**Модификация нейронной сети U-Net   
в задаче сегментации многоканальных спутниковых снимков**

**Аннотация** – в данной статье представлены результаты работы сверточной нейронной сети для сегментации спутниковых четырехканальных снимков. Входные изображения содержат синий, зеленый, красный и ближний инфракрасный спектры. Алгоритм предназначен для обнаружения зданий. Была использована модификация нейронной сети U-Net с двумя кодировщиками. Значения коэффициента Соренсена и индекса Жакара были рассчитаны для 16 различных городских районов.

**Ключевые слова** – сегментация изображений, спутниковые снимки, сверточная нейронная сеть, глубокое обучение.

Введение

Спутник – это объект, который вращается на околоземной орбите для выполнения задач по получению глобальной информации о Земле, военной разведки, обеспечения спутниковой связи, метеорологических наблюдений и т.д. В настоящее время существует множество спутников, способных делать снимки высокого разрешения поверхности Земли. В таблице 1 представлены некоторые из них.

Таблица 1. Некоторые известные спутники

|  |  |  |
| --- | --- | --- |
| **Модель спутника** | **Год запуска** | **Стоимость снимков** |
| Landsat 8 | 2013 | Бесплатно |
| Sentinel | 2014 | Бесплатно |
| MODIS | 2002 | Бесплатно |
| WorldView | 2016 | Платно |
| QuickBird | 2001 | Платно |
| GeoEye-1 | 2008 | Платно |
| IKONOS | 1999 | Платно |
| Jilin-1 | 2015 | Платно |
| SPOT-6 | 2012 | Платно |
| Gaofen-2 | 2014 | Платно |
| TripleSat | 2015 | Платно |

Получение спутниковых снимков Земли включает в себя следующие особенности:

* Спутники дважды фотографируют одну и ту же территорию. Первое изображение получается черно-белым, детализированным, второе – цветным, с разрешением ниже среднего. Невозможно с первой попытки сделать цветные фотографии с высоким разрешением, потому что свет преломляется в атмосфере Земли. Кроме того, цветной снимок необходимо растянуть, поскольку в цифровом виде изображение имеет меньший размер, после чего оба снимка объединяются в единую фотографию – итоговый результат дистанционного зондирования данной территории.
* Космические камеры различают цвета по-разному, поэтому исходные спутниковые изображения не выглядят как естественные фотографии из-за дифракции и рассеяния света в атмосфере Земли. Чтобы сделать цвета нормальными для человеческого восприятия, необходимо выполнить их коррекцию.
* Условия съемки и тип камеры вызывают эффект смещения. В этом случае необходимо реализовать устранение искажений исходного изображения.

Спутники делают тысячи снимков каждый день, в связи с чем разработка методов их автоматической обработки очень актуальна. Современные спутники способны делать фотографии с пространственным разрешением не более 3 м/пиксель, благодаря чему на них можно обнаружить такие небольшие объекты, как здания, свалки и т.д. Развитие технологий позволило применять методы глубокого обучения для сегментации спутниковых изображений [1].

В данной статье представлен метод сегментации спутниковых снимков на основе сверточной нейронной сети. Такие сети способны в режиме реального времени обнаруживать и классифицировать объекты. Сверточные нейронные сети показали свою эффективность в различных задачах компьютерного зрения [2].

Автоматическая сегментация является важной частью предварительной обработки аэрофотоснимков. На сегодняшний день сверточные нейронные сети широко используются для решения этой задачи. В частности, алгоритм U-Net показал свою эффективность в сегментации медицинских изображений [3] и спутниковых снимков [4]. Основным преимуществом этой архитектуры является то, что U-Net может показывать хорошие результаты даже при обучении на небольшом наборе данных. Структура Г-Туе параллельна, поэтому графические процессоры являются идеальным инструментом для работы с ней [5].

Существуют некоторые требования к сегментации спутниковых изображений [6]:

* Размер и тип зданий могут существенно варьироваться.

