



ТОМСКИЙ  
ПОЛИТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ



## Модели полносверточных нейронных сетей для семантической сегментации поврежденных союзным короедом деревьев кедра на многоспектральных снимках с БПЛА

Н. Г. Марков: профессор ОИТ ТПУ, д.т.н.

И.А. Керчев: с.н.с ИМКЭС СО РАН, к.б.н.

К. А. Маслов: аспирант ОИТ ТПУ

К.Р. Мачука: аспирант ОИТ ТПУ

О.С. Токарева: доцент ОИТ ТПУ, к.т.н.

В.В. Осипова: доцент ОИТ ТПУ, к.т.н.

16 ноября  
2022

# Введение

В последние годы в России активно проявился агрессивный лесной вредитель - союзный короед *Ips amitinus* (Eichhoff), заселяющий деревья сосны сибирской кедровой *P. Sibirica* (далее для краткости - кедр). Обширные очаги усыхания насаждений, вызванные вредителями, влекут за собой значимые экономические и экологические последствия.

Наличие таких негативных факторов приводит к необходимости оперативного мониторинга кедровников и выявления очагов короеда путем анализа полученных с БПЛА снимков. Для проведения анализа снимков сегодня актуальна разработка новых моделей полносверточных нейронных сетей, позволяющих решать с высокой точностью задачи сегментической сегментации (по пиксельной классификации) деревьев кедр на многоспектральных снимках.

# Цель работы

Разработка, программная реализация и исследование полносверточных нейронных сетей для семантической сегментации (классификации по пикселям) деревьев кедра, поврежденных союзным короедом, на многоспектральных снимках с БПЛА

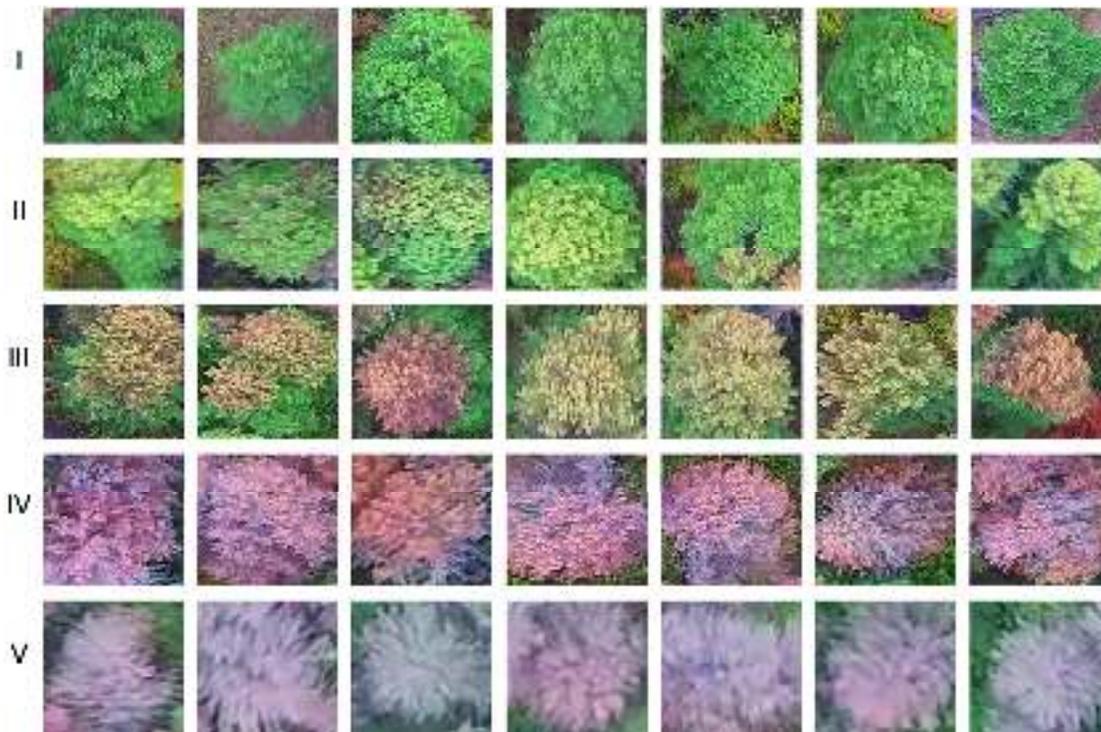
## Задачи

- Подготовить по исходным данным (многоспектральные и RGB снимки) выборки для обучения и исследования нейронных сетей
- Разработать модели глубокого обучения\*:
  - полносверточная сеть U-Net (модификация классической U-Net)
  - полносверточная сеть MN-U-Net
- Провести исследование разработанных моделей

