

ИССЛЕДОВАНИЕ МОДЕЛИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ МО-U-NET С РОБАСТНОЙ ФУНКЦИИ ПОТЕРЬ ХЬЮБЕРА В УСЛОВИЯХ ЗАШУМЛЕННЫХ ИЗОБРАЖЕНИЙ

Малкин А.Ю.

Томский политехнический университет, аспирант ИШИТР

e-mail: aym27@tpu.ru

Аннотация

Исследуется эффективность модели сверточной нейронной сети Mo-U-Net при наличии импульсных помех на изображениях. Модель обучается с применением робастной функции потерь (РФП) Хьюбера при решении задачи попиксельной классификации зашумленных помехами изображений пораженных деревьев пихты. Результаты исследований позволили рекомендовать РФП Хьюбера к использованию при обучении этой модели для работы с зашумленными изображениями.

Ключевые слова: сверточная нейронная сеть, Mo-U-Net, пихта сибирская, беспилотный летательный аппарат, семантическая сегментация, импульсные помехи на изображениях, робастная функция потерь Хьюбера.

Введение

В настоящее время остаются актуальными задачи оперативного выявления пораженных насекомыми-вредителями хвойных деревьев. Одной из таких задач является задача обнаружения пораженных уссурийским полиграфом деревьев пихты сибирской по изображениям, полученным с беспилотных летательных аппаратов (БПЛА). Анализ таких изображений с БПЛА представляет интерес для работников лесной отрасли с целью оперативного проведения фитосанитарных мероприятий в пораженных вредителями хвойных лесах [1].

Ручная разметка деревьев по классам состояния здоровья (степени поражения вредителями) деревьев на изображениях является затратной по времени задачей. Поэтому всё большую популярность приобретает применение моделей сверточных нейронных сетей (СНС) для обнаружения пораженных деревьев и классификации их по классам здоровья [2, 3].

Известно, что в результате съемки фотокамерой, установленной на БПЛА, возможно появление на изображениях импульсных помех (выбросов) разной физической природы. Это сказывается на итоговом качестве сегментации изображений с помощью моделей СНС. Поэтому актуальными являются задачи исследования устойчивости моделей СНС к таким помехам. Известно, что для получения удовлетворительных результатов сегментации необходимо при обучении модели СНС применять робастные функции потерь, устойчивые к выбросам, что в итоге может повысить помехоустойчивость модели СНС [4].

Целью данной работы является исследование эффективности модели СНС Mo-U-Net, являющейся модификацией известной модели СНС U-Net, при семантической сегментации зашумленных импульсными помехами изображений пораженных уссурийским полиграфом деревьев пихты сибирской. Обучение модели Mo-U-Net ведется при использовании робастной функции потерь Хьюбера.

Задача исследования устойчивости модели Mo-U-Net к импульсным помехам

В данной работе исследуется модель СНС Mo-U-Net, которая является модификацией классической модели U-Net [5] и подробно описана в работе [6]. Модель Mo-U-Net применяется для семантической сегментации изображений с импульсными помехами пораженных уссурийским полиграфом деревьев пихты.

Ставится задача оценки устойчивости модели Mo-U-Net с робастной функцией потерь (РФП) Хьюбера при различных параметрах импульсных помех. Выбор в пользу данной функции потерь сделан по причине её широкой изученности в ряде других работ, подтверждающих её робастность [7–9]. Из работы [4] следует, что РФП Хьюбера описывается следующим выражением:

$$\rho(z) = \begin{cases} \frac{1}{2}z^2, & |z| \leq \beta \\ \beta|z| - \frac{1}{2}\beta^2, & |z| > \beta \end{cases} \quad (1)$$

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

$$\rho'(z) = \begin{cases} z, & |z| \leq \beta \\ -\beta, & z < -\beta \\ \beta, & z > \beta \end{cases} \quad (2)$$

Формула (2) является непрерывной на множестве вещественных чисел $\mathbb{R}$, что позволяет использовать РФП Хьюбера (1) в качестве функции потерь в алгоритме обратного распространения ошибки при обучении модели СНС Mo-U-Net.

