

Классификация лесных насаждений с близкими спектральными характеристиками по RGB-изображениям сверхвысокого пространственного разрешения (на примере осиновых и березовых насаждений)

¹Калашников Р.А., ^{1,2}Пестунов И.А., ³Рылов С.А., ²Ермаков Н.Б., ⁴Радчиков А.Н.

¹ ФИЦ ИВТ

² Никитский ботанический сад – Национальный научный центр РАН

³ Хакасский государственный университет им. Н. Ф. Катанова

⁴ ООО «Беспилотные технологии»

г. Новосибирск

Актуальность и цель работы

- Ручное выделение и классификация объектов на изображениях высокого и сверхвысокого пространственного разрешения – трудоемкий процесс.
- Значительная часть информации об объекте содержится не только в спектральных, но и в семантических признаках (цветовые и текстурные признаки, особенности формы, признаки отношения и т.п.)
- Классические методы анализа изображений применительно к данной задаче малоэффективны.

Цель работы – исследовать возможность распознавания осиновых и березовых насаждений с помощью нейросетевой семантической сегментации RGB-изображений сверхвысокого пространственного разрешения.

Область практического применения: задачи контроля и оценки состояния лесной растительности, а также оценка стоимости аренды лесных площадей.

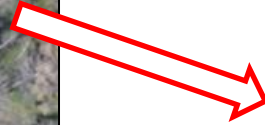
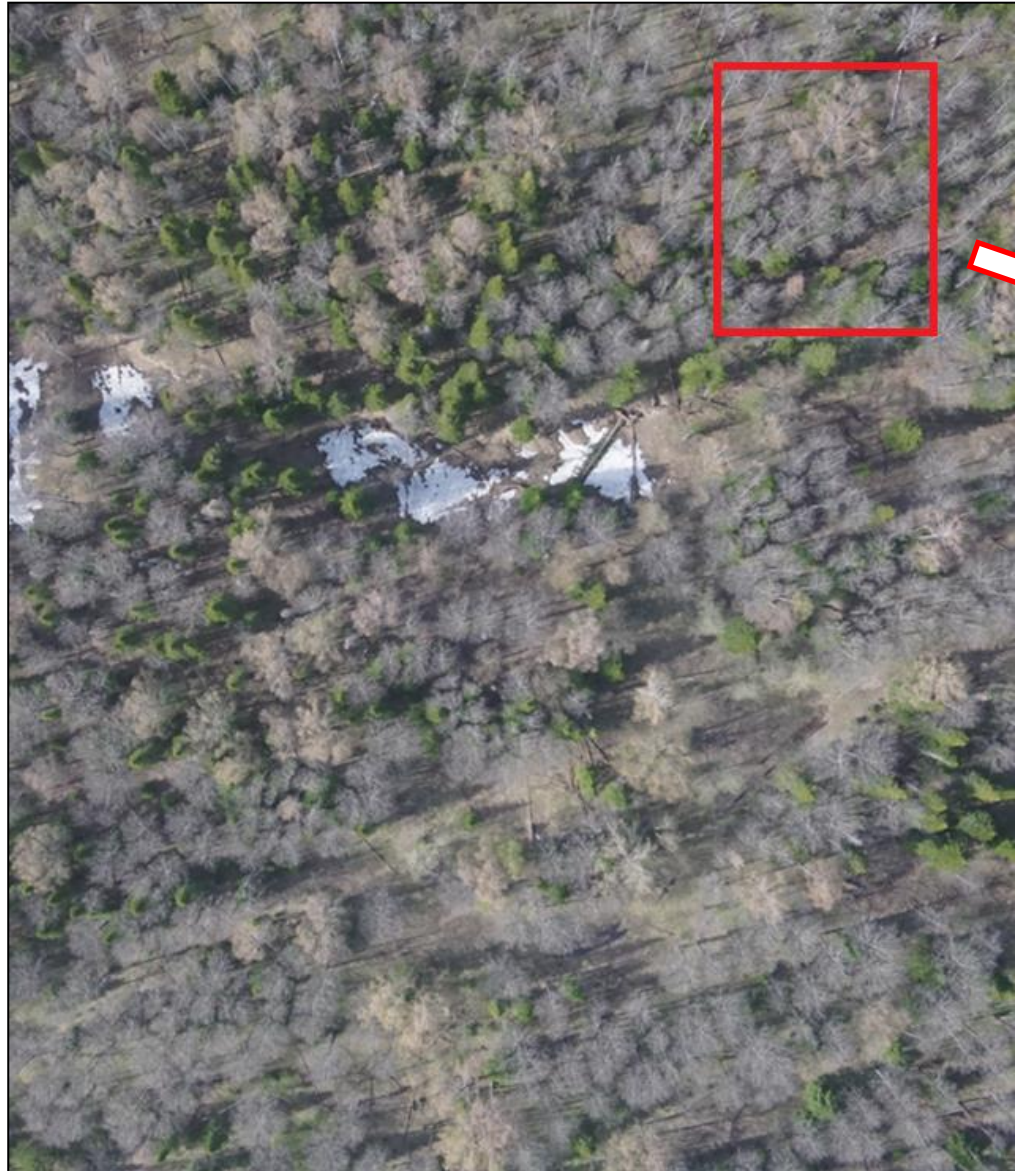
Объект исследования и исходные данные

Исследовались лесные насаждения в районе п. Аламбай Алтайского края.
Для обработки использовалось 68 RGB-изображений размером 6000x4000 пикселей с разрешением ~5 см, полученных с помощью БПЛА с высоты 250 м в **апреле** 2021 г.