*\*Для разработки моделей глубокого обучения используются библиотеки tensorflow и keras*

# Классы деревьев кедра

RGB-композиция снимков с БПЛА DJI Phantom 3 Standard



Классы деревьев *P. sibirica*, повреждённых союзным короедом:

**I** – Условно здоровое

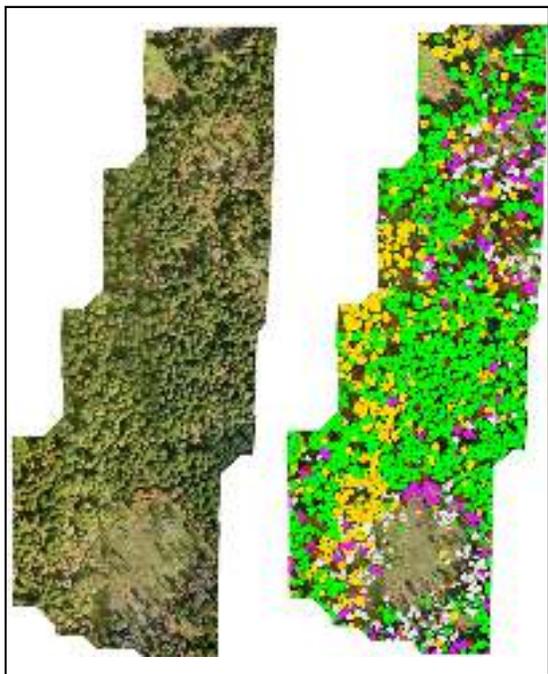
**II** – Свежезаселенное

**III** – С усохшей вершиной

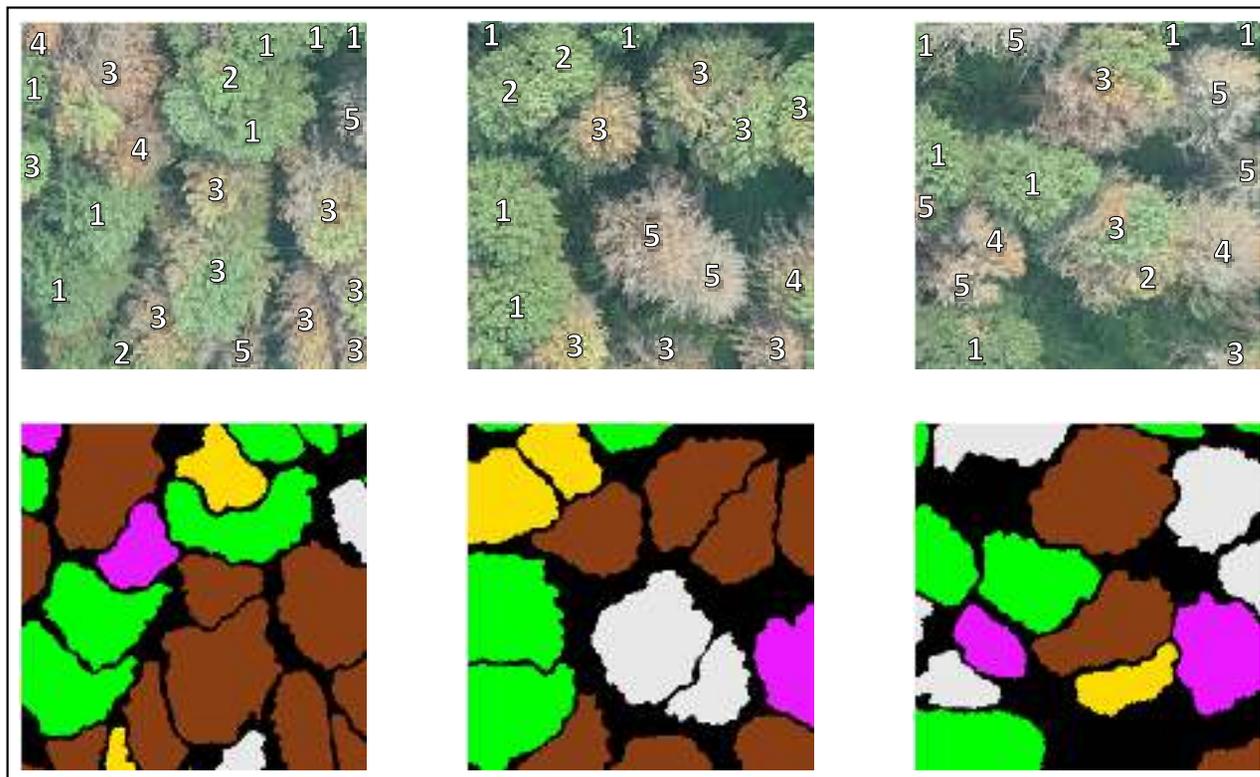
**IV** – Свежий сухостой

**V** – Старый сухостой

# Разметка исходных снимков экспертами



Классификация проведена экспертами Института мониторинга климатических и экологических систем СО РАН



- |  |  |
|--|--|
|  0 Фон              |  3 С усохшей вершиной |
|  1 Условно здоровое |  4 Свежий сухостой    |
|  2 Свежезаселенное  |  5 Старый сухостой    |

# Обучающие и валидационные выборки

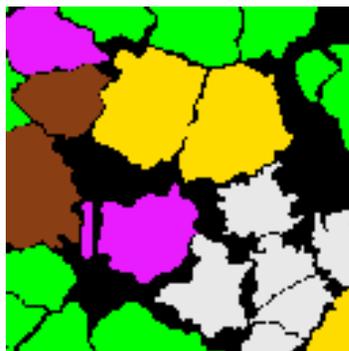
Путем нарезки исходных снимков на фрагменты 256x256 пикселей было получено:

## Выборка по RGB снимкам

- 4101 обучающий фрагмент (256x256x3)
