

# Нейросетевые технологии поиска целевых объектов на снимках дистанционного зондирования Земли

Н.С. Абрамов<sup>1</sup>, А.А. Талалаев<sup>1</sup>, В.П. Фраленко<sup>1</sup>, О.Г. Шишкин<sup>1</sup>, В.М. Хачумов<sup>1</sup>

<sup>1</sup>ИПС им. А.К. Айламазяна РАН, Петра Первого 4 «а», с. Веськово, Ярославская область, Россия, 152021

**Аннотация.** В работе решены многоклассовая и одноклассовая задачи поиска и классификации целевых объектов на снимках дистанционного зондирования Земли. Для повышения эффективности распознавания разработаны инструменты подготовки обучающих выборок, оптимального конфигурирования и применения нейронных сетей глубокого обучения с использованием технологий высокопроизводительных вычислений. Выполнен сравнительный анализ полученных результатов.

## 1. Введение

На сегодняшний день в области обработки данных дистанционного зондирования Земли (ДЗЗ) наблюдается всплеск активности: создаются новые программные комплексы, модернизируются методы обработки изображений высокого разрешения. Текущая ситуация характеризуется совершенствованием оборудования космических аппаратов (КА) и наземных станций управления, расширением функциональных возможностей и спектра выполняемых задач по обработке изображений. В сферу применения данных КА входят мониторинг лесных, сельскохозяйственных и арктических зон, анализ природных катаклизмов, охрана окружающей среды, государственная безопасность и др. Растущие объемы пополняемых данных ДЗЗ существенно повысили требования к скорости и качеству обработки информации. В последнее время все чаще применяют искусственные нейронные сети (ИНС) и технологии высокопроизводительных вычислений.

Авторами настоящей работы создан научный и практический задел в области решения различных задач на основе интеллектуальной обработки снимков ДЗЗ (мультиспектральных, панхроматических, цветных), в том числе поиска ригидных объектов и зон интереса с применением разработанного спектрографического подхода и обобщенной метрики (пожары, разливы рек, оценка ледовой обстановки и др.) [1-6]. В предлагаемой работе представлены результаты новых углубленных исследований, связанных с применением современных сверточных нейронных сетей (СНС) для обработки панорамных полноцветных снимков ДЗЗ, полученных с беспилотных летательных аппаратов (БПЛА), предложены методы и инструменты для повышения их эффективности и быстродействия на задаче поиска и распознавания объектов военной техники с необходимой полнотой и точностью, которая остается до сих пор нерешенной даже при обилии программных средств. Современная постановка задачи поиска и распознавания объекта нейронной сетью включает в себя этапы выбора типа, настройки параметров ИНС и подготовки входных данных. В рамках проведенного исследования рассмотрены многоклассовая и одноклассовая задачи. Под первой

задачей понимается поиск и распознавание объектов нескольких классов одновременно. Вторая задача предусматривает поиск нейронной сетью объектов одного класса.

## 2. Методы, программно-инструментальные средства и результаты обработки изображений с использованием ИНС

Для обработки изображений ДЗЗ использовались два вида СНС: сверточная нейронная сеть из библиотеки «nnForge» [7] и сеть класса «Darknet» [8]. Обе реализации отличаются поддержкой разнообразных видов слоев, при этом обеспечивается гибкая настройка, имеется возможность менять структуру под свои задачи. Кроме этого, сеть второго вида не только классифицирует целевые объекты, но и сообщает об их позициях на снимке. Отличительная особенность рассматриваемых СНС – поддержка ускорения вычислений с использованием графических процессорных устройств (ГПУ) как при обучении, так и при эксплуатации. Копии обученных ИНС распределяются между имеющимися ГПУ, на которых данные обрабатываются независимо и асинхронно. Для организации вычислительного процесса был использован специальный программный комплекс проектирования нейросетевых прикладных систем [9-10]. Для повышения качества классификации реализовано специальное средство автоматизированной подготовки обучающих выборок. Предварительно человек-эксперт готовит снимки с установленным маркером прозрачности, где пиксели фона устанавливаются невидимыми с помощью управления альфа-каналом. На рисунке 1 показан исходный фрагмент снимка ДЗЗ, справа этот же фрагмент после изменения альфа-канала. Для удобства снимки с объектами разных классов раскладываются в разные директории.

