Выделение горнодобывающих карьеров на спутниковых снимках   
с помощью алгоритмов глубокого обучения

**Анна Островская 1), Владимир Хрящев 2), Александр Семенов 2)**

1) Российский университет дружбы народов (РУДН), 117198 Москва, Россия, ул. Миклухо-Маклая д. 6, e-mail: leon.ivanovsky@yahoo.com ostrovskaya\_aa@rudn.university

2) Ярославский государственный университет им. П.Г. Демидова, 150003 Ярославль, Россия,   
ул. Советская д. 14, email: v.khryashchev@uniyar.ac.ru, agsedov@gmail.com

***Аннотация:*** *Целью нашего исследования была разработка алгоритма глубокого обучения для автоматического выделения горнодобывающих карьеров на спутниковых снимках. Для анализа качества разработанных алгоритмов использовался коэффициент подобия Дайса, сопоставляющий друг с другом полученные алгоритмом результаты и области, ранее размеченные экспертами. В статье было показано, как сверточные нейронные сети, реализованные на современных графических процессорах, могут использоваться для обнаружения объектов.*

***Ключевые слова*:** компьютерное зрение, глубокое обучение, сегментация спутниковых снимков, детектирование горнодобывающих карьеров.

# 1. ВВЕДЕНИЕ

На сегодняшний день проблема детектирования объектов на спутниковых снимках высокого разрешения находится в фокусе научных исследований. Автоматическая сегментация изображений позволяет выделять интересующие области на снимке. Большинство подходов по решению данной проблемы предполагают разработку алгоритмов глубокого машинного обучения.

В машинном обучении задача сегментации изображений обычно переформулируется как задача классификации на уровне пикселей. Наиболее простой (и медленный) подход к решению этой проблемы – ручная сегментация изображений. Тем не менее, это трудоемкий и долгий процесс, который подвержен ошибкам. Автоматическая сегментация позволяет обеспечить обработку изображений «на лету» сразу его после получения. Сегментация спутниковых снимков находит свое применение в области городского планирования, рекультивизации заброшенных объектов, для контроля вырубки лесов и т.д.

В данной статье представлены результаты исследования работы сверточной нейронной сети, способной обнаруживать и классифицировать объекты в режиме реального времени. Особенность таких сетей заключается в том, что дескрипторы изображений формируются самостоятельно, в процессе обучения. Подобного рода алгоритм превосходит по качеству традиционные методы машинного обучения. Структура сверточной нейронной сети параллельна и как следствие она отлично подходит для обучения на графических процессорах, которые состоят из тысячи ядер, которые вместе способны решать множество задач одновременно [1]. Сверточные нейронные сети стали повсеместно использоваться в компьютерном зрении с тех пор, как алгоритм AlexNet [2] победил в конкурсе ImageNet Challenge: ILSVRC 2012 [3]. В последние годы были предложены различные алгоритмы глубокого обучения, нацеленные на сегментацию изображений.

Один из самых успешных алгоритмов сегментации основан на использовании пирамидальных нейронных сетях. Пирамидальные сети показали приемлемые результаты детектирования объектов на спутниковых снимках [4]. Использование пирамидальных сетей позволило получить значение индекса Жаккара равное 0.49 для сегментации изображений со спутника DeepGlobe [5].

В статье [6] представлена архитектура U-Net - особый тип пирамидальных сетей, который получила свое развитие при решении задачи сегментации биомедицинских изображений. Позже эта модель была применена для классификации спутниковых изображений на уровне пикселей [7]. Архитектура U-Net использует пропускаемые соединения для объединения признаков, полученных на разных этапах работы сети. Используя архитектуру U-Net, авторы статьи [8] получили значение коэффициента Дайса равное 0,75 для детектирования зданий на спутниковых снимках.

Эта статья состоит из шести частей. Первая часть посвящена использованию сверточных нейронных сетей в машинном обучении и особенностям задачи сегментации спутниковых снимков. В ней также приведен обзор работ, связанных с задачей детектирования объектов на изображениях высокого разрешения. Второй раздел посвящен доступным базам данных спутниковых снимков. В третьей части описываются разработанная архитектура сверточной нейронной сети для выделения горнодобывающих карьеров на спутниковых снимках и некоторые особенности их обучения. В четвертой части представлены результаты численных экспериментов для разработанного алгоритма. В заключении подводятся итоги исследования. И наконец, в конце содержится список литературы.

