

НЕЙРОННАЯ СЕТЬ U-NET В ЗАДАЧАХ РАСПОЗНАВАНИЯ КЛАССОВ ЗЕМНОЙ ПОВЕРХНОСТИ И ОБЛАКОВ

Н.А. Бирюков^{1*}, М.П. Лапчинская¹

¹ Московский государственный университет геодезии и картографии, Москва, Россия

* 4096414@gmail.com

Цитирование: Бирюков А.Н., Лапчинская М.П. Нейронная сеть U-Net в задачах распознавания классов земной поверхности и облаков // Известия вузов «Геодезия и аэрофотосъемка». 2021. Т. 65. № 6. С. 673-679. DOI:10.30533/0536-101X-2021-65-6-673-679

Ключевые слова. геоинформационные системы, машинное обучение, распознавание образов, нейронные сети, цифровая обработка снимков, семантическая сегментация.

Аннотация

В настоящее время нейронные сети активно внедряются в геоинформационные системы для автоматизации процессов дешифрирования и проведения анализа с использованием методов машинного обучения. Использование методов нейронных сетей для задач распознавания классов на космо- и аэрофотоснимках обусловлено постоянно растущим количеством данных ДЗЗ и аэрофотоснимков, позволяющим проводить обучение и тестировать результирующие модели. В статье рассматриваются архитектуры нейронных сетей, используемых в ГИС, и обосновывается метод нейронных сетей для автоматизации распознавания классов земной поверхности и облаков, являющийся оптимальным на космо- и аэрофотоснимках. В статье сравниваются результаты распознавания трёх архитектур U-Net: U-Net с метрикой оценки “точность”, U-Net с метрикой “коэффициент Жаккара” и U-Net с метрикой “коэффициент Дайса”. Для сравнения используются разные массивы цифровых изображений и распознаваемые классы: спутниковые снимки WorldView 3 с распознаваемым классом “лесные массивы”, аэрофотоснимки взволнованной морской поверхности с классом “рябь на ряби” и аэрофотоснимки облачности с классом “облака”. При сравнении нейронных сетей, используемых в ГИС, выявлены особенности, влияющие на возможность их использования в задачах распознавания классов земной поверхности и облаков: архитектура RetinaNet не позволяет отразить сложную метрику пространственных объектов, а Mask R-CNN затрачивает больше времени на обработку данных при обучении, чем U-Net. Результаты распознавания моделей U-Net интерпретированы и представлены в виде таблицы. Приведённое в статье исследование показывает, что U-Net с метрикой “точность” обладает наибольшей обучаемостью и способна распознавать все рассматриваемые классы, хоть и с наличием зернистости. В то время как модели U-Net с метриками “коэффициент Жаккара” и “коэффициент Дайса” показывают лучшие результаты при распознавании класса “облака”.

1 Введение

Развитие методов машинного обучения, в частности методов нейронных сетей, и их внедрение в различные системы связано с постоянно растущим количеством информации [1] и увеличением вычислительных мощностей пользовательских устройств (компьютеров, ноутбуков и др.).

Применительно к методам нейронных сетей для автоматизации задач, решаемых с помощью геоинформационных систем, большую роль играет рост количества данных ДЗЗ и аэрофотоснимков.

В данной статье представлено исследование и описание оптимального метода нейронных сетей, используемого в геоинформационных системах, для задач распознавания классов земной поверхности и облаков, позволяющего в дальнейшем автоматизировать процесс векторизации.

2 Материалы и методы

2.1 Спецификация данных для машинного обучения в ГИС

Для распознавания объектов на космо- и аэрофотоснимках при машинном обучении с учителем требуются некоторый образец (*англ.* – sample) и эталон (*англ.* – truth). В качестве sample выступают аэрофотоснимки и данные ДЗЗ главным образом в формате GeoTIFF и пространственные объекты.

