

ИСПОЛЬЗОВАНИЕ МЕТОДОВ АУГМЕНТАЦИИ ДАННЫХ В ЗАДАЧЕ ДЕТЕКТИРОВАНИЯ КАРЬЕРОВ ПРИ ПОМОЩИ НЕЙРОННОЙ СЕТИ ГЛУБОКОГО ОБУЧЕНИЯ DEEPLABV3+

А. М. Сорока, А. В. Хлебко

Белорусский государственный университет, Минск, Беларусь

E-mail: soroka.a.m.@gmail.com, antonkhlebko@gmail.com

Разработан метод детектирования карьеров на спутниковых снимках с использованием нейронной сети глубокого обучения deeplabv3+. Решена задача бинарной семантической сегментации спутниковых снимков для определения границ карьеров на снимке. Исходными данными для обучения нейронной сети глубокого обучения были 965 размеченных карьеров на спутниковых снимках Республики Беларусь. Итоговая точность составила 83.76%.

Ключевые слова: нейронные сети глубокого обучения; семантическая сегментация; дистанционное зондирование; определение границ карьеров.

Введение. Задачи дистанционного зондирования поверхности как правило характеризуются относительно небольшим количеством исходных данных, что ограничивает использование методов глубокого обучения в такого рода задачах.

Для улучшения характеристик методов глубокого обучения широко используются методы аугментации данных, которые модифицируют обучающие примеры таким образом, чтобы количество данных примеров было достаточным для построения нейронных сетей глубокого обучения. В то же время методы аугментации могут приводит к значительным искажениям семантики исходных данных, что в свою очередь может снизить характеристики методов глубокого обучения в данных задачах.

Настоящая работа посвящена исследованию влияния методов аугментации данных на метрику качества на примере задачи детектирования границ карьеров на спутниковом снимке с помощью нейронных сетей глубокого обучения (НСГО) путем решения задачи семантической сегментации спутникового снимка на два класса: «карьер», «фон».

Нейронная сеть глубокого обучения deeplabv3+. Для решения задачи детектирования границ карьеров использовалась полносверточная НСГО deeplabv3+. НСГО deeplabv3+ состоит из двух частей: кодер и декодер. В качестве кодера используется сверточная НСГО [3] со следующей за ней пространственной пирамидой [4], которая обеспечивает сохранение пространственных признаков для лучшего восстановления изображения и, как следствие, получения более четких границ.

Декодер состоит из последовательных слоев повышения дискретизации и слоев свертки, необходимых для корректировки результатов с предыдущих слоев.

Процедура обучения нейронной сети глубокого обучения. В качестве исходных данных использованы 965 размеченных на спутниковых снимках карьеров, из-за недостаточного объема данных в качестве декодера сети использовалась НСГО Resnet, обученная на наборе данных ImageNet.

Для оценки точности полученного решения использовалось значение mIoU (формула 1) при ограниченном количестве итераций обучения с использованием процедуры скользящего контроля с разделением исходного набора данных на обучающий и набор данных для контроля с количеством изображений 715 и 250, соответственно.

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{P_i \cap L_i}{P_i \cup L_i}, \quad (1)$$

где P_i - множество пикселей, которые сеть отметила, как относящиеся к i -ому классу, L_i - множество пикселей, относящихся к i -ому классу из эталонной разметки, N — число классов.

Так как данные не являются сбалансированными по количеству пикселей каждого класса, используется функция потерь, заданная формулой 2, в которой весовые коэффициенты для каждого класса определены экспериментальным путем [4]:

$$L = \sum_{i=1}^W \sum_{j=1}^H \sum_{n=1}^N A_n T_{ijn} \ln \left(\frac{O_{ij}}{\sum_{l=1}^W \sum_{k=1}^H O_{lk}} \right), \quad (2)$$

где W – ширина изображения, H – высота изображения, N – число классов, A_n – экспериментально подобранный для каждого класса, весовой коэффициент, T_{ijn} – значение пикселя в оригинальной разметке, O_{ij} – элемент выхода НСГО.

