

Мачука Мендоса Кристиан Родриго (Эквадор)
Томский политехнический университет, г. Томск
Научный руководитель: Марков Николай Григорьевич,
д-р техн. наук, профессор

СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ СНИМКОВ ПОВРЕЖДЕННЫХ ВРЕДИТЕЛЯМИ ХВОЙНЫХ ДЕРЕВЬЕВ С ИСПОЛЬЗОВАНИЕМ ТРАНСФЕРНОГО ОБУЧЕНИЯ

Аннотация. Обнаружение повреждённых насекомыми - вредителями деревьев на ранних стадиях является актуальной задачей. В работе представлены две модели полносверточных нейронных сетей на основе классической модели U-Net и использовано трансферное обучение для ускорения процесса их обучения и повышения качества семантической сегментации снимков поврежденных деревьев с беспилотных летательных аппаратов. Приведены результаты исследования эффективности таких моделей полносверточных нейронных сетей.

Введение

Вспышки массового размножения насекомых-вредителей наносят непоправимый вред хвойным лесам. Поэтому актуальными являются оперативный мониторинг таких лесов с помощью беспилотных летательных аппаратов (БПЛА) и семантическая сегментация снимков деревьев, собранных с использованием БПЛА, для выявления деревьев на начальных стадиях поражения вредителями.

Семантическая сегментация снимков относится к задаче присвоения меток классов объектов каждому пикселю изображения. Присвоение таких меток является сложной задачей. В последние годы семантическая сегментация снимков выполняется путем использования сверточных нейронных сетей (СНС). Однако, для обучения СНС требуется большое количество эталонных снимков высокого качества, которые сложно создать из-за высокой стоимости съемки и разметки.

В работе предлагаются две модели полносверточных нейронных сетей для семантической сегментации снимков поврежденных вредителями деревьев кедра *P. sibirica* на основе классической СНС U-Net [1]. Формируются два набора данных: первый набор для полностью контролируемого обучения двух предлагаемых моделей СНС, а второй - для обучения этих моделей на основе трансферного обучения (ТО). Проводятся исследования эффективности обученных моделей.

Исходные данные для исследований

В качестве исходных данных использовались RGB-изображения крон деревьев *P. sibirica*, полученные с помощью БПЛА и собранные

в панорамы. Изображения имеют пространственное разрешение около 3 см и 10 см в зависимости от высоты съемки. Разметка изображений для создания эталонных карт сегментации (масок) является сложной задачей. При ее решении экспертами - этмологами выделялось пять классов жизненного состояния деревьев *P. Sibirica*, поврежденных союзным короедом: 1 – «условно здоровое»; 2 – «свежезаселенное»; 3 – «с усохшей вершиной»; 4 – «свежий сухостой»; 5 – «старый сухостой» (рисунок 1). К классу 0 – «фон» относились деревья других пород и иные объекты земной поверхности.