Алгоритм должен обнаруживать объекты любого размера. Использование нескольких кодировщиков в структуре нейронной сети может решить эту проблему.

* Разделение объектов с высокой плотностью локации.

Алгоритм должен быть оштрафован за плохое разделение объектов друг от друга во время обучения, чтобы улучшить качество выходной маски. Это достигается путем тщательного выбора функции потерь.

* Обученная модель должна быть инвариантна к поворотам.

Эта проблема может быть решена путем увеличения числа обучающих данных с помощью аугментации.

* Аэрофотоснимки имеют разное пространственное разрешение.

Важно создать алгоритм, который способен детектировать объекты на любых спутниковых снимках.

* Алгоритмы должны быть устойчивы к шуму.

Спутниковые снимки снимаются в разных погодных ситуациях. На некоторых фотографиях может присутствовать шум: дымка, туман, блики от отражающих поверхностей.

В данной статье представлены результаты обучения сверточной нейронной сети для сегментации многоканальных спутниковых снимков с целью обнаружения на них зданий. Для решения поставленной задачи предлагается использовать модификацию архитектуры U-Net с использованием двух кодировщиков. Данное исследование продолжает исследования, приведенные в статьях [7,8].

Оставшаяся часть статьи организована следующим образом. Второй раздел посвящен некоторым доступным базам данных спутниковых снимков. Архитектура разработанной нейронной сети приведена в третьей части. В четвертом разделе представлены результаты численных экспериментов на большой базе спутниковых снимков различных городских районов. И наконец, в заключении подводятся итоги исследования.

Базы спутниковых снимков

Разработанная нейронная сеть была предварительно обучена на изображениях базы данных Spacenet. Для создания 11-битных восьмиканальных фотографий использовались спутники WorldView-2 и WorldView-3. База данных Spacenet содержит различные наборы снимков с уже выделенными зданиями [9]. Для предварительной подготовки модели нами использовался набор снимков г. Хартум, Судан. Примеры изображений из базы данных Spacenet показаны на рис. 1.

После этого разработанная модель прошла дообучение на изображениях 16 различных регионов Российской Федерации приватной базы данных. Каждое такое изображение имеет черно-белую маску, размеченную экспертами вручную. Этот набор данных охватывает около 30 квадратных километров. Четырехканальные снимки приватной базы данных с пространственным разрешением 3 м/пиксель содержат синий, зеленый, красный и ближний инфракрасный (NIR) цветовые каналы.



Рис. 1. Примеры изображений из базы данных SpaceNet

Разработанный алгоритм машинного обучения требует на вход изображения размером 256 × 256 пикселей. Для обучения сверточной нейронной сети каждый снимок и соответствующая ему маска разрезалась на две непересекающиеся полосы, каждая из которых была разделена на участки размером 256 × 256 пикселей с шагом в 128 пикселей.

Для увеличения обучающего набора были применены три типа преобразований:

* Поворот на 90, 180, 270 градусов и зеркальное отражение. В результате, обучающий набор данных увеличился в 8 раз.
* Хроматические искажения. Изображения были переведены из цветового пространства RGB в цветовое пространство HSV. К координатам HSV также были добавлены небольшие случайные величины. Для ближнего инфракрасного канала вместо хроматического искажения к значению канала добавлялись случайные величины из интервала [-0,06, +0,06], после чего производилась нормализация канала.
* Случайные сдвиги, масштабы и повороты на малый угол.

В результате таких преобразований, обучающий набор данных содержал 9784 батча, каждый из которых состоял из 16 изображений размером 256 × 256 пикселей.

Сверточная нейронная сеть

В данной статье представлена модификация известной архитектуры U-Net. Ее классическая структура описана в статье [3].

