

СЕГМЕНТАЦИЯ ДЕРЕВЬЕВ ПИХТЫ НА ЗАШУМЛЕННЫХ СНИМКАХ С БПЛА С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ U-NET

Малкин А.Ю.

Томский политехнический университет, ИШИТР, гр. 8ВМ22, e-mail: malkin@tpu.ru

Аннотация

Решается задача сегментации зашумленных снимков крон пораженных уссурийским полиграфом деревьев пихты сибирской с помощью модифицированной модели сверточной нейронной сети (СНС) U-Net. Исследуется эффективность этой модели СНС при сегментации зашумленных снимков в случаях использования робастной функции потерь Коши и функции потерь Focal Loss. Выявлена зависимость точности классификации от амплитуды выбросов на зашумленных снимках для каждой из функций потерь.

Ключевые слова: сверточная нейронная сеть U-Net, пихта сибирская, беспилотный летательный аппарат, семантическая сегментация зашумленных снимков, робастная функция потерь Коши

Введение

В настоящее время актуальной является проблема раннего выявления хвойных деревьев, пораженных насекомыми-вредителями. Так, например, важно знать состояние здоровья деревьев кедр, пораженных союзным короедом, или деревьев пихты сибирской, заселенных уссурийским полиграфом. Для оперативного мониторинга хвойных лесов и выявления очагов размножения таких вредителей все чаще начинают применять беспилотные летательные аппараты (БПЛА), позволяющие получить снимки высокого разрешения [1]. Для анализа таких снимков путем семантической сегментации (попиксельной классификации) хвойных деревьев и определения степени их поражения (состояния здоровья) применяются сверточные нейронные сети (СНС) [2,3].

Снимки, получаемые с БПЛА, могут иметь шум, чаще всего в виде аномальных выбросов. Такие шумы приводят к искажению результатов анализа снимков, проводимого с помощью СНС. Поэтому для получения адекватных результатов семантической сегментации снимков необходимо при обучении СНС применять специальные функции потерь, устойчивые к выбросам на снимках, – робастные функции.

Целью данной работы является исследование эффективности модифицированной модели СНС U-Net при сегментации зашумленных снимков в случаях использования робастной функции потерь Коши и функции потерь Focal Loss. Полученные результаты исследований являются новыми.

Задача исследования модифицированной модели U-Net при сегментации зашумленных снимков деревьев пихты

Известно, что снимки хвойных лесов, получаемые с помощью установленной на БПЛА фото или кинокамеры, довольно часто имеют шумы в виде выбросов разной физической природы. Поэтому для получения адекватных результатов семантической сегментации таких снимков при использовании СНС необходимо при ее обучении применять специальные функции потерь, устойчивые к выбросам.

Ставится задача исследования эффективности модифицированной модели СНС U-Net [4] при сегментации зашумленных снимков в случаях использования робастной функции потерь Коши и функции потерь Focal Loss. Для проведения исследований необходимо использовать датасет, созданный по дешифрованным сотрудникам ИМКЭС СО РАН и ТПУ снимкам деревьев пихты сибирской, полученным с помощью БПЛА. В работе [3] на снимках определяются пять классов состояния пораженных уссурийским полиграфом деревьев пихты: «живые», «отмирающие», «свежий сухостой», «старый сухостой», «фон». Датасет включает фрагменты снимков с БПЛА и соответствующие этим фрагментам маски (получены в результате дешифрирования исходных снимков) размером 256x256 пикселей цветового пространства RGB. Он поделён на три выборки: обучающая (2004 фрагмента), валидационная (672 фрагмента) и тестовая (91 фрагмент).

Для выявления эффективности модифицированной модели U-Net при последовательном использовании функций потерь Focal Loss и Коши было произведено зашумление 24.4 % площади пикселей каждого фрагмента (16000 зашумленных пикселей) обучающей выборки. Пиксели для зашумления на фрагменте выбирались случайным образом с использованием генератора случайных

чисел с равномерным законом распределения. Амплитуда выброса определяет, во сколько раз увеличивается яркость зашумлённого пикселя относительно его предыдущего значения. На рис. 1 показан пример исходного фрагмента (без зашумления) и три фрагмента при различных амплитудах выбросов.