В исследованиях будут определяться значения коэффициента β функции потерь, при котором модель СНС Mo-U-Net позволит обеспечить лучшее качество сегментации изображений деревьев пихты. Для исследований используются следующие значения коэффициента β : (0, 0,1, 0,2, ..., 9,0].

Для оценки качества сегментации каждого класса деревьев применяется метрика IoU_c:

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

где TP_c – это истинно-положительные предсказания, FP_c – ложно-положительные предсказания, FN_c – ложно-отрицательные предсказания. Для оценки точности классификации деревьев также используется интегрированная метрика mean Intersection over Union (mIoU), рассчитываемая как среднее значение IoU_c по всем C классам:

$$\text{mIoU} = \frac{1}{N} \sum_{C=1}^N \text{IoU}_c, \quad (5)$$

где N – число классов деревьев. В нашем случае N равно пяти. Обучение и валидация модели завершается, когда значение метрики mIoU на валидационной выборке не превысит ранее достигнутое максимальное значение ее по прошествии 5 эпох обучения подряд.

При проведении исследований будет использоваться исходный датасет с размеченными изображениями пораженных деревьев пихты, описанный в работе [3]. Исходными изображениями при его формировании были RGB-изображения, полученные с БПЛА, оснащенного фотокамерой, и имеющие разрешение 0,1 м. Авторами [3] проведена ручная разметка изображений деревьев пихты по следующим классам их здоровья: «живые», «отмирающие», «свежий сухостой», «старый сухостой». Пятым классом в датасете является «фон». Таким образом, экспертами решена задача мультиклассификации изображений деревьев пихты. В результате формирования исходного датасета были подготовлены 2004 фрагмента с размером 256x256 пикселей цветового пространства RGB обучающей выборки, 672 и 91 фрагмент в, соответственно, валидационной и тестовой выборках.

Формирование датасетов с зашумленными изображениями деревьев пихты

На основе исходного датасета для проведения исследований созданы три новых датасета, содержащих изображения с искусственно внесенными импульсными помехами. В первом таком датасете зашумляются фрагменты обучающей выборки, а во втором датасете в дополнение к зашумлению обучающих фрагментов аналогично зашумляются фрагменты валидационной выборки. Третий датасет содержит зашумленные фрагменты по всем трем выборкам. Для каждого фрагмента, подлежащего зашумлению, случайным образом по дискретному равномерному закону распределения выбираются пиксели в количестве 25% от общего количества пикселей фрагмента. Для выбранных пикселей в каждом из RGB-каналов происходит увеличение амплитуды яркости в 1,5 раза. На рисунке 1 представлено зашумление импульсными помехами 25% площади обучающего фрагмента при амплитуде помех, равной 1,5.

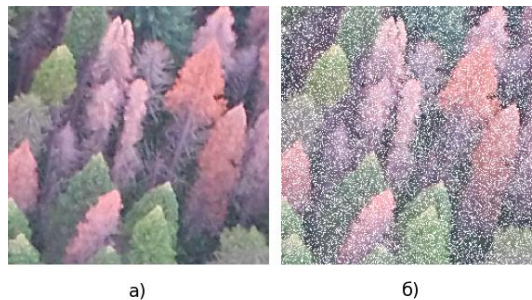


Рис. 1. Визуализация зашумления фрагмента снимка пихты: а – исходные данные. б – фрагмент с увеличением амплитуды яркости пикселей в 1,5 раза.

На языке программирования Python 3 с применением фреймворка PyTorch программно реализованы модель СНС Mo-U-Net и РФП Хьюбера. Модель обучена и валидирована с применением представленных ранее четырех датасетов изображений деревьев пихты. Получены следующие результаты.

Результаты исследований и их обсуждение

В случае работы модели СНС Mo-U-Net с датасетом, не содержащим импульсные помехи на фрагментах, было установлено, что максимальное значение метрики mIoU на валидационной выборке достигается моделью в случае значения коэффициента β , равного 7,6, и составляет 0,7706. Метрика mIoU на тестовой выборке в таком случае составляет 0,6473. На рисунке 2 представлены результаты валидации и тестирования модели, обученной с применением различных значений коэффициента функции потерь.

