

# ОБУЧЕНИЕ МОДЕЛЕЙ ПОЛНОСВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ПРИ РЕШЕНИИ ЗАДАЧИ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ СНИМКОВ ПОВРЕЖДЕННЫХ ВРЕДИТЕЛЯМИ ХВОЙНЫХ ДЕРЕВЬЕВ

Мачука М.К.Р.

Томский политехнический университет, ИШИТР, аспирант, гр. А2-36,  
e-mail: kristianrodrigo1@tpu.ru

## Введение

Общеизвестно, что вспышки массового размножения насекомых-вредителей наносят непоправимый вред хвойным лесам. Поэтому актуальным является оперативный мониторинг таких лесов с помощью беспилотных летательных аппаратов (БПЛА) и выявление по его результатам деревьев на начальных стадиях повреждения вредителями. При этом снимки с БПЛА крон деревьев обычно имеют высокое и сверхвысокое пространственное разрешение, что позволяет в результате их анализа с помощью моделей глубокого обучения оценивать состояние поврежденных хвойных деревьев [1].

В работе рассматривается задача обучения трех моделей полносверточных нейронных сетей, основанных на архитектуре U-Net [2]. Модели сетей используются для семантической сегментации снимков с БПЛА деревьев сосны сибирской кедровой *Pinus sibirica* (Du Tour) (далее- *P. Sibirica*), поврежденных союзным короедом *Ips amitinus* (Eichhoff) [3].

## Исходные данные

В качестве исходных данных использовались RGB-снимки, сделанные на штатную фотокамеру БПЛА DJI Phantom 3 Standard. При анализе снимков с БПЛА деревьев *P. sibirica*, поврежденных союзным короедом, экспертами выделялись пять классов их жизненного состояния [3]: «Условно здоровое», «Свежезаселенное», «С усохшей вершиной», «Свежий сухостой» и «Старый сухостой». Шестым классом являлся «Фон» – деревья других пород и иные объекты земной поверхности. Каждое из деревьев экспертом путем визуального дешифрирования относилось к одному из классов. Кроме разметки деревьев по классам осуществлялось оконтуривание их крон. В итоге получены эталонные карты сегментации.

Для формирования обучающей и валидационной выборок исходные снимки деревьев *P. sibirica* были нарезаны на фрагменты размером 256 x 256 пикселей путем сдвига предыдущего фрагмента на 128 пикселей. Всего было получено 10455 обучающих и 2880 валидационных фрагментов. Тестовая выборка формируется по отдельному снимку.

## Модели полносверточных сетей

Для решения задачи мультиклассификации деревьев *P. sibirica* разработаны три новые модели полносверточных нейронных сетей, каждая из которых основана на известной модели полносверточной сети U-Net, позволяющей вести семантическую сегментацию изображений (попиксельную классификацию). В исходную модель U-Net вносились такие изменения: изменены размеры входных и выходных тензоров, функция активации ReLU заменена на функцию ELU, слой пакетной нормализации применялся после каждой нелинейности, а слои свертки не изменяли размеры карт признаков. В итоге была получена модифицированная модель U-Net. Далее на ее основе и предложенной в работе [1] идеи введения вместо модели U-Net ансамбля из трех таких моделей различной 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 (256x256x3, 128x128x3 и 64x64x3), также имеются три выхода разных масштабов, а процедура Dropout заменена на пространственную процедуру Dropout. Третья предложенная модель — multihead-residual-U-Net (МН-Res-U-Net), обладает всеми особенностями модели МН-U-Net и дополнительно включает остаточные блоки [4].