Качество снимков зависит во многом от спутника или носителя и используемой съёмочной аппаратуры. Каждый спутник, ведущий съёмку на орбите Земли, имеет разное количество каналов, которые могут быть задействованы при обучении, а каждая съёмочная аппаратура – свои съёмочные характеристики, влияющие на итоговое изображение. Пространственные объекты представлены двумя составляющими: метрикой или координатным описанием и семантикой – смысловым содержанием. Эти данные подготавливаются в ГИС, таких как КБ Панорама, ArcGIS, QGIS и могут быть представлены для обработки и обучения в виде файлов различного разрешения (*.csv, *.shp, *.geojson и т.д.).

При выполнении процесса векторизации используются космо- и аэрофотоснимки, представленные в цифровой форме. В таком случае образцами являются снимки, представленные в виде n двумерных массивов со значениями яркости пикселей, где n – используемое количество каналов снимка. Истинные значения могут быть представлены в цифровой форме как растровые данные, представляющие собой изображение того же разрешения, что и образец, каждый пиксель которого имеет определённое значение:

- 0 или 255, преобразуемые путём бинаризации в 0 или 1 для двух классов, где вторым классом являются все объекты, не входящие в первый класс;

- фиксируемые пользователем значения от 0 до 255 (возможно использование трёх каналов для получения RGB изображения маски с целью лучшего восприятия), преобразуемые в соответствующие классам значения для трёх и более классов.

При машинном обучении без учителя используются только образцы.

2.2 Методы нейронных сетей

2.2.1 Методы нейронных сетей в геоинформационных системах

Методы нейронных сетей основываются на построении алгоритмов, имитирующих деятельность нейронов человеческого мозга при решении конкретных задач [2,3]. В ГИС используются глубокие нейронные сети, показавшие высокую точность результатов предсказаний в других областях науки и производства [4].

При использовании методов нейронных сетей для задач распознавания имеют значения результаты, получаемые обученными моделями сетей и их пригодность для такого рода задач.

На текущий момент времени в ГИС наиболее популярны для решения задач распознавания и детектирования объектов: U-Net, MaskR-CNN, RetinaNet. Рассмотрим данные нейронные сети с точки зрения получаемых результатов:

RetinaNet – свёрточная нейронная сеть для детектирования объектов.

Результатом, получаемым после обучения нейронной сети, является набор якорных блоков (якорных рамок) указанных пользователем классов, что позволяет решать с помощью данной нейронной сети задачи распознавания объектов на космо- и аэрофотоснимках [5].

Нейронная сеть **Mask R-CNN** реализована в ArcGIS PRO для объектной сегментации.

Результатом предсказаний модели, получаемой при обучении Mask R-CNN, являются блоки, внутри которых заключены объекты, относящиеся к классу и маска, в которой все объекты одного класса, отличающиеся друг от друга, и имеют разные значения пикселей [6].

U-Net – свёрточная нейронная сеть, реализуемая в ArcGIS для семантической сегментации [4]. Она состоит из сужающегося пути и расширяющегося пути (*англ.* – contraction path and expansive path). Исходная архитектура U-Net [4] представлена на Рисунке 1:

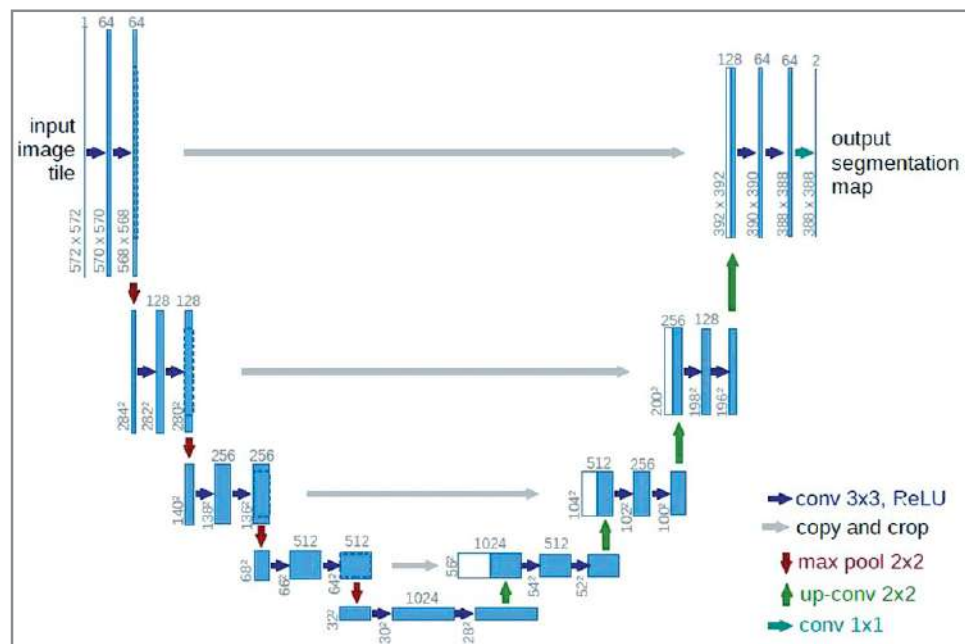


Рис. 1. Архитектура U-Net.

С течением времени происходит улучшение возможностей U-Net. Например, введён параметр, известный как “Padding” (рус. – набивка нулями), который отвечает за размерность результатов, получаемых с помощью свёрточных слоёв, позволяющий сохранить пространственный размер исходного массива, поступающего на входной слой. Сохранение пространственного размера является необходимым при решении задач распознавания по космо- и аэрофотоснимкам в современных условиях. Результат, получаемый с помощью обученной модели U-Net, представляет собой маску для изображения, в которой каждому пикселю соответствует значение 0 и 1 при классификации двух классов или фиксированные значения в диапазоне от 0 до 255 при классификации трёх и более классов [4].

При рассмотрении нейронных сетей, используемых в ГИС, получен вывод, что RetinaNet и Mask R-CNN не являются оптимальными для задач распознавания:

- RetinaNet непригодна для поставленной задачи в связи с тем, что якорные блоки не позволяют отразить сложную метрику классов земной поверхности и облаков [5].
- Mask R-CNN используется для случаев, когда необходимо в рамках одного конкретного класса найти все объекты и отделить отличающиеся друг от друга. Таким образом, при попытке семантической сегментации (распознавании классов) вместо желаемой бинарной маски со значениями 0 и 1 будет получена маска с фиксированными значениями для объектов в рамках одного класса. Это ведёт к необходимости отнесения всех объектов к одному классу на уже готовой маске, что в свою очередь

требует дополнительных операций и уменьшает скорость обработки.

2.2.2 Свёрточная нейронная сеть U-Net в задачах распознавания

Среди рассмотренных нейронных сетей, используемых в ГИС, U-Net является оптимальным решением для распознавания классов земной поверхности и облаков по следующим основным причинам:

- 1) U-Net позволяет вести обучение на больших массивах данных, состоящих из космо- и аэрофотоснимков с разными разрешениями и яркостными характеристиками, в результате чего может быть обеспечена устойчивость получаемой модели к разным яркостным значениям пикселей для одного и того же класса на разных снимках;
- 2) процесс предобработки данных для обучения и постобработки результатов предсказаний U-Net является наименее ресурсоёмким по сравнению с другими методами нейронных сетей, используемыми в ГИС (Mask R-CNN);
- 3) среди сетей для семантической сегментации U-Net показывает наилучшие с точки зрения точности и скорости работы результаты [7];
- 4) получаемая в ходе обучения модель может производить обработку и получение предсказаний последовательно, по нескольким аэрофото- и космоснимкам;
- 5) модели, получаемые при обучении на выборке небольших размеров, позволяют производить достаточно точную сегментацию, что говорит об их высоком качестве [7].

