiley-IEEE*.
19. Xu, F., Jin, Y. Q., & Moreira, A. (2016). A preliminary study on SAR advanced information retrieval and scene reconstruction. *IEEE Geoscience and Remote Sensing Letters, 13(10)*, 1443–1447.
20. Hie, H., Wang, S., Lie, K., Lin, S., & Hou, B. (2014). Multilayer feature learning for polarimetric synthetic radar data classification. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*.
21. Geng, J., Fan, J., Wang, H., Ma, X., Li, B., & Chen, F. (2015). High-resolution SAR image classification via deep convolutional autoencoders. *IEEE Geoscience and Remote Sensing Letters, 12(11)*, 2351–2355.
22. Geng, J., Wang, H., Fan, J., & Ma, X. (2017). Deep supervised and contractive neural network for SAR image classification. *IEEE Transactions on Geoscience and Remote Sensing, 55(4)*, 2442–2459.
23. Lv, Q., Dou, Y., Niu, X., Xu, J., Xu, J., & Xia, F. (2015). Urban land use and land cover classification using remotely sensed SAR data through deep belief networks. *Journal of Sensors, 2015*, P. 538063.
24. Hou, B., Kou, H., & Jiao, L. (2016). Classification of polarimetric SAR images using multi-layer autoencoders and superpixels. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(7)*, 3072–3081.
25. Zhang, L., Ma, W., & Zhang, D. (2016). Stacked sparse autoencoder in PolSAR data classification using local spatial information. *IEEE Geoscience and Remote Sensing Letters, 13(9)*, 1359–1363.
26. Qin, F., Guo, J., & Sun, W. (2017). Object-oriented ensemble classification for polarimetric SAR imagery using restricted Boltzmann machines. *Remote Sensing Letters, 3*, 204–213.
27. Zhao, Z., Jiao, L., Zhao, J., Gu, J., & Zhao, J. (2017). Discriminant deep belief network for high-resolution SAR image classification. *Pattern Recognition, 61*, 686–701.
28. Zhang, L., Lu, D., & Moon, W. M. (2014). PolSAR Image Classification based on QCEA-optimized BP Neural Network. *CGU – CSSS*.
29. Adam, A., Grammatikopoulos, L., Karras, E., Protopapadakis, E., & Karantzas, K. (2019). A semantic 3D point cloud segmentation approach based on optimal view selection. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 6th International Workshop LowCost 3D – Sensors, Algorithms, Applications: Vol. XLII-2/W17*. Strasbourg.
30. Riemenschneider, H., Bdis-Szomor, A., Weissenberg, J., & Gool, L. (2014). Learning where to classify in multi-view semantic segmentation. *In Proceedings European Conference on Computer Vision*.
31. Shao, Z., Zhang, L., & Wang, L. (2017). Stacked Sparse Autoencoder Modeling Using the Synergy of Airborne LiDAR and Satellite Optical and SAR Data to Map Forest Above-Ground Biomass. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(12)*, 5569–5582.
32. Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation. *Proceedings – 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* (pp. 77–85).
33. Li, J., Chen, B. M., & Lee, G. (2018). SO-Net: Self-Organizing Network for Point Cloud Analysis. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 9397–9406).
34. Weiss, U., Biber, P., Laible, S., Bohlmann, K., & Zell, A. (2010). Plant species classification using a 3D LIDAR sensor and machine learning. *9th International Conference on Machine Learning and Applications, ICMLA 2010* (pp. 339–345).

