

УДК 004:621.396.96:551.26

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПОИСКА ГИПЕРБОЛИЧЕСКИХ ОСЕЙ СИНФАЗНОСТИ ГЕОРАДИОЛОКАЦИОННЫХ СИГНАЛОВ

¹Петрова Е.А., ²Соколов К.О., ²Прудецкий Н.Д., ¹Ведехин К.Э.¹ФГАОУ ВО «Северо-Восточный федеральный университет им. М.К. Аммосова»,
Якутск, e-mail: ea.petrova@s-vfu.ru;²Институт горного дела Севера им. Н.В. Черского Сибирского отделения
Российской академии наук, Якутск, e-mail: k.sokolov@ro.ru

На сегодняшний день весьма актуальными являются задачи, связанные с созданием программных алгоритмов для анализа и интерпретации данных георадиолокационных исследований. Подобные автоматизированные системы позволяют быстро и с минимальными затратами делать предположения о свойствах горных пород, строить разрезы и выделять геологические слои с высокой степенью достоверности. Современные программные системы с использованием методов машинного обучения успешно решают задачи классификации и обнаружения объектов на изображениях. В статье приведены результаты исследования возможности применения моделей сверточной нейронной сети для поиска образов гиперболических осей синфазности на изображениях георадиолокационных радарограмм. Обучение моделей проводилось на примере данных георадиолокационных зондирований массива многолетнемерзлых горных пород. Для повышения эффективности обучения нейронной сети к набору данных были применены некоторые методы, предоставляемые библиотекой Albumentations пакета PyTorch: повышение контрастности и яркости, размытие и т.п. Использование данных методов позволило также обойтись сравнительно малым объемом данных для обучения. Итоги тестирования обученной модели по поиску осей синфазностей сигналов показали точность обнаружения 50% по метрике mAP с пороговым значением IoU, равным 0,5. Результаты данного исследования будут использованы авторами в программном приложении, созданном для исследований электрофизических свойств горных пород.

Ключевые слова: метод георадиолокации, гиперболические оси синфазности, оси дифракции, глубокие нейронные сети, сверточные нейронные сети

USING CONVOLUTIONAL NEURAL NETWORKS TO SEARCH FOR HYPERBOLIC AXES OF IN-PHASE GEORADOLocation SIGNALS

¹Petrova E.A., ²Sokolov K.O., ²Prudetskiy N.D., ¹Vedekhin K.E.¹M.K. Ammosov North-Eastern Federal University, Yakutsk, e-mail: ea.petrova@s-vfu.ru;²Institute of Mining of the North named after N.V. Cherskiy of the Siberian Branch
of the Russian Academy of Sciences, Yakutsk, e-mail: k.sokolov@ro.ru

To date, the tasks related to the creation of software algorithms for the analysis and interpretation of georadolocation research data are very relevant. Such automated systems make it possible to make assumptions about the properties of rocks quickly and at minimal cost, to build sections and to isolate geological layers with a high degree of reliability. Modern software systems using machine learning methods successfully solve the problems of classification and detection of objects in images. The article presents the results of a study of the possibility of using convolutional neural network models to search for images of hyperbolic axes of common phase in images of geo-radar radarograms. The training of the models was carried out on the example of the data of geo-radar sounding of an array of permafrost rocks. To increase the efficiency of neural network training, some methods provided by the Albumentations library of the PyTorch package were applied to the data set: increasing contrast and brightness, blurring, etc. Using these methods also made it possible to dispense with a relatively small amount of training data. The results of testing the trained model for finding the axes of common-mode signals showed a detection accuracy of 50% according to the mAP metric with an IoU threshold value of 0.5. The results of this study will be used by the authors in a software application created for the study of the electrophysical properties of rocks.