- 4101 обучающая маска (256x256x3)
- 1025 валидационных фрагментов (256x256x3)
- 1025 валидационных масок (256x256x3)



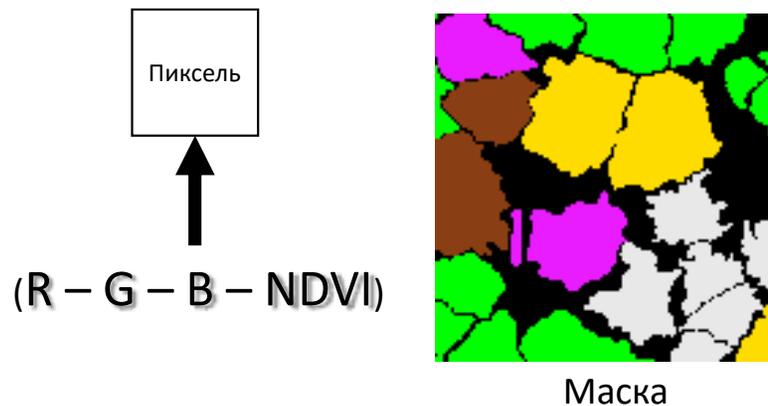
RGB Фрагмент



Маска

## Выборка по многоспектральным снимкам

- 4101 обучающий фрагмент (256x256x4)
- 4101 обучающая маска (256x256x3)
- 1025 валидационных фрагментов (256x256x4)
- 1025 валидационных масок (256x256x3)



# NDVI

**NDVI** (*Normalized Difference Vegetation Index*) - нормализованный относительный индекс растительности - простой количественный показатель количества фотосинтетически активной биомассы (обычно называемый вегетационным индексом). Один из самых распространенных и применяемых индексов для решения задач, использующих количественные оценки растительного покрова.

Вычисляется по следующей формуле:

$$NDVI = \frac{NIR - RED}{NIR + RED},$$

где,

NIR - интенсивность отраженного света в ближнем инфракрасном диапазоне

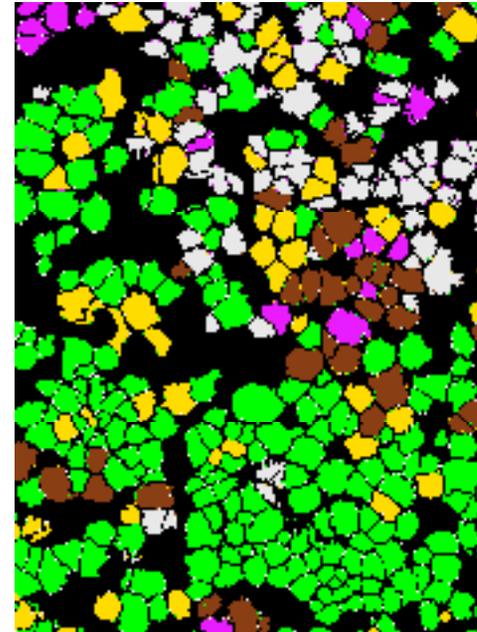
RED - интенсивность отраженного света в красном диапазоне