# 2. БАЗЫ СПУТНИКОВЫХ СНИМКОВ

Базы изображений являются важнейшей составляющей успешного обучения и оценки качества работы алгоритмов. На сегодняшний день существует несколько доступных баз спутниковых снимков.

База данных IKONOS [10] содержит 11-битные и 8-битные цветные изображения в формате GEOTIFF со спутника IKONOS. Каждое изображение базы данных IKONOS имеет пространственное разрешение 0,82 м/пиксель. Снимки были сняты с углом наклона от земли

# Inria_mini.png

**Рис. 1. – Примеры изображений из базы данных IKONOS**

# C:\Users\User\Downloads\DSTL.png

**Рис. 2. – Примеры изображений из базы данных DSTL**

# C:\Users\User\Downloads\DSTL.png

**Рис. 3. – Примеры изображений из базы данных DSTL**

до солнца не более 15° и с облачностью менее 15%. Примеры изображений из базы данных IKONOS показаны на рис. 1.

Набор данных DSTL содержит 50 спутниковых снимков в формате GEOTIFF размером Впервые эта база данных была представлена в конкурсе Kaggle «DSTL Satellite Imagery Feature Detection» [13]. Изображения этой базы данных разбиты на 10 различных классов: «здания», «искусственные сооружения», «дороги», «трассы», «деревья», «сельскохозяйственные угодья», «водные пути», «стоячая вода», «большие транспортные средства» (например грузовые автомобили или автобусы) и «небольшие транспортные средства» (автомобили, микроавтобусы или велосипеды). Примеры из базы изображений DSTL показаны на рис. 3.

Для нашего исследования были использованы 14 цветных спутниковых снимка из приватной базы данных российских городов. Каждый снимок размером порядка 8192x8192 пикселей имеет пространственное разрешение 0.5 м/пиксель. Изображения этой базы данных покрывают территорию трех российских городов: Москвы, Ярославля и Рыбинска. Примеры снимков из этой базы показаны на рис. 4.

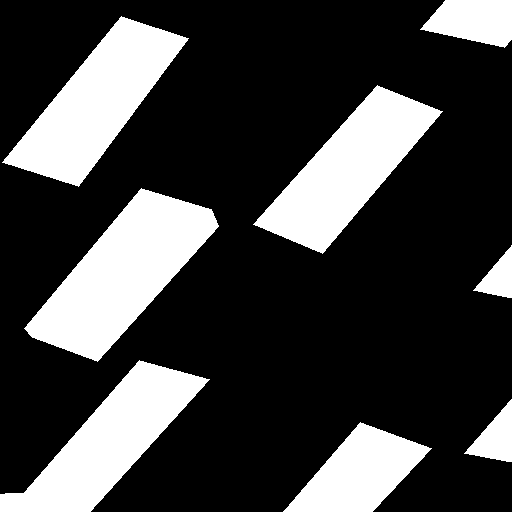
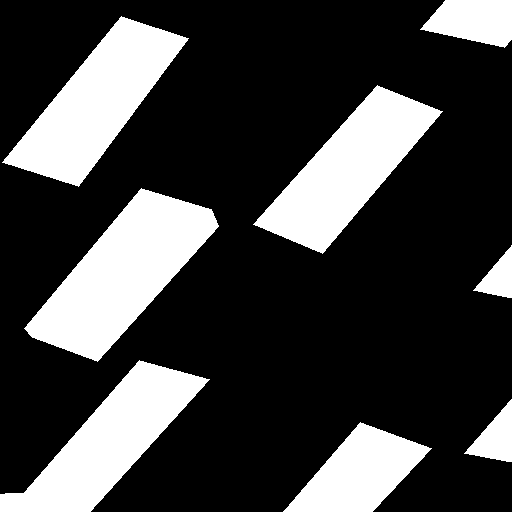
# 3. СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

В данной статье представлена разработанная глубокая сверточная сеть U-Net [7]. Это специальная архитектура, направленная на быстрое и качественное обнаружение выделение различных объектов на снимках высокого разрешения [14]. Исследование, приведенное в данной работе, продолжает исследования, представленное в статьях [15, 16].