Keywords: gpr, hyperbolic axes of in-phase, diffraction axes, deep neural networks, convolutional neural networks

Метод георадиолокации в настоящее время является наиболее передовым и производительным методом при решении задач инженерно-геологических исследований. Основанием метода служит изучение отклика объектов зондируемой среды, различающихся электрическим сопротивлением и диэлектрической проницаемостью [1]. В георадиолокации одним из классических способов определения диэлектрической проницаемости среды является использование осей синфазности, образованных

локальными неоднородностями, залегающими в ее толще. Такие объекты образуют на изображениях радарограмм выраженные линии гиперболических форм. По кривизне гиперболической кривой и ее пространственно-временному положению на профиле радарограммы вычисляется диэлектрическая проницаемость среды. Это позволяет строить предположения о геологическом строении массива горных пород выше объекта дифракции, его структуре и глубине залегания.

Цель исследования – автоматизация обнаружения гиперболических осей синфазности в результатах георадиолокационных измерений массива мерзлых горных пород.

С точки зрения исследований георадиолокационные данные являются достаточно сложным объектом. Данная сложность часто обусловлена присутствием различного вида шума и помех. Вследствие чего линии осей синфазности не всегда четко визуализируются на изображениях георадиолокационных радарограмм. Поэтому обработка таких данных часто требует значительного времени. Несмотря на то, что задача интерпретации данных георадиолокации частично автоматизирована, основная нагрузка лежит на плечах оператора-геофизика. Таким образом, автоматизация процедур обработки и интерпретации данных георадиолокационных исследований является важной задачей не только для сокращения времени обработки, но и для получения достоверного результата.

Искусственные нейронные сети глубокого обучения успешно зарекомендовали себя в распознавании образов на изображениях, даже достаточно зашумленных. В данной работе для локализации гиперболических образов на изображениях радарограмм были применены искусственные нейронные сети на основе алгоритмов Faster R-CNN. Faster R-CNN – архитектура детектирования образов, разработанная в компании Microsoft, основана на сверточных нейронных сетях с оригинальной методикой системы якорей. Несмотря на появление в последние годы новых алгоритмов, показывающих более низкий процент ошибок, модели на основе Faster R-CNN стабильно не уступают им и держатся в лидерах.

Схема алгоритма Faster R-CNN [2] приведена на рис. 1.

На вход сети подается тензор изображения любого размера, который проходит через сверточные слои (Conv Layers). Из каждого слоя извлекаются признаки (Feature Maps), которые проходят через сеть регионов – Region Proposal Networks (RPN) – сеть региональных предложений. RPN – это центральный элемент Faster R-CNN. Распознанные регионы подаются затем на полносвязные слои, вместе с признаками (Feature Maps), отобранными после генерации регионов признаков. Последнее позволяет эффективно увеличивать скорость обучения сети без потери качества. В итоге на полносвязных слоях каждый прямоугольный регион получает свою оценку принадлежности объекту, также уточняются координаты регионов.

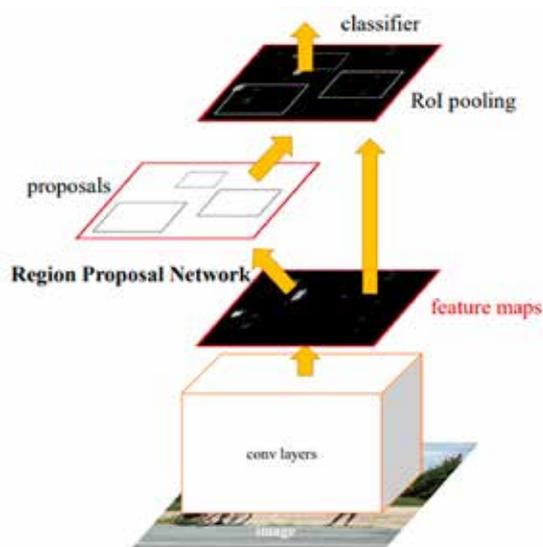


Рис. 1. Архитектура Faster R-CNN. Центральное звено – модуль RPN

Материалы и методы исследования

Для обучения и тестирования модели нейронной сети были использованы модули из программного пакета PyTorch – платформы машинного обучения с открытым исходным кодом [3]. Модули PyTorch позволяют создавать модели данных, основанные на известных архитектурах нейронных сетей глубокого обучения: resnet, alexnet, squeezeNet, vgg16, densenet, inception и т.д.