## Особенности обучения и валидации предлагаемых моделей

При обучении предлагаемых моделей полносверточных сетей применялись две функции потерь: mIoU\* - метрика mIoU в мягкой формулировке [5] и focal-loss [6]. Функция потерь mIoU\* имеет параметр сглаживания меток и параметр, устраняющий большие по модулю значения градиента этой функции и точки сингулярности. Функция Focal-loss основана на функции потерь кросс-энтропии, но в нее

включен модулирующий параметр, который позволяет сокращать потери от простых примеров и увеличивать значение корректирующих примеров в случаях неправильной классификации деревьев *P.sibirica*. При обучении использовался известный алгоритм оптимизации Adam. Более того, при обучении моделей также применялась техника аугментации «на лету», позволяющая расширять обучающую выборку синтезированными фрагментами непосредственно во время обучения.

Для определения оптимальных гиперпараметров при обучении и валидации предлагаемых моделей использовался один из методов глобальной оптимизации – байесовская оптимизация. Байесовская оптимизация позволяет строить стохастическую модель функции отображения из десяти значений гиперпараметров (параметры, определяющие структуру нейронной сети и как сеть обучается). Путем итеративного включения информации, полученной при предыдущих оценках функции потерь, байесовская оптимизация позволяет выбрать оптимальный набор гиперпараметров для следующей оценки. Для итоговых моделей каждой из СНС были определены такие значения параметров, при которых функция потерь показала наилучшее значение из сотен, рассчитанных по валидационной выборке.

Для оценки эффективности предложенных моделей использовалась метрика Intersection over Union для каждого класса деревьев (IoUc). Значения IoUc, превышающие 0.5, соответствуют приемлемому качеству классификации. Использовалась также метрика mean Intersection over Union (mIoU), рассчитываемая как среднее значение IoUc по всем классам деревьев *P.sibirica*.

### Результаты обучения и валидации предлагаемых моделей

Из визуального анализа результатов обучения и валидации модели U-Net, приведенных в качестве примера на рис. 1, следует, что эта модель для каждой из функций потерь успешно воспроизводит границы между деревьями и сегментирует большую долю крон.

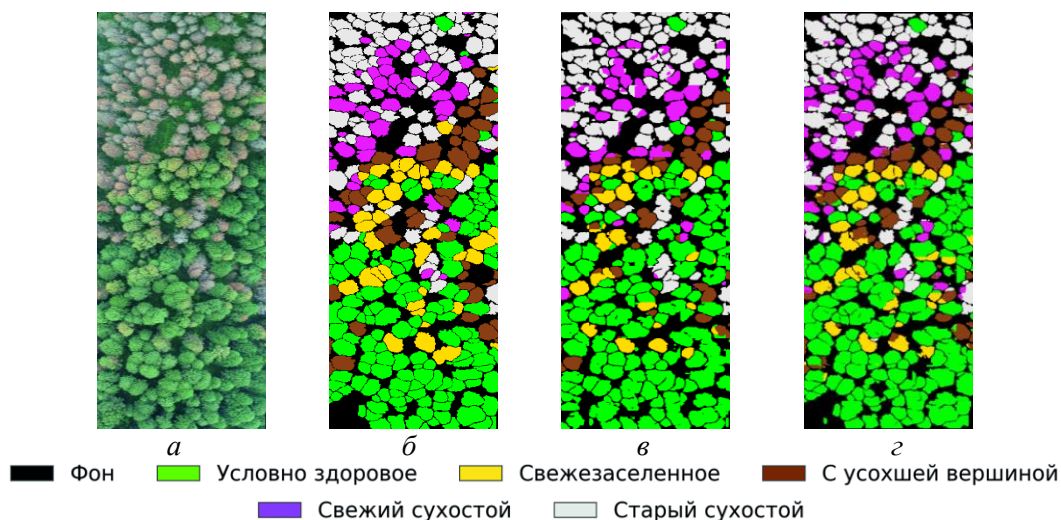


Рис. 1. Пример сегментации деревьев *P. Sibirica* на снимке: а — исходный снимок; б — эталонная карта сегментации этого снимка, полученная экспертами; в — результат валидации U-Net для случая функции потерь mIoU\*; г — результат валидации U-Net для focal-loss