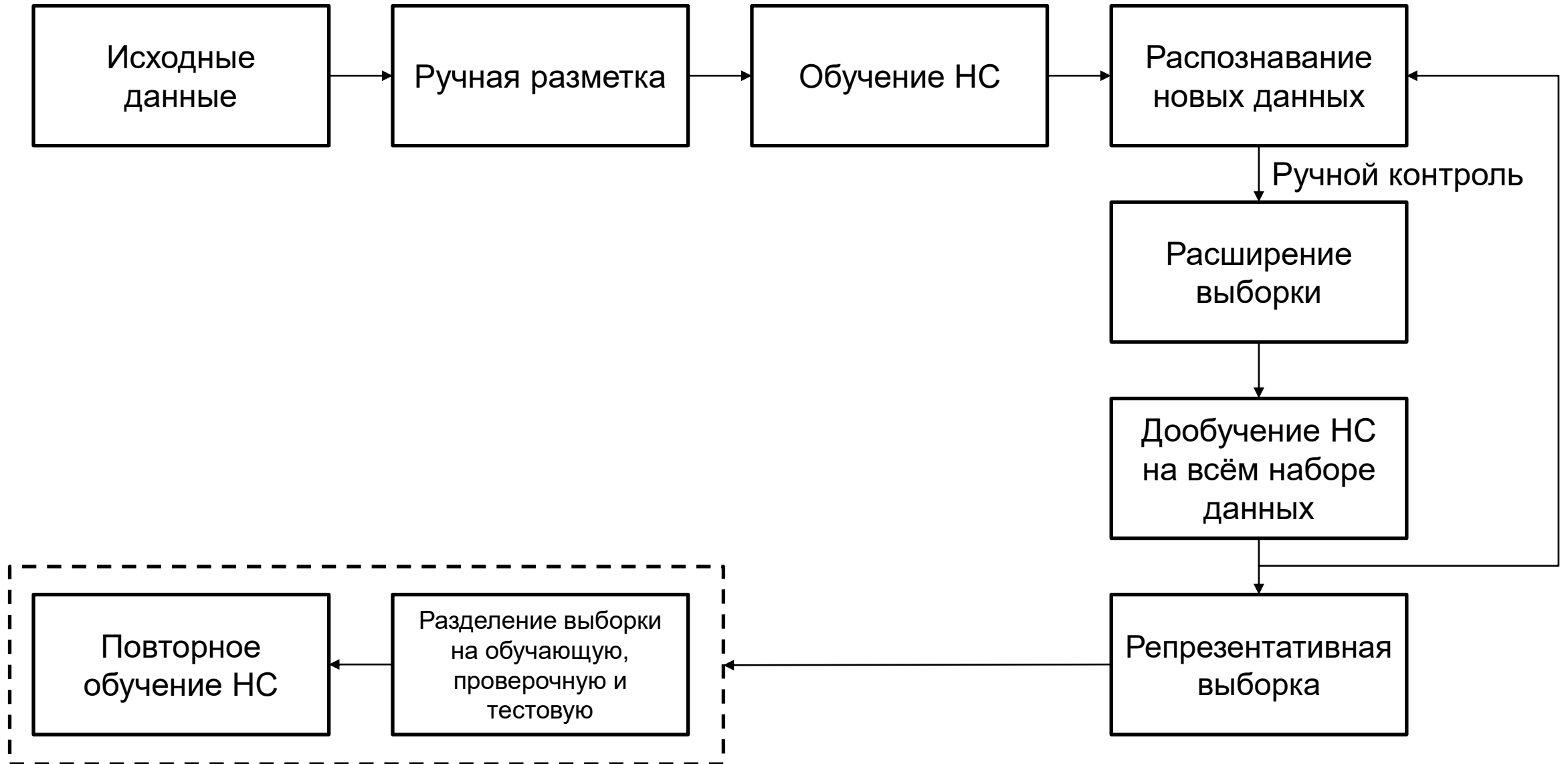
Пример исходного RGB-изображения размером 6000x4000 пикселей

Сложность проблемы



Увеличенный фрагмент изображения

Схема полуавтоматического расширения обучающей выборки с использованием нейросетей



Кластер ФИЦ ИВТ

Каждый из трёх узлов:

- HPE Apollo 6500 G10+
- 2 x AMD EPYC 7452 (32 Cores) 2,35 GHz
- 1 TB RAM
- 8 x Nvidia A100 80GB SXM4



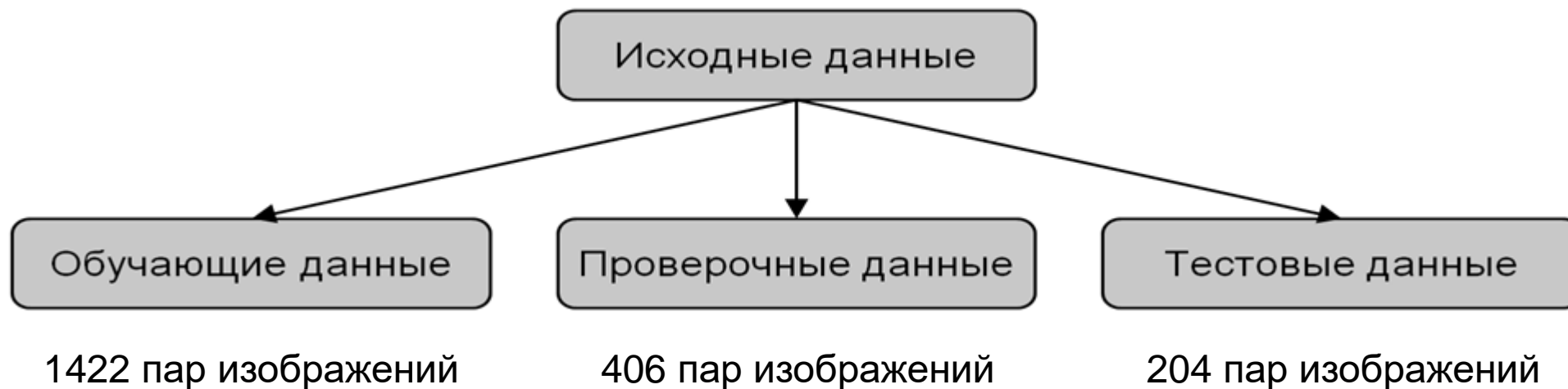
NVIDIA A100 TENSOR CORE GPU SPECIFICATIONS (SXM4 AND PCIe FORM FACTORS)

	A100 80GB PCIe	A100 80GB SXM
FP64	9.7 TFLOPS	
FP64 Tensor Core	19.5 TFLOPS	
FP32	19.5 TFLOPS	
Tensor Float 32 (TF32)	156 TFLOPS 312 TFLOPS*	
BFLOAT16 Tensor Core	312 TFLOPS 624 TFLOPS*	
FP16 Tensor Core	312 TFLOPS 624 TFLOPS*	
INT8 Tensor Core	624 TOPS 1248 TOPS*	
GPU Memory	80GB HBM2e	80GB HBM2e
GPU Memory Bandwidth	1,935GB/s	2,039GB/s
Max Thermal Design Power (TDP)	300W	400W***
Multi-Instance GPU	Up to 7 MIGs @ 10GB	Up to 7 MIGs @ 10GB
Form Factor	PCIe dual-slot air cooled or single-slot liquid cooled	SXM
Interconnect	NVIDIA® NVLink® Bridge for 2 GPUs: 600GB/s ** PCIe Gen4: 64GB/s	NVLink: 600GB/s PCIe Gen4: 64GB/s