Набор данных был подготовлен в созданном авторами приложении для разметки регионов и формирования xml-файлов описаний. Структура файлов описаний была создана на основе формата Pascal VOC [4] – один из рекомендуемых форматов для обучения моделей нейронных сетей детектирования объектов. Координаты ограничивающей рамки кодировались четырьмя значениями в пикселях: x_{min} и y_{min} – координаты верхнего левого угла ограничивающей рамки, x_{max} и y_{max} – координаты нижнего правого угла. Размеры изображений были установлены равными 640x640 пикселей с разрешением 72 dpi.

При обучении сети к тренировочной выборке были применены некоторые методы, предоставляемые библиотекой Albumentations [5] пакета PyTorch:

- RandomBrightnessContrast (порог яркости = 0.5, порог контрастности = 1, $p = 0.1$);
- RandomGamma ($p = 0.2$).

Использование методов аугментации помогает бороться с переобучением и повышает производительность глубоких нейронных сетей.

Для создания набора данных было использовано 330 изображений радарограмм, на каждой из которых были размечены от 2 до 10 регионов. Возможности функций библиотеки Albumentations были применены также для увеличения объема набора данных. Такие трансформации, как Blur, HorizontalFlip, RandomBrightnessContrast, RandomGamma, позволили расширить набор данных приблизительно в 5 раз. Изображения и файлы с разметкой были разбиты на обучающую и тренировочную выборку в соотношении 6:1.

Для обучения были взяты предварительно обученные модели, что позволило сократить время обучения по сравнению с обучением с нуля. Были использованы следующие модели сверточных нейронных сетей:

- ConvNext;
- DarkNet;
- EfficientNet;
- MbV3;
- SqueezeNet;
- MobileNetV3;
- Nano;
- ResNet50.

Для корректировки весов сети в качестве оптимизаторов обучения моделей были исследованы два наиболее распространенных метода: SGD – стохастический градиентный спуск и Adam – адаптивная оценка момента. Оптимизатор SGD был выбран с использованием скользящего среднего: импульс $\text{momentum} = 0.9$ и Nesterov Accelerated Gradient в значении True, что обеспечивает быструю сходимость и уменьшает колебания. Шаг обучения был установлен минимальным: $\text{learning rate} = 0.001$. Для Adam были установлены $\text{lr} = 0.001$ и $\text{weight_decay} = 0.0005$. Был выбран оптимизатор SGD, который по результатам обучения показал лучшую сходимость.

Обучение моделей проводилось с использованием видеокарты NVIDIA GeForce RTX 3060. Настраиваемые параметры обучения:

- количество эпох – 100;
- размер пакетов – 6;
- алгоритм оптимизации – SGD ($\text{learning rate} = 0.01$; $\text{momentum} = 0.9$; $\text{Nesterov} = \text{true}$).

Для анализа результатов обучения была использована оценка точности предсказания регионов по метрике mAP (mean Average Precision) [6] – популярная метрика измерения точности алгоритмов детекторов объектов, таких как Faster R-CNN, SSD и т.д.

$$mAP = \int_0^1 p(r) dr .$$

Для оценки mAP используются следующие вычисления: precision – сколько положительных предсказаний оказались верными; recall – сколько положительных предсказаний были верными среди всех истинно верных случаев.

$$\text{precision} = TP / (TP + FP),$$

$$\text{recall} = TP / (TP + FN)$$

где TP (true positive) – предположение о принадлежности объекта классу было верным, FP (false positive) – предположение о принадлежности объекта классу было неверным, FN (false negative) – предположение о непринадлежности объекта классу было неверным.

