



**ТОМСКИЙ
ПОЛИТЕХ**



СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ ПОВРЕЖДЕННЫХ ДЕРЕВЬЕВ ПИХТЫ СИБИРСКОЙ НА ЗАШУМЛЕННЫХ СНИМКАХ С БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Малкин Артем Юрьевич

Марков Николай Григорьевич

12 ноября 2024

Актуальность работы



- Вспышки массового распространения насекомых-вредителей наносят непоправимый вред лесам. Актуальна задача раннего выявления очагов вредителей хвойных деревьев.
- Существует проблема отсутствия современного инструментария для оперативного проведения дистанционного мониторинга и анализа полученных данных.
- Частичное её решение – применение беспилотных летательных аппаратов (БПЛА) с аппаратурой для съёмки, применение сверточных нейронных сетей (СНС) для анализа полученных снимков.
- По причине несовершенства аппаратуры возможно зашумление снимков. Для уменьшения влияния необходимы модели СНС, устойчивые к помехам.

Цель и задачи

Цель работы – исследовать точность семантической сегментации (попиксельной классификации) пораженных уссурийским полиграфом деревьев пихты сибирской (далее – пихты) на зашумленных снимках с БПЛА, выполняемой с помощью модифицированной модели СНС U-Net.

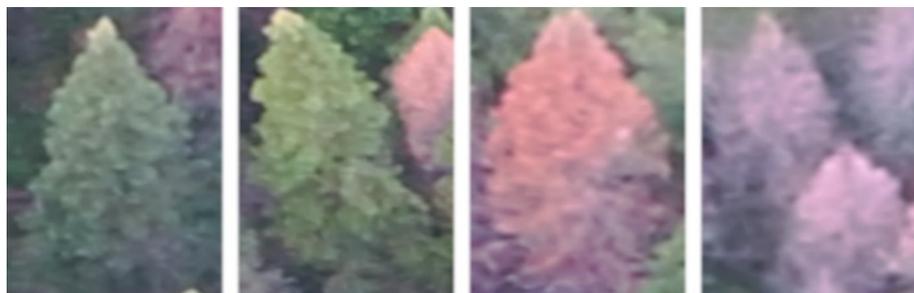
Задачи:

1. Разработать датасет с зашумленными импульсными помехами фрагментами снимков деревьев пихты для семантической сегментации моделью нейронной сети.
2. Спланировать эксперименты по мультиклассификации зашумленных снимков при использовании модифицированной модели U-Net:
 - в случае робастной функции потерь (РФП) Коши;
 - в случае функции потерь (ФП) Focal Loss.

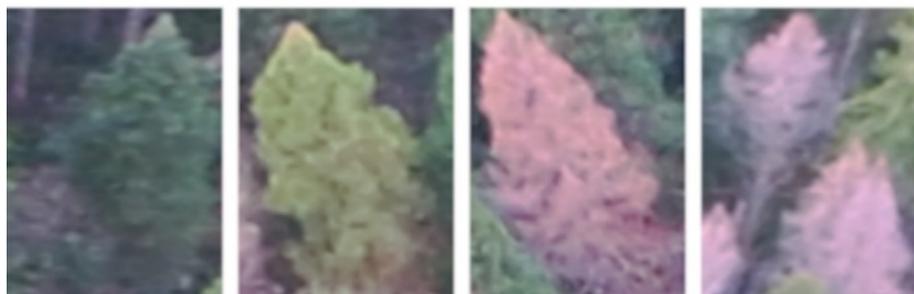
Исследуются две модифицированные модели U-Net с разными функциями потерь.

3. Программно реализовать модифицированные модели U-Net.
4. Провести две серии экспериментов по выявлению помехоустойчивости обученных моделей U-Net при различных параметрах зашумления фрагментов датасета.

Задача мультиклассификации деревьев пихты сибирской



Пример исходных снимков деревьев пихты, пораженных уссурийским полиграфом:



(а) – живые деревья

(б) – отмирающие деревья

(в) – свежий сухостой

(г) – старый сухостой



Оценка состояния здоровья (степени повреждения вредителями) деревьев сводится к решению задачи мультиклассификации.