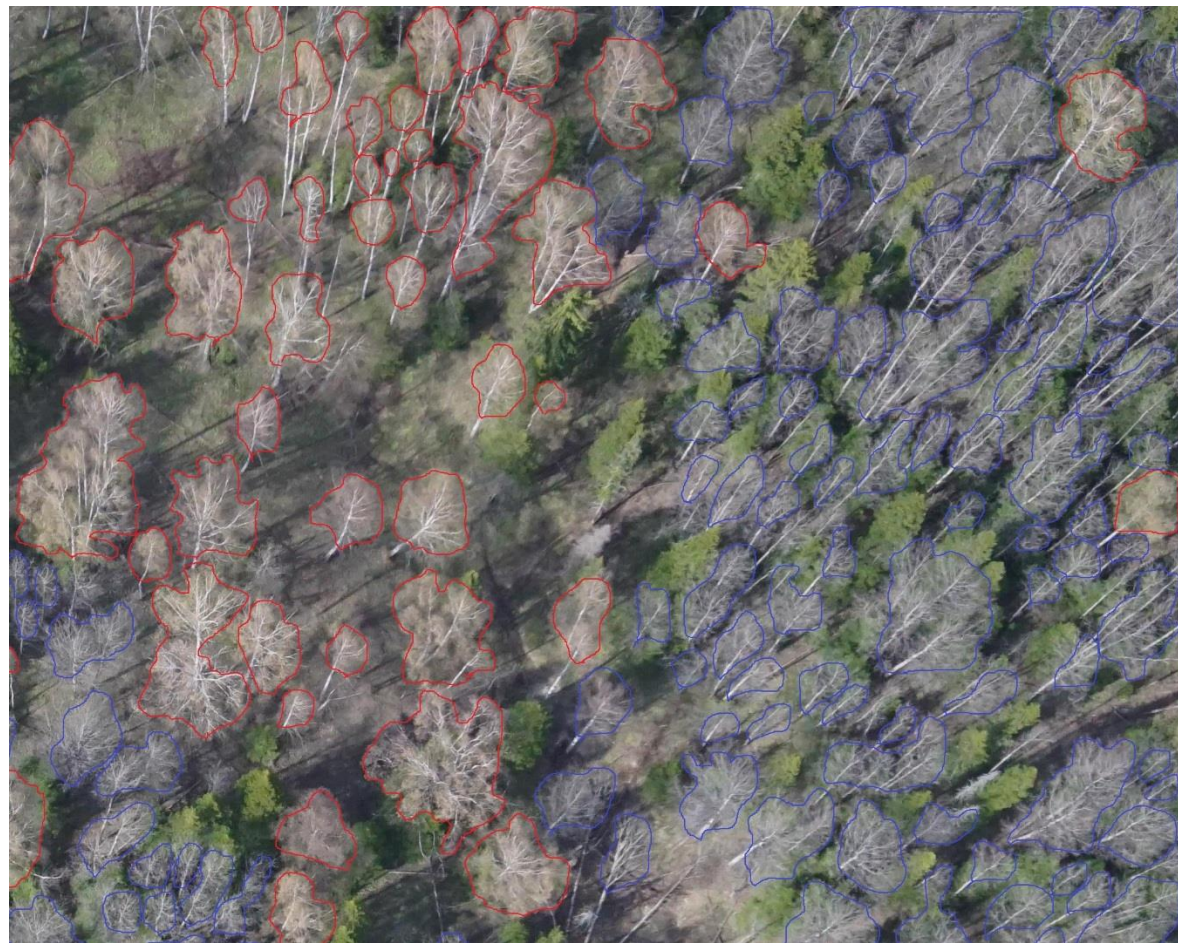
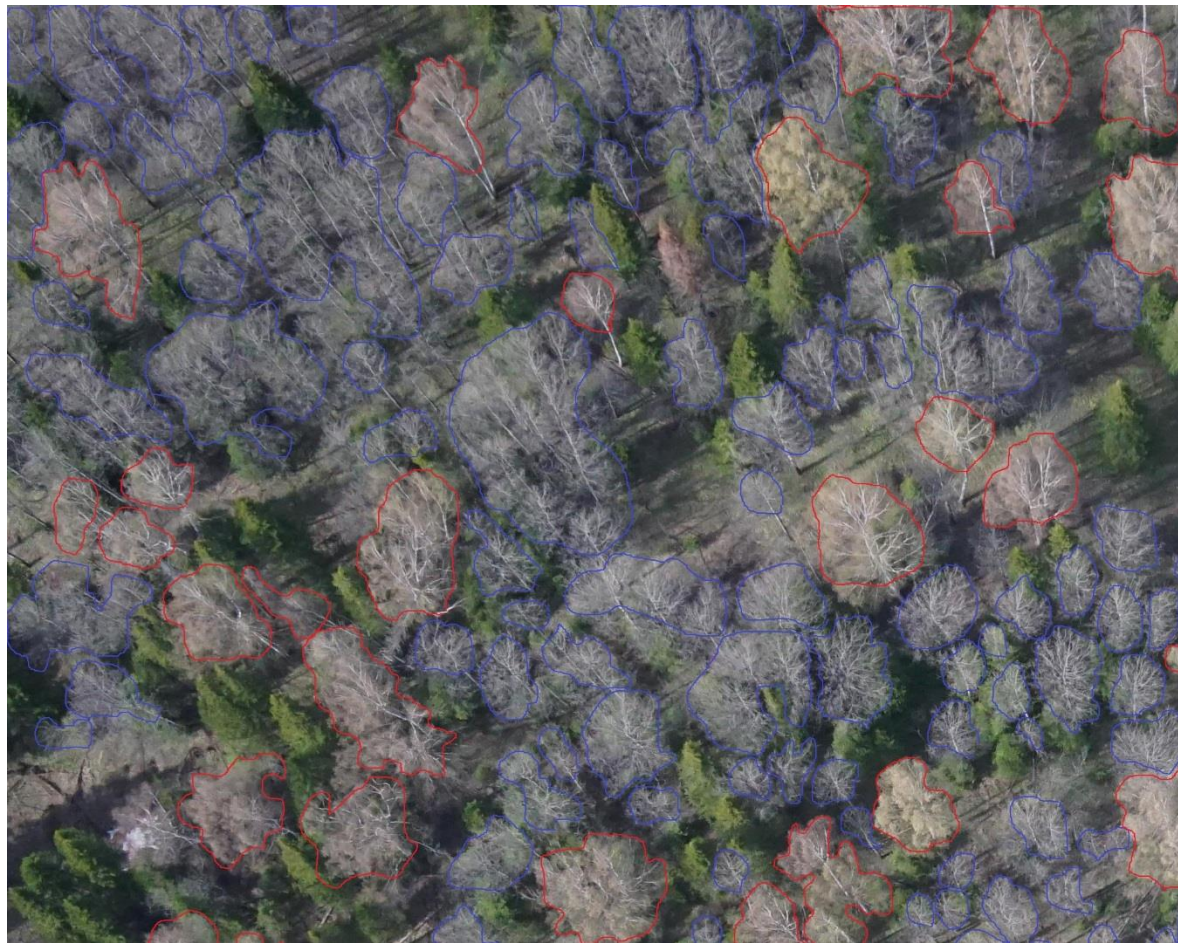
Формирование обучающей выборки

Работа, выполнявшаяся ранее, проводилась с использованием 3000 фрагментов размером 256x256 пикселей. Однако, фрагменты размером 256x256 пикселей как правило не позволяли покрыть интересующий объект целиком.