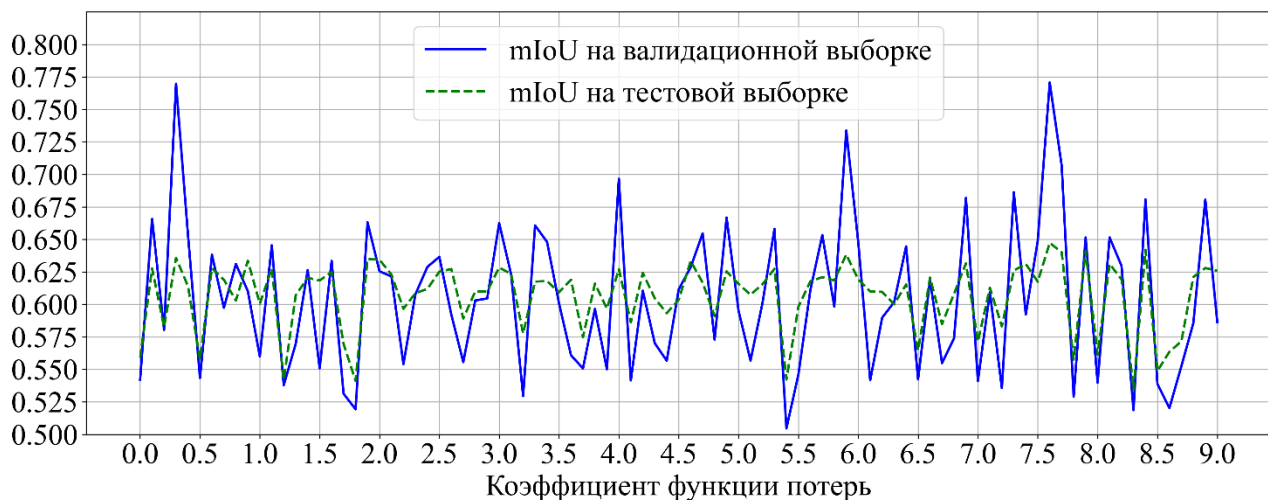


Рис. 2. Валидация и тестирование модели при отсутствии импульсных помех в датасете

Аналогичные результаты получены в результате обучения, валидации и тестирования модели Mo-U-Net на новых датасетах, содержащих импульсные помехи на фрагментах тех или иных выборках этих датасетов. Так, при использовании первого датасета и обучении модели на фрагментах с импульсными помехами обучающей выборки квазиоптимальное значение коэффициента РФП Хьюбера равно 1,5. В таком случае модель в результате валидации и тестирования достигает значений метрики mIoU, соответственно, 0,6456 и 0,6339.

В результате обучения и валидации модели на зашумленных фрагментах второго датасета было установлено, что при коэффициенте β , равном 1,4, модель СНС Mo-U-Net демонстрирует максимальное значение метрики mIoU на валидационной выборке, которое составляет 0,7325. Метрика mIoU на тестовой выборке составляет 0,6307.