(а)

(б)

(в)

(г)

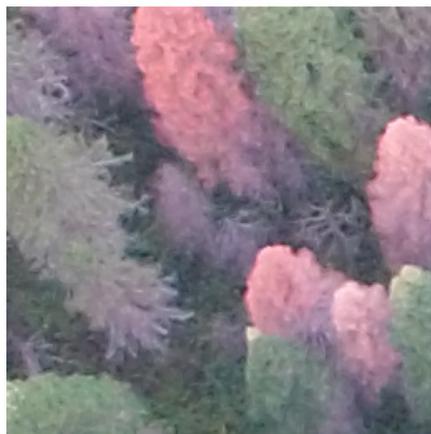
Исходный датасет по снимкам крон деревьев ПИХТЫ

- Разметка снимков экспертом (получение эталонных карт сегментации) и нарезка их на фрагменты проведена сотрудниками ИМКЭС СО РАН и ТПУ.
- Обучающая выборка: 2004 фрагмента 256x256 пикселей.
- Валидационная выборка: 672 фрагмента 256x256 пикселей.
- Тестовая выборка: 91 фрагмент 256x256 пикселей.
- 4 класса (состояния) деревьев и Фон, всего – 5 классов.

Зашумление фрагментов снимков ПИХТЫ

Зашумление ведётся двумя способами:

- изменение площади зашумления при фиксированной амплитуде, равной 1.5;
- увеличение амплитуды яркости пикселя при фиксированной площади зашумления фрагмента, равной 25%.



а)



б)



в)



г)

Визуализация зашумления снимка пихты:

а) исходный фрагмент; б) зашумление 12.5% площади фрагмента; в) зашумление 25% площади фрагмента; г) зашумление 50% площади фрагмента; при неизменной амплитуде зашумления, равной 1.5

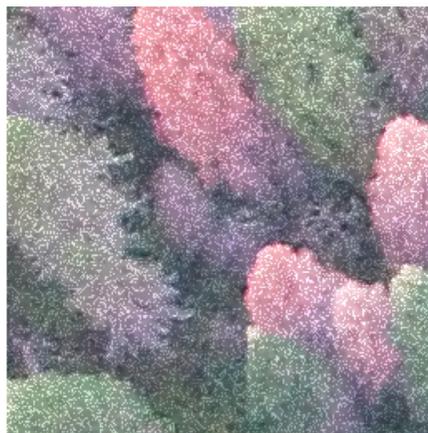
Зашумление фрагментов снимков пихты

Зашумление ведётся двумя способами:

- изменение площади зашумления при фиксированной амплитуде, равной 1.5;
- **увеличение амплитуды яркости пикселя при фиксированной площади зашумления фрагмента, равной 25%.**



а)



б)



в)