# Тестовый участок

В качестве тестового участка использовалось изображение 2048x1536 пикселей

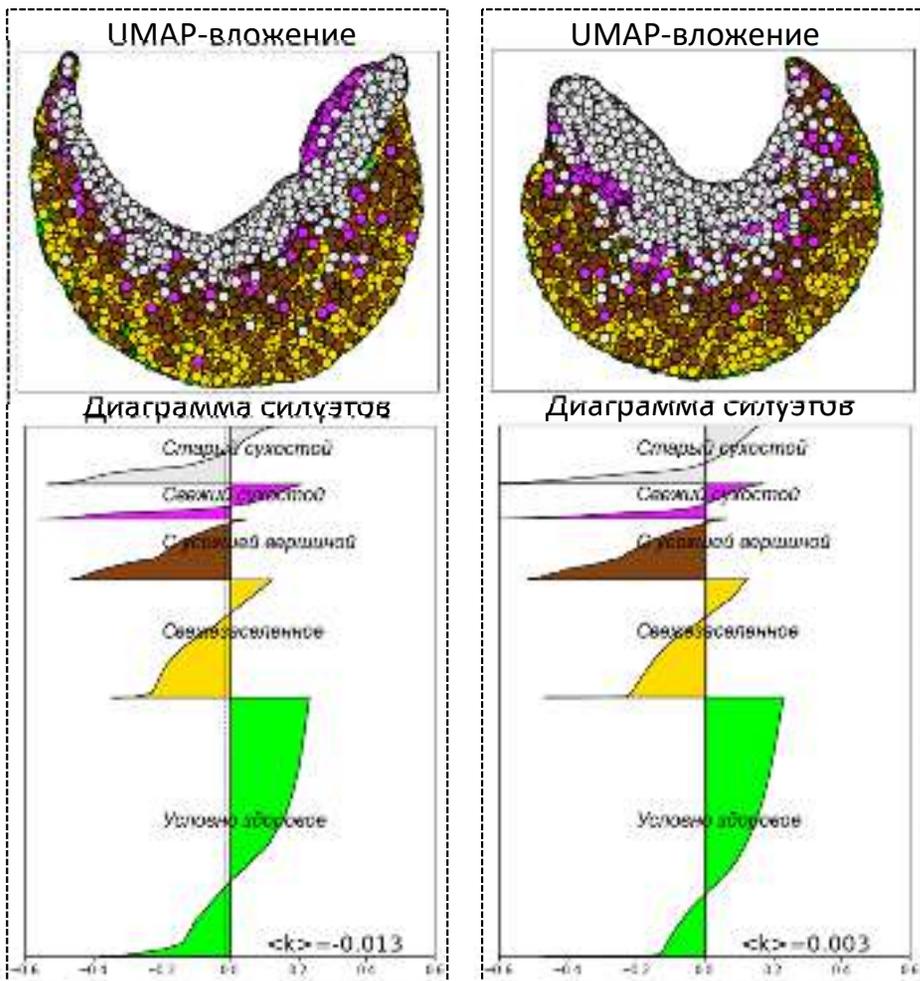


Тестовый участок (RGB-композиция)