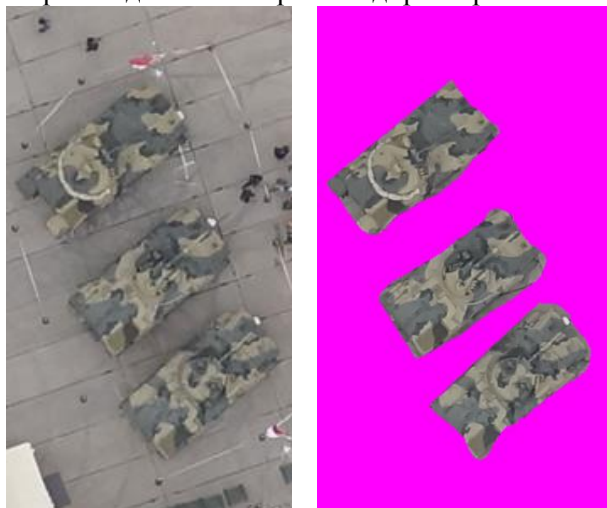


Рисунок 1. Изображения: оригинальное и с маской из альфа-канала.

Для различных настроек ИНС из библиотеки «nnForge» (конфигураций и характеристик слоев) осуществляется работа с промасштабированными копиями этих изображений. Коэффициент масштабирования выбирается так, чтобы каждый целевой объект помещался в отдельное сканирующее окно, размер которого совпадает с размером входного окна ИНС. Эксперименты проводились над изображениями военной техники размером 6000x4000 пикселей, сделанными с высоты 300 метров на международной выставке «Russia Arms Expo – 2015» (RAE-2015).

Экспериментально выбрана следующая архитектура СНС из библиотеки «nnForge»:

- слой извлечения контраста с окном Гауссиана размером 9x9 пикселей [11]; исходный размер окна с данными: 39x39;
- слой свертки с картами признаков 6x6 (всего 136 карт), для нормализации используется модуль гиперболического тангенса, после обработки размер окна 34x34;
- слой усреднения (average subsampling) с маской 2x2, после обработки размер окна 17x17;
- слой свертки с картами признаков 6x6 (всего 272 карты), после обработки размер окна 12x12;

- слой усреднения с маской 2x2, после обработки размер окна 6x6;
  - слой свертки с картами признаков 6x6 (всего 544 карты), после обработки размер окна 1x1;
  - dropout-слой с регулируемой вероятностью отключения связей между нейронами (экспериментально установлена равной 0.05);
  - слой свертки с картами признаков размером 1x1, количество карт признаков соответствует числу выделяемых классов, применяется функция активации типа гиперболический тангенс.
- При распознавании сканирующее окно перемещается по изображению с шагом в один пиксель и обрабатывается нейронной сетью. В результате последовательной обработки всего изображения имеем раскрашенную карту, где целевые объекты отделены от фона. Пример исходного изображения и результат его обработки приведен на рисунке 2.



**Рисунок 2.** Исходное изображение и результат его обработки.

Для обучения было использовано 3.626 млн. объектов, автоматически извлеченных из 73 снимков. Получены следующие результаты. Полнота классификации: фон – 0.9976, техника – 0.9354. Точность классификации: фон – 0.9392, техника – 0.9974. Время обучения: 24 часа на одной Nvidia Geforce GTX 1060 и с использованием одного ядра CPU Intel Core i7 6850K (из имеющихся 6 ядер, 3.6 @ 4.0 ГГц). Время обработки десяти панорамных снимков на одном ГПУ: 3051 с; на двух ГПУ: 1640 с.

Разработчики сети типа «Darknet» в результате многочисленных экспериментов подобрали весьма удачную архитектуру, такую, что она работает на разных обучающих/тестовых выборках [12, 13]. Сеть в определенных пределах устойчива к тому, что на вход могут подаваться изображения разного размера, подвергнутые геометрическим искажениям. Основным регулируемым параметром является размер входного слоя СНС.

Разработанный в рамках настоящего исследования программно-инструментальный комплекс включает программы «YOLORotate», «YOLOAnchors» и «YOLOGetObjects».

Программа «YOLORotate» предназначена для преобразования изображений в формат, подходящий для обучения ИНС типа «YOLO v2». Данные, необходимые для подготовки серии проведенных экспериментов, включают множество панорамных снимков, сделанных с БПЛА. На них присутствует четыре класса целевых объектов: «БМП» (боевые машины пехоты), «Военные автомобили», «САУ» (самоходные артиллерийские установки) и «Танки», информация о координатах и размерах каждого из объектов предварительно собрана и сохранена в текстовых файлах. Каждое такое изображение имеет размер не менее чем 832x832 пикселя, где все целевые объекты занимают относительно малую часть снимка. «YOLORotate» осуществляет поворот изображений из обучающей выборки с заданным шагом, например, в 15 градусов. При этом вырезается максимально возможное количество изображений-фрагментов с целевыми объектами. Всего были автоматически созданы: 4361 фрагмент для обучающей выборки и 1173 фрагмента для тестовой выборки. Гарантируется, что на каждом таком фрагменте имеется хотя бы один целевой объект. Для того, чтобы обеспечить расположение целевых объектов в случайных позициях получаемых изображений-фрагментов, используется