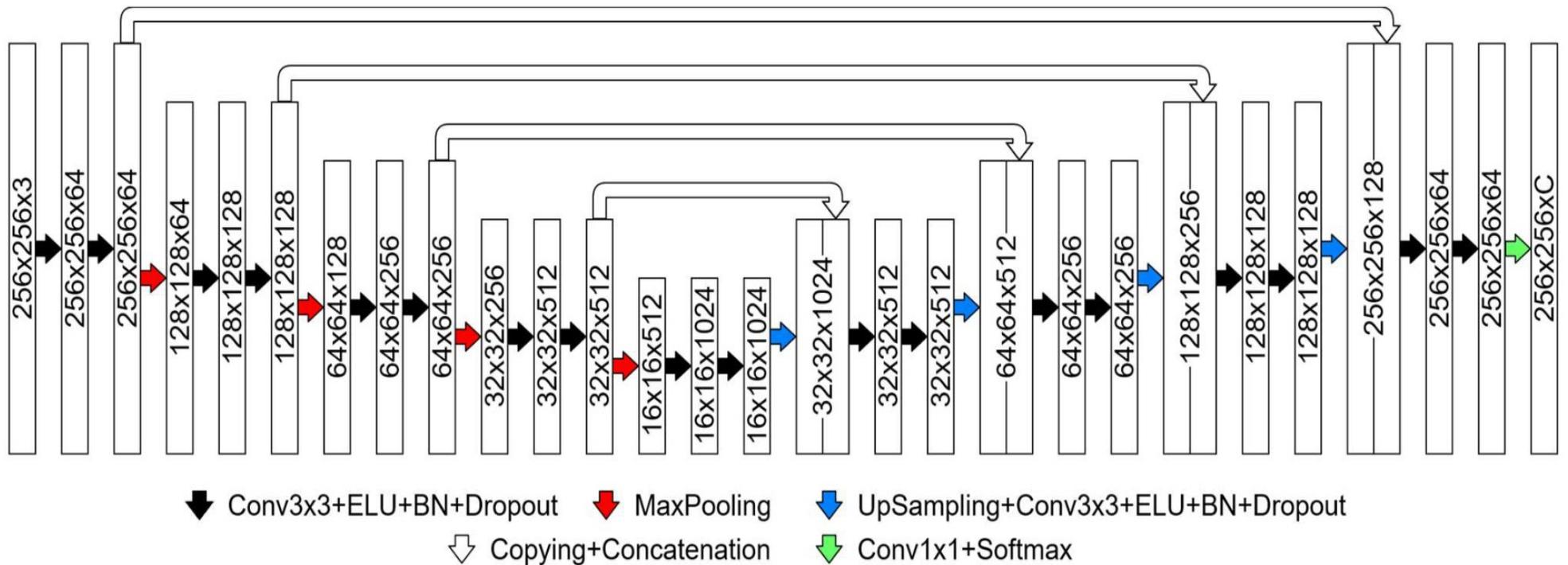


г)

Визуализация зашумления снимка пихты:

а) исходный фрагмент; б) увеличение амплитуды в 1.5 раза; в) увеличение амплитуды в 2.0 раза; г) увеличение амплитуды в 3.0 раза; при неизменной площади зашумления, равной 25% площади фрагмента

Архитектура модифицированной модели U-Net



РФП Коши

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

где $z \in [0, 1]$ – величина ошибки предсказания, β – коэффициент функции потерь

ФП Focal Loss

$$\rho(z) = -\alpha * (1 - z)^\gamma * \log(z),$$

где $z \in [0, 1]$ – величина ошибки предсказания, α и γ – коэффициенты функции потерь

Коэффициенты функций потерь изменяются в следующих пределах: [3; 9].