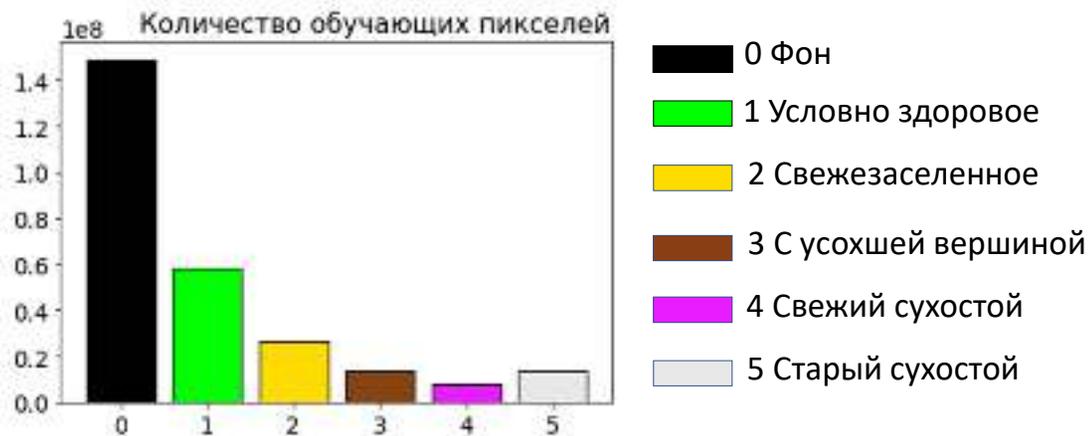
Результат разметки тестового участка экспертами

# Разведочный анализ обучающих выборок



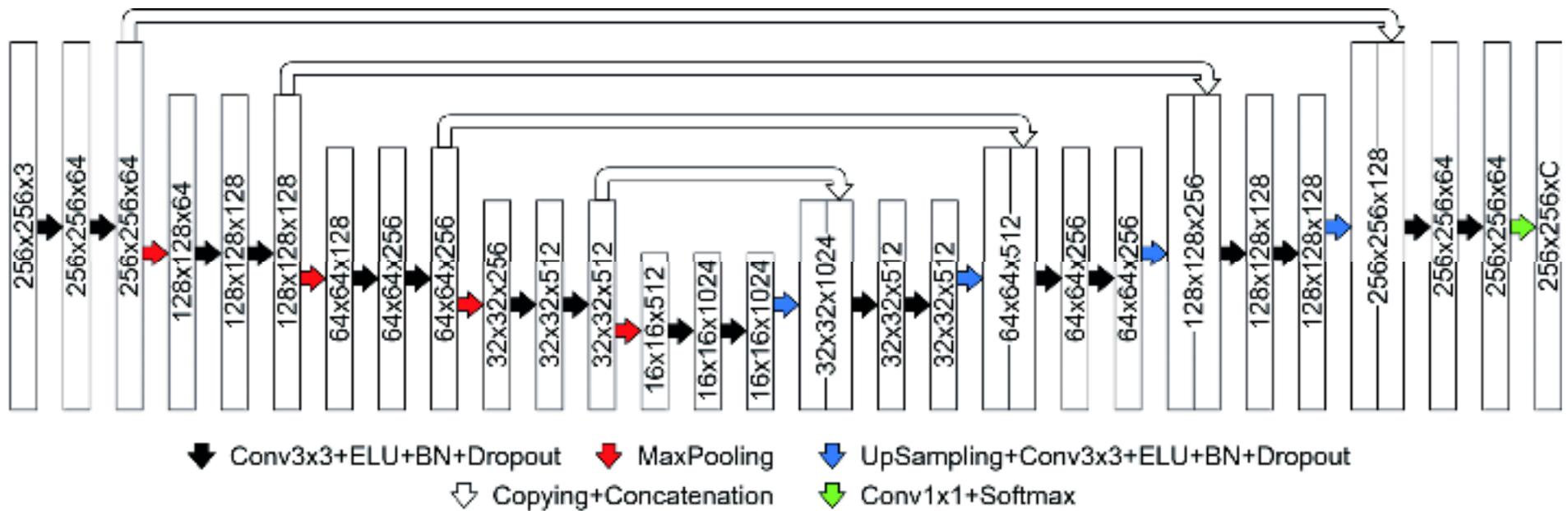
RGB снимки

Многоспектральные снимки



- Значительное преобладание пикселей фона
- Наименее представлены деревья кедра класса «Свежий сухостой»
- Большой количество пикселей соответствуют отрицательные значения коэффициента силуэтов
- Необходимо использовать модели глубокого обучения, способные учитывать текстуры и формы объектов, для решения поставленной задачи классификации деревьев кедра

# Архитектура полносверточной сети U-Net





# Особенности подготовки обучающих выборок

- По горизонтали и вертикали снимки перекрывают 128 пикселей предыдущего снимка, это позволяет увеличить обучающую выборку и смягчить непоследовательную сегментацию на границах снимков.
- Была применена аугментация <<на лету>> для улучшения обобщающей способности полносверточных сетей, в нее включались: повороты по горизонтали, извлечение случайных участков, изменения масштаба, яркости и контраста и эластические преобразования фрагментов.
- К снимкам была применена нормализация для повышения эффективности и стабильности моделей нейросетей.

