

## Модели U-Net для семантической сегментации повреждённых деревьев сосны сибирской кедровой на снимках с БПЛА

Н. Г. Марков<sup>1</sup>, К. А. Маслов<sup>1</sup>, И. А. Керчев<sup>2</sup>, О. С. Токарева<sup>1</sup>

<sup>1</sup> *Национальный исследовательский Томский политехнический университет  
Томск, 634050, Россия*

*E-mails: markovng@tpu.ru, kam20@tpu.ru, ost@tpu.ru*

<sup>2</sup> *Институт мониторинга климатических и экологических систем СО РАН  
Томск, 634055, Россия*

*E-mail: ivankerchev@gmail.com*

В трёх регионах Западной Сибири: Томской, Кемеровской и Новосибирской областях — с 2019 г. наблюдается стремительное усыхание древостоев сосны сибирской кедровой *Pinus sibirica* Du Tour, вызванное новым чужеродным вредителем — союзным короедом *Ips amitinus* Eichh. Успешность мероприятий по борьбе с короедом напрямую зависит от своевременности выявления заселённых им деревьев. Выявить их можно по характерным признакам вершинного усыхания, которое малозаметно при наземном обследовании. Использование беспилотных летательных аппаратов (БПЛА) обеспечивает высокую оперативность наблюдений и получение снимков крон деревьев со сверхвысоким пространственным разрешением. Статья посвящена разработке моделей глубокого обучения U-Net и их исследованию при решении задачи семантической сегментации деревьев *P. sibirica*, повреждённых союзным короедом, на снимках с БПЛА. Для анализа состояния деревьев экспертами было выделено пять классов жизненного состояния *P. sibirica*: «условно здоровое», «свежезаселённое», «с усохшей вершиной», «свежий сухостой» и «старый сухостой». Деревья других видов и оставшиеся объекты на изображениях относились к классу «фон». В качестве исходных данных использовались снимки с квадрокоптера DJI Phantom 3 Standard, полученные в июле 2019 г. Для решения задачи сегментации изображений с БПЛА был предложен ряд моделей глубокого обучения: полносвёрточная сеть U-Net на основе известной модели U-Net и две её модификации — multihead-U-Net (МН-U-Net) и multihead-residual-U-Net (МН-Res-U-Net). МН-U-Net — ансамбль из трёх моделей U-Net различной глубины, причём модели делят между собой часть весов и пытаются одновременно анализировать изображение в различных масштабах. Модель МН-Res-U-Net обладает всеми свойствами МН-U-Net и дополнительно использует остаточные блоки. Результаты исследований показали, что все модели успешно классифицируют пиксели пяти классов из шести: модели U-Net и МН-Res-U-Net хорошо сегментируют все классы, кроме класса «свежезаселённое», а модель МН-U-Net — все классы, кроме класса «свежий сухостой». Основную сложность для сегментации представили классы промежуточного состояния деревьев. Однако МН-U-Net справляется с сегментацией промежуточного класса «свежезаселённое», U-Net — класса «с усохшей вершиной», а МН-Res-U-Net — класса «свежий сухостой».

**Ключевые слова:** беспилотный летательный аппарат, глубокое обучение, полносвёрточная сеть, U-Net, семантическая сегментация, сосна сибирская кедровая, *Pinus sibirica* Du Tour, союзный короед, *Ips amitinus* Eichh.

Одобрена к печати: 13.01.2022

DOI: 10.21046/2070-7401-2022-19-1-65-77

### Введение

Во многих странах чужеродные или инвазионные насекомые-дендрофаги становятся серьёзной угрозой естественным лесам, лесным плантациям и лесной промышленности (Bradshaw et al., 2016; Liu et al., 2021). Союзный короед *Ips amitinus* Eichh. — представитель фауны центрально-европейских горных хвойных лесов, к настоящему времени широко распространился во многих странах Европы (Økland et al., 2019). В 2019 г. этот стволовой вредитель был впервые обнаружен на юго-востоке Западной Сибири. Здесь он начал повреждать деревья

новой для него кормовой породы — сосны сибирской кедровой *Pinus sibirica* Du Tour (Керчев и др., 2021a; Kerchev et al., 2019). Стремительное расселение чужеродного вредителя требует использования эффективных методов выявления и учёта повреждённых деревьев в насаждениях. Раннее выявление деревьев на начальных стадиях заселения вредителем выступает залогом успешности фитосанитарных мероприятий. Характерная для союзного короеда поэтапность заселения веток на дереве, начиная с самой вершины, приводит к тому, что сначала дехромация хвои наблюдается в верхней части кроны, в то время как на нижних незаселённых ветках хвоя ещё 1–2 года может оставаться зелёной. При наземных лесопатологических обследованиях сомкнутость крон значительно затрудняет оценку состояния как отдельных деревьев, так и насаждения в целом. Использование беспилотных летательных аппаратов (БПЛА) может обеспечить большую оперативность наблюдений, с высоким пространственным разрешением, обеспечивающим достаточную детализацию повреждений отдельных деревьев (Керчев и др., 2021a).