Из таблицы 1, где приведены результаты сегментации на валидационной выборке, видно, что все три модели полносверточных сетей для каждой из функций потерь продемонстрировали высокое качество классификации пикселей классов «Фон», «Условно здоровое» и «Старый сухостой». Однако при обучении этих моделей с использованием функции потерь mIoU\* классы промежуточного состояния деревьев представляют сложность для сегментации. Относительно невысокое качество сегментации деревьев этих классов обосновывается их недостаточной представленностью в обучающей выборке. Поскольку функция потерь focal-loss разработана для обучения нейронных сетей в случае возникновения проблемы несбалансированных по классам выборок, то результаты, полученные при использовании такой функции при решении нашей задачи, указывают на высокое качество классификации всех классов деревьев *P. sibirica*.

Метрики качества сегментации деревьев *P. Sibirica* на валидационной выборке

Модель	Функция потерь	IoUc						mIoU
		Фон	Условно здоровое	Свежеселённое	С усохшей вершиной	Свежий сухостой	Старый сухостой	
U-Net	m IoU*	0.90	0.82	0.77	0.79	0.79	0.84	0.82
МН-U-Net		0.87	0.74	0.64	0.68	0.59	0.77	0.71
МН-Res-U-Net		0.84	0.77	0.67	0.70	0.51	0.74	0.71
U-Net	Focal-loss	0.91	0.85	0.83	0.85	0.84	0.86	0.86
МН-U-Net		0.90	0.85	0.82	0.84	0.84	0.86	0.85
МН-Res-U-Net		<b>0.93</b>	<b>0.90</b>	<b>0.87</b>	<b>0.89</b>	<b>0.89</b>	<b>0.89</b>	<b>0.89</b>

Как показано на рис. 2, для каждой из моделей при обучении и валидации с использованием функции потерь focal-loss требуется меньшее число эпох, чем в случае функции потерь mIoU\*. Это означает, что при использовании функции потерь focal-loss довольно значительно сокращается время на вычисление моделей и, в итоге, на обучение.

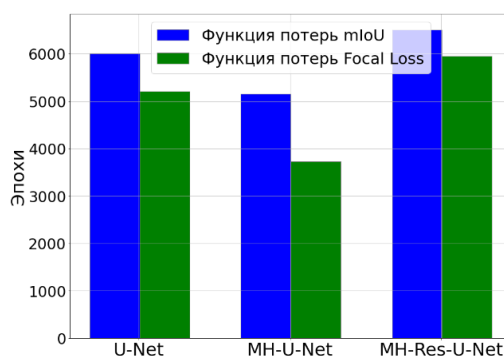


Рис. 2. Количество эпох при обучении моделей СНС для ста итераций байесовской оптимизации

## Заключение

Рассмотрено обучение трех предложенных моделей полносверточных нейронных сетей для классификации деревьев *P. sibirica*, поврежденных союзным короедом *Ips amitinus*, на снимках с БПЛА. Исследования показали, что применение функции потерь focal-loss при обучении позволяет этим моделям успешно распознавать деревья всех классов.

## Список использованных источников

1. 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.
2. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. – URL: <https://arxiv.org/pdf/1505.04597> (дата обращения 12.11.2022).
3. Керчев И.А., Кривец С.А., Бисирова Э.М., Смирнов Н.А. Распространение союзного короеда *Ips amitinus* в Западной Сибири // Российский журнал биологических инвазий. – 2021. – № 4. – С.77–84.
4. He K., Zhang X., Ren S., Sun J., Deep Residual Learning for Image Recognition. [Электронный ресурс]. – URL: <https://arxiv.org/abs/1512.03385> (дата обращения 21.11.2022).
5. Bertels J., Eelbode T., Berman M., Vandermeulen D., Maes F., Bisschops R., Blaschko M. Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice. – URL: <https://arxiv.org/abs/1911.01685> (дата обращения 18.01.2023)
6. Lin T., Goyal P., Girshick R., He K., Dollar P. Focal Loss for Dense Object Detection. – URL: <https://arxiv.org/pdf/1708.02002> (дата обращения 20.12.2022).