Метрики IoU_c , $mIoU$

Метрика IoU_c для каждого класса c деревьев может быть вычислена по формуле:

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c},$$

где TP_c , FP_c и FN_c – количество, соответственно, истинно-положительных, ложноположительных и ложно-отрицательных решений для деревьев класса c .

Для оценки точности классификации деревьев используется интегрированная метрика mean Intersection over Union ($mIoU$), рассчитываемая как среднее значение IoU_c по всем c классам.

$$mIoU = \frac{1}{C} \sum_{c=1}^C IoU_c$$

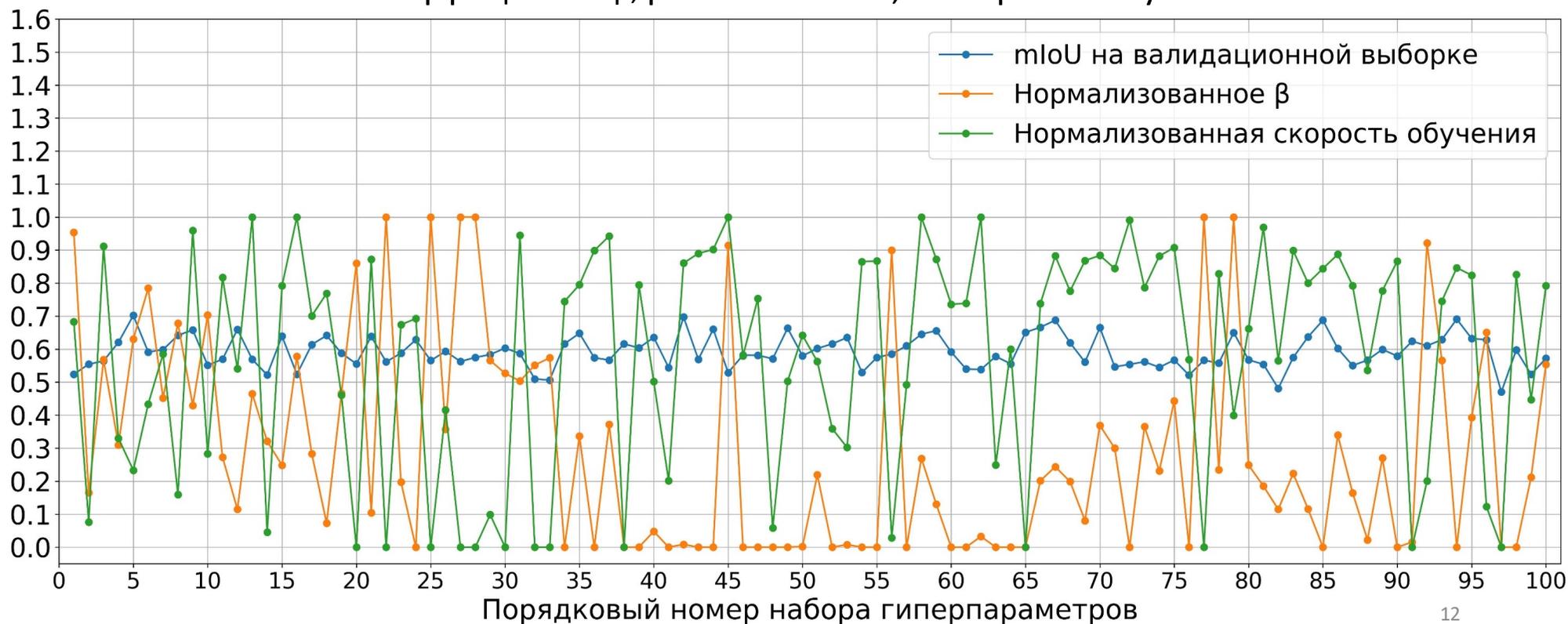
Значения метрик IoU_c и $mIoU$, превышающие 0,5, соответствуют высокой точности классификации деревьев.