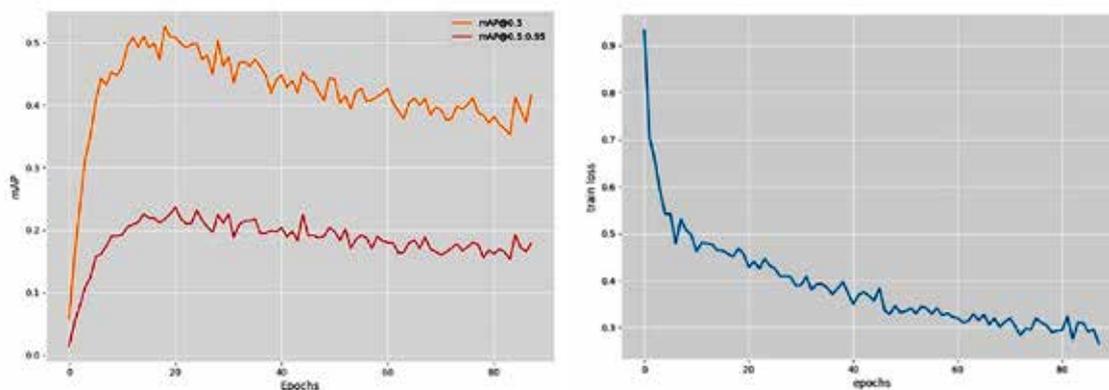
IoU (Intersection over Union) – процент перекрытия предсказанной областью его реальной области нахождения. Для расчета точности модели были определены пороги IoU: до 0.5 и от 0.5 до 0.95.

Результаты исследования и их обсуждение

Расчет точности детектирования объектов оценивался на изображениях валидационного набора данных, не участвовавших в обучении. По результатам обучения лучший показатель точности детектирования по оценке mAP показала предварительно обученная модель ResNet50 FPN V2 (рис. 2).

Глубокие нейронные сети позволяют решать задачи повышения эффективности работы сети, но при увеличении количества сверточных слоев часто возникает проблема: стабильно возрастающая точность через несколько эпох начинает резко падать. Модели классов ResNet (Residual Network) [7] были разработаны для решения этой проблемы. Главной особенностью архитектуры сетей ResNet является наличие «остаточного обучения» – дополнительная передача признаков через два следующих слоя. Это дает возможность сохранять более детальную информацию со слоев (рис. 3).

Анализ результатов тестирования обученной модели показал обнаружение множественных близко расположенных/наложенных друг на друга, а также отображенных лишь частично гиперболических осей синфазности. Пример успешного обнаружения искомых объектов приведен на рис. 4, а. Однако большинство размеченных регионов набора данных содержат размытые, малоконтрастные образы гиперболических осей, которые не были локализованы. Это объясняет небольшой процент точности, равный 50% по метрике mAP (рис. 4, б).



а)

б)

Рис. 2. Результаты обучения сети: а) оценка точности по метрике mAP; б) график ошибки обучения тренировочного набора данных

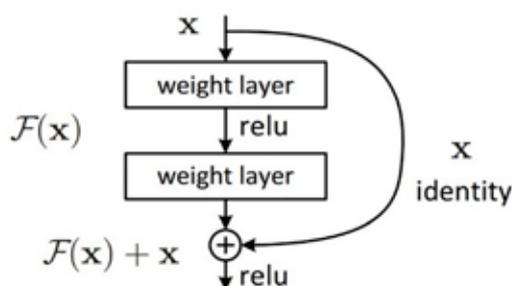
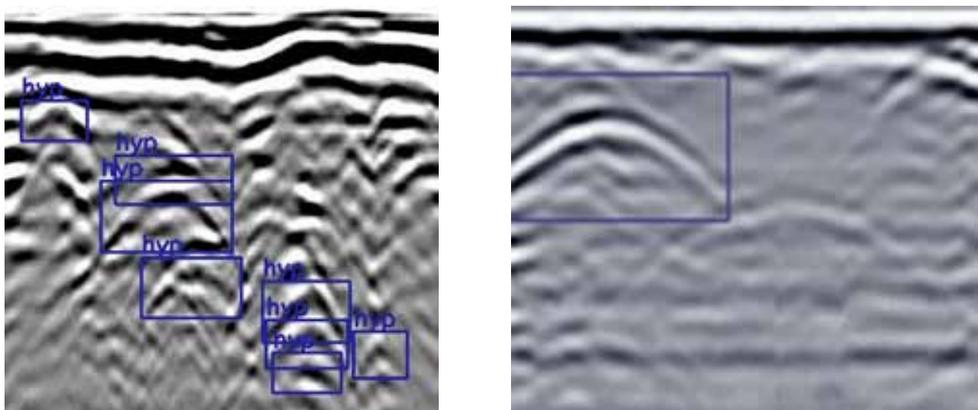