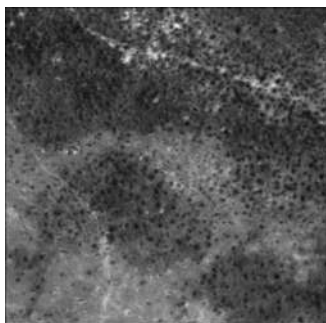
2.3 Выборки цифровых изображений, используемые для распознавания классов земной поверхности и облаков

Для анализа моделей использовались следующие распознаваемые классы и выборки цифровых изображений:

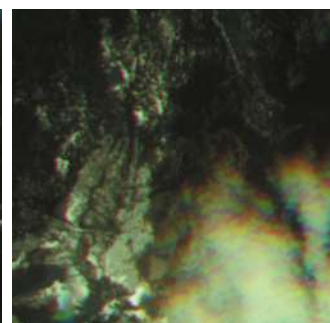
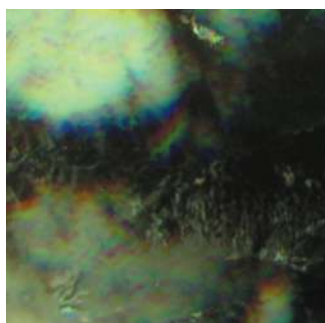
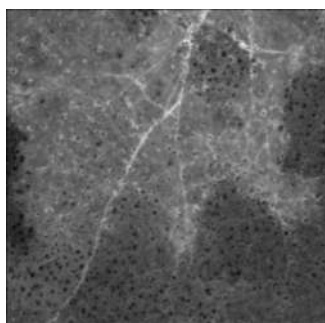
- Космоснимки компании DigitalGlobe (450 снимков), снятые с помощью спутника WorldView3, пространственное разрешение сенсора в Надире – 1.24 м, динамический диапазон 11 бит на пиксель. Распознаваемый класс – лесные массивы. Снимки имеют разные яркостные характеристики распознаваемых объектов класса. Примеры фрагментов изображения представлен на Рисунке 2:
- Аэрофотоснимки взволнованной морской поверхности (200 снимков), масштаб 1 : 160.

Распознаваемый класс – рябь на ряби, занимающий большую часть снимков. Для формирования выборки использовалась аугментация новых снимков и масок по 5 исходным. При обучении на снимках взволнованной морской поверхности имеет место корреляция яркостных характеристик пикселей, усложняющая процесс обучения. Пример фрагмента цифрового изображения с распознаваемым классом представлены на Рисунке 3:

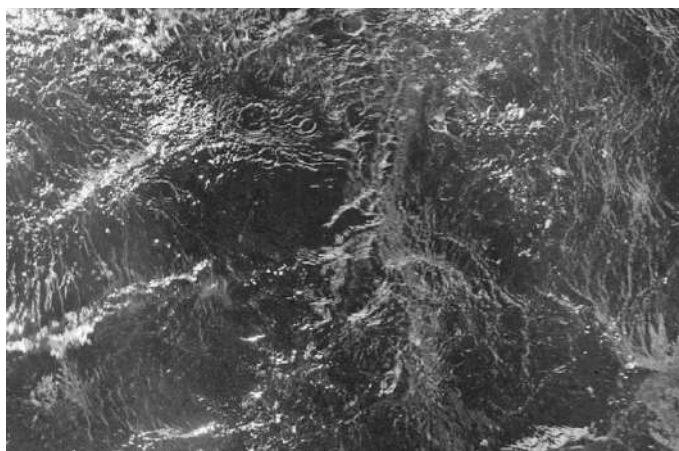
- Массив цифровых снимков облачности (92 снимка). Распознаваемый класс – облака, высота съёмки 2000 – 3000 м, масштаб от 1 : 10 000 до 1 : 15 000. На снимках присутствует дымка, усложняющая процесс обучения и распознавания. Фрагменты цифрового изображения с распознаваемым классом представлены на Рисунке 4:



▲ Рис. 2. Фрагменты цифрового изображения класса “лесные массивы”.



▲ Рис. 3. Фрагмент цифрового изображения класса “рябь на ряби”.



◀ Рис. 4. Фрагменты цифрового изображения класса “облака”.

3 Результаты

В рамках сравнения использовались следующие модели обоснованной архитектуры U-Net:

- 1) U-Net с “padding” и метрикой оценки “Точность”;
- 2) U-Net с “padding” и метрикой оценки “Коэффициент Дайса”;
- 3) U-Net с “padding” и метрикой оценки “Коэффициент Жаккара”.