Рис. 1. Визуализация зашумления фрагмента снимка пихты: а – исходные данные; б – увеличение амплитуды яркости пикселя в 1,5 раза; с – увеличение амплитуды яркости пикселя в 2,0 раза; d – увеличение амплитуды яркости пикселя в 3,0 раза

Модифицированная модель U-Net

Для проведения исследований выбрана модификация классической полносвёрточной нейронной сети U-Net, обученной при различных функциях потерь. Выбор в пользу классической архитектуры U-Net как основы для модификации был сделан по причине её эффективности при работе с медицинскими изображениями [4]. Данная архитектура была модифицирована сотрудниками ТПУ для решения задачи семантической сегментации изображений крон хвойных деревьев [3]. Ключевыми изменениями являются:

- размер входного тензора составляет 256x256x3;
- свёртки не уменьшают размеры карт признаков;
- применяется пакетная нормализация (batch normalization);
- используется функция активации ELU;
- выходной тензор вычисляется 5 свёртками с фильтрами размером 1x1, что позволяет классифицировать пиксели 5 классов (четыре класса состояний пораженных уссурийским полиграфом деревьев пихты и фон).

Выбор робастной функции потерь

В работе [5] приводится исследование нескольких робастных функций потерь: Хьюбера, Эндрюса, Тьюки, Коши, Уэлша, Мешалкина и других. Исследование заключалось в определении эффективности этих функций потерь при решении задачи классификации набора одномерных данных «Ирисы Фишера» при наличии выбросов. Робастность функций потерь позволяет уменьшить влияние выбросов на величину потерь при обучении нейронной сети очередными данными, если они выбиваются из нормального распределения выборки.

Для обучения СНС была выбрана робастная функция потерь Коши по причине её широкой изученности в ряде других работ, подтверждающих её робастность [6, 7]. Из работы [5] следует, что функция Коши описывается следующим выражением:

$$\rho(z) = \ln\left(\frac{1}{2} * \left(\frac{z}{\beta}\right)^2 + 1\right), \quad (1)$$

где z – величина ошибки, β – свободный коэффициент. Производная функции Коши представлена формулой:

$$\rho'(z) = \frac{z}{\frac{z^2}{2} + \beta^2} \quad (2)$$

Из формулы (2) следует, что описываемая ею функция является непрерывной на множестве вещественных чисел $\mathbb{R}$, что позволяет использовать выражение (1) в качестве функции потерь в алгоритме обратного распространения ошибки при обучении СНС.

Для сравнительного анализа используется хорошо себя зарекомендовавшая функция потерь Focal Loss, описываемая следующим выражением [8]:

$$\rho(z) = -\alpha^* (1-z)^{\gamma} \ln(z), \quad (3)$$

где z – величина ошибки, α и γ – коэффициенты при функции Focal Loss.

Обучение и валидация СНС на зашумлённых снимках

С помощью байесовского алгоритма оптимизации выполняется поиск квазиоптимальных комбинаций гиперпараметров при обучении моделей СНС на выборке с различной амплитудой выбросов при зашумлении пикселей, занимающих 24.4% от площади каждого фрагмента. Количество таких комбинаций составляет 100. Исследуемыми гиперпараметрами являются: скорость обучения в диапазоне $[10^{-4}; 10^{-2}]$, параметры β (для функции Коши) в диапазоне $[3.0; 9.0]$, α и γ (для Focal Loss) в диапазоне $[3.0; 9.0]$. При обучении очередной модели на очередном наборе гиперпараметров берётся во внимание значение метрики mIoU, которая рассчитывается как среднее по значениям IoU для каждого из исследуемых классов деревьев:

$$IoU = \frac{TP}{TP + FP + FN}, \quad (4)$$

где TP – это истинно-положительные предсказания, FP – ложно -положительные предсказания, FN – ложно - отрицательные предсказания.

$$mIoU = \frac{1}{N} \sum_{c=1}^N IoU_c, \quad (5)$$

где N – число классов деревьев. Обучение моделей на очередном наборе гиперпараметров происходит до тех пор, пока не возникает переобучение модели, которое выражается в неувеличении метрики mIoU после 11 эпох обучения подряд.