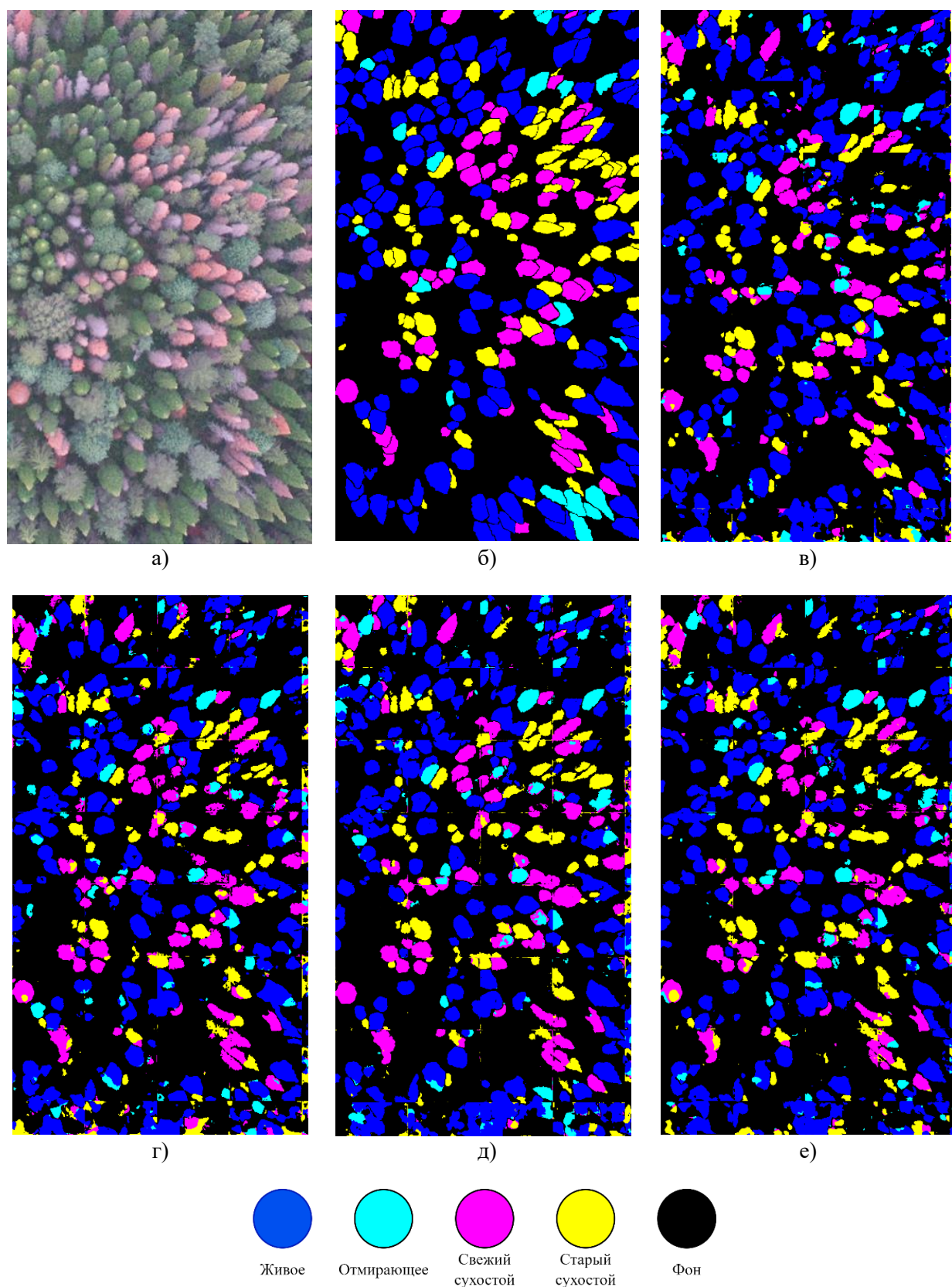
В случае наличия импульсных помех на всех фрагментах третьего датасета квазиоптимальным значением коэффициента функции потерь является 2,4. В таком случае качество сегментации валидационной выборки по метрике mIoU составляет 0,7397, на тестовой выборке – 0,6287.

В таблице 1 представлена сводная информация о результатах исследования модели Mo-U-Net, полученных с использованием каждого из трех датасетов. Из таблицы 1 следует, что в каждом отдельном случае зашумления выборок датасета рекомендуется осуществлять подбор квазиоптимального значения коэффициента β РФП Хьюбера. В случае, если датасет не содержит импульсных помех на фрагментах, обученная и валидированная модель СНС Mo-U-Net демонстрирует наибольшие значения метрики mIoU на валидационной и тестовой выборках относительно рассматриваемых случаев зашумления фрагментов выборок в каждом из новых датасетов. Тем не менее, метрика mIoU на тестовой выборке не уменьшается более чем на 0,02 в результате работы модели Mo-U-Net с датасетами, содержащими импульсные помехи на фрагментах. Более того, можно говорить о плавном уменьшении метрики mIoU на тестовой выборке с увеличением числа выборок с зашумленными фрагментами. Все это позволяет считать, что использование РФП Хьюбера при обучении модели Mo-U-Net не позволяет последней резко снизить устойчивость к импульсным помехам при решении задачи мультиклассификации изображений деревьев пихты.