35. Yao, X., Guo, J., Hu, J., & Cao, Q. (2019). Using deep learning in semantic classification for point cloud data. *IEEE Access*.
36. Briechele, S., Krzystek, P., & Vosselman, G. (2019). Semantic labeling of als point clouds for tree species mapping using the deep neural network PointNet++. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*, 42(2/W13), 951–955.
37. Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Proceedings Advances in Neural Information Processing Systems*.
38. Landrieu, L., & Simonovsky, M. (2018). Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City.
39. Yang, M., & Förstner, W. (2010). Plane Detection in Point Cloud Data. *Proceedings of the 2nd International Conference on Machine Control Guidance: Vol. 1* (pp. 95–104). Bonn.
40. Li, C., Zaheer, M., Zhang, Y., Poczos, B., & Salakhutdinov, R. R. (2019). Point cloud GAN. *Deep Generative Models for Highly Structured Data, DGS@ICLR 2019 Workshop*.
41. Marulanda, F. G., Libin, P., Verstraeten, T., & Nowé, A. (2018). IPC-Net: 3D point-cloud segmentation using deep inter-point convolutional layers. *International Conference on Tools with Artificial Intelligence, ICTAI* (pp. 293–301).
42. Uy, M. A., & Lee, G. (2018). PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 4470–4479).
43. Zhou, Y., & Tuzel, O. (2018). VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 4490–4499).
44. Yi, L., Zhao, W., Wang, H., Sung, M., & Guibas, L. J. (2019). GSPN: Generative shape proposal network for 3D instance segmentation in point cloud. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 3942–3951).
45. Yang, Y., Feng, C., Shen, Y., & Tian, D. (2018). FoldingNet: Point Cloud Auto-Encoder via Deep Grid Deformation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 206–215).
46. Yotsumata, T., Sakamoto, M., & Satoh, T. (2020). Quality improvement for airborne lidar data filtering based on deep learning method. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2020*, 355–360.
47. Gülch, E., & Obrock, L. S. (2020). Automated semantic modelling of building interiors from images and derived point clouds based on deep learning methods. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2020*, 421–426.
48. Tchapmi, L. P., Choy, C. B., Armeni, I., Gwak, J. Y., & Savarese, S. (2017). Segcloud: Semantic segmentation of 3d point clouds. *International Conference on 3D Vision (3DV)*.
49. Dai, A., Chang, A. X., Savva, M., Halbe, M., Funkhouser, T., & Niener, M. (2017). Scannet: Richly annotated 3D reconstructions of indoor scenes. *Proceedings Computer Vision and Pattern Recognition (CVPR)*.
50. Riegler, G., Ulusoy, A. O., & Geiger, A. (2017). Oct-net: Learning deep 3d representations at high resolutions. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
51. Su, H., Jampani, V., Sun, D., Maji, S., Kalogerakis, E., Yang, M., & Kautz, J. (2018). SPLATNet: Sparse Lattice Networks for Point Cloud Processing. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 2530–2539).
52. Graham, B. (2015). Sparse 3d convolutional neural networks. *British Machine Vision Conference*.
53. Adams, A., Baek, J., & Davis, M. A. (2010). Fast high-dimensional filtering using the permutohedral lattice. *In proceedings Computer Graphics Forum*, 28, 753–762.
54. Hermosilla, P., Ritschel, T., Vazquez, P. P., Vinacua, A., & Ropinski T. (2018). Monte-Carlo convolution for learning on non-uniformly sampled point clouds. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2018)*.
55. Li, Y., Rui, B., Mungchao, S., Wei, W., Xinhan, D., & Baoquan, C. (2018). PointCNN: Convolution On X-Transformed Points. *NeurIPS 2018*.
56. Pan, H., Liu, S., Liu, Y., & Tong, X. (2018). Convolutional neural networks on 3D surfaces using parallel frames. *Arxiv preprint*, arXiv: 1808.04952.
57. Tatarchenko, M., Park, J., Koltun, V., & Zhou, Q. (2018). Tangent convolutions for dense prediction in 3D. *CVPR*.

58. Gall, Y. L., Thomas, H., Goulette, F., Deschaud, J., & Marcotegui, B. (2018). Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods. *2018 International Conference on 3D Vision (3DV)*. Verone.
59. Choy, C. B., Xu, D., Gwak, J., Chen, K., & Savarese, S. (2018). 3D-r2n2: A unified approach for single and multi-view 3D object reconstruction. *In Proceedings of the European Conference on Computer Vision (ECCV)*.
60. Zeng, A., Song, S., Niener, M., Fisher, M., Xiao J., & Funkhouser, T. (2017). 3Dmatch: Learning the matching of local 3D geometry in range scans. *CVPR*. 2017.
61. Zhao, Y., Li, X., Huang, H., Zhang, W., Zhao, S., Makkie, M., Zhang, M., Li, Q., & Liu, T. (2019). 4D Modeling of fMRI Data via Spatio-Temporal Convolutional Neural Networks (ST-CNN). *IEEE Transactions on Cognitive and Developmental Systems*.
62. Choy, C., Gwak, J., & Savarese, S. (2019). 4D Spatio-Temporal ConvNet: Minkowski Convolutional Neural Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
63. Abrosimov, M. A., & Brovko, A. V. (2017) A method for normalizing a point cloud to be processed using an artificial neural network. In *Informatsionno-kommunikatsionnye tekhnologii v nauke, proizvodstve i obrazovanii ICIT-2017 [Information and communication technologies in science, production and education ICIT-2017]* (pp. 262–268). Voronezh [in Russian].
64. Babaev, A. M. (2019). Neural Network Technologies for 3D Object Recognition. *Mezhdunarodnyi zhurnal gumanitarnykh i estestvennykh nauk [International Journal of Humanities and Natural Sciences]*, 39(12-2), 74–76 [in Russian]
65. Kazdorf, S. Ya., & Pershina, Zh. S. (2019). Semantic Segmentation Algorithm for 3D Scenes. *Cloud of Science*, 6(3), 451–461 [in Russian].
66. Aoki, Y., Goforth, H., Srivatsan, R. A., & Lucey, S. (2019). Pointnetlk: Robust & efficient point cloud registration using PointNet. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 7156–7165).
67. Zaganidis, A., Sun, L., Duckett, T., & Cielniak, G. (2018). Integrating Deep Semantic Segmentation into 3-D Point Cloud Registration. *IEEE Robotics and Automation Letters*, 3(4), 2942–2949.
68. Neidhart, H., & Sester, M. (2003). Identifying building types and building clusters using 3D-laser scanning and GIS-data. *Machine Learning*.
69. Zhang, B., Huang, S., Shen, W., & Wei, Z. (2019). Explaining the PointNet: What Has Been Learned Inside the PointNet? *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 71–74).

Author details

Aleksey A. Kolesnikov – Ph. D., Associate Professor, Department of Cartography and Geoinformatics.

Received 14.03.2022

© A. A. Kolesnikov, 2022