Анализ результатов исследования моделей СНС, обученных на зашумлённых фрагментах

В результате обучения и валидации получены восемь моделей СНС: каждая из них обучалась со своим набором, представляющим из себя ту или иную функцию потерь (робастная функция Коши или Focal Loss), отсутствие или наличие зашумления обучающих фрагментов, а также различная амплитуда выбросов пикселей (увеличение яркости в 1.5, 2.0, 3.0 раза). Эти модели показали максимальное значение метрики mIoU на валидационной выборке (зашумление её не проводилось), их значения метрики mIoU на тестовой выборке (зашумление не проводилось) представлены в таблице 1.

Таблица 1

Результаты исследования моделей СНС в случае функций потерь Коши и Focal Loss

Функция потерь	Количество зашумлённых пикселей на обучающих фрагментах	Амплитуда выбросов	α	γ	β	mIoU на тестовой выборке
Коши	0	0,0	-	-	4,356	0,660
Коши	16000	1,5	-	-	6,166	0,652
Коши	16000	2,0	-	-	6,065	0,635

Коши	16000	3,0	-	-	6,796	0,598
Focal loss	0	0,0	4,433	2,699	-	0,670
Focal loss	16000	1,5	6,460	8,809	-	0,628
Focal loss	16000	2,0	4,044	3,002	-	0,656
Focal loss	16000	3,0	5,886	7,324	-	0,570

Из таблицы 1 следует, что зашумление в 24.4 % от площади каждого из обучающих фрагментов приводит к уменьшению метрики mIoU на тестовой выборке, причем в случае модели, обученной с применением функции потерь Коши, падение метрики относительно модели, обученной на фрагментах без зашумления, составило от 0.008 до 0.062, а для модели с применением функции Focal Loss падение метрики mIoU составляет от 0.014 до 0.100 относительно модели, обученной на фрагментах без зашумления. Отсюда можно сделать вывод, что робастная функция потерь Коши даёт более высокие результаты сегментации зашумлённых снимков, чем функция потерь Focal Loss.