генератор псевдослучайных чисел. Анализ показал, что при выбранном размере исходных данных 832x832 пикселя в кадр попадает до 20 целевых объектов. Программа «YOLOAnchors» использует метод k-средних [14] для выявления ширины и высоты типичных целевых объектов на выходном окне нейронной сети. В результате работы программы находится несколько таких пар размеров, используемых в дальнейшем при обучении ИНС.

Программа «YOLOGetObjects» осуществляет сегментацию изображений на фрагменты размером 832x832 пикселя, при этом обеспечивается частичное пересечение, каждое следующее окно захватывает четверть от предыдущего (как по горизонтали, так и по вертикали). Всего на один панорамный снимок приходится 70 фрагментов. Далее все фрагменты независимо обрабатываются с использованием ГПУ и процессора общего назначения. На следующем шаге осуществляется объединение информации о всех найденных целевых объектах. Пример результата использования обученной ИНС показан на рисунке 3.



Рисунок 3. Результат обработки снимка многоклассовой ИНС.

Во всех экспериментах применялось пакетное обучение, при котором следующий шаг подстройки весовых коэффициентов основывается на информации о результатах обработки ограниченной группы изображений обучающей выборки. Каждая группа изображений на новой эпохе обучения формируется случайным образом, предпочтение отдается тем группам, в которых встречаются представители всех классов целевых объектов. Использование пакетного обучения позволяет повысить качество работы нейронной сети и отказаться от ресурсоемкого dropout-слоя [15]. Лучший размер пакета подбирается экспериментально для каждой решаемой задачи.

В таблицах 1-3 приведены уточняющие характеристики и результаты проведенных экспериментов при решении многоклассовой задачи – одновременного поиска и распознавания объектов четырех классов.

Таблица 1. Результаты тестирования ИНС с входным окном 416x416 пикселей.

		6 изображений в пакете	14 изображений в пакете
Доля найденных целевых объектов	БМП	0.9346	0.9731
	Военные автомобили	0.9880	0.9983
	САУ	1.0000	1.0000
	Танки	0.9967	0.9934
Усредненная доля найденных целевых объектов		0.9798	0.9912
Полнота	БМП	0.7500	0.6885
	Военные автомобили	0.9811	0.9949
	САУ	0.9006	0.9655

	Танки	0.9058	0.9107
Нормализованная точность	БМП	0.9118	0.8947
	Военные автомобили	0.8627	0.9541
	САУ	0.9990	1.0000
<i>F1</i> -мера	Танки	0.8561	0.7672
	БМП	0.8230	0.7782
	Военные автомобили	0.9181	0.9741
	САУ	0.9473	0.9825
	Танки	0.8802	0.8328
Усредненная <i>F1</i> -мера		0.8922	0.8919
Отношение количества найденных целевых объектов к их общему количеству		0.9869	0.9948
Время обучения, часы		1.18	2.44
Время обработки изображения размером 6000x4000 пикселей, секунды			2.13

**Таблица 2.** Результаты тестирования ИНС с входным окном 832x832 пикселя.

		6 изображений в пакете	14 изображений в пакете
Доля найденных целевых объектов	БМП	0.9615	0.9538
	Военные автомобили	0.9983	0.9974
	САУ	1.0000	1.0000
	Танки	0.9967	0.9934
Усредненная доля найденных целевых объектов		0.9891	0.9862
Полнота	БМП	0.5846	0.6077
	Военные автомобили	0.9657	0.9657
	САУ	0.9290	0.9432
	Танки	0.9421	0.9322
Нормализованная точность	БМП	0.8630	0.9004
	Военные автомобили	0.9462	0.9160
	САУ	0.9927	0.9724
	Танки	0.7123	0.7484
<i>F1</i> -мера	БМП	0.6970	0.7256
	Военные автомобили	0.9558	0.9402
	САУ	0.9598	0.9576
	Танки	0.8113	0.8303
Усредненная <i>F1</i> -мера		0.8560	0.8634
Отношение количества найденных целевых объектов к их общему количеству		0.9945	0.9925
Время обучения, часы		4.78	11.11
Время обработки изображения размером 6000x4000 пикселей, секунды			5.50

