

ПРОГРАММНОЕ СРЕДСТВО ДЛЯ СЕГМЕНТАЦИИ СПУТНИКОВЫХ СНИМКОВ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ DEEPLABV3+

Антоненко А.В., студент

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Перцев Д.Ю. – канд. техн. наук, доцент

Описано программное средство для попиксельной классификации спутниковых снимков. Обсуждается модель DeepLabV3+ с энкодером ResNet-50, анализируются результаты работы модели.

Спутниковые снимки применяются при мониторинге окружающей среды, в сельском хозяйстве и картографии [1]. Ручной анализ таких изображений трудоёмок, поэтому возникает потребность в автоматизации с помощью нейросетевых методов. В данной работе для попиксельной классификации спутниковых снимков используется архитектура DeepLabV3+ [2] с энкодером ResNet-50 [3], предобученным на ImageNet. Модуль ASPP извлекает контекст на нескольких масштабах, а декодер восстанавливает пространственные границы объектов.

Энкодер ResNet-50 состоит из свёрточных остаточных блоков, каждый из которых реализует обходное (skip) соединение, позволяющее градиенту проходить через десятки слоёв без выраженного затухания. Это особенно важно при обучении на спутниковых снимках, где требуется одновременно учитывать локальные детали (границы объектов, тонкие линейные структуры) и крупномасштабный контекст сцены. В архитектуре DeepLabV3+ последние блоки энкодера работают с атрозными свёртками (dilated convolutions) с коэффициентами расширения 6, 12 и 18, что увеличивает рецептивное поле без дополнительного понижения пространственного разрешения.

Ключевым элементом модели выступает модуль ASPP (Atrous Spatial Pyramid Pooling), который параллельно применяет несколько атрозных свёрток с разными коэффициентами расширения и ветвь глобального усредняющего пулинга. После этого признаки объединяются и проходят последующую обработку, благодаря чему сеть одновременно учитывает мелкие текстурные признаки и глобальные пространственные зависимости.

Декодер DeepLabV3+ принимает высокоуровневые признаки из ASPP и низкоуровневые карты из ранних слоёв энкодера, после чего объединяет их через конкатенацию и последовательность свёрток 3×3 . Такое объединение позволяет восстанавливать чёткие границы классов и уменьшать ошибки на стыках объектов. На выходе используется поточечная свёртка 1×1 , формирующая карту вероятностей для каждого класса земного покрова. Далее применяется операция выбора максимальной вероятности по каналному измерению, что даёт итоговую попиксельную разметку снимка.

Модель обучалась на семи классах земного покрова: фон, город, облака, поля, песок, деревья, вода. Функция потерь состоит из трёх слагаемых: кросс-энтропия ($\alpha = 0,3$), индекс Жаккара ($\beta = 0,5$) и фокальная потеря ($\gamma = 0,2$). Веса классов рассчитаны обратно пропорционально частоте встречаемости, что компенсирует дисбаланс между доминирующими и редкими категориями. Аугментации нарастают постепенно – от простых отражений и поворотов на 90° в начале обучения до масштабирования, яркостных искажений и упругих деформаций к концу.

Обучение выполнялось на GPU NVIDIA GeForce RTX 5060 с 8 ГБ видеопамяти. Установлен эффективный батч 16 (физический $4 \times$ аккумуляция 4), включена смешанная точность (формат bfloat16), выбран планировщик OneCycleLR с начальным значением LR, равным $5 \cdot 10^{-4}$. Всего было запланировано 70 эпох при разрешении 384×384 ; однако ранняя остановка сработала на 53-й эпохе.

Лучший средний IoU составил 0,8679 на 29-й эпохе, а показатель F-score – 0,9220. Наивысшую точность IoU показал класс «Песок» со значением 0,98, в то время как для классов «Город» и «Облака» этот показатель составил около 0,88. Слабее всего распознается «Фон» со значением 0,46.

Для практического применения создан интерфейс пользователя с использованием PyQt6, в котором реализована загрузка снимков, пакетная обработка, отображение масок, а также статистики по классам. Поддерживается ансамбль нескольких чекпоинтов.

Использование DeepLabV3+ с комбинированной функцией потерь обеспечило высокую точность классификации различных категорий земного покрова. Разработанный интерфейс на PyQt6 оптимизирует анализ спутниковых данных, объединяя обработку и визуализацию статистики, а ансамблирование моделей гарантирует устойчивость системы к искажениям в реальных условиях.

Список использованных источников:

1. Chen, L.-C. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation / L.-C. Chen, Y. Zhu, G. Papandreou // *Proceedings of the ECCV*. – 2018. – P. 801–818.
2. Iakubovskii, P. Segmentation Models Pytorch [Электронный ресурс] / P. Iakubovskii. – Режим доступа : https://github.com/qubvel/segmentation_models.pytorch. – Дата доступа: 25.03.2026.
3. He, K. Deep Residual Learning for Image Recognition / K. He [et al.] // *Proceedings of the CVPR*. – 2016. – P. 770–778.