В рассмотренном случае число классов $N = 2$, так как в задаче только два класса: «карьер» и «фон».

Для определения оптимальных весовых коэффициентов проведено экспериментальное исследование зависимости точности решения от весовых коэффициентов A_n при фиксированном количестве эпох, равным 50. Оптимальное значение вектора весовых коэффициентов составило [1, 5].

Влияние аугментации на процесс обучения. Для определения влияния методов аугментации на точность предложенного решения, было проведено экспериментальное исследование, которое состояло из серии экспериментов. Используемые методы аугментации [6] и полученная с их помощью точность представлены в таблице.

Операция горизонтального и вертикального отражения применялась к изображениям с вероятностью 0.5. Масштабирование применялось к каждому изображению со случайным коэффициентом масштабирования от 0.75 до 2. Поворот изображений осуществлялся на случайный угол в диапазоне от 0° до 90°. Изменение параметров данных методов аугментации не приводило к изменению точности решения. Изменение параметров следующих методов аугментации наложение полупрозрачного прямоугольника, изменение контраста изображения, эластичная деформация [7], приводило к существенному изменению точности решения. Параметры данных методов выбираются случайным образом из заданных диапазонов.

Таблица

Зависимость точности решения от используемых методов аугментации

Метод аугментации	mIoU, %
Отражение, масштабирование, вырезание части изображения	75.13
Поворот	77.01
Наложение полупрозрачного прямоугольника	79.02
Изменение контраста изображения	79.66
Эластичная деформация	77.35

Заключение. Рассмотрено влияние различных методов аугментации на метрику качества.

Для решения проблемы несбалансированности данных была предложена формула для расчёта значения функции потерь с использованием весовых коэффициентов классов для увеличения вклада пикселей, принадлежащих к классу «карьер». Коэффициенты классов «фон» и «карьер», в результате эксперимента по определению лучших коэффициентов, были выбраны 1 и 5.

Для улучшения точности были использованы следующие методы аугментации данных:

1. Отражение, масштабирование, вырезание части изображения.

Данные методы аугментации позволили получить метрику качества mIoU 75.13%.

2. Вращение.

Улучшение mIoU составило 2%.

3. Наложение полупрозрачного прямоугольника.

Улучшение mIoU составило 2%.

4. Изменение контраста.

Улучшение mIoU составило 0.66%.

5. Эластичная деформация.

Данный метод аугментации не улучшил метрику точности, поэтому не был использован в дальнейшем.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Chen L., Papandreou G., Kokkinos I. et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2018. Vol. 40. № 4. P. 834–848.
2. Long J., Shelhamer E., Darrell T. Fully Convolutional Networks for Semantic Segmentation // 2015 IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 3431–3440. DOI: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965)
3. Albawi S., Mohammed T. A., Al-Zawi S. Understanding of a convolutional neural network // 2017 International Conference on Engineering and Technology (ICET). 2017. DOI: [10.1109/ICEngTechnol.2017.8308186](https://doi.org/10.1109/ICEngTechnol.2017.8308186)
4. Chen L., Zhu Y., Papandreou G. et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation // ECCV 2018: Computer Vision – ECCV 2018. P. 833–851. DOI: https://doi.org/10.1007/978-3-030-01234-2_49
5. Xia K., Yin H., Qian P. et al. Liver Semantic Segmentation Algorithm Based on Improved Deep Adversarial Networks in Combination of Weighted Loss Function on Abdominal CT Images // IEEE Access. 2019. Vol. 7. P. 96349-96358.
6. Fawzi A., Samulowitz H., Turaga D., Frossard P. Adaptive data augmentation for image classification // IEEE International Conference on Image Processing (ICIP). 2016. P. 3688–3692. DOI: [10.1109/ICIP.2016.7533048](https://doi.org/10.1109/ICIP.2016.7533048)
7. Castro E., Cardoso J. S., Pereira J. C. Elastic deformations for data augmentation in breast cancer mass detection // 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). 2018. P. 230–234. DOI: [10.1109/BHI.2018.8333411](https://doi.org/10.1109/BHI.2018.8333411)