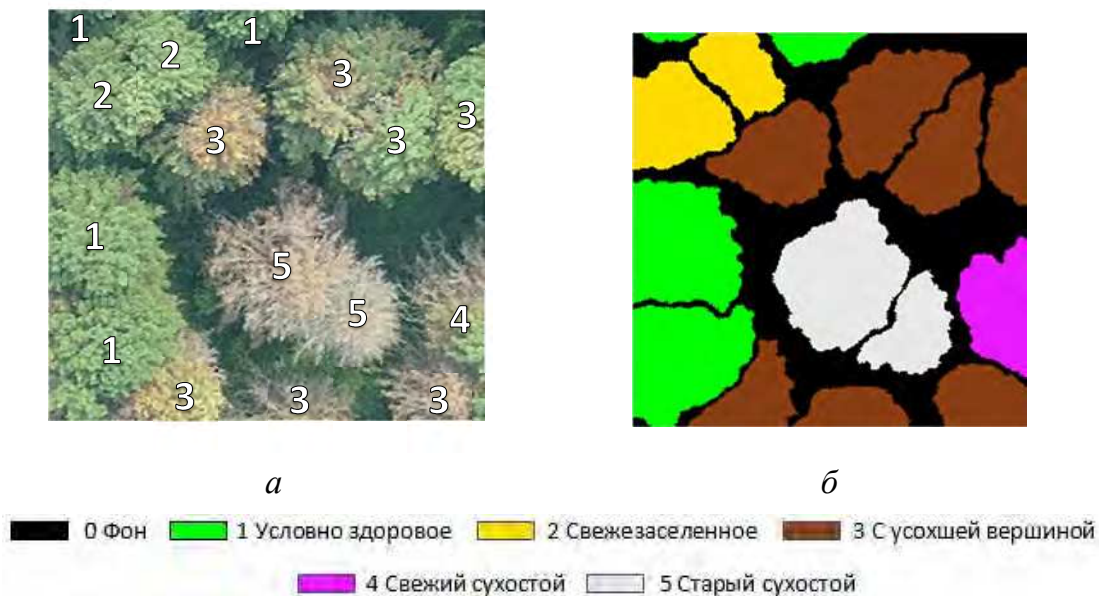


Рис. 1. Пример разметки фрагмента панорамы, выполненной экспертами: а – исходный фрагмент (каждое из деревьев отнесено визуальнo к одному из 5 классов); б – сформированная эталонная маска

Наборы данных

Для обучения предложенных моделей были сформированы два набора данных. Панорамы для первого набора данных были получены в июле 2019 г., а панорамы для второго набора данных – в мае 2021 г.

Для формирования обучающей и валидационной выборок первого набора данных панорамы и эталонные карты сегментации нарезались на фрагменты размером 256*256 пикселей со сдвигом 128 пикселей. Всего было получено 10 445 обучающих и 2880 валидационных фрагментов. Второй набор данных был сформирован таким же образом, но без сдвига фрагментов. В результате было получено 996 обучающих и 540 валидационных фрагментов. В качестве тестовой выборки использовались фрагменты изображения размером 2048*2819 пикселей.

Модели полносверточных сетей

Для решения задачи классификации деревьев *P. sibirica* были разработаны две новые модели полносверточных нейронных сетей, каждая из которых основана на классической модели U-Net. В исходную модель U-Net были внесены следующие изменения: размеры входных и выходных тензоров были изменены, функция активации ReLU была заменена функцией ELU, слой пакетной нормализации применялся после каждой нелинейности, а слои свертки не изменяли размер карты признаков. В результате была получена модифицированная модель U-Net. На ее основе с использованием предложенного в [2] подхода к формированию ансамбля из трех U-Net разной глубины была разработана вторая модель полносверточных сетей – multihead-U-Net, сокращенно МН-U-Net. По сравнению с первой предложенной моделью она имеет существенные изменения: вместо одного используются три входа разных масштабов ($256*256*3$, $128*128*3$ и $64*64*3$) и имеются три выхода разных масштабов, а процедура Dropout заменена на пространственную процедуру Dropout.

Обучение предложенных моделей СНС

Для обучения предложенных моделей была применена аугментация «на лету». Это позволило расширить каждую обучающую выборку путем создания модифицированных выборок, полученных случайными преобразованиями исходных примеров.

Поскольку в обучающих выборках присутствует большой дисбаланс классов деревьев, была выбрана функция потерь Focal Loss [3]. Эта функция является улучшенной версией Cross-Entropy Loss и включает в себя два модулирующих параметра, которые уменьшают потери от простых примеров и увеличивают значение потерь в случаях неправильной классификации примеров деревьев *P. sibirica*.

Для определения наилучших гиперпараметров предложенных моделей СНС использовалась байесовская оптимизация. Этот подход к оптимизации позволяет построить стохастическую модель функции отображения из двенадцати значений гиперпараметров, определяющих структуру нейронной сети, коэффициент скорости обучения сети, степень модификации изображений при их аугментации и значения модуляции функции потерь.

Трансферное обучение моделей СНС

Трансферное обучение – это повторное использование предварительно обученной модели для решения новой задачи. Такой подход сокращает требуемые вычислительные ресурсы и объем размеченных данных, необходимых для обучения новых моделей. Поскольку вторая обучающая выборка значительно меньше первой, а сезонные и световые условия получения снимков для этих выборок отличаются,

тогда ТО подходит для решения задачи семантической сегментации деревьев. Поэтому весовые коэффициенты моделей, обученных с помощью первого набора данных, предлагается использовать для обучения моделей на второй небольшой выборке.

Результаты исследования эффективности моделей СНС

Количественная оценка эффективности предложенных моделей СНС проводилась путем применения метрики Intersection over Union (IoU), также известной как коэффициент Жаккара. Значения IoU_c для каждого с класса деревьев и среднего по всем классам mIoU, превышающие 0,5, можно считать приемлемым качеством сегментации.