Рис. 3. Строительный блок сети ResNet



а)

б)

Рис. 4. Примеры результатов тестирования: а) с максимальным количеством локализованных образов; б) с нелокализованными образами

Заключение

Обработка результатов георадиолокационных исследований является сложной технической задачей из-за наличия на радарограммах большого количества помех и шума различного характера: переотра-

жение, рассеяние волн и пр. Состав исследуемого грунта также влияет на картину сигналов на радарограмме: например, во влажных глинистых грунтах радиоволны быстро затухают. По этим причинам задача достоверного определения состава под-

почвенной среды традиционно решается в комплексе с другими дорогостоящими методами: сейсморазведка, бурение и пр.

На сегодняшний день существует достаточно много успешных исследований в использовании нейронных сетей для автоматизации обработки георадарных данных. Авторы работы также применили алгоритмы искусственных нейронных сетей глубокого обучения для интерпретации результатов георадиолокационных измерений.

Полученное значение эффективности модели составило 50% при IoU = 0.5 и 20% при допуске IoU от 0.5 до 0.95, что является показателем малой успешности результатов обучения по метрике mean Average Precision. В данное время авторы продолжают работу по исследованию моделей сверточных нейронных сетей для повышения процента точности обнаружения. В частности, ведется работа по увеличению количества и качества данных и подбору/настройке параметров и гиперпараметров сети. Кроме того, планируется деление общего класса набора данных на несколько классов в зависимости от расположения осей синфазности на временной оси. Данное изменение объясняется тем, что характер рисунка осей синфазности при увеличении глубины сканирования меняется: изображение становится размытым, бледным и нечетко визуализируемым. Такие изображения лучше выделить в отдельный класс.

Объем обучающего набора данных был расширен благодаря использованию методов аугментации программного пакета PyTorch. Кроме того, за счет применения некоторых методов аугментации также и при обучении модели удалось повысить процент точности детектирования. Например, показатель точности изменился с 25 до 40% уже только

при добавлении функции яркости/контрастности (RandomBrightnessContrast).

При достижении оптимальной точности локализации искомым объектам результаты работы планируется использовать авторами в приложении, предназначенном для исследования электрофизических свойств грунта. Таким образом, результаты данного исследования помогут эффективно и на более качественном уровне проводить работы по выявлению свойств горных пород и построению разрезов геологических слоев.

Список литературы

1. Петрова Е.А., Соколов К.О. Автоматизация выделения гиперболических осей синфазности георадарных сигналов, дифрагированных от локальных неоднородностей // Современные наукоемкие технологии. 2020. № 11. С. 61–67.
2. Shaoqing Ren et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in neural information processing systems*. 2015. Т. 28. [Электронный ресурс]. URL: <https://proceedings.neurips.cc/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html> (дата обращения: 14.12.2022).
3. Ян Пойнтер. Программируем с PyTorch: Создание приложений глубокого обучения // СПб.: Питер, 2020. 256 с.
4. Everingham M., Winn J. The PASCAL visual object classes challenge 2012 (VOC2012) development kit. *Pattern Anal. Stat. Model. Comput. Learn., Tech. Rep.* 2012. Т. 2007. [Электронный ресурс]. URL: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/devkit_doc.pdf (дата обращения: 14.12.2022).
5. Buslaev A. et al. Albumentations: fast and flexible image augmentations. *Information*. 2020. Т. 11. No. 2. [Электронный ресурс]. URL: <https://www.mdpi.com/2078-2489/11/2/125> (дата обращения: 14.12.2022).
6. Padilla R., Netto S.L., Da Silva E.A.B. A survey on performance metrics for object-detection algorithms // 2020 international conference on systems, signals and image processing (IWSSIP). IEEE, 2020. [Электронный ресурс]. URL: <https://ieeexplore.ieee.org/abstract/document/9145130> (дата обращения: 14.12.2022).
7. He K. et al. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. С. 770–778. [Электронный ресурс]. URL: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html (дата обращения: 28.12.2022).