## Функция потерь

В качестве функции потерь использовалась модифицированная усредненная метрика mIoU (модифицированный усредненный коэффициент Жаккара):

$$\text{IoU}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c + \text{FP}_c}, \quad \text{mIoU} = \frac{1}{C} \sum_{c=1}^C \text{IoU}_c \quad (1)$$

$$\text{mIoU}(\mathbf{T}, \mathbf{P}) \approx \frac{1}{C} \sum_{c=1}^C \left( \frac{\sum_{i=1}^H \sum_{j=1}^W \text{LS}(\mathbf{T})_{ic} \cdot \mathbf{P}_{ic} + \theta_s}{\sum_{i=1}^H \sum_{j=1}^W [\text{LS}(\mathbf{T})_{ic} + (1 - \text{LS}(\mathbf{T})_{ic}) \cdot \mathbf{P}_{ic}] + \theta_s} \right) \rightarrow \max, \quad \text{LS}(\mathbf{T}) = (1 - \theta_{ls}) \cdot \mathbf{T} + \frac{\theta_{ls}}{C} \quad (2)$$

$$\text{Loss} = -\text{mIoU}(\mathbf{T}, \mathbf{P}) \rightarrow \min \quad (3)$$

где  $\mathbf{T}$  – тензор с эталонным выходом,  $\mathbf{P}$  – действительный выход нейронной сети,  $H$  – высота снимка,  $W$  – ширина снимка,  $C$  – количество классов,  $\theta_s > 0$  – сглаживающий параметр, устраняющий большие по модулю значения градиента функции потерь и точку сингулярности,  $0 \leq \theta_{ls} < 1$  – параметр сглаживания меток

# Гиперпараметры моделей сетей

$\theta_{lr}$  – скорость обучения,

$\theta_d$  – коэффициент дропаута,

$\theta_s$  – сглаживание целевой функции,

$\theta_{ls}$  – сглаживание меток,

$\theta_z$  – изменение масштаба фрагментов,

$\theta_b$  – изменение яркости фрагментов,

$\theta_c$  – изменение контраста фрагментов,

$\theta_\alpha$  и  $\theta_\sigma$  – параметры эластических деформаций,

$\theta_{rf}$  – уменьшение скорости обучения при выходе «на плато»,

$\theta_{rp}$  – количество эпох для подтверждения выхода «на плато»