Для сравнительного анализа выполнено распознавание с помощью указанных моделей U-Net на каждом из массивов данных, численные результаты которого представлены в таблице 1

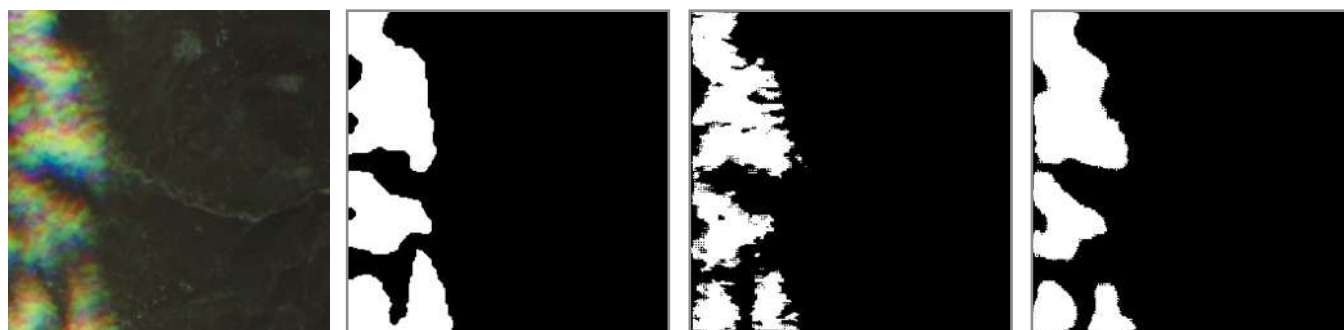
Тип модели	Распознаваемый класс								
	Рябь на ряби			Лесные массивы			Облака		
	Макс. вероятность	Пороговое значение	Точность	Макс. Вероятность	Пороговое значение	Точность	Макс. вероятность	Пороговое значение	Точность
	1	2	3	4	5	6	7	8	9
U-Net (Точность)	0,868	0,750	0,695	0,490	0,675	0,634	0,615	0,800	0,836
U-Net (Коэффициент Жаккара)	1,000	0,800	0,196	1,000	0,730	0,367	0,999	0,950	0,654
U-Net (Коэффициент Дайса)	1,000	0,800	0,291	1,000	0,700	0,478	1,000	0,930	0,723

Таблица 1. Результаты тестирования моделей U-Net.

4 Обсуждение результатов

Проведённый анализ показал, что:

- Модели U-Net с метрикой “точность” показывают лучшие результаты при распознавании классов “лесные массивы” и “рябь на ряби” относительно других обучаемых моделей. Кроме того, они показывают сравнимые с другими моделями результаты при распознавании класса “облака”. Это говорит о хорошей обучаемости данных моделей (Табл. 1).
- Модели U-Net с метрикой “точность” позволяют отражать сложную метрику распознаваемых объектов классов, хотя и с наличием незначительной цифровой зернистости (Табл. 1, столбец “Точность”).
- Модели с метриками “коэффициент Жаккара” и “коэффициент Дайса” хорошо показывают себя при распознавании класса “облака”, однако не способны обучиться на иных указанных массивах данных (Табл. 1 и Рис. 5).



а) исходное цифровое изображение облачности

б) эталон для класса “облака”

в) результат распознавания U-Net с метрикой “точность”

г) результат распознавания U-Net с метрикой “коэффициент Жаккара”

Рис. 5. Результаты распознавания класса “облака” на снимке облачности моделями U-Net с разными метриками.

В перспективе планируется провести исследование функций потерь, подходящих для моделей с метриками “коэффициент Жаккара” [8] и “коэффициент Дайса”, с целью улучшения их обучаемости.

5 Выводы

Применение U-Net с метрикой “точность” для распознавания классов земной поверхности и U-Net с метриками “коэффициент Жаккара” и “коэффициент Дайса” для распознавания облаков позволит в дальнейшем производить автоматизированное дешифрирование с использованием других классов

и выборочных цифровых изображений, подконтрольное выполняющему его специалисту. Кроме того, данная работа обращает внимание на возможность применения методов нейронных сетей, внедрённых в геоинформационные системы, для решения задач машинного обучения, в которых ведётся работа с пространственными данными и распознаваемыми классами, запечатлёнными на аэрофото- и космоснимках [9, 10].