Планирование экспериментов

- Планируется две серии экспериментов: по одной для каждой модели U-Net. Первая серия – с использованием РФП Коши, вторая – с использованием ФП Focal Loss.
- В каждой из серий проводится 11 экспериментов, в которых используются **зашумлённые фрагменты обучающей выборки** с разной долей зашумления по площади (от 0% до 50%) и разной амплитудой ($A = 1.5, 2.0, 3.0$) импульсных помех (выбросов).
- Дополнительно исследуется влияние различных комбинаций **зашумления обучающей, валидационной и тестовой выборок** на точность классификации. Фрагменты будут иметь зашумленную площадь 25% от площади фрагмента и амплитуду выбросов 1.5.
- Применяется алгоритм байесовской оптимизации при формировании наборов гиперпараметров.

РФП Коши. Обучение на фрагментах без зашумления

Максимальное значение метрики mIoU на валидационной выборке 0.7884 достигнуто при значении коэффициента β , равном 6.5205, и скорости обучения 0.0031.



ФП Focal Loss. Обучение на незашумлённых фрагментах

Максимальное значение метрики mIoU на валидационной выборке 0.7140 достигнуто при значении коэффициентов α и γ , равных 4.4331 и 3.1041, и скорости обучения 0.0064.



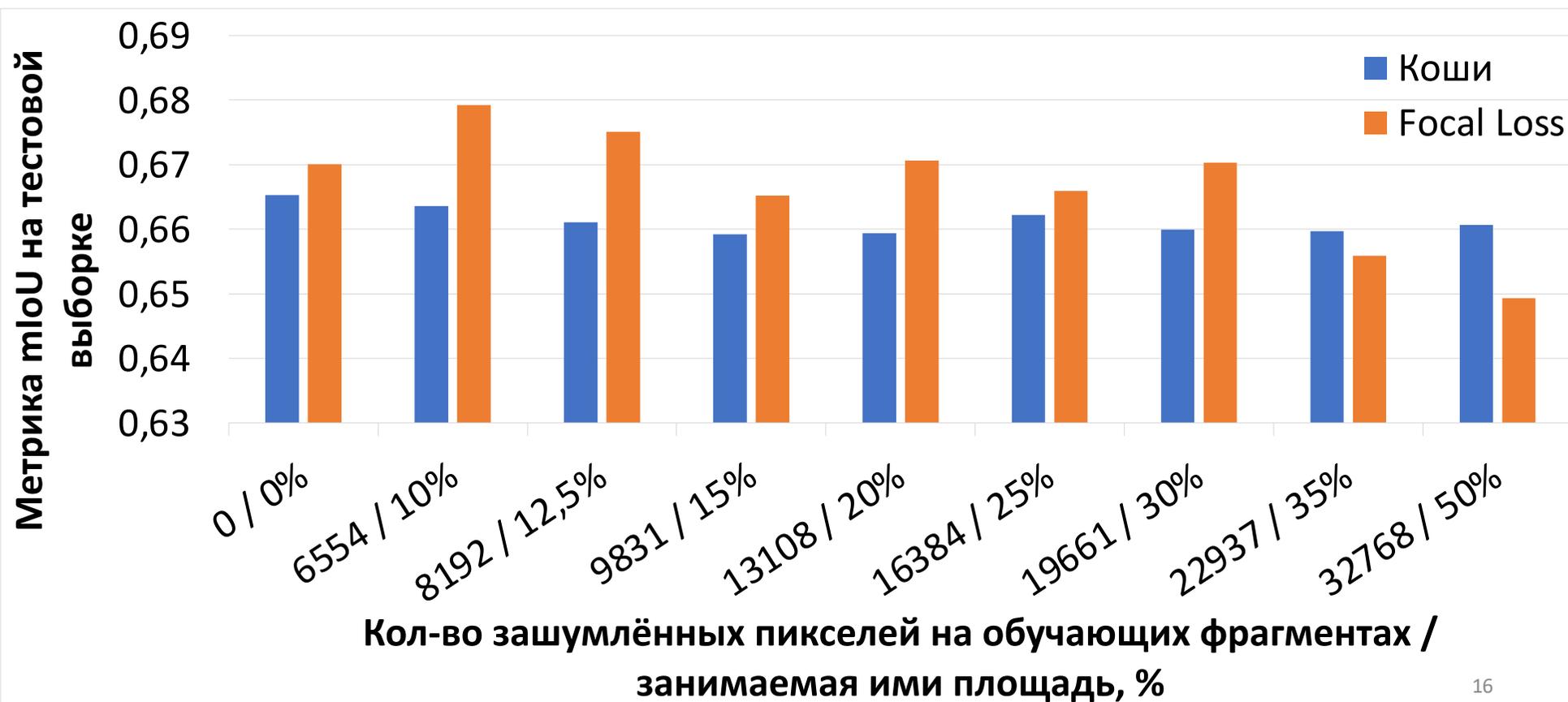
РФП Коши. Точность классификации деревьев по метрике $mIoU$ в зависимости от параметров зашумления обучающих фрагментов (случай постоянной амплитуды выбросов)

Занимаемая зашумлёнными пикселями площадь обучающих фрагментов, %	Амплитуда имп. помех (выбросов)	Скорость обучения	Коэффициент β	$mIoU$ на тестовой выборке
0%	0	0.0031	6.5205	0.6653
10%	1.5	0.0098	3.1708	0.6636
12,5%	1.5	0.0073	8.1773	0.6611
15%	1.5	0.0029	8.9910	0.6592
20%	1.5	0.0090	3.0058	0.6594
25%	1.5	0.0086	3.1708	0.6622
30%	1.5	0.0071	4.9932	0.6599
35%	1.5	0.0061	8.9851	0.6597
50%	1.5	0.0010	6.6832	0.6607

ФП Focal Loss. Точность классификации деревьев по метрике mIoU в зависимости зашумления обучающих фрагментов (случай постоянной амплитуды выбросов)

Занимаемая зашумлёнными пикселями площадь обучающих фрагментов, %	Амплитуда имп. помех (выбросов)	Скорость обучения	Коэффициент		mIoU на тестовой выборке
			α	γ	
0%	0.0	0.0064	4.4331	3.1041	0.6701
10%	1.5	0.0067	3.1987	6.7168	0.6792
12,5%	1.5	0.0032	6.7553	3.1520	0.6751
15%	1.5	0.0043	6.2227	5.5577	0.6652
20%	1.5	0.0028	7.7003	7.1951	0.6706
25%	1.5	0.0070	6.1240	4.9331	0.6659
30%	1.5	0.0078	3.2237	5.6579	0.6703
35%	1.5	0.0086	5.5296	6.4003	0.6559
50%	1.5	0.0032	5.0804	3.6445	0.6493

Значения метрики mIoU на тестовой выборке, показанные модифицированными моделями U-Net в результате обучения на зашумлённых фрагментах ($A = 1.5$)