**Таблица 3.** Результаты тестирования ИНС с входным окном 416x416 пикселей.

		6 изображений в группе	14 изображений в группе
Доля найденных целевых объектов	БМП	0.9615	0.9500
	Военные автомобили	1.0000	0.9983
	САУ	1.0000	1.0000
	Танки	0.9934	0.9983
Усредненная доля найденных целевых объектов		0.9887	0.9867

Полнота	БМП	0.7846	0.7538
	Военные автомобили	0.9923	0.9966
	САУ	0.9432	0.9473
	Танки	0.9223	0.9438
Нормализованная точность	БМП	0.8970	0.9230
	Военные автомобили	0.9421	0.9166
	САУ	1.0000	1.0000
	Танки	0.8510	0.8616
$F1$ -мера	БМП	0.8371	0.8299
	Военные автомобили	0.9665	0.9549
	САУ	0.9708	0.9729
	Танки	0.8852	0.9008
Усредненная $F1$ -мера		0.9149	0.9146
Отношение количества найденных целевых объектов к их общему количеству		0.9945	0.9937
Время обучения, часы		4.69	10.37
Время обработки изображения размером 6000x4000 пикселей, секунды		3.98	

В таблице 4 приведены сравнительные результаты обработки тестовой выборки при решении одноклассовой и многоклассовой задач. Время обучения выбранной конфигурации сети на каждом из четырех классов объектов военной техники составило 6.84, 13.6 и 26.8 часов при работе с группами по 28, 56 и 112 изображений. В одноклассовом случае ИНС работает только с одним классом, у пользователя есть возможность выбрать лучший вариант – сеть, обучавшуюся с использованием группы изображений оптимального размера. В последнем столбце таблицы 4 собраны лучшие коэффициенты для смешанного режима, когда для отдельных классов, например, для «БМП», используется обученная на пакете из 56 изображений, для класса «Танки» – из 112. Среднее время обработки изображения размером 6000x4000 пикселей одной отдельной одноклассовой нейронной сетью такое же, что и у многоклассовой нейронной сети, то есть четыре секунды.

**Таблица 4.** Экспериментальные данные, многоклассовая и одноклассовая задачи.

Класс объектов	Доля найденных объектов и размер группы изображений					
	Многоклассовая задача			Одноклассовая задача		
	6	14	28	56	112	/
БМП	0.9615	0.9500	0.8577	0.9115	0.8538	0.9115
Военные автомобили	1.0000	0.9983	0.9057	0.8885	0.8954	0.9057
САУ	1.0000	1.0000	0.8276	0.9635	0.9473	0.9635
Танки	0.9934	0.9983	0.9174	0.9455	0.9521	0.9521
Усредненная доля найденных целевых объектов	0.9887	0.9867	0.8771	0.9273	0.9121	0.9332
Время обучения одной отдельной нейронной сети, часы	4.69	10.37	6.84	13.6	26.8	–

Результаты экспериментов подтвердили эффективность применения одноклассовых нейронных сетей. Однако, обучение комплекса таких сетей требует больше вычислительных ресурсов, чем для одной многоклассовой сети. Увеличение сложности задачи при том же количестве карт признаков приводит к снижению для одноклассовой ИНС усредненной доли найденных целевых объектов.

### 3. Выводы

В статье представлены результаты исследований, связанных с применением современных сверточных нейронных сетей для обработки панорамных полноцветных аэроснимков ДЗЗ. С применением СНС «nnForge» и «Darknet» решены многоклассовая и одноклассовая задачи поиска и классификации целевых объектов. Для повышения эффективности распознавания разработаны методы подготовки обучающих выборок, оптимального конфигурирования и применения высокопроизводительных вычислений. Выполненный сравнительный анализ показал, что одноклассовый подход имеет преимущество в качестве распознавания, но проигрывает по времени работы. В целом следует отметить, что возможности сверточных нейронных сетей позволяют решать одновременно задачи поиска (локализации) и распознавания объектов на снимках ДЗЗ с высокой точностью и полнотой.