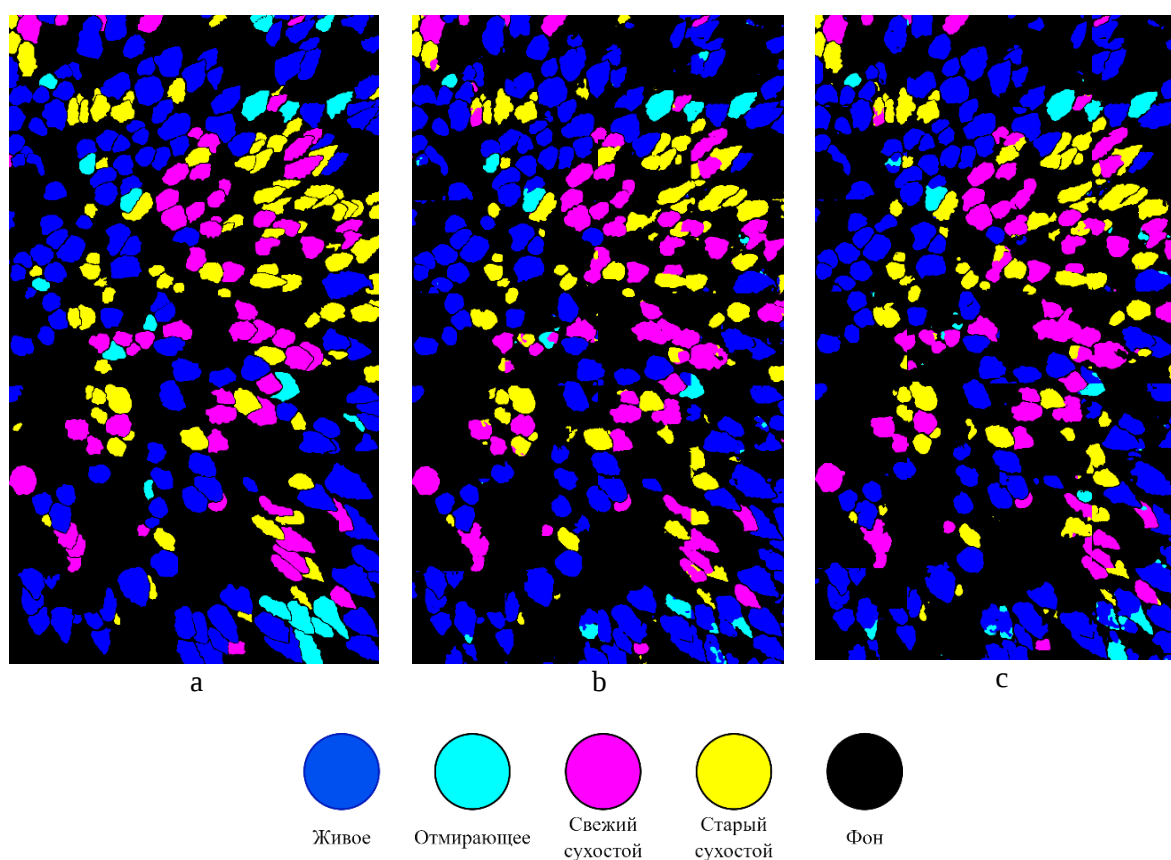


Рис. 2. Сегментация тестового снимка: а – эталонная разметка экспертами;
 б – сегментация моделью CHC U-Net, обученной с применением функции потерь Коши;
 с – сегментация моделью CHC U-Net, обученной с применением функции потерь Focal Loss

На рис. 2 представлены визуальные результаты сегментации тестового снимка крон пихты моделями CHC, обученными на зашумлённых фрагментах (амплитуда выбросов – 1.5 от яркости пикселя) с применением робастной функции потерь Коши и функции потерь Focal Loss. Видно, что классификация деревьев класса «Отмирающие», которые наименее представлены в датасете, выполняется хуже, чем для других классов деревьев. Это наблюдается для каждой функции потерь. При этом модель CHC, обученной с применением функции Коши, выполняет задачу лучше модели CHC, обученной с использованием функции потерь Focal Loss.

На рис. 3 представлены зависимости метрики mIoU, полученные на тестовой и валидационных выборках, от амплитуды выбросов яркости у пикселей на обучающих фрагментах. По этим результатам можно сделать два вывода: функция потерь Focal Loss позволяет обучить модель на данных без

зашумления с более высокой метрикой mIoU на тестовой выборке в сравнении с робастной функцией потерь Коши (0.670 против 0.660), а модели, обученные с применением функции потерь Focal Loss, демонстрируют, что значения метрик mIoU на тестовой выборке превышают таковые на валидационной выборке.