Для анализа изображений, полученных в результате съёмки лесных массивов с БПЛА, обычно используют различные традиционные методы классификации (Chenari et al., 2017; Gini et al., 2018; Lee et al., 2019). В работе (Onishi, Ise, 2018) предлагается классифицировать деревья различных пород с использованием методов глубокого обучения. В статье (Safonova et al., 2019) для классификации поражённых вредителями деревьев пихты применены различные модели свёрточных нейронных сетей (СНС). На наш взгляд, для анализа состояния поражённых деревьев *P. sibirica* перспективными среди моделей СНС могут оказаться полносвёрточные сети U-Net, позволяющие решать задачи семантической сегментации изображений (Ronneberger et al., 2015). Именно с помощью таких сетей получены обнадёживающие результаты классификации поражённых уссурийским полиграфом деревьев пихты сибирской *Abies sibirica* Ledeb. (Керчев и др., 2021б).

Цель данной работы состоит в создании и исследовании новых моделей глубокого обучения U-Net для семантической сегментации деревьев *P. sibirica*, повреждённых союзным короедом, на снимках с БПЛА.

### Задача анализа состояния деревьев *P. sibirica* на снимках с БПЛА

При анализе на снимках деревьев *P. sibirica*, повреждённых союзным короедом, экспертами выделяется пять классов их жизненного состояния (Керчев и др., 2021a):

- класс I — «условно здоровое»: дерево имеет ярко-зелёную крону без признаков дехромации или пожелтения, дерево условно здоровое;
- класс II — «свежезаселённое»: крона зелёная, но верхушечные ветви начинают приобретать бледный зелёно-жёлтый оттенок. На этом этапе жуки под корой завершают своё развитие или (в весенний период) находятся в подстилке, в летний период они могут спуститься в нижние ярусы кроны;
- класс III — «с усохшей вершиной»: верхушка имеет выраженный жёлто-бурый оттенок, нижняя часть кроны ярко- или бледно-зелёная. В весенний и осенний периоды жуки находятся на стадии зимовки в подстилке либо весной после массового лёта и летом осваивают нижний ярус ветвей или осваивают вершины соседних деревьев; ствол этих деревьев может осваиваться сопутствующими видами вредителей;
- класс IV — «свежий сухостой»: вся крона буро-рыжего цвета, а макушка дерева может нести ветви без хвои; дерево покинуто союзным короедом, нижняя часть ствола заселена шестизубчатым короедом или усачами;
- класс V — «старый сухостой»: крона не несёт хвои или имеются остатки неосыпавшейся рыже-бурой хвои на самых нижних ветках; союзный короед покинул дерево в предыдущем вегетационном сезоне, в древесине могут сохраняться личинки усачей.

На *рис. 1* (см. с. 67) представлены фотографии крон деревьев *P. sibirica* перечисленных классов, полученные с помощью фотокамеры, установленной на БПЛА.



Рис. 1. Классы деревьев *P. sibirica*, повреждённых союзным короедом

В качестве исходных данных использовались фотоснимки, сделанные на штатную камеру БПЛА (квадрокоптер DJI Phantom 3 Standard). Снимки участков насаждений были собраны в панорамы, всего их было получено четыре. Снимки для первой из них были сделаны 11 июля 2019 г. с высоты 350 м, что обеспечило пространственное разрешение около 10 см. Снимки для второй, третьей и четвёртой панорам были сделаны 23 июля 2019 г. с высоты 150 м, что соответствует разрешению около 3 см. Для нивелирования разницы в размерах деревьев в кадре из-за разной высоты съёмки панорамы были отмасштабированы. Ставилась задача оценки по этим снимкам состояний деревьев *P. sibirica*, имеющих разную степень повреждения союзным короедом, соответствующую одному из пяти вышеперечисленных классов. Шестым классом был «фон» — деревья других пород и иные объекты земной поверхности. Каждое из деревьев на снимках относилось визуальным дешифрированием к одному из классов экспертом-энтомологом, участвовавшим в наземном обследовании участков. Второй эксперт проводил проверку результатов разметки снимков первым экспертом и осуществлял оконтуривание крон деревьев *P. sibirica*. Итоговым результатом оконтуривания стали эталонные карты сегментации.