Благодарности

Авторы благодарны рецензентам за оценку работы и ценные замечания.

Библиография

1. Ховатов Н. Э. Информационные технологии в электронике / Н. Э. Ховатов // ModernScience. 2021. № 3-1. С. 490-495.
2. Mitchell T. M. Machine learning. New York: McGraw Hill Science/Engineering/Math, 1997. 432 p.
3. Amr T. Hands-on machine learning with scikit-learn and scientific python toolkits. Birmingham: Packt Publishing Ltd., 2020. 342 p.
4. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation // Proceeding of the Medical Image Computing and Computer-Assisted Intervention (MICCAI) 18th International Conference. Part III. 5-9 October. Munich. 2015. P. 234-241. URL: <https://arxiv.org/pdf/1505.04597.pdf> (дата обращения 10.06.2021).
5. Focal loss for dense object detection [Электронный ресурс]. / T. Lin, P. Goyal, R. Girshick, K. He, P. Dollár // IEEE. 2018. Vol. 42. No. 2. P. 318-327 URL: <https://arxiv.org/pdf/1708.02002.pdf> (дата обращения 12.11.2021). DOI: 10.1109/TPAMI.2018.2858826
6. Mask R-CNN [Электронный ресурс] / K. He, G. Gkioxari, P. Dollár, R. Girshick // Proceeding of International Conference on Computer Vision (ICCV). 2017. URL: <https://arxiv.org/pdf/1703.06870.pdf> (дата обращения 10.06.2021).
7. Емельянов А. В. Анализ методов семантической сегментации изображений на основе нейронных сетей / А. В. Емельянов // Математические методы в технике и технологиях (ММТТ). 2019. Т. 12-1. С. 195-201.
8. Guo, Y. Cloud detection for satellite imagery using attention-based U-Net convolutional neural network / Y. Guo, X. Cao, B. Liu, M. Gao // Symmetry. 2020. Vol. 12. No. 6. P. 1-15. DOI: 10.3390/sym12061056
9. Lynda, N. O. Application of deep learning in satellite image-based land cover mapping in Africa / N. O. Lynda, E. I. Chukwukalo, N. A. Nnanna, A. A. Aminu // International Journal of Advanced Computer Science and Applications. 2021. Vol. 12. No. 9. P. 418-428. DOI: 10.14569/IJACSA.2021.0120948
10. Varga, O. G. Validation of visually interpreted CORINE land cover classes with spectral values of satellite images and machine learning / O. G. Varga, Z. Kovács, L. Bekő, P. Burai, Z. C. Szabó, I. Holb, S. Ninsawat, S. Szabó // Remote Sensing. 2021. Vol. 13. No. 5. P. 1-24. DOI: 10.3390/rs13050857

Бирюков Никита Андреевич

Магистратура МИИГАиК 1 курс; кафедра информационно-измерительных систем (ИС); факультет геоинформатики и информационной безопасности (ФГиИБ)
ORCID: 0000-0001-9856-4696

Маргарита Петровна Лапчинская

Кафедра информационно-измерительных систем (ИС); факультет геоинформатики и информационной безопасности (ФГиИБ);
Кандидат технических наук; Доцент по кафедре информационно-измерительных систем
ORCID: 0000-0002-7354-8283

Поступила 26.11.2021. Рецензия получена 01.12.2021. Принята к публикации 20.12.2021

U-NET NEURAL NETWORK IN EARTH SURFACE CLASSES AND CLOUDS RECOGNITION TASKS

Nikita A. Biryukov^{1*}, Margarita P. Lapchinskaya¹

¹ Moscow State University of Geodesy and Cartography, Moscow, Russia

* 4096414@gmail.com

Citation: Biryukov NA, Lapchinskaya MP. U-Net neural network in earth surface classes and clouds recognition tasks. *Izvestia vuzov. Geodesy and Aerophotosurveying*. Moscow. 2021;65(6): 673-679. (In Russian). DOI:10.30533/0536-101X-2021-65-6-673-679

Keywords. geoinformation systems, machine learning, pattern recognition, neural networks, semantic segmentation.