# Качественные результаты сегментации

Тестовый Участок

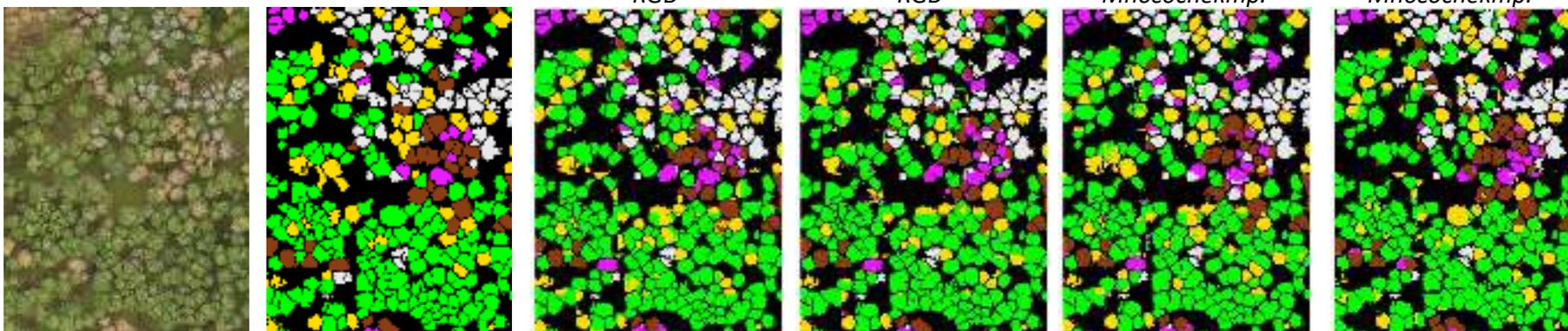
Эталон

U-Net  
*RGB*

MH-U-Net  
*RGB*

U-Net  
*Многоспектр.*

MH-U-Net  
*Многоспектр.*



0 Фон

1 Условно здоровое

2 Свежезаселенное

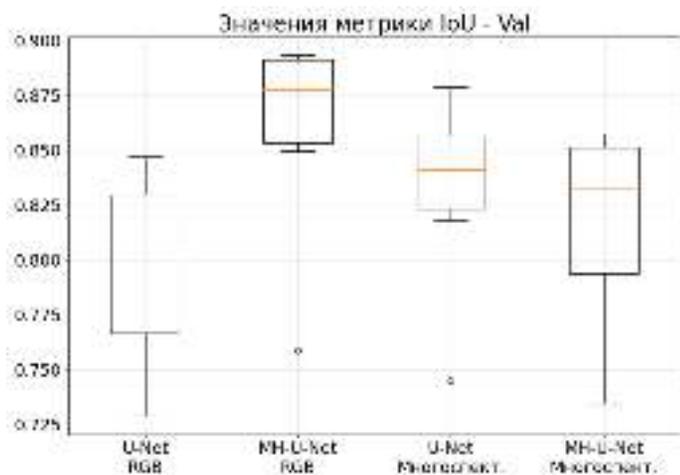
3 С усохшей вершиной

4 Свежий сухостой

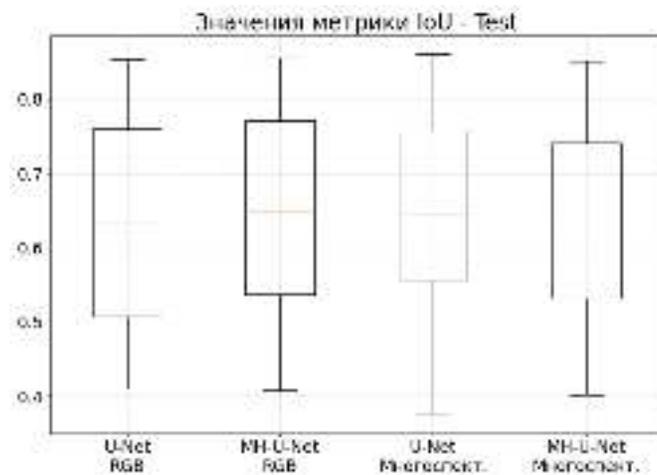
5 Старый сухостой

# Количественные результаты сегментации

Модель СНС		IoU					mIoU		
		Фон	Условно здоровое	Свежеза- селённое	С усохшей вершиной	Свежий сухостой	Старый сухостой	Val	Test
U-Net	RGB	0.8528	0.7573	0.4279	0.5869	0.4101	0.7628	0.7995	0.6330
MH-U-Net		0.8557	<b>0.7593</b>	0.4283	0.6480	0.4073	<b>0.7827</b>	<b>0.8579</b>	<b>0.6469</b>
U-Net	Многогл.	<b>0.8607</b>	0.7395	<b>0.4737</b>	0.6390	0.3744	0.7783	0.8306	0.6443
MH-U-Net		0.8510	0.7028	0.4016	<b>0.6619</b>	<b>0.4276</b>	0.7790	0.8163	0.6373



KruskalResult(statistic=7.146, pvalue=0.0673)



KruskalResult(statistic=0.255, pvalue=0.968)

# Заключение

- Предложены и исследованы две модели полносверточных нейронных сетей для классификации изображений деревьев кедра, поврежденных союзным короедом, на снимках с БПЛА. Показано, что обе модели сетей успешно распознают на снимках с БПЛА деревья классов «Условно здоровое», «С усохшей вершиной», «Старый сухостой» и «Фон» ( $IoU > 0.5$ ). Хуже распознают деревья классов «Свежезаселённое» и «Свежий сухостой» ( $IoU < 0.5$ ).
- Сети, обученные на многоспектральных снимках, показывают более точную классификацию четырех из шести классов деревьев, чем обученные на RGB снимках. Однако значения метрики mIoU и теста Краскела-Уоллиса не показывают преимущества многоспектральных снимков.
- Значение mIoU для тестового участка ниже, чем для валидационной выборки, в основном из-за относительно небольшой площади тестового снимка и несбалансированности классов.
- Эффективность предложенных моделей может быть улучшена путем увеличения объема выборки для обучения или при использовании дополнительных спектральных каналов.

Спасибо за ваше внимание