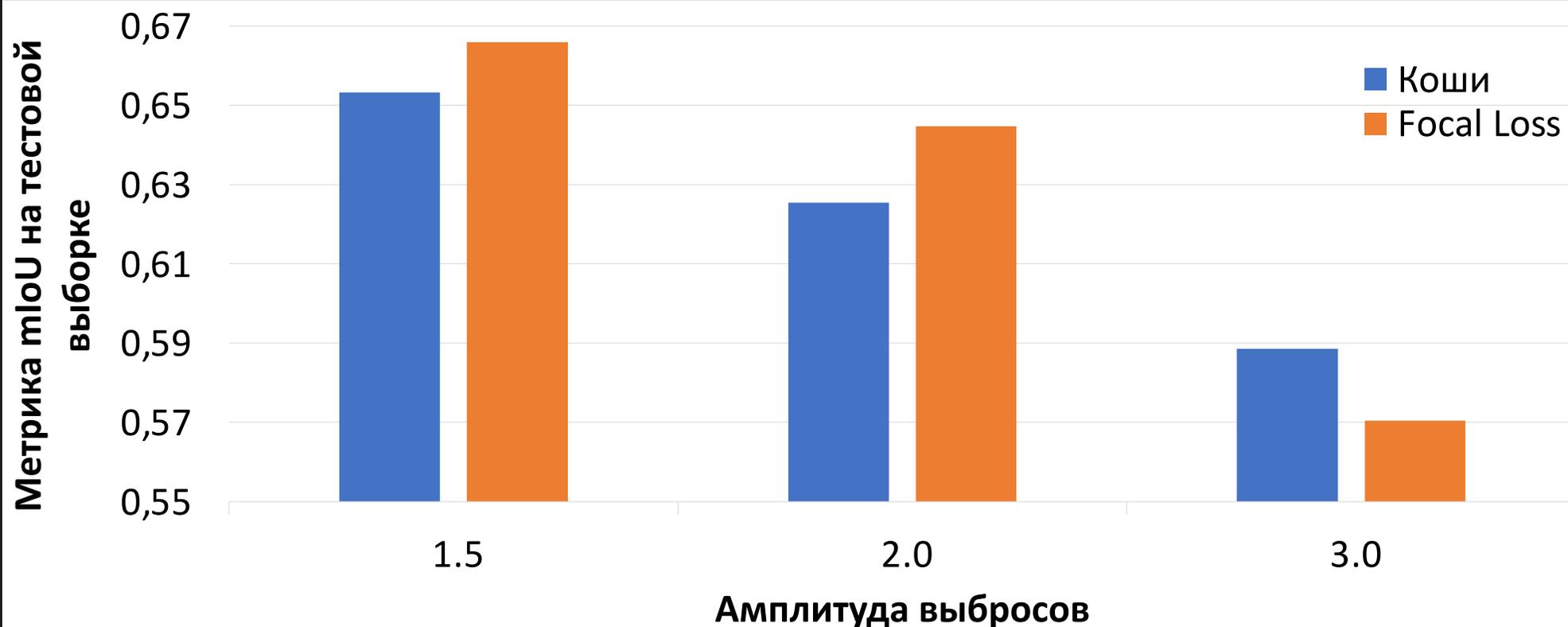
РФП Коши. Зависимость точности классификации деревьев по метрике mIoU от параметров зашумления обучающих фрагментов (случай постоянной площади выбросов)

Занимаемая зашумлёнными пикселями площадь обучающих фрагментов, %	Амплитуда имп. помех (выбросов)	Скорость обучения	Коэффициент β	mIoU на тестовой выборке
25%	1.5	0.0086	3.1708	0.6533
	2.0	0.0010	6.0652	0.6254
	3.0	0.0023	6.7963	0.5885

