

Распознавание и сегментация земной поверхности с использованием свёрточных нейронных сетей

Я.А. Шуринова
Самарский национальный
исследовательский университет
им. академика С.П. Королева
Самара, Россия
yana.shurinova009@gmail.com

А.А. Белоусов
Самарский национальный
исследовательский университет
им. академика С.П. Королева
Самара, Россия
belousov.aa@ssau.ru

Аннотация — В работе исследуется влияние различных факторов на результат обучения свёрточных нейронных сетей при решении задачи сегментации спутниковых изображений. Рассматриваются разные подходы к решению задачи, в том числе различные конфигурации нейронных сетей и интерпретации исходной задачи. Установлено, что нейронная сеть архитектуры U-Net с кодировщиком ResNet-34, использующей в функции потери комбинацию CCE и мягкий индекс Жаккара в равных пропорциях показывает наилучшие результаты для рассматриваемой задачи, также установлено, что использование функции активации LeakyReLU дает лучший результат для расширенной архитектуры U-Nets.

Ключевые слова—сегментация спутниковых изображений, машинное обучение, свёрточные нейронные сети.

I. ВВЕДЕНИЕ

В настоящее время решение задач распознавания и сегментации изображений земной поверхности может быть использовано при строительстве дорог для более качественного их расположения в зависимости от грунта, также при построении карт местности [1]. В качестве подзадач данной темы встречаются следующие: моделирование паводков - нахождение воды/водоемов на изображении, используется для предотвращения различных утечек и наводнений; другая подзадача – это нахождение лесов на изображении – задача, решение которой может быть использовано для предотвращения лесных пожаров, например детектирование очагов возгорания среди лесных угодий и их сегментирование [2, 3].

В данной работе рассматривается решение задач распознавания и сегментации изображений земной поверхности с помощью свёрточных нейронных сетей. Исследуются различные архитектуры нейронных сетей, а также влияние некоторых факторов на результат.

II. РАСПОЗНАВАНИЕ И СЕГМЕНТАЦИЯ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Данная работа содержит в себе исследование влияния таких факторов, как функция активации, функция потери, архитектура на результат обучения нейронной сети.

В данной работе исследования проводились на наборе данных, сформированном для международных соревнований DEEPGLOBE [4]. Всего в наборе данных присутствует около 100000 изображений: спутниковых изображений и их масок, которые представляет собой сегментированное изображение, где определенным

цветом обозначена конкретная область, например, голубым-города, синим-водоёмы и др. Максимальное число классов, на которое изображение может быть сегментировано, равняется 7: водоёмы, города, пастбища, леса, пустоши, земли с/х, неизвестная область. Данные для обучения, подбора гиперпараметров и тестирования поделены в отношении 70:20:10.

Для решения поставленной задачи используются свёрточные нейронные сети. Одна из них носит название расширенная архитектура U-Nets и часто используется для семантической сегментации [5]. Были исследованы три конфигурации при различных функциях активации: ReLU, LeakyReLU, SiLU, результаты данных нейронных сетей на тестовой выборке с использованием метрики mIoU показаны в сводной таблице результатов 1. В качестве функции потери использована категориальная перекрестная энтропия.

Лучший результат показала архитектура, использующая функцию активацию LeakyReLU, это объясняется тем, что функция LeakyReLU не подвержена проблеме, характерной для функции активации ReLU, связанной с обнулением некоторых весов. Более того, можно заметить, что значение метрики mIoU для функций активации ReLU и SiLU значительно ниже, причем ReLU на отрицательной оси принимает нулевое значение, а SiLU немонотонна. В тоже время функция активации LeakyReLU, для которой значение метрики mIoU более чем в два раза выше по сравнению с предыдущими функциями активации, монотонна и не равна нулю на отрицательной оси. Таким образом, отрицательные значения карт признаков несут в себе важную информацию.

Вторая исследуемая архитектура – архитектура U-Net с кодировщиком ResNet-34 [6]. Данная архитектура предполагает использование сочетания категориальной перекрестной энтропии и мягкого индекса Жаккара/Lovasz-softmax в качестве функции потери. Для исследования была выбрана первая комбинация, общая формула функции потери представлена ниже.

$$L(w) = (1-a) \cdot CCE(w) + a \cdot L'(w),$$

где L' – функция Жаккара, CCE - категориальная перекрестная энтропия, a - параметр-константа.

Были обучены нейронные сети данной архитектуры при функции активации ELU и разных a . Результаты представлены в таблице I. Наилучший результат при $a = 0,5$, что означает применение в равном объеме функции потери категориальной перекрестной энтропии и мягкого индекса Жаккара (Soft Jaccard index). Можно сделать вывод о том, что значения функции потери Жаккара, для которой в равной степени важно количество предсказанных пикселей для каждого из классов, и значения функции потери CCE , для которой

важна доля предсказанных пикселей вне зависимости от их класса, должны быть в равной степени значимы для достижения лучшего результата.

Третья исследуемая архитектура – Deeplabv3+ [7]. Улучшенная за счет выделения декодировщика версия архитектуры Deeplabv3. Результат также представлен в таблице I.

