

Решение задачи распознавания объектов и инцидентов на фотоматериалах, полученных с беспилотных летательных аппаратов с использованием методов глубокого обучения

А.Ю.Белякова¹, Ю.Д.Беляков², П.С.Замятин²

¹Иркутский государственный аграрный университет имени А.А. Ежевского, Иркутск

²Санкт-Петербургский государственный университет, Санкт-Петербург

Аннотация: В статье рассмотрены методы и алгоритмы дообучения сверточной нейронной сети VGG16 для решения задачи распознавания объектов на снимках с БПЛА (беспилотных летательных аппаратов). При отсутствии необходимого количества исходной информации предлагается работать на аугментированном наборе данных. В статье представлена архитектура нейронной сети и рассмотрено ее действие на конкретном примере. При разработке сервиса, выполняющего загрузку изображения и вывод результатов работы модели, использовался фреймворк Flask, обучение моделей происходило с использованием облачного сервиса Google Colab на основе Jupyter Notebook.

Ключевые слова: глубокое обучение, нейронная сеть, обработка снимков БПЛА, распознавание объектов, аугментация.

Постановка задачи

В нефтедобывающей промышленности важно следить за состоянием эксплуатируемых объектов. Однако это бывает проблематично из-за больших площадей. Поэтому выгодно автоматизировать данный процесс, путем сбора изображений с помощью БПЛА (беспилотных летательных аппаратов), и автоматическим определением аномалий.

В данной работе было проведено дообучение сверточной нейронной сети VGG16 [1, 2] для решения задачи распознавания объектов на снимках с БПЛА.

Исходные данные

Для решения поставленной задачи был собран набор из 166 снимков земной поверхности, снятых с летательных аппаратов в высоком качестве (3840x2160). Каждое из изображений относится к одному из 6 классов: вскрытие грунта (на изображении присутствуют признаки вскрытия грунта),

скважины без разлива (на изображении присутствуют нормально функционирующие нефтяные скважины), скважины с разливом (на изображении присутствуют нефтяные скважины, рядом с которыми заметны темные пятна от разлива нефти), спец. техника (на изображении присутствует спец. техника, совершающая какие-либо работы), другое (на изображении не выявлено никаких из перечисленных аномалий, как правило это снимки чистого рельефа).

Наиболее часто встречающаяся проблема в данной области - недостаток данных. Для решения поставленной задачи были представлены 44 фотографии с аномалиями и 122 без таковых. Естественно, этого количества недостаточно для обучения нейронной сети, но высокое разрешение фотографий позволяло применять различные методы для увеличения размеров тренировочной выборки [3, 4].

Методы и решение

Аугментация данных (data augmentation) – это метод создания дополнительных обучающих данных из имеющихся данных. Для достижения хороших результатов глубокие сети должны обучаться на очень большом объеме данных [5, 6]. Следовательно, если исходный обучающий набор содержит ограниченное количество изображений, необходимо выполнить аугментацию, чтобы улучшить результаты модели.

Был применен следующий алгоритм аугментации [7]:

1. Фотографии с аномалиями разбивались на 16 изображений, разбиение происходило с небольшим сдвигом, для того, чтобы избежать разрезание ключевых объектов (рис.1).
2. Далее, вручную, полученные снимки разбивались на 6 классов. На данном этапе кроется вторая трудность - не всегда достоверно известно есть ли на фотографии аномалия.

3. Последний шаг - отзеркаливание изображений, для того, чтобы модель училась определять объекты на снимках, снятых с разных ракурсов.