На первом этапе проводился разведочный анализ исходных данных, сформированных в виде трёх панорам снимков и используемых в последующем для обучения и валидации предлагаемых моделей глубокого обучения.

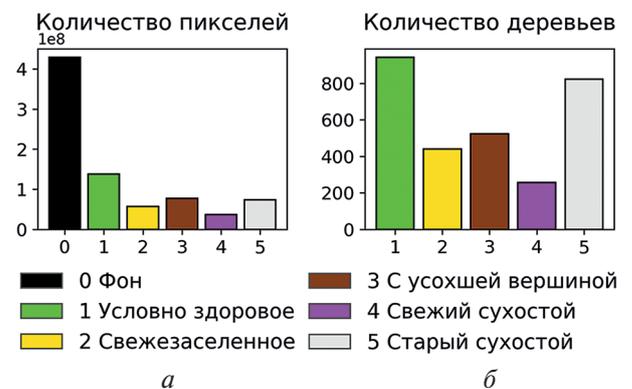
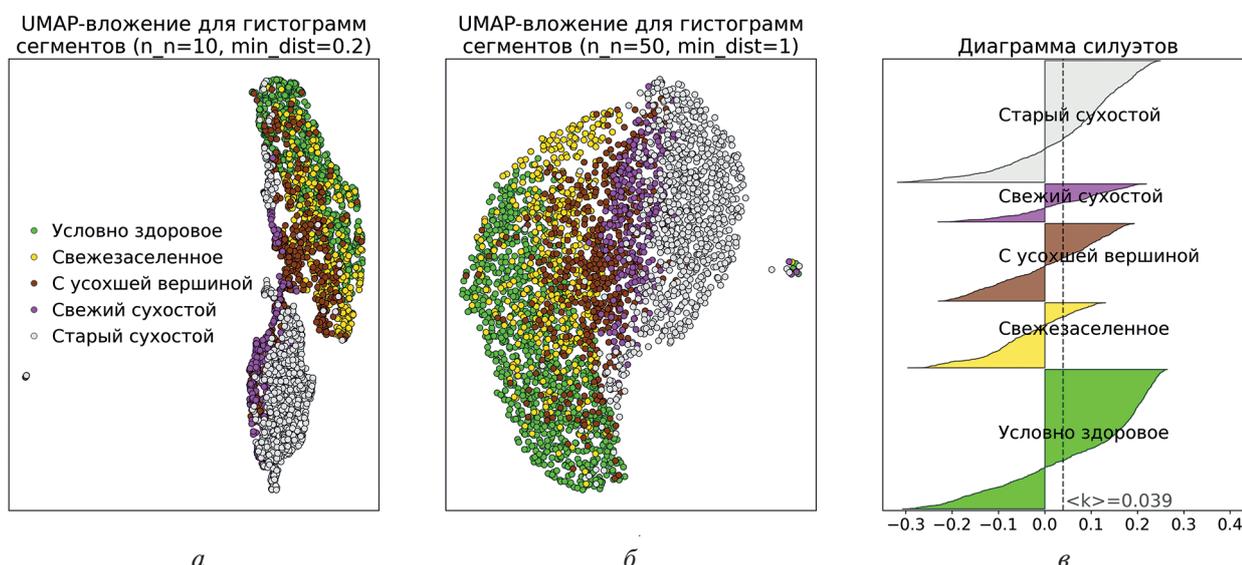


Рис. 2. Количество по классам в обучающей выборке: а — пикселей; б — деревьев *P. sibirica*

Четвёртая панорама была использована в качестве тестовой при проведении исследований этих моделей. На *рис. 2* (см. с. 67) представлены результаты разведочного анализа в виде гистограмм, позволяющих оценить сбалансированность наборов данных в обучающей выборке. Видно, что имеется проблема: значительное преобладание пикселей класса «фон» и слабая представленность таких классов деревьев *P. sibirica* в обучающей выборке, как «свежезаселённое» и «свежий сухостой».

Анализ локальных особенностей распределения гистограмм яркости в трёх каналах на UMAP-вложении (*англ.* Uniform Manifold Approximation and Projection) (*рис. 3а*) (McInnes et al., 