Все разработанные сверточные нейронные сети были реализованы с помощью библиотеки Tensorflow. TensorFlow - это программная библиотека с открытым исходным кодом для высокопроизводительных

# U-Net.png

**Рис. 1. – Архитектура сети U-Net**

**Рис. 7. – Примеры нарезанных изображений и соответствующих им масок**

численных расчетов. Этот фреймворк используется также в области машинного обучения, например для создания нейронных сетей. Он предоставляет множество реализаций широко используемых структурных блоков нейронной сети, а также готовых инструментов для предварительной обработки изображений [17].

Как показано на рис.5, U-Net состоит из двух частей: кодировщика (слева) и декодировщика (справа). Кодировщик представляет собой нейронную сеть с типичной архитектурой сверточной нейронной сети, состоящей из четырех блоков. Каждый такой блок состоит из двух сверточных слоев (conv) с фильтром 3 × 3, с примененной к каждому из них функции активации ReLU, а также операции понижения дискретизации с помощью оператора maxpooling с размером фильтра 2 × 2 и шагом 2. Декодировщик содержит такое же количество блоков, как и кодировщик. Каждый блок декодировщика состоит из операций повышения дискретизации с помощью оператора upsampling с размером фильтра 2 × 2 и объединения (merging) с соответствующим набором признаков от кодировщика, двух сверточных слоев с фильтром 3 × 3 и примененной к каждому из них функции активации ReLU. Последний слой сети выполняет операцию свертки с фильтром размера 1x1 для соотносения каждого пикселя определенному классу. В итоге, сеть имеет 19 сверточных слоев, 18 функций активации ReLU, 4 операций понижения дискретизации, 4 операций upsampling и 4 операций слияния признаков.

В качестве алгоритма численной оптимизации был выбран оптимизатор адаптивной оценки моментов (Adam). Этот оптимизирующая функция объединяет лучшие подходы от градиентного спуска и импульсных оптимизаторов, показывая быструю сходимость для большинства задач машинного обучения [19]. В качестве функции потерь была выбрана бинарная кросс-энтропия. Классификатор заканчивал свое обучение после выполнения 60 эпох.

# 4. РЕЗУЛЬТАТЫ

Обучение и тестирование разработанных моделей было осуществлено на спутниковых снимках из приватной базы данных. Поскольку изначально спутниковые снимки выглядели очень темными, для каждого из них выполнялись нормализация цветовых каналов и эквивализация гистограммы яркости.

Информация о местоположении карьеров была получена из json-файлов и сгенерирована в виде черно-белых масок, в которых каждый пиксель, относящийся к карьерам, окрашивался в белый цвет. Примеры спутниковых снимков и соответствующих им масок показаны на рис. 7.

Традиционный подход при сегментации изображений основан на использовании частей снимков, которые соответствуют входу сверточной нейронной сети. Разработанные алгоритмы требуют на вход данные размером 512 × 512 пикселей, поэтому перед обучением моделей, каждый спутниковый снимок и соответствующая ему маска нарезались на части соответствующего размера с помощью метода скользящего окна с шагом 256. Подготовленная выборка данных содержала 2063 картинки размером 512 × 512 пикселей. Для поставленной задачи брались только 2 класса: «карьеры» и «не карьеры».

Разработанные сверточные нейронные сети были обучены и протестированы на суперкомьютере NVIDIA DGX-1, предоставленным Центром искусственного интеллекта и цифровой экономики ЯрГУ им. П.Г. Демидова.