Аугментация

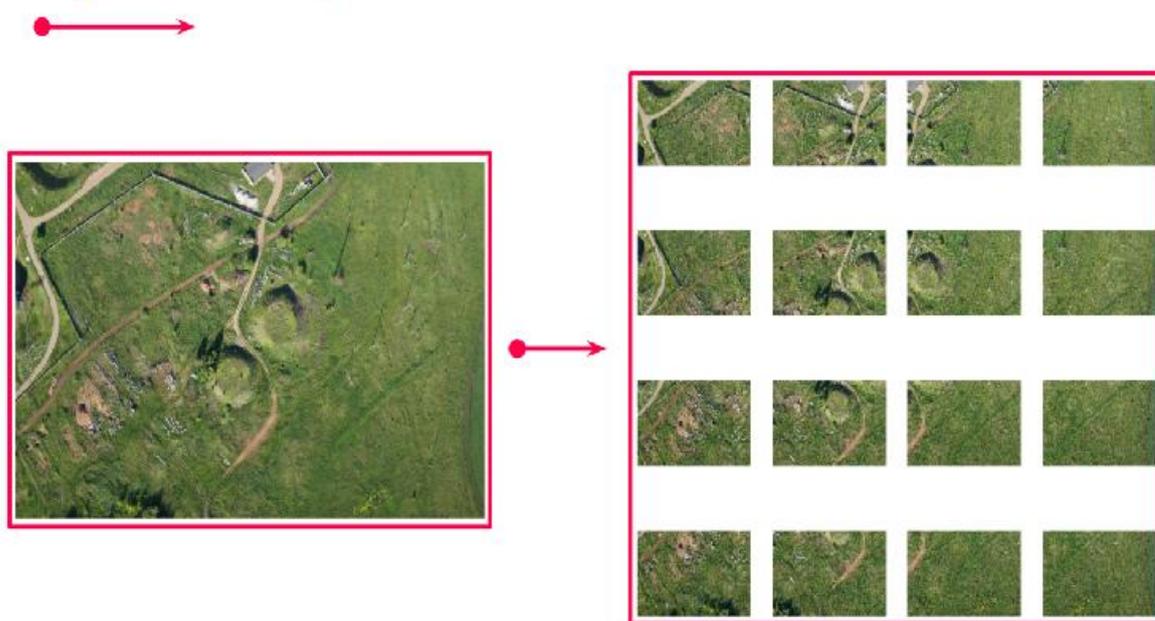


Рис.1. – Аугментация снимка

Таким образом, был увеличен объем данных до 3600 изображений, которые разбили на 6 категорий, а именно: скважины с разливами и без разливов, вскрытие грунта, складирование материала, спец. техника и фотографии без аномалий.

Обучение моделей происходило с использованием облачного сервиса Google Colab на основе Jupyter Notebook. Для обучения модели использовался фреймворк Keras для создания веб-приложения фреймворк Flask.

Для классификации объектов мы выбрали модель VGG16 с бинарным выходным слоем и дообучили ее на аугментированном наборе данных (рис.2).

Для каждого из классов была обучена модель бинарной классификации по принципу One-vs-Rest [8, 9]. В результате было получено 6 моделей бинарной классификации.