ТАБЛИЦА I. РЕЗУЛЬТАТЫ НЕЙРОННЫХ СЕТЕЙ

Архитектура	Доп.	Loss	mIoU
U-Nets	ReLU	1,1956	0,1628
U-Nets	LeakyReLU	0,7267	0,4057
U-Nets	SiLU	2,4397	0,1811
U-Net с ResNet-34	$a = 0,25$	0,5751	0,3916
U-Net с ResNet-34	$a = 0,5$	0,5833	0,4580
U-Net с ResNet-34	$a = 0,75$	0,7032	0,3991
Deeplabv3+	-	0,8102	0,4187

В качестве дополнительного исследования был рассмотрен тот же набор данных исходных спутниковых изображений, а маски были пропущены через фильтр Гаусса. Пример матрицы вероятностей для одного из классов до и после применения фильтра Гаусса показан на рис. 1.

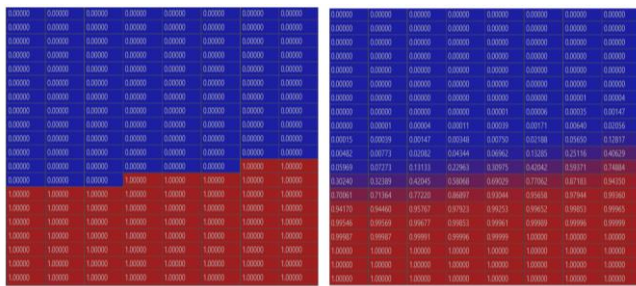


Рис. 1. Пример матрицы вероятностей для одного из классов до и после применения фильтра Гаусса

Результаты нейронных сетей на полученном наборе данных показали значения ниже предыдущих, за счет того, что вероятности принадлежности пикселя к определенному классу на границах стали менее четкими.

Таким образом, наилучший результат показала нейронная сеть с архитектурой U-Net с кодировщиком ResNet-34 с комбинацией категориальной перекрестной энтропией и мягким индексом Жаккара в равных пропорциях в качестве функции потерь. Значение метрики mIoU в данном случае равно 0,4580.

Пример использования полученной нейронной сети на самостоятельно сделанных изображениях при помощи Google Maps показан на рис. 2.

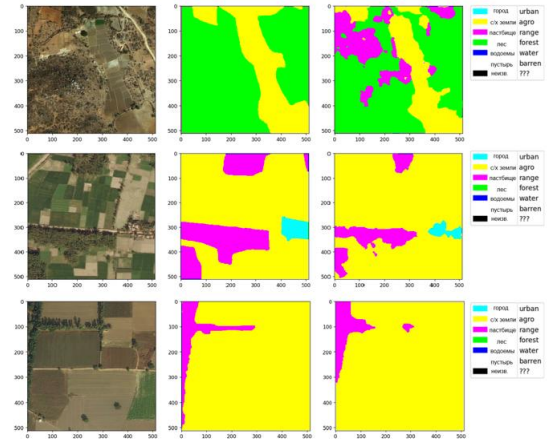


Рис. 2. Пример работы лучшей модели нейронной сети

Слева на рисунке 2 показан исходный спутниковый снимок, по центру – эталонная маска, справа выход нейронной сети.

Стоит сказать, что выходное изображение обрабатывается фильтром. Это необходимо для того, чтобы избавиться от небольших погрешностей.

ЗАКЛЮЧЕНИЕ

В данной работе проведены исследования влияния функции активации, архитектуры нейронной сети, применение фильтра Гаусса на результат распознавания и сегментации спутниковых изображений. Установлено, что применение фильтра Гаусса на наборе данных отрицательно влияет на результат обучения и тестирования нейронной сети; наиболее подходящей конфигурацией архитектуры нейронной сети среди U-Nets, U-Net с кодировщиком ResNet-34, Deeplabv3+ является нейронная сеть со второй архитектурой и линейной комбинацией SSE и мягким индексом Жаккара в качестве функции потерь. Значение метрики mIoU равно 0,458.

ЛИТЕРАТУРА

- [1] Kuo, T.S. Deep Aggregation Net for Land Cover Classification / T.S. Kuo, K.S. Tseng, J.W. Yan, Y.C. Liu, Y.C. Wang // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. – 2018. – P. 252-256.
- [2] Tian, C. Dense fusion classmate network for land cover classification / C. Tian, C. Li, J. Shi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. – 2018. – P. 192-196.
- [3] Seferbekov, S. Feature pyramid network for multi-class land segmentation / S. Seferbekov, A. Buslaev, V. Iglovikov, A. Shvetsy // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. – 2018. – P. 272-275.
- [4] Land cover classification [Electronic resource]. – Access mode: <https://www.kaggle.com/datasets/balraj98/deeplabv3-land-cover-classification-dataset> (20.12.2022).
- [5] Ghosh, A. Stacked U-Nets for Ground Material Segmentation in Remote Sensing Imagery / A. Ghosh, M. Ehrlich, S. Shah, L. Davis, R. Chellappa // Computer Society Conference on Computer Vision and Pattern Recognition Workshops. – 2018. – P. 252-254.
- [6] Rakhlin, A. Land Cover Classification from Satellite Imagery With U-Net and Lovasz-Softmax Loss / A. Rakhlin, A. Davydov, S. Nikolenko // Computer Society Conference on Computer Vision and Pattern Recognition Workshops. – 2018. – P. 262-266.
- [7] Rakshit, S. Multiclass semantic segmentation using DeepLabV3+ / S. Rakshit [Electronic resource]. – Access mode: https://keras.io/examples/vision/deeplabv3_plus/ (20.12.2022).