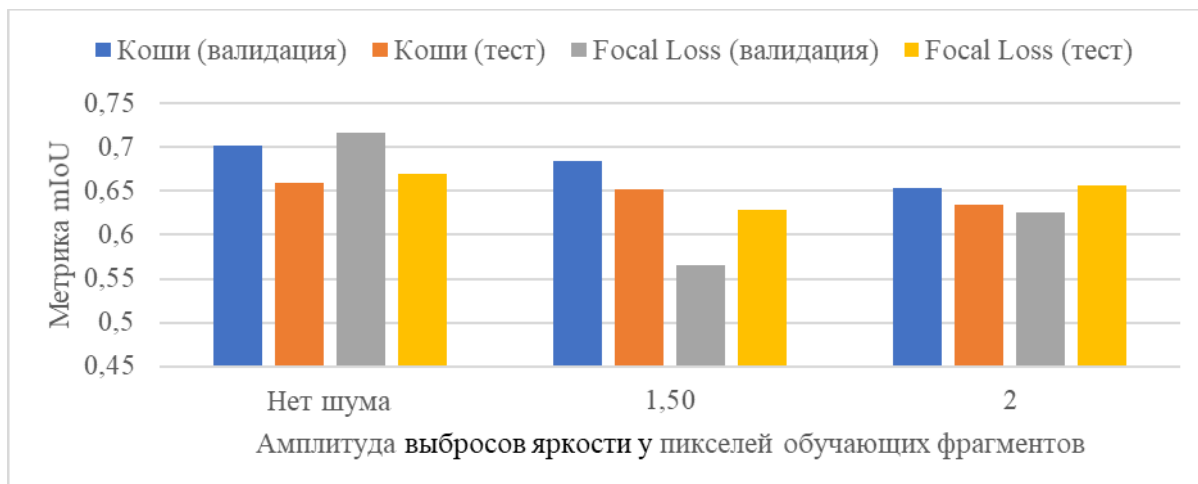


Рис. 3. Зависимость метрики mIoU на валидационной и тестовой выборках от амплитуды выбросов у пикселей фрагментов

Влияние зашумления выборок фрагментов на качество сегментации деревьев

Для исследования метрики mIoU на валидационной и тестовой выборках при зашумлении обучающей, валидационной, тестовой выборок в различных комбинациях используются модели СНС на основе U-Net, обученные с применением робастной функции потерь Коши и функции потерь Focal Loss. Происходит обучение, валидация, тестирование моделей, у которых сохранены весовые коэффициенты, полученные в ходе исследования влияния зашумления 24.4 % площади обучающих фрагментов с амплитудой выбросов 1.5 (таблица 1). На рис. 4 представлены значения метрики mIoU на валидационной и тестовой выборках, показанные моделями СНС, обученными с применением функции потерь Focal Loss и робастной функции потерь Коши для зашумления фрагментов различных вариантов выборок.

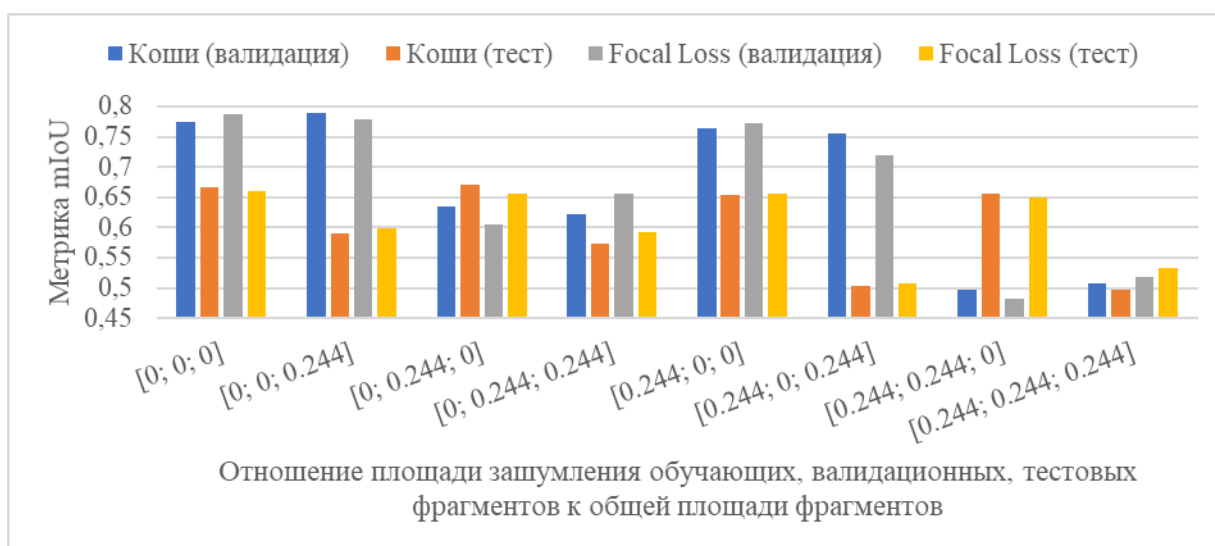


Рис. 4. Значения метрики mIoU для моделей СНС, обученных с применением функции потерь Коши и функции потерь Focal Loss, при различных комбинациях зашумления фрагментов из обучающей, валидационной, тестовой выборок