Формирование используемой в данной работе обучающей выборки, содержащей фрагменты размером 512x512 пикселей, производилось вручную с использованием веб-сервиса Supervisely и предложенного алгоритма расширения выборки. Таким образом, итоговая выборка состояла из 1063 фрагментов, размеченных вручную, и 969 фрагментов, полученных с помощью алгоритма расширения данных.



Примеры ручной разметки изображений



Примеры фрагментов из обучающей выборки



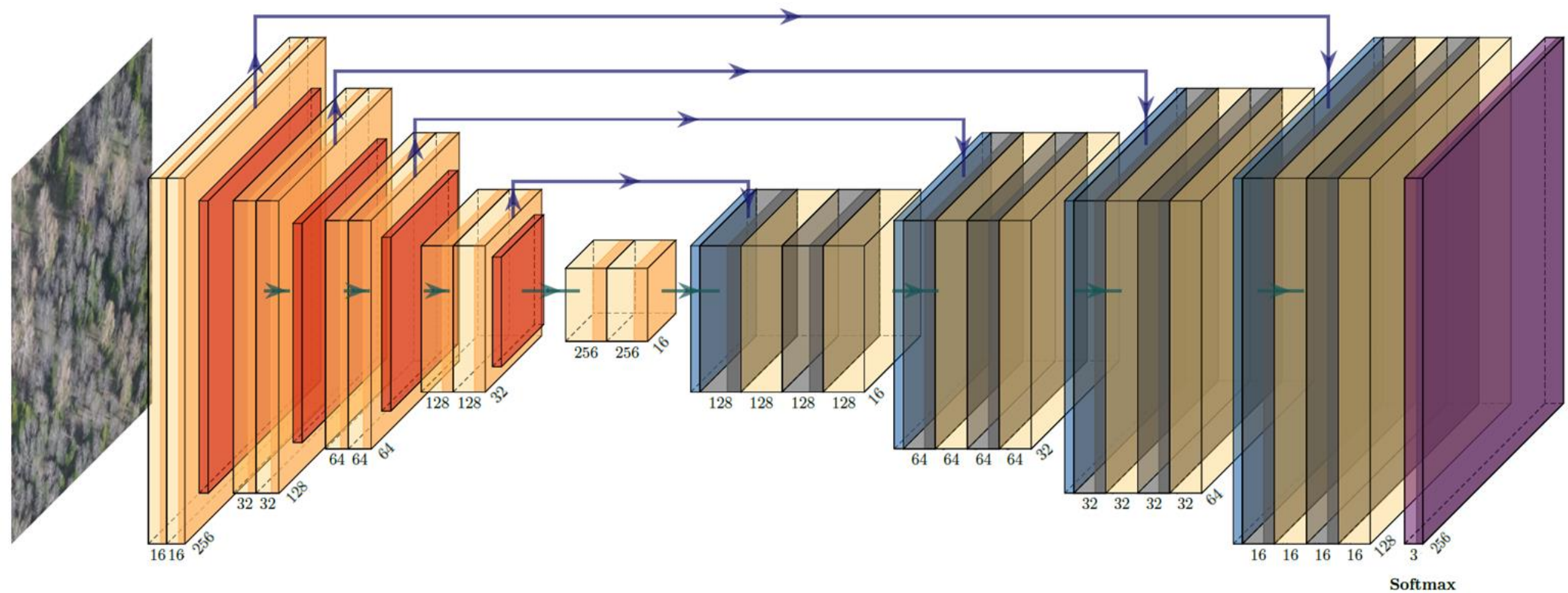
Примеры пар фрагментов и их масок с размерами 256x256 и 512x512 пикселей

Аугментация данных

Аугментация обучающих данных производилась с помощью:

- отображения относительно вертикальной оси;
- отображения относительно горизонтальной оси;
- поворота на углы, кратные 90° ;
- растяжения / сжатия по произвольной оси.

Попытка применения различных яркостных преобразований для аугментаций приводила к ухудшению результатов распознавания объектов, поэтому такие преобразования не использовались.



Архитектура ранее использованной СНС

Параметры модели сверточной нейронной сети

Параметры модели:

- В качестве метрики выбрана функция Intersection over Union:

$$IoU = \frac{\text{Площадь пересечения областей}}{\text{Площадь объединения областей}}$$

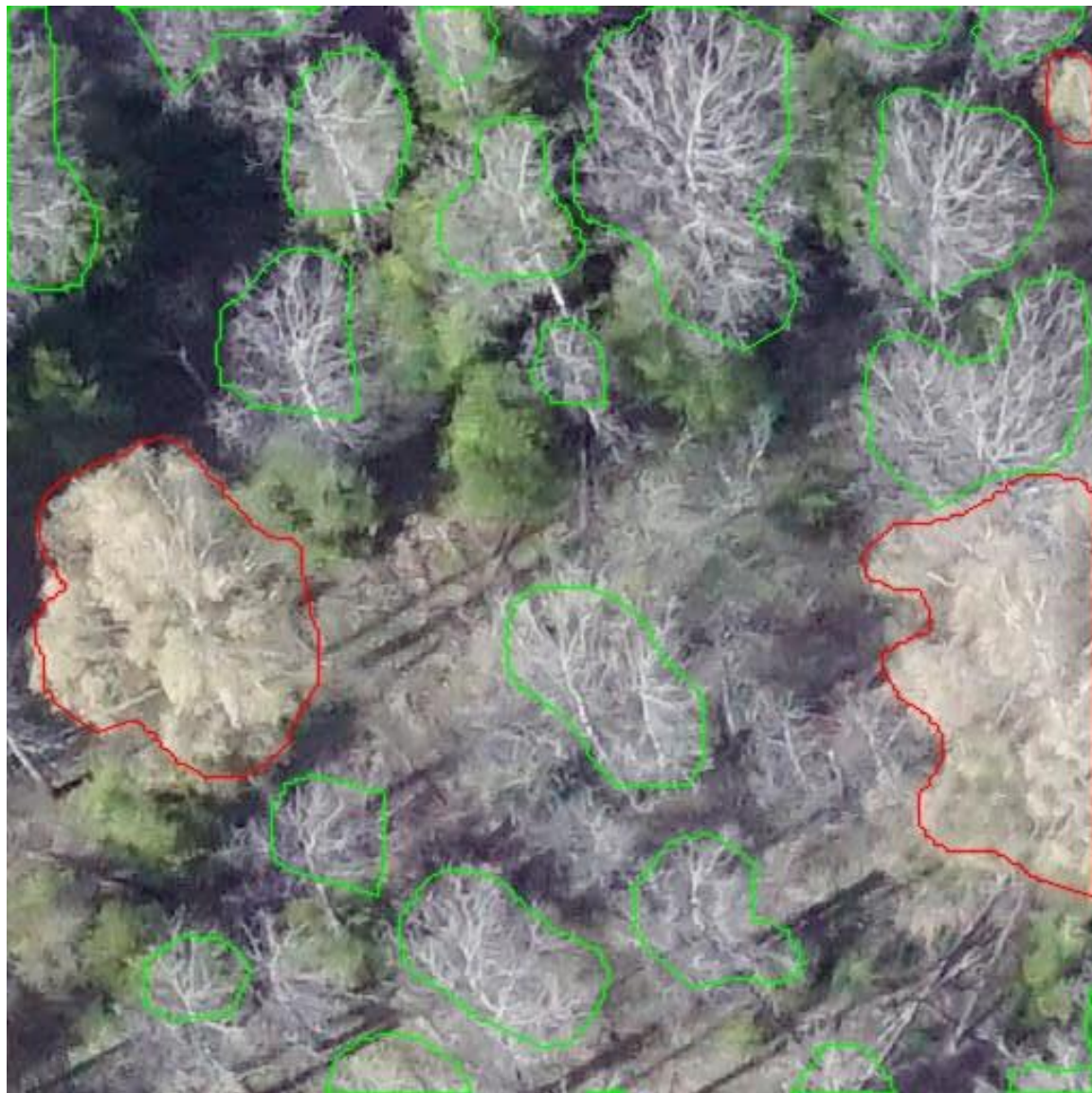
- В качестве функции потерь использована функция потерь Дайса:

$$Dice = \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2},$$

где p_i - пиксели, сегментированные сетью; g_i - пиксели маски.

- В качестве оптимизатора – Adam со скоростью обучения равной 10^{-4} .
- Мини-пакеты размером 16.
- Функции активации $ReLU(x) = \max(0, x)$.

Улучшение качества разметки с использованием НС

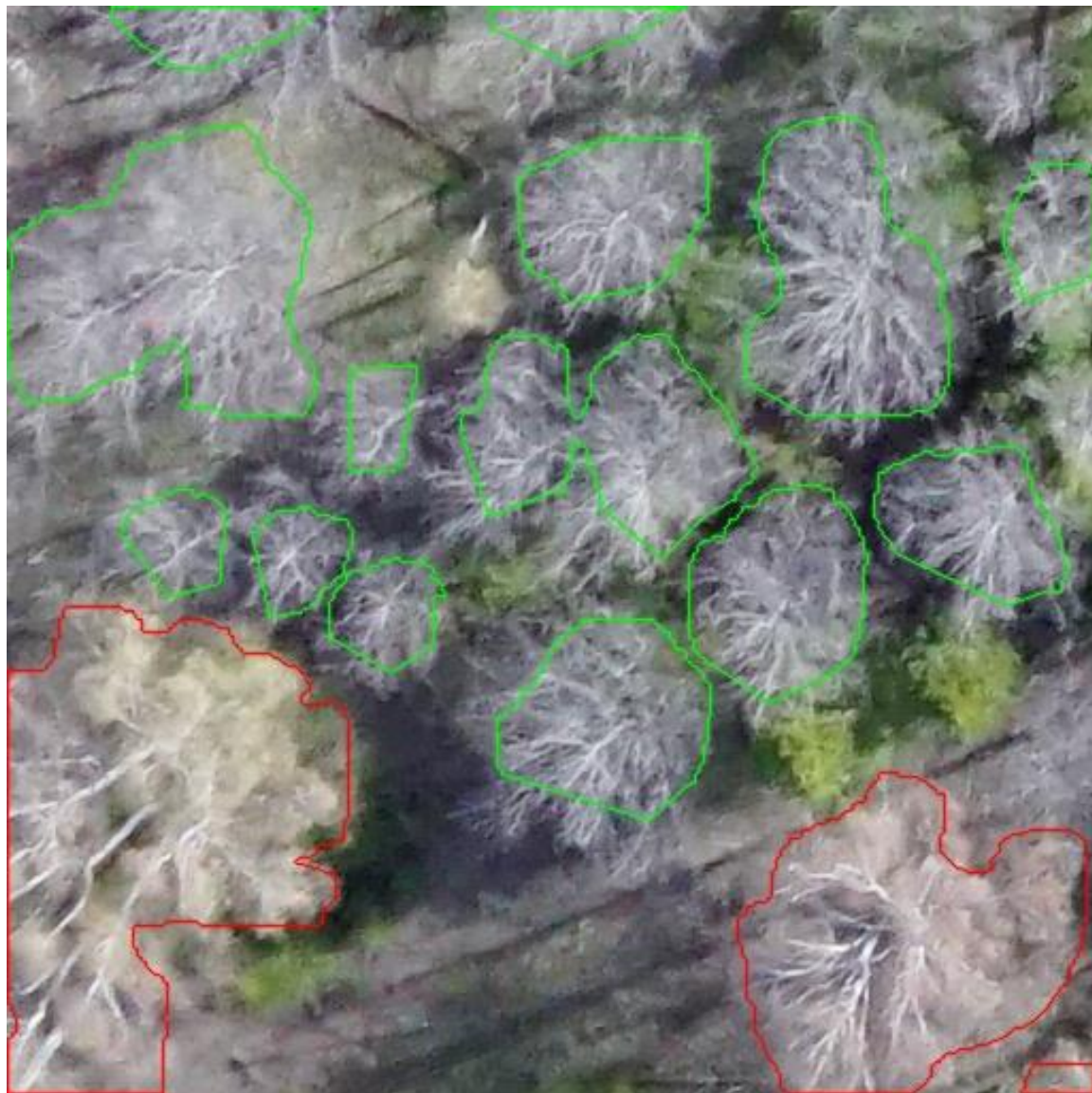


Исходная разметка



Улучшенная разметка

Улучшение качества разметки с использованием НС



Исходная разметка

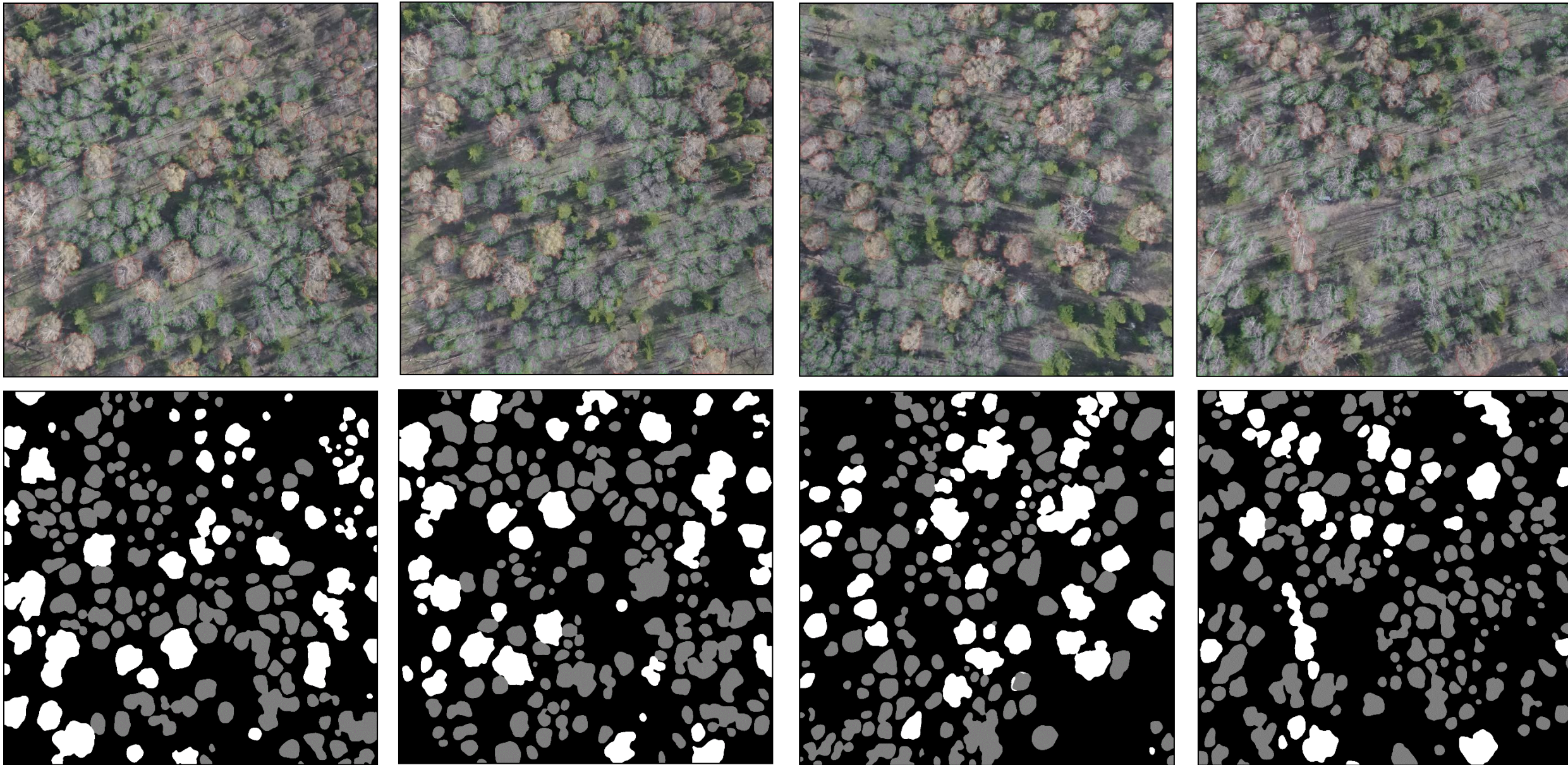


Улучшенная разметка

Результаты сравнительного анализа