### 4. Литература

- [1] Фраленко, В.П. Анализ спектрографических текстур данных дистанционного зондирования Земли // Искусственный интеллект и принятие решений. – 2010. – № 2. – С. 11-15.
- [2] Фраленко, В.П. Интеллектуальный анализ аэрокосмических снимков с применением высокопроизводительных вычислителей // Материалы конференции «Искусственный интеллект: проблемы и пути решения» (14-15 марта 2018 г.), 2018.
- [3] Абрамов, Н.С. Методы, модели и программные средства обработки данных космического мониторинга Арктической зоны / Н.С. Абрамов, А.Ю. Агроник, Ю.Г. Емельянова, А.В. Латышев, А.А. Талалаев, В.П. Фраленко, М.В. Хачумов // Авиакосмическое приборостроение. – 2017. – № 7. – С. 38-51.
- [4] Фраленко, В.П. Локализация и классификация военной техники в потоке изображений от БПЛА // Материалы конференции «Фундаментальная наука – Армии» в рамках Третьего международного военно-технического форума «АРМИЯ-2017» (23 августа 2017 г.), 2017.
- [5] Лебедев, А.С. Экспериментальное исследование задачи поиска сложных ригидных объектов в системе обработки космических снимков «ПС НСКиД» / А.С. Лебедев, В.П. Фраленко, Г.С. Чэн, Г.Л. Чжан // Современные проблемы науки и образования. – 2015. – № 1 [Электронный ресурс]. – Режим доступа: <https://www.science-education.ru/ru/article/view?id=18607> (17.11.2018).
- [6] Хачумов, В.М. Перспективы построения высокопроизводительной системы обработки данных дистанционного зондирования Земли / В.М. Хачумов, В.П. Фраленко, Chen Guo Xian, Zhang Guo Liang // Программные системы: теория и приложения. – 2015. – № 1. – С. 121-133.
- [7] nnForge by milakov [Электронный ресурс]. – Режим доступа: <http://milakov.github.io/nnForge> (17.11.2018).
- [8] Redmon, J. YOLO9000: Better, Faster, Stronger / J. Redmon, A. Farhadi. – [Electronic resource]. – Access mode: <https://arxiv.org/abs/1612.08242> (17.11.2018).
- [9] Талалаев, А.А. Комплекс инструментальных средств для проектирования нейросетевых прикладных систем / А.А. Талалаев, В.П. Фраленко // Научно-технический вестник Поволжья. – Казань: Научно-технический вестник Поволжья. – 2013. – № 4. – С. 237-243.
- [10] Талалаев, А.А. Архитектура комплекса конвейерно-параллельной обработки данных в гетерогенной вычислительной среде / А.А. Талалаев, В.П. Фраленко // Вестник Российского университета дружбы народов. Серия Математика. Информатика. Физика. – 2013. – № 3. – С. 113-117.
- [11] Pinto, N. Why is real-world visual object recognition hard? / N. Pinto, D.D. Cox, J.J. DiCarlo // PLoS Comput Biol. [Electronic resource]. – Access mode: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0040027> (17.11.2018).
- [12] Everingham, M. The pascal visual object classes (voc) challenge / M. Everingham, L. Van Gool, C.K. Williams, J. Winn, A. Zisserman // International Journal of Computer Vision. – 2010. – Vol. 88. – P. 303-338.

- [13] Lin, T.-Y. Microsoft coco: Common objects in context / T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, C.L. Zitnick // European Conference on Computer Vision, Springer. – 2014. – P. 740-755.
- [14] Celebi, M.E. A comparative study of efficient initialization methods for the k-means clustering algorithm / M.E. Celebi, H.A. Kingravi, P.A. Vela // Expert Systems with Applications. – 2013. – Vol. 40. – P. 200-210.
- [15] Srivastava, N. Dropout: a simple way to prevent neural networks from overfitting / N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov // Journal of Machine Learning Research. – 2014. – Vol. 15. – P. 1929-1958.

### **Благодарности**

Работа выполнена при поддержке РФФИ (проекты № 18-29-03011-мк «Исследование и разработка новых методов и технологий для задач интеллектуального анализа и оптимизации обработки больших потоков данных дистанционного зондирования Земли» и № 17-29-07003-офи\_м «Разработка методов и моделей динамического планирования поведения и иерархического интеллектуального управления движением беспилотных летательных аппаратов в условиях неопределенной среды при ограничениях на вычислительные ресурсы»).

## **Neural network technologies to search for targets in Earth remote sensing images**

**N.S. Abramov<sup>1</sup>, V.M. Khachumov<sup>1</sup>, A.A. Talalaev<sup>1</sup>, V.P. Fralenko<sup>1</sup>, O.G. Shishkin<sup>1</sup>**

<sup>1</sup>Program Systems Institute of RAS, Yaroslavl region, Pereslavsky district, Veskovo village, Peter the First street, 4 A, Russia, 152021

**Abstract.** In the paper multi-class and single-class problems of searching and classifying target objects in remote sensing images of the Earth are solved. To improve recognition efficiency, tools have been developed for preparing training samples, optimal configuration, and using deep learning neural networks using high-performance computing technologies. A comparative analysis of the results performed.