Как правило, качество работы алгоритмов сегментации обычно оценивается специальными метриками, осуществляющими сравнение предсказанных и экспертных масок изображений между собой. Для оценки разработанного алгоритма использовался коэффициент подобия Дайса (dice). Этот показатель принимает значения из отрезка [0, 1] и фактически является бинарной мерой сходства двух множеств. Коэффициент Серенсена рассчитывается по следующей формуле:

где - мощность пересечения, а - сумма мощностей экспертной разметки и предсказаний [19]. Для поставленной задачи числитель и знаменатель можно вычислить по следующим формулам

, ,

где - значения вероятностей принадлежности пикселей к тому или иному классу экспертной разметки и предсказаний соответственно. Зависимость значения функции потерь от числа выполненных эпох обучения для разработанного алгоритма глубокого обучения показана на рис. 8.

# 5. ЗАКЛЮЧЕНИЕ

Для выделения горнодобывающих карьеров на изображениях высокого разрешения была разработана глубокая сверточная сеть U-Net. Это специальная архитектура, направленная на быструю и качественную сегментацию спутниковых снимков.

Численные эксперименты для разработанной модели проводились на изображениях из приватной базы данных. Для оценки разработанного алгоритма использовался коэффициент подобия Дайса.

Статья была подготовлена при финансовой поддержке Министерства образования Российской Федерации в рамках научного проекта № 14.575.21.0167 (идентификатор RFMEFI57517X0167). Авторы также благодарны Центру Искусственного Интеллекта Ярославского государственного университета им. П.Г. Демидова за обеспечение доступа к суперкомпьютеру NVIDIA DGX-1.

# 6. СПИСОК ЛИТЕРАТУРЫ

[1] Я. Гудфеллоу, И. Бенджио, А. Курвилль, *Глубокое обучение*. М.: ДМК Пресс, 2017, 652 с.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *In Advances in neural information processing systems*, 2012, pp. 1097–1105.

[3] ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC2012), Web: http://image-net.org/challenges/LSVRC/2012/.

[4] S. Seferbekov, V. Iglovikov, A. Buslaev, A. Shvets. Feature Pyramid Network for Multi-Class Land Segmentation. Web: https://arxiv.org/pdf/1806.03510.pdf.

[5] DeepGlobe. CVPR 2018 – Satellite Challenge, Web: http://deepglobe.org.

[6] 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. 234–341.

[7] G. Chhor, C. Bartolome Aramburu, I. Bougdal-Lambert. Satellite Image Segmentation for Building Detection using U-net. Web: http://cs229.stanford.edu/proj2017/final-reports/5243715.pdf. [8] A. Chaurasia, E. Culurciello. LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. Web: https://arxiv.org/pdf/1707.03718.pdf.

[9] The Cambridge-driving Labeled Video Database (CamVid), Web: http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/.

[10] IKONOS Satellite Images, Web: https://www.satimagingcorp.com/gallery/ikonos/.

[11] QuickBird 60cm Global High-Resolution Satellite Imagery, Web: http://www.landinfo.com/QuickBird.htm.

[12] DigitalGlobe satellite, Web: https://www.digitalglobe.com.

[13] DSTL Satellite Imagery Feature Detection, Web: https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection.

[14] Дж. Вандер Плас, *Python для сложных задач: наука о данных и машинное обучение*. СПб.: Питер, 2018. 576 с.

[15] Khryashchev V.V., Priorov A.L., Pavlov V.A., Ostrovskaya A.A. Deep learning for region detection in high-resolution aerial images // Proceedings of 16-th IEEE East-West Design & Test Symposium (EWDTS’2016), Kazan, Russia, September 14 - 17, 2018. P. 792-796.

[16] V. Khryashchev, L. Ivanovsky, V. Pavlov, A. Ostrovskaya and A. Rubtsov, "Comparison of Different Convolutional Neural Network Architectures for Satellite Image Segmentation. *Proceedings of the 23rd Conference of Open Innovations Association FRUCT’23, Bologna, Italy.* 2018, pp. 172-179.

[17] А. Джулли, С. Пал, *Библиотека Keras – инструмент глубокого обучения*. М.: ДМК Пресс, 2018, 294 с.

[18] Н. Шакла, Машинное обучение & Tensorflow. СПб: Питер, 2019, 336 с.

[19] D. P. Kingma, J. Ba. Adam: A Method for Stochastic Optimization. Web: https://arxiv.org/abs/1412.6980.