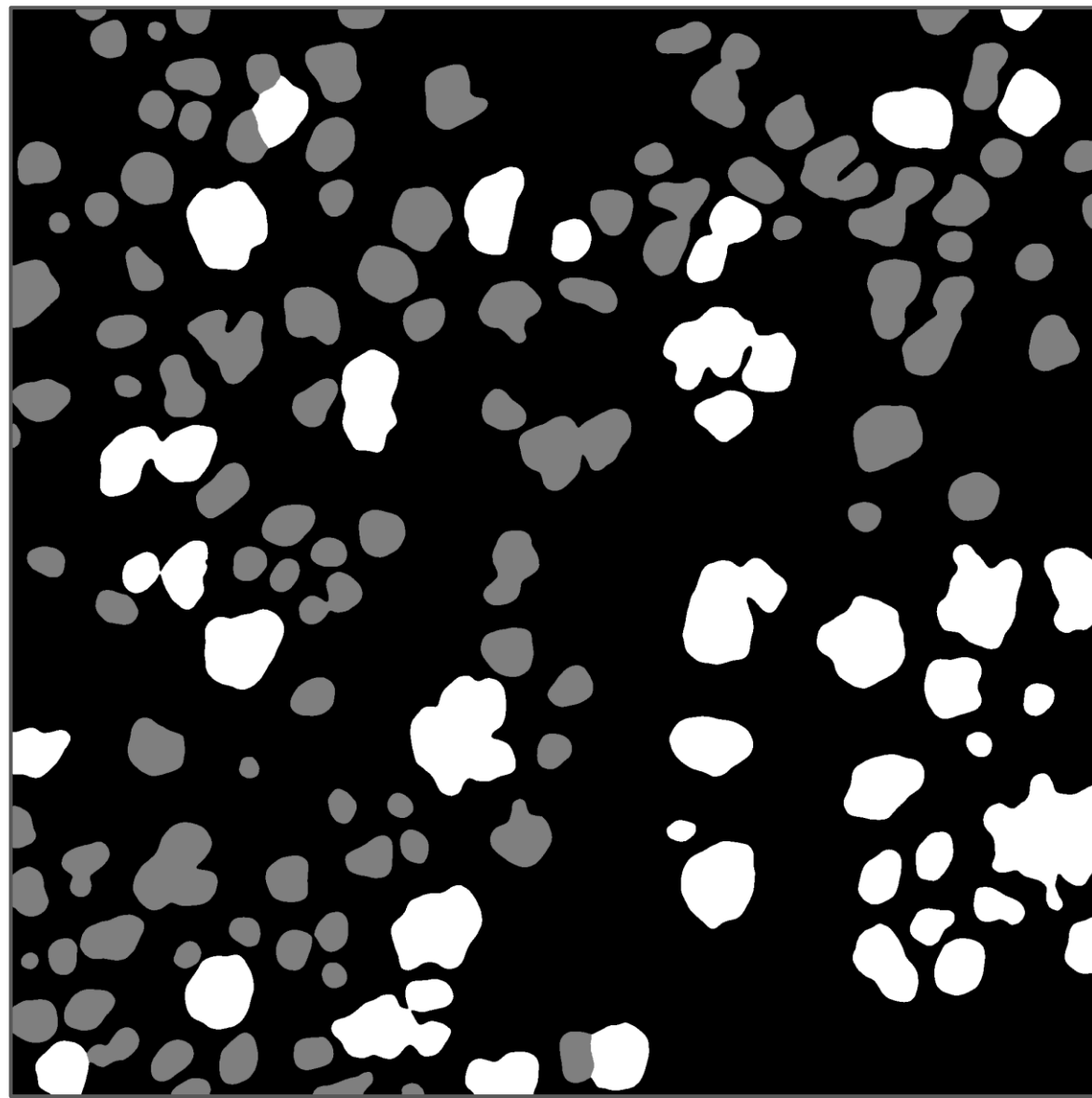
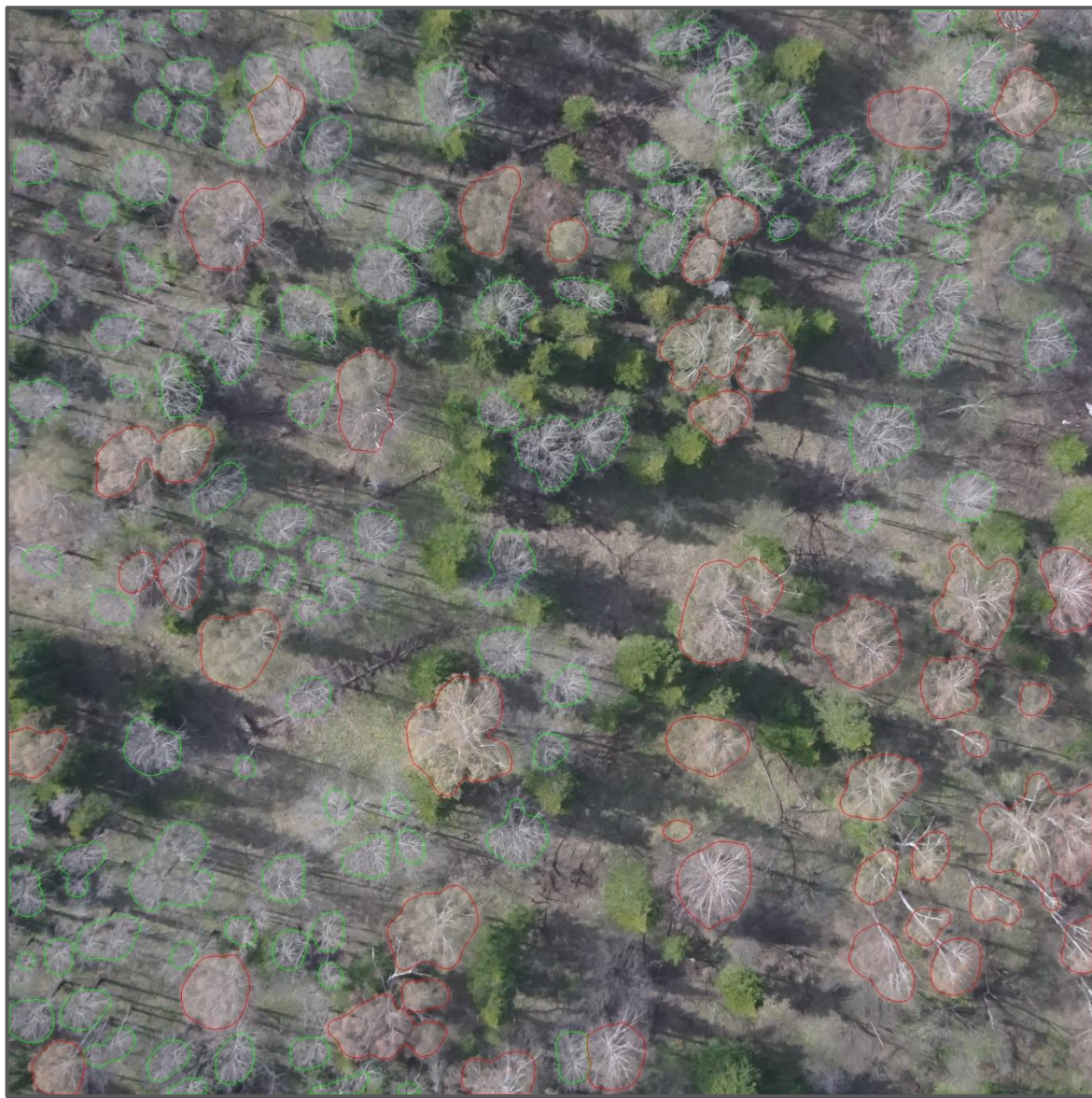
Данные	Обучающие	Проверочные	Тестовые	Обучающие	Проверочные	Тестовые
Размер фрагментов (пикс.)	256x256			512x512		
Значение метрики	0.922	0.817	0.784	0.967	0.882	0.859
Значение функции потерь	0.041	0.104	0.128	0.018	0.064	0.077