Таблица 1. Значения метрики mIoU в результате обучения и валидации модели Mo-U-Net в случае зашумления различных выборок датасета

Отношение площади выбросов к площади фрагментов выборки			Коэффициент функции потерь	Значение метрики mIoU	
Обучающая	Валидационная	Тестовая		Валидационная выборка	Тестовая выборка
0	0	0	7,6	0,7706	0,6473
0,25	0	0	1,5	0,6456	0,6339
0,25	0,25	0	1,4	0,7325	0,6307
0,25	0,25	0,25	2,4	0,7397	0,6287

На рисунке 3 для визуального анализа представлены результаты сегментации изображения тестового участка моделью Mo-U-Net, полученные при проведении описанных исследований. По полученным картам сегментации видно, что классификация деревьев класса «Отмирающие», которые наименее представлены в подготовленных датасетах, выполняется с худшей точностью, чем для других классов деревьев. Из рисунка 3д видим, что модель СНС Mo-U-Net, обученная и валидированная на фрагментах, содержащих импульсные помехи, выполняет задачу мультиклассификации лучше, чем в результате ее работы с другими датасетами с зашумленными фрагментами (рисунок 3г, е).



*Рис. 3. Сегментация тестового изображения моделью Mo-U-Net:
а) исходные данные, б) эталонная разметка экспертом, в) случай отсутствия импульсных помех в датасете, г) случай импульсных помех на обучающих фрагментах, д) случай импульсных помех на обучающих и валидационных фрагментах, е) случай импульсных помех на всех трёх выборках датасета*

Заключение

Проведено исследование модели СНС Мо-U-Net с РФП Хьюбера при решении задачи мультиклассификации зашумленных импульсными помехами изображений деревьев пихты. Показано, что зашумление фрагментов любой из выборок датасетов приводит к необходимости определения квазиоптимального значения коэффициента β РФП Хьюбера. В случае отсутствия импульсных помех на фрагментах деревьев пихты в датасете модель Мо-U-Net обеспечивает максимально возможное качество сегментации изображений деревьев пихты.

Установлено, что использование РФП Хьюбера обеспечивает небольшое и плавное снижение устойчивости модели Мо-U-Net при увеличении числа выборок с зашумленными фрагментами. Это позволяет рекомендовать к применению РФП Хьюбера при обучении модели Мо-U-Net для решения задач мультиклассификации изображений с импульсными помехами.

Список использованных источников

1. Шагалиев Р.Д. Идентификация сухостойных деревьев с помощью искусственного интеллекта на основе данных дистанционного зондирования Земли / Р.Д. Шагалиев, Е.А. Богдан, А.Ф. Галямов, Л.Н. Белан, О.И. Ишкинина, Г.Г. Валиев // Вестник СВФУ. Серия «Науки о Земле». – 2024. – № 4. С. 138-149. DOI: 10.25587/2587-8751-2024-4-138-149
2. Safonova A. et al. Detection of fir trees (*Abies sibirica*) damaged by the bark beetle in unmanned aerial vehicle images with deep learning // Remote sensing. – 2019. – Т. 11. – № 6. – С. 643.
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. 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.
5. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. – Springer international publishing, 2015. – С. 234-241.
6. Марков Н.Г., Мачука К. Модели и методы глубокого обучения для решения задач дистанционного мониторинга лесных ресурсов // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2024. – Т. 335. – № 6. – С. 55–74. DOI: 10.18799/24131830/2024/6/4600
7. Gupta D., Hazarika B. B., Berlin M. Robust regularized extreme learning machine with asymmetric Huber loss function // Neural Computing and Applications. – 2020. – Т. 32. – №. 16. – С. 12971-12998.
8. Balasundaram S., Prasad S. C. Robust twin support vector regression based on Huber loss function // Neural Computing and Applications. – 2020. – Т. 32. – №. 15. – С. 11285-11309.
9. Wang Y., Hu S., Wu S. Object tracking based on huber loss function // The Visual Computer. – 2019. – Т. 35. – С. 1641-1654.