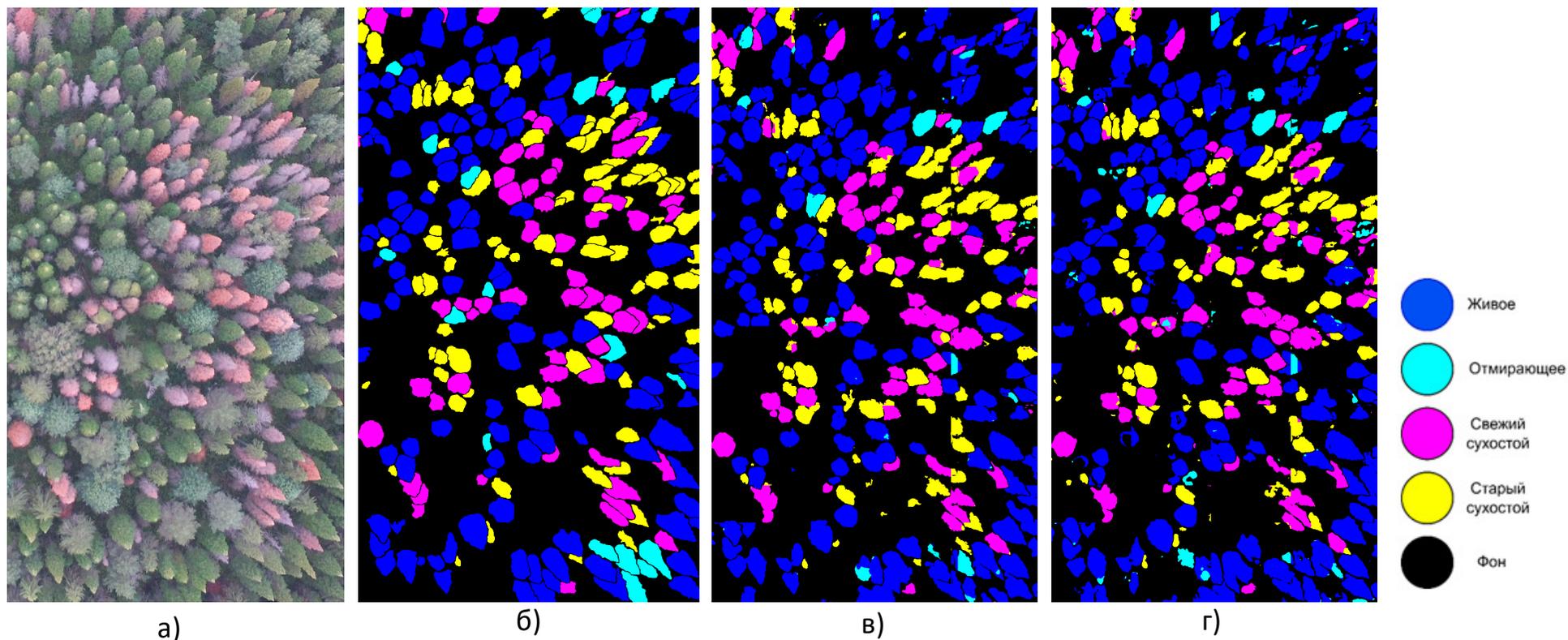
ФП Focal Loss. Зависимость точности классификации деревьев по метрике mIoU от параметров зашумления обучающих фрагментов (случай постоянной площади выбросов)

Занимаемая зашумлёнными пикселями площадь обучающих фрагментов, %	Амплитуда имп. помех (выбросов)	Скорость обучения	Коэффициент		mIoU на тестовой выборке
			α	γ	
25%	1.5	0.0070	6.1240	4.9331	0.6659
	2.0	0.0010	4.0445	3.0102	0.6447
	3.0	0.0010	5.8859	7.3226	0.5705

Результаты тестирования модифицированных моделей U-Net после обучения на зашумлённых фрагментах с различной амплитудой выбросов в случае доли зашумления 25%