Примеры работы сверточной нейронной сети



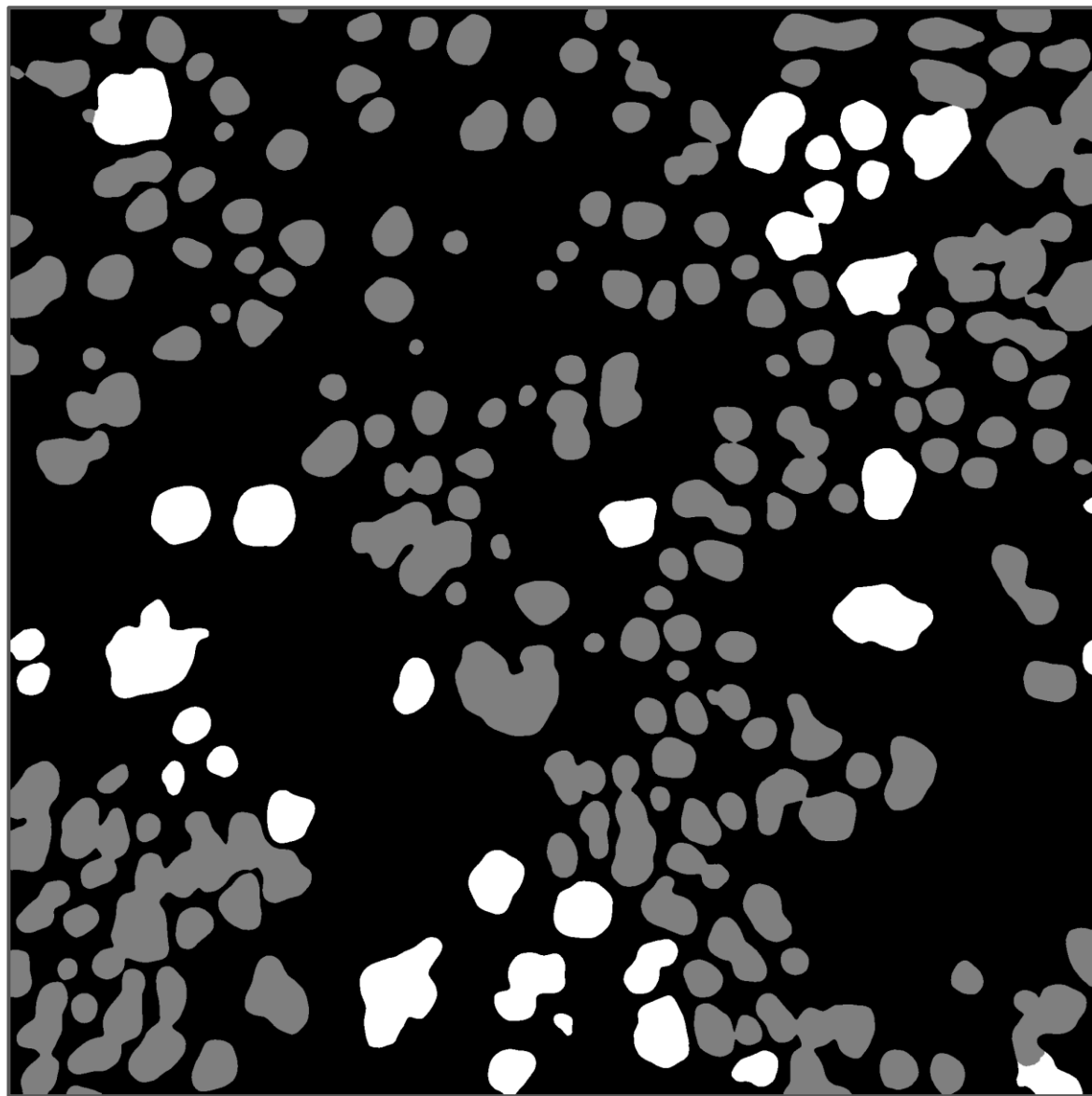
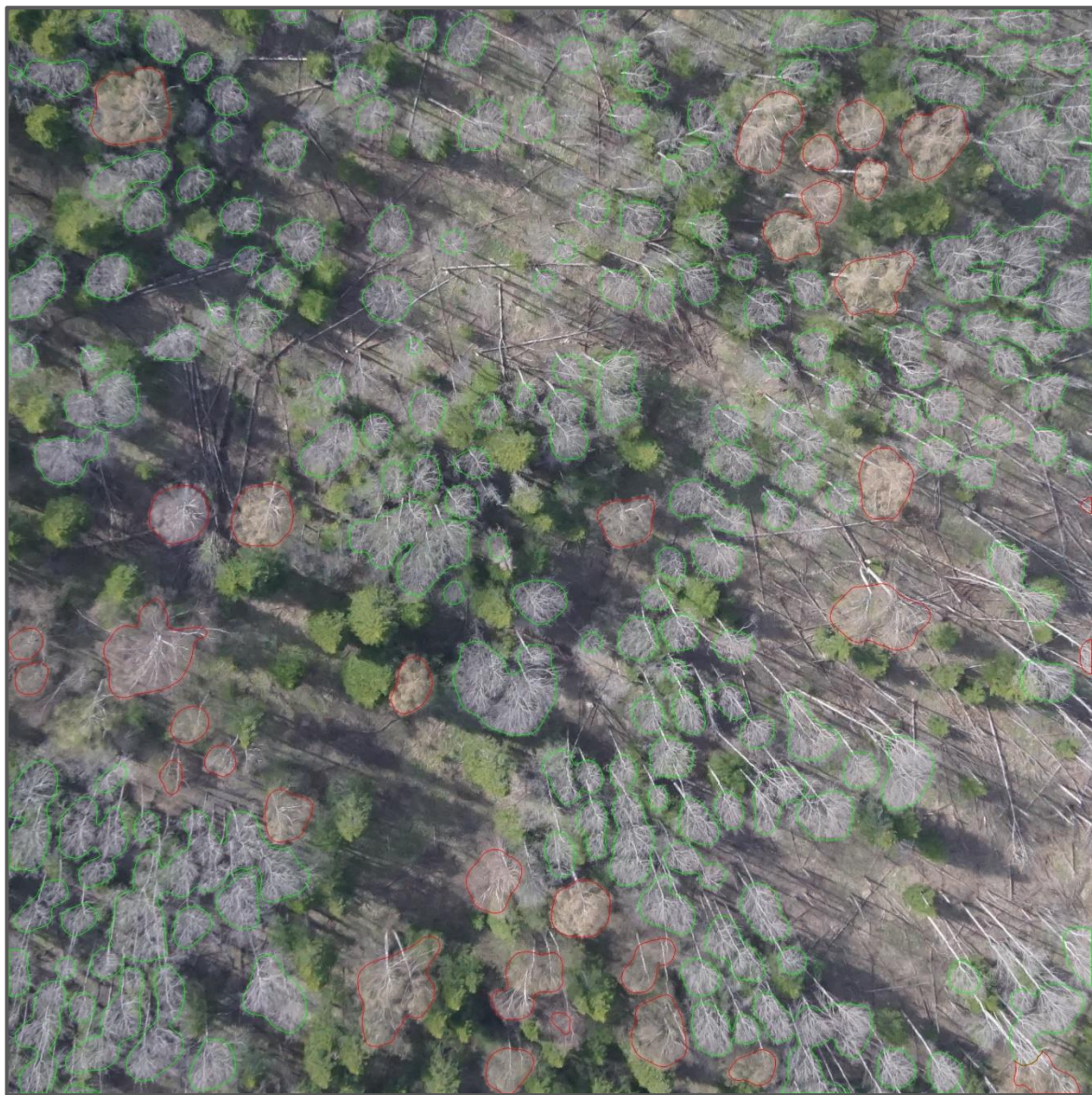
Примеры работы алгоритма на изображениях размером 2048x2048 пикселей. На сегментированных изображениях белым цветом обозначены березовые насаждения, серым – осиновые.

Примеры работы сверточной нейронной сети



Результаты предсказаний: 13,79% - осиновые насаждения, 25,14% - берёзовые насаждения, 61,07% - фон. 18

Примеры работы сверточной нейронной сети



Результаты предсказаний: 21,82% - осиновые насаждения, 12,97% - берёзовые насаждения, 65,21% - фон. 19

Результаты

- Показана возможность распознавания осиновых и березовых насаждений с помощью нейросетевой семантической сегментации RGB-изображений сверхвысокого пространственного разрешения.
- Выполнена разметка и аугментация исходных изображений.
- Предложена сверточная нейронная сеть на основе архитектуры U-Net.
- Предложен алгоритм полуавтоматического расширения обучающей выборки.
- Проведена сравнительная оценка полученной модели с классической. Предложенная модель обеспечила точность сегментации ~96% на обучающей, ~88% на проверочной и ~85% на тестовой выборках.
- Предложенная модель показала результаты на тестовой выборке на 7% выше классической архитектуры U-Net.
- Данный алгоритм может быть использован в задачах, связанных с оценкой стоимости аренды порубочных площадей.

Спасибо за внимание!

Исследование выполнено за счет гранта Российского научного фонда № 22-17-20012, <https://rscf.ru/project/22-17-20012/> при паритетной финансовой поддержке Правительства Республики Хакасия