Из результатов, приведенных на рис. 4, следует, что наличие шума на обучающих фрагментах не приводит к сильному уменьшению метрики mIoU на валидационной и тестовой выборках, если фрагменты последних выборок не зашумлены. Однако при зашумлении фрагментов обучающей и валидационной выборок метрика mIoU на тестовой выборке уменьшается на 0.01, а при зашумлении фрагментов обучающей и тестовой выборок метрика mIoU для тестовой выборки уменьшается на 0.17. Наличие шума на фрагментах валидационной выборки не влияет на метрику mIoU на тестовой выборке и наоборот.

Заключение

Проведено исследование эффективности модифицированных моделей СНС U-Net, обученных с применением двух функций потерь: Коши и Focal Loss. В результате исследования было определено, что при зашумлении в 24.4 % от площади каждого фрагмента обучающей выборки качество сегментации фрагментов тестовой выборки моделью СНС, обученной с применением функции Коши, выше относительно модели СНС, обученной с применением функции потерь Focal Loss. Выявлена зависимость точности классификации крон деревьев от амплитуды выбросов на зашумлённых снимках: при увеличении амплитуды выбросов происходит уменьшение метрики mIoU.

В результате зашумления в 24.4 % от площади каждого фрагмента обучающей, валидационной и тестовой выборок и при амплитуде выбросов, равной 1.5 в различных комбинациях зашумленных выборок было установлено отсутствие зависимости метрики mIoU на тестовой выборке при зашумлении валидационной выборки и метрики mIoU на валидационной выборке при зашумлении тестовой выборки. Уменьшение метрики mIoU на тестовой (до 0.17) и валидационной (до 0.28) выборках происходит, если соответствующая выборка и обучающая выборка были зашумлены, причем этот результат не зависит от выбранной при обучении СНС функции потерь.

Список использованных источников

1. Денисов С.А., Домрачев А.А., Елсуков А.С. Опыт применения квадрокоптера для мониторинга возобновления леса // Вестник Поволжского государственного технологического университета. Сер.: Лес. Экология. Природопользование. – 2016. – № 4 (32). – С. 34–46. – DOI: 10.15350/2306- 2827.2016.4.34.
2. Shasha Lu: Examining the influencing factors of forest health, its implications on rural revitalization // ScienceDirect: сайт. – 2021. – URL: <https://www.sciencedirect.com/science/article/pii/S026483772032603X>.
3. Керчев И.А., Маслов К.А., Марков Н.Г., Токарева О.С. Семантическая сегментация поврежденных деревьев пихты на снимках с беспилотных летательных аппаратов // Современные проблемы дистанционного зондирования Земли из космоса. – 2021. – Т. 18. – №. 1. – С. 116–126.
4. Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation // arXiv: сайт. – 2015. – URL: <https://arxiv.org/abs/1505.04597>.
5. Сивак М.А. Робастное обучение нейронных сетей с простой архитектурой для решения задач классификации: диссертация на соискание учёной степени кандидата технических наук: 05.13.17. – Новосибирский государственный технический университет, Новосибирск, – 2022 – 117 с.
6. Thamsanqa Mlotshwa, Heinrich van Deventer, Anna Sergeevna Bosman. Cauchy Loss Function: Robustness Under Gaussian and Cauchy Noise // arXiv: сайт. – 2023. – URL: <https://arxiv.org/abs/2302.07238>.
7. Xuelong Li, Quanmao Lu, Yongsheng Dong, Dacheng Tao. Robust subspace clustering by Cauchy loss function // arXiv: сайт. – 2019. – URL: <https://arxiv.org/abs/1904.12274>.
8. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollár. Focal Loss for Dense Object Detection // arXiv: сайт. – 2017. – URL: <https://arxiv.org/abs/1708.02002>.