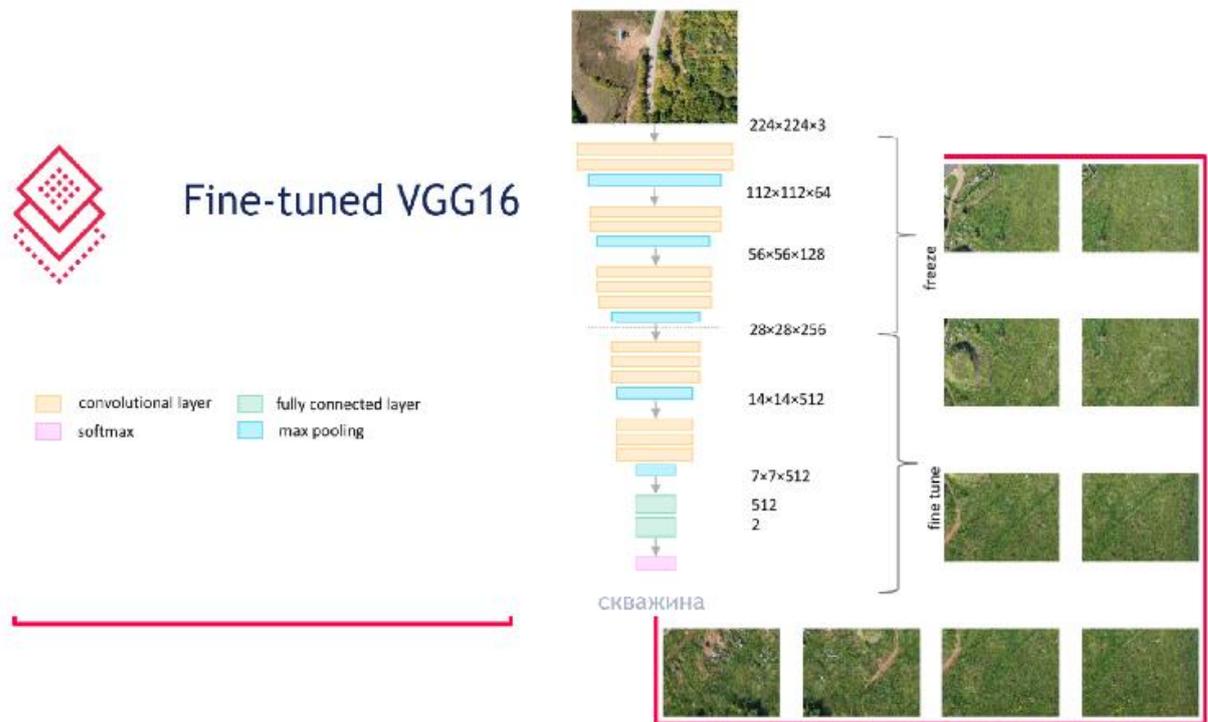


Рис.2. – Архитектура нейронной сети

На обучающих и валидационных данных была достигнута точность 85-90 %, на тестовых данных F-мера варьировалась от 0.75 до 0.9 в зависимости от класса. Как видно из рисунка 3, метрики для некоторых из классов хуже, чем для других. Это происходит по ряду причин. Для класса со скважинами без разливов, довольно мало примеров, также зачастую на фотографии присутствуют тени и естественные потемнения на земле, рядом со скважинами, что затрудняет классификацию. Для класса со специальной техникой сеть слишком сильно привязалась к наличию дорог, поэтому часто встречаются ошибки второго рода.

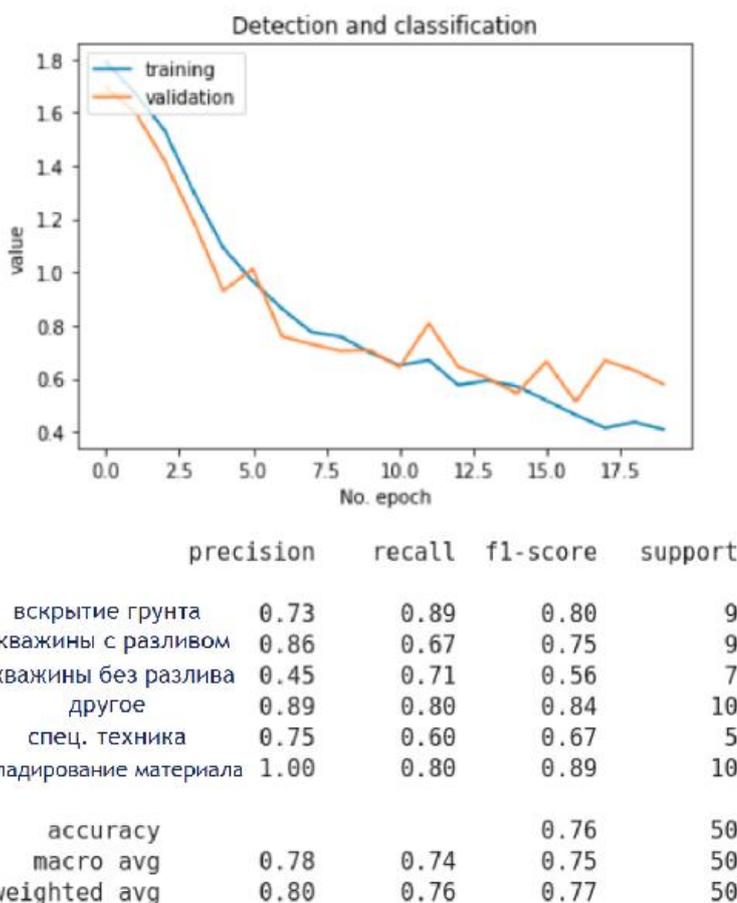


Рис.3. – Метрики качества

Каким образом можно улучшить разработанную модель:

- Во-первых, это новые данные, чем больше данных, тем точнее модель будет определять объекты.
- Во-вторых, нужна корректная разметка данных экспертами.
- В-третьих, это использование графического процессора, оно позволит ускорить обучение и обработку изображений.
- В-четвертых, тестирование других моделей, которые возможно решают данную задачу лучше.