Первоначальная структура сети U-Net была изменена: для каналов RGB и NIR использовались два разных кодировщика (рис. 2). Выходы кодировщиков объединялись с соответствующими слоями декодировщика. В итоге модифицированная нейронная сеть имела 38 сверточных слоев, 37 функций активации ReLU, 37 операций пакетной нормализации, 1 функцию активации сигмоид, 10 операций понижения дискретизации, 5 операций повышения дискретизации и 11 операций слияния.



Рис. 2. Архитектура модифицированной сети U-Net с двумя кодировщиками

Численные результаты

Качество алгоритмов сегментации оценивалось с помощью коэффициента Дайса () и индекса Жакара (). Эти величины рассчитывались по следующим формулам

 

 

 

 (4)

где и - значения пикселей экспертной разметки и предсказаний соответственно.

Во время предварительного обучения модели на наборе данных Spacenet коэффициент Дайса достиг значения 0,84, а индекс Жакара 0,77. Значения коэффициента Дайса и индекса Жакара для спутниковых снимков 16 регионов Российской Федерации для модифицированной сети U-Net с предварительным обучением (pre-trained model) и без него (without pre-training) показаны на рис. 3.

Результаты сегментации спутникового изображения из приватной базы данных показаны на рис. 5. Для него коэффициент Дайса и индекс Жакара оказались равными 0,78 и 0,67 соответственно.



Рис. 3. Значения коэффициента Дайса и индекса Жакара для разработанной модели с предварительным обучениеми без него



Рис. 4. Результаты сегментации спутникового изображения из приватной базы данных

Заключение

В статье были продемонстрированы результаты обучения и тестирования сверточной нейронной сети, предназначенной для сегментации спутниковых изображений. Для работы с четырехканальными снимками была разработана модифицированная архитектура U-Net с двумя кодировщиками. Алгоритм глубокого обучения был предварительно обучен на наборе данных Spacenet. Значения коэффициента Дайса и индекса Жаккара на данных из приватной базы были равными 0,78 и 0,67 соответственно.

Будущие направления исследования могут включать в себя:

* Категоризацию зданий по размеру и площади. Для различных типов зданий могут быть применены различные кодировщики.
* Использование границ объектов в качестве дополнительной опции для их детектирования.

Разработанный алгоритм глубокого обучения может быть использован для оценки уровня урбанизации различных регионов и отслеживания строительства крупных объектов.

Благодарности

Статья была подготовлена при финансовой поддержке Министерства образования Российской Федерации в рамках научного проекта № 14.575.21.0167 (идентификатор RFMEFI57517X0167).

Литература

1. Zhang, Liangpei, Lefei Zhang, and Bo Du. “Deep learning for remote sensing data: A technical tutorial on the state of the art.” IEEE Geoscience and Remote Sensing Magazine (2016): 22-40.
2. S. Seferbekov, V. Iglovikov, A. Buslaev, A. Shvets. Feature Pyramid Network for Multi-Class Land Segmentation. Web: https://arxiv.org/pdf/1806.03510.pdf.
3. O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, vol. 9351, 2015, pp. 234–241.
4. Zhang, Z., Liu, Q., & Wang, Y. (2018). “Road extraction by deep residual u-net”. IEEE Geoscience and Remote Sensing Letters, 15(5), pp. 749-753.
5. A. Gulli, S. Pal., *Deep Learning with Keras*, Packt Publishing, 2017, 320 p.
6. X. X. Zhu et al., “Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources,” in IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, Dec. 2017.
7. V. V. Khryashchev, V. A. Pavlov, A. Priorov and A. A. Ostrovskaya, "Deep Learning for Region Detection in High-Resolution Aerial Images," 2018 IEEE East-West Design & Test Symposium (EWDTS), Kazan, 2018, pp. 1-5.
8. L. Ivanovsky, V. Khryashchev, V. Pavlov and A. Ostrovskaya, "Building Detection on Aerial Images Using U-NET Neural Networks," 2019 24th Conference of Open Innovations Association (FRUCT), Moscow, Russia, 2019, pp. 116-122.
9. SpaceNet Database, Web: http://explore.digitalglobe.com/spacenet