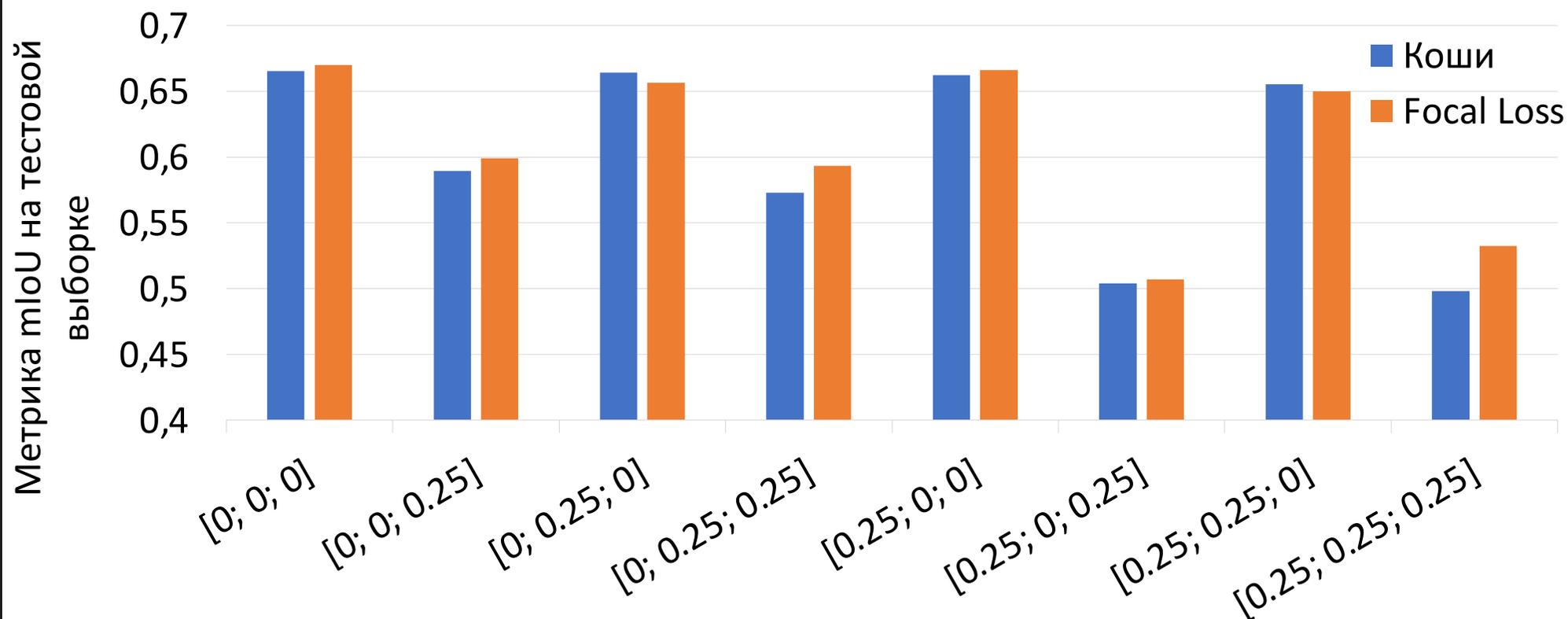


Сегментация тестового снимка по результатам экспериментов (зашумление 25% площади обучающих фрагментов, амплитуда выбросов – 1.5)



Сегментация тестового снимка: а) исходный снимок; б) эталонная карта; результаты сегментации с помощью моделей СНС, обученной с применением: в) РФП Коши; г) ФП Focal Loss

Значения метрики mIoU на тестовой выборке, полученные в зависимости от комбинаций зашумленных выборок



Отношение площади зашумления **[обучающих; валидационных; тестовых]** фрагментов к общей площади фрагментов ($A = 1.5$)

Заключение

Решены поставленные задачи подготовки датасета и модификации моделей U-Net. По полученным результатам исследований можно сделать следующие выводы:

- Применение РФП Коши в модифицированной модели U-Net позволяет повысить точность классификации поражённых деревьев пихты на сильно зашумлённых фрагментах обучающей выборки (доля выбросов – не менее 35% или амплитуда выбросов составляет 3.0). В других случаях параметров зашумления фрагментов при обучении модифицированной модели U-Net рекомендуется применять ФП Focal Loss.
- Точность классификации поражённых деревьев пихты значительно зависит от зашумления фрагментов обучающей и тестовой выборок. Зашумление фрагментов валидационной выборки оказывает слабое влияние на точность классификации. Наихудшая точность классификации будет в случае зашумления всех трёх выборок.
- Полученные результаты исследований и сделанные по ним выводы имеют научную ценность, поскольку указывают на принципиальную возможность использования модифицированных моделей U-Net при анализе зашумлённых импульсными помехами снимков хвойных деревьев.



**ТОМСКИЙ
ПОЛИТЕХ**



СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ ПОВРЕЖДЕННЫХ ДЕРЕВЬЕВ ПИХТЫ СИБИРСКОЙ НА ЗАШУМЛЕННЫХ СНИМКАХ С БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Малкин Артем Юрьевич

Марков Николай Григорьевич

12 ноября 2024