

ИСПОЛЬЗОВАНИЕ АУГУМЕНТАЦИЙ ПРИ СЕГМЕНТАЦИИ ЗДАНИЙ НА СПУТНИКОВЫХ СНИМКАХ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ U-NET

Жук А. О., Абламейко С. В.

*Белорусский государственный университет, Минск, Беларусь,
e-mail: ales.zhook@gmail.com*

Введение

Обнаружение и выделение зданий на спутниковых снимках – актуальная задача для различных приложений: построение карт местности, развитие инфраструктуры города, поиск незаконно построенных объектов. Обработать вручную большое число изображений потребует немало времени и ресурсов. Поэтому разрабатываются алгоритмы автоматической сегментации спутниковых снимков. Задача автоматического выделения усложняется плохими погодными условиями, разнообразием форм и цветов выделяемых конструкций.

Разработано большое число алгоритмов по выделению и сегментации зданий, многие из которых базируются на нейронных сетях. Классическая нейронная сеть для сегментации объектов – U-NET. Впервые была применена в 2015 году для сегментации медицинских изображений [1]. Часто используется модифицированная сеть, например, в работе [2] продемонстрирована возможность использования предобученной сети. Описана нейронная сеть U-NET, у которой в качестве расширяющего пути использовалась VGG11, обученная на весах ImageNet [3].

Мы, в нашей работе, предлагаем внести ещё дополнительные модификации. Добавить на этапе обучения такие трансформации как добавление шума, изменение яркости и контраста изображения, трансформация перспективы. И показываем, как это позволило улучшить результат сегментации.

Обучающее множество

Для решения задачи сегментации зданий использовалось множество, описанное в [4]. Обучающее множество (аналогично и тестовое) содержит 180 цветных трёхканальных изображений различных поселений размера 5000 x 5000 пикселей с пространственным разрешением 0.3 метра. Пример изображения и его маски приведён на рисунке 1.

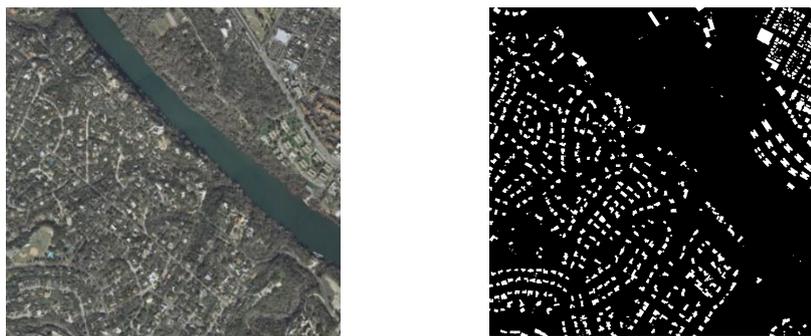


Рис. 1. Пример изображения обучающего множества

Следует отметить, что изображения в обучающем множестве и множестве для тестирования содержат снимки разных городов. Обучение нейронной сети на снимках одних городов, а тестирование на других позволяет понять как алгоритм адаптируется для других данных похожей природы.

Архитектура сети и обучение

В данной работе использовалась архитектура, сходная с сетью, описанной в [2].

Параметры обучения сети:

1. Исходное множество было разбито на два: обучающее 150 и валидационное 30 изображений соответственно. На каждой итерации из каждого изображения обучающего множества случайным образом вырезается сегмент размера 768 x 768, все такие сегменты группируются в батчи и передаются на вход сети. Размер батча был выбран 8.

2. В качестве функции стоимости использовался focal loss [5].

3. На этапе создания обучающего множества дополнительно производились аугментации изображений, описанные в следующем разделе.

4. Мы использовали алгоритм оптимизации Adam [6] с коэффициентом обучения 0.0001. Во время тренировки модели коэффициент обучения несколько раз уменьшали.