Таблица 1

Метрики качества сегментации деревьев P. Sibirica на валидационной выборке первого набора данных

Модель	IoU _c						mIoU
	Фон	Условно здоровое	Свежеселённое	С усохшей вершиной	Свежий сухой	Старый сухой	Val
U-Net	0.91	0.85	0.83	0.85	0.84	0.86	0.86
MH-U-Net	0.90	0.85	0.82	0.84	0.84	0.86	0.85

Из таблицы 1 видно, что обе модели СНС на первом наборе данных продемонстрировали высокое качество сегментации пикселей всех классов деревьев. Модифицированная сеть U-Net показывает незначительно более высокую эффективность, чем модель MH-U-Net.

Так как при обучении моделей на обучающей и валидационной выборках первого набора были получены результаты высокого качества, то полученные при этом оптимальные весовые коэффициенты моделей использовались далее в качестве начальных весов при обучении предложенных моделей на втором наборе данных. В таблице 2 представлены результаты сегментации деревьев на втором наборе без использования ТО и с его применением.

Таблица 2

Метрики качества сегментации деревьев P. Sibirica на валидационной выборке второго набора данных

Семантическая сегментация без использования трансферного обучения							
Мо- дель	IoUc						mIoU
	Фон	Условно здоро- вое	Свежеза- селённое	С усох- шей вер- шиной	Свежий сухо- стой	Старый сухо- стой	Val.
U-Net	0.819	0.640	0.314	0.468	0.404	0.603	0.541
MH- U-Net	0.813	0.636	0.257	0.479	0.379	0.598	0.527
Семантическая сегментация с использованием трансферного обучения							
Мо- дель	IoUc						mIoU
	Фон	Условно здоро- вое	Свежеза- селённое	С усох- шей вер- шиной	Свежий сухо- стой	Старый сухо- стой	Val.
U-Net	0.832	0.638	0.292	0.523	0.376	0.638	0.551
MH- U-Net	0.834	0.618	0.335	0.503	0.421	0.629	0.557

Аналогичным образом оценивалось качество сегментации деревьев на тестовой выборке. Полученные результаты приведены в таблице 3. Из таблиц 2 и 3 следует, что качество сегментации двух промежуточных классов деревьев даже при использовании ТО невысокое. Однако использование ТО повышает качество сегментации по метрике mIoU и значительно сокращает время обучения моделей (рисунок 2).

Таблица 3

Метрики качества сегментации деревьев P. Sibirica на тестовой выборке

Семантическая сегментация без использования трансферного обучения							
Мо- дель	IoUc						mIoU
	Фон	Условно здоро- вое	Свежеза- селённое	С усохшей вершиной	Свежий сухо- стой	Старый сухо- стой	Test
U-Net	0.793	0.705	0.259	0.543	0.449	0.779	0.588
MH- U-Net	0.791	0.702	0.187	0.557	0.428	0.7452	0.568
Семантическая сегментация с использованием трансферного обучения							
Мо- дель	IoUc						mIoU
	Фон	Условно здоро- вое	Свежеза- селённое	С усохшей вершиной	Свежий сухо- стой	Старый сухо- стой	Test
U-Net	0.808	0.715	0.296	0.569	0.436	0.771	0.599
MH- U-Net	0.812	0.679	0.328	0.559	0.440	0.758	0.596



Рис. 2. Время, необходимое для обучения моделей

Заключение

В работе рассмотрено исследование эффективности двух предложенных моделей полносверточных нейронных сетей при семантической сегментации поврежденных деревьев *P. sibirica* на снимках с БПЛА. Показало, что использование трансферного обучения значительно сокращает время обучения этих моделей и повышает качество семантической сегментации деревьев на снимках.

СПИСОК ЛИТЕРАТУРЫ

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [Электронный ресурс]. 2015. URL: <https://arxiv.org/pdf/1505.04597> (дата обращения 20.01.2023).
2. Марков Н.Г., Маслов К.А., Керчев И.А., Токарева О.С. Модели U-Net для семантической сегментации повреждённых деревьев сосны сибирской кедровой на снимках с БПЛА // Современные проблемы дистанционного зондирования Земли из космоса. 2022. № 1. С. 65–77.
3. Lin T., Goyal P., Girshick R., He K., Dollar P. Focal Loss for Dense Object Detection [Электронный ресурс]. 2018. URL: <https://arxiv.org/pdf/1708.02002> (дата обращения 20.01.2023).