Acknowledgements. The authors thank the reviewers for their assessment of the work and valuable comments.

Abstract. This article is devoted to defining a neural network method for Earth surface classes and clouds recognition automation, used in GIS and is optimal for satellite and aerial images. This article contains the description of the data for machine learning in GIS and neural network methods included in GIS such as RetinaNet and U-Net. Three U-Net architectures with different metrics are compared: U-Net with accuracy, Jaccard and Dice coefficients. In order to conduct the comparative analysis of the results three classes and datasets are used: “woodlands” class for WorldView 3 satellite images, “ripple on ripple” for aerial images of the rough sea surface and “clouds” for aerial images of cloudiness. Basing on the analysis of U-Net, Mask R-CNN and RetinaNet, it can be concluded that it is possible to use these neural network methods in solving the tasks of recognition of Earth surface classes and clouds: Mask R-CNN consumes more time on the processing than U-Net, and RetinaNet does not recognize complex metric of the described classes and only gives anchor boxes. U-Net main features optimize this method for the tasks considered. The results of U-Net model recognition on testing raster images are also presented in the article. The provided recognition results demonstrate that U-Net with “accuracy” steadily learns on available datasets and recognizes all classes, thus with the presence of grain, and shows that models with “Jaccard Coefficient” and “Dice Coefficient” can only recognize class “clouds” and do it better compared to U-Net with “accuracy”.

References

1. Khovатов NE. Information technologies in electronics. *Modern Science*. [Modern Science], 2021;3-1: 490-495. (In Russian)
2. Mitchell TM. *Machine learning*. New York: McGraw Hill Science/Engineering/Math; 1997.
3. Amr T. *Hands-on machine learning with scikit-learn and scientific python toolkits*. Birmingham: Packt Publishing Ltd.; 2020.
4. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation // *Proceeding of the Medical Image Computing and Computer-Assisted Intervention (MICCAI) 18th International Conference. Part III*. Munich. 2015. p. 234-241. URL: <https://arxiv.org/pdf/1505.04597.pdf> (accessed 10 June 2021).
5. Lin T, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. *IEEE*. 2018;42(2): 318-327, available at: <https://arxiv.org/pdf/1708.02002.pdf> (accessed 12 November 2021). DOI: 10.1109/TPAMI.2018.2858826
6. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *Proceeding of International Conference on Computer Vision (ICCV)*. 2017; available at: <https://arxiv.org/pdf/1703.06870.pdf> (accessed at 10 June 2021).
7. Emel'yanov AV. Semantic segmentation methods analysis on neural networks basis. *Matematicheskie metody v tekhnike i tekhnologiyakh (MMTT)* [Mathematical methods in engineering and technology (MMET)]. 2019;12-1: 195-201. (In Russian)
8. Guo Y, Cao X, Liu B, Gao M. Cloud detection for satellite imagery using attention-based U-Net convolutional neural network. *Symmetry*. 2020;12(6): 1-15. DOI: 10.3390/sym12061056
9. Lynda NO, Chukwukalo EI, Nnanna NA, Aminu AA. Application of deep learning in satellite image-based land cover mapping in Africa. *International Journal of Advanced Computer Science and Applications*. 2021;12(9): 418-428. DOI: 10.14569/IJACSA.2021.0120948
10. Varga OG, Kovács Z, Bekő L, Burai P, Szabó ZC, Holb I, Ninsawat S, Szabó S. Validation of visually interpreted CORINE land cover classes with spectral values of satellite images and machine learning. *Remote Sensing*. 2021;13(5): 1-24. DOI: 10.3390/rs13050857

Nikita A. Biryukov

Master's degree; Year 1; Department of Information-Measuring Systems; Faculty of Geoinformatics and Information Security

ORCID: 0000-0001-9856-4696

Margarita P. Lapchinskaya

Department of Information-Measuring Systems; Faculty of Geoinformatics and Information Security; PhD in Engineering; Assistant professor

ORCID: 0000-0002-7354-8283