Аугментации

Стандартные аугментации данных, которые используются в задачах сегментации спутниковых снимков – это повороты на углы кратные 90 градусам и отражения. В данной работе мы предлагаем добавить дополнительные трансформации изображений. Ниже перечислен итоговый набор аугментаций:

- Поворот на случайный угол кратный 90 градусам
- Вертикальное отражение
- Горизонтальное отражение
- Добавление Гауссовского шума
- Случайное изменение hsv в некотором диапазоне.
- Случайное изменение яркости и контраста изображения.
- Трансформация перспективы изображения

На каждой итерации обучения для конкретного изображения вероятность того, что некоторая аугментация будет применена составляла 0.25.

Для тестирования и реализации аугментаций мы воспользовались библиотекой описанной в работе [7].

Эксперименты и результаты

Для проведения экспериментов мы воспользовались возможностями облачных вычислений, предоставляемых Google Colaboratory [8]. Нам была выделена видеокарта Tesla P100 с 16GB памяти.

Результат оценивался с помощью метрик: точность и коэффициент Жаккара.

$$A = \frac{1}{n * m} \sum_{i,j}^{n,m} \begin{cases} 1, \text{ где } y_{ij} = \bar{y}_{ij} \\ 0 \end{cases} \quad (1)$$

$$J = \frac{1}{n*m} \sum_{i,j=1}^{n,m} \frac{y_{ij} * \bar{y}_{ij}}{y_{ij} + \bar{y}_{ij} - y_{ij} * \bar{y}_{ij}}, \quad (2)$$

где y_{ij} – истинные значения пикселя, \bar{y}_{ij} – значение, предсказанное моделью, $n * m$ – размер изображения.

В таблице ниже приведено сравнение результатов нескольких экспериментов: с использованием стандартных аугментаций, дополнительных аугментаций и решения [98].

Табл. 1. Результаты экспериментов

Модель, описанная выше + стандартные аугментации						
	Беллингхем	Блумингтон	Инсбрук	Сан-Франциско	Тироль	Общий
коэф. Жаккара	69.03	73.44	74.50	75.02	76.54	74.12
Точность	96.89	97.45	96.88	91.31	97.87	96.08
Модель, описанная выше + дополнительные аугментации						
коэф. Жаккара	69.95	75.19	75.46	77.29	77.69	75.78
Точность	96.96	97.61	97.06	92.26	98.00	96.38
Решение [9]						
коэф. Жаккара	69.75	72.04	74.64	74.55	77.40	73.91
Точность	96.77	97.13	96.83	91.14	97.92	95.96

Исходя из метрик выше можно сделать вывод, что для улучшения результата выделения дискретных объектов на спутниковых снимках, следует использовать различные трансформации изображений.

Литература

1. Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox. // MICCAI. – Springer, 2015 – pp. 234–241.
2. Iglovikov, V., TerausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation / V. Iglovikov, A. Shvets, – arXiv preprint, 2018 – 5p.
3. Russakovsky, O., ImageNet Large Scale Visual Recognition Challenge / O. Russakovsky, J. Deng, H. Su, et al. // International Journal of Computer Vision – Springer / 2015 – pp. 211-252
4. Maggiori, E., Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark / E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez. // IEEE International Symposium on Geoscience and Remote Sensing Symposium (IGARSS), July 23-28 – IEEE, 2017 – pp. 3226– 3229
5. Lin, T., Focal Loss for Dense Object Detection / T. Lin, P. Goyal, R. Girshick, K. He and P. Dollár. // IEEE International Conference on Computer Vision (ICCV), Venice, 22-29 Oct. – IEEE, 2017 – pp. 2999-3007
6. Kingma, D., Adam: A method for stochastic optimization / D. Kingma, J. Ba // International Conference on Learning Representations., Banff, Canada, April 14-16, 2014 – 15p.
7. Buslaev, A., Albumentations: Fast and flexible image augmentations / A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin // Information – MDPI, 2020 – Vol. 11. – 4 p.
8. Google Colaboratory. [Electronic resource]. – Mode of access: <https://colab.research.google.com>. Date of access: 06.09.2020.
9. Girard, N., Polygonal Building Segmentation by Frame Field Learning / Nicolas Girard, Dmitriy Smirnov, Justin Solomon, and Yuliya Tarabalka, – arXiv preprint, 2020 – 30p.