Рассмотрим работу нейронной сети на конкретном примере (рис. 4, 5). Нейронная сеть разбивает фотографию на части, и присваивает каждой части категорию, если на исходной фотографии присутствует несколько аномалий, то модель их выявит.



Рис. 4. – Снимок со складированием материала

Для сегментации была использована тепловая карта, полученная на основе весов модели по методу Class Activation Mapping (CAM) [10]. Такой подход позволяет наглядно увидеть, что модель действительно находит признаки классов, а не запоминает данные или угадывает.

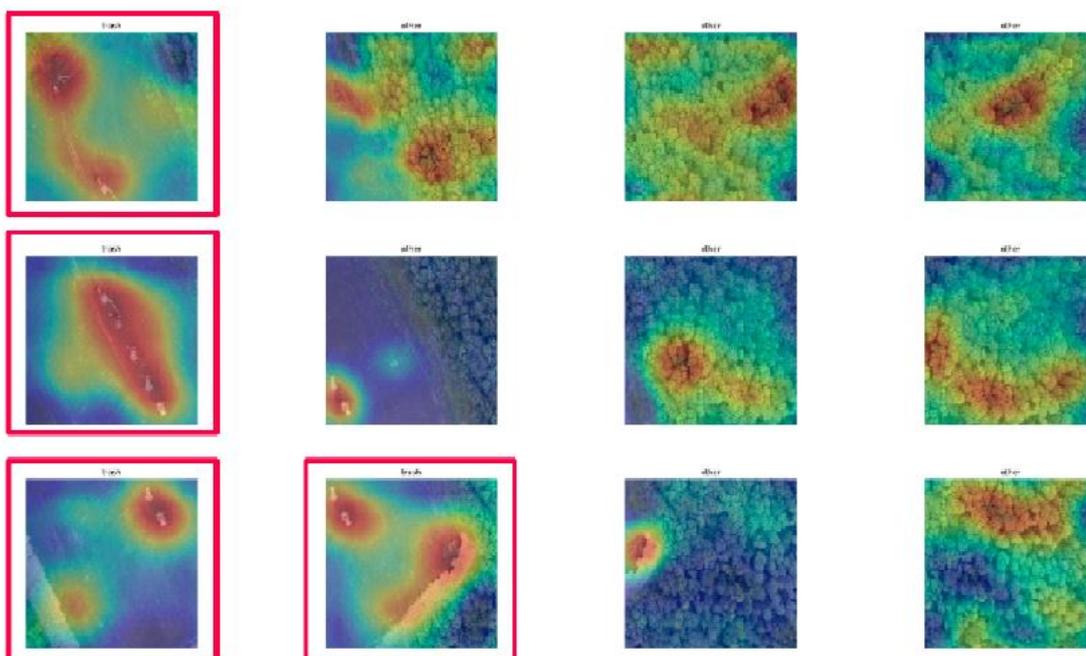


Рис.5. – Результат работы модели

На тестовом наборе данных модель отлично распознала скважины с разливом, без него и вскрытие грунта.

Результаты

С помощью методики создания дополнительных обучающих данных аугментации увеличено количество исходных данных.

Реализована задача дообучения нейронной сети, на основе распознавания и классификации аномалий на изображениях.

Разработан полностью функционирующий сервис, выполняющий загрузку изображения и вывод результатов работы модели.

Литература

1. Simonyan K., Zisserman A. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition CoRR abs, 5. URL: arxiv.org/pdf/1409.1556.pdf

2. Пучков Е.В. Сравнительный анализ алгоритмов обучения искусственной нейронной сети // Инженерный вестник Дона. 2013. №4. URL: ivdon.ru/ru/magazine/archive/n4y2013/2135

3. Конарев Д. И., Гуламов А. А. Повышение точности предварительно обученных нейронных сетей путём тонкой настройки // Материалы конференции Информационные технологии в управлении. 2020. С. 209-212. URL: itc.etu.ru/assets/files/itc-2020/papers/209.pdf

4. Кустикова В.Д. Методы глубокого обучения и примеры практического применения для решения задач компьютерного зрения // Сборник тезисов участников форума "Наука будущего - наука молодых". 2017. С. 102-103. URL: elibrary.ru/item.asp?id=32571150

5. Горбацевич В.С., Князь В.А., Мельниченко М.А., Степаньянц Д.Г. Методы глубокого обучения для автоматического построения трехмерных моделей по аэрофотоснимкам // Вестник компьютерных и информационных технологий. 2019. №2. С. 23-29. URL: elibrary.ru/item.asp?id=37134717

6. Алёшин С.П., Бородина Е.А. Нейросетевое распознавание классов в режиме реального времени // Инженерный вестник Дона. 2013. №1. URL: ivdon.ru/ru/magazine/archive/n1y2013/1494
7. Моисеева Е.Д. Аугментация изображений // Российская наука: Тенденции и возможности. 2020. С. 129-132. URL: elibrary.ru/item.asp?id=42794332
8. Rifkin R., Aldebaro K. 2004. In Defense of One-Vs-All Classification. Journal of Machine Learning Research. №vol, 5: 101–141. URL: jmlr.org/papers/volume5/rifkin04a/rifkin04a.pdf
9. Коротеев М.В. Обзор некоторых современных тенденций в технологии машинного обучения. // E-Management. 2018. №1. С. 26-35. URL: doi.org/10.26425/2658-3445-2018-1-26-35
10. Zhou, B., Khosla A., Lapedriza A., Oliva A. and Torralba A., 2016. Learning Deep Features for Discriminative Localization. CVPR: 2921-2929. URL: cnnlocalization.csail.mit.edu/Zhou_Learning_Deep_Features_CVPR_2016_paper.pdf

References

1. Simonyan K., Zisserman A. CoRR abs. 2014. URL: arxiv.org/pdf/1409.1556.pdf
2. Puchkov E.V. Inzhenernyj vestnik Dona. 2013. №4. URL: ivdon.ru/ru/magazine/archive/n4y2013/2135
3. Konarev A. A. Gulamov Materialy konferencii Informacionnye tehnologii v upravlenii. 2020. p. 209-212. URL: itc.etu.ru/assets/files/itc-2020/papers/209.pdf
4. Kustikova V.D. Sbornik tezisov uchastnikov foruma "nauka budushhego - nauka molodyh". 2017. p. 102-103. URL: elibrary.ru/item.asp?id=32571150
5. Gorbacevich V.S., Knjaz' V.A., Mel'nichenko M.A., Stepan'janc D.G. Vestnik komp'juternyh i informacionnyh tehnologij. 2019. №2. p. 23-29. URL: elibrary.ru/item.asp?id=37134717



6. Aljoshin S.P., Borodina E.A. Inzhenernyj vestnik Dona. 2013. №1. URL: ivdon.ru/ru/magazine/archive/n1y2013/1494
7. Moiseeva E.D. Rossijskaja nauka: Tendencii i vozmozhnosti. 2020. p. 129-132. URL: elibrary.ru/item.asp?id=42794332
8. Rifkin R., Aldebaro K. Journal of Machine Learning Research. 2004. №5. p. 101–141. URL: jmlr.org/papers/volume5/rifkin04a/rifkin04a.pdf
9. Koroteev M.V. E-Management. 2018. №1. pp. 26-35. URL: doi.org/10.26425/2658-3445-2018-1-26-35.
10. Zhou, B., Khosla A., Lapedriza, A. Oliva A. and Torralba A., 2016. CVPR: 2921-2929. URL: cnnlocalization.csail.mit.edu/Zhou_Learning_Deep_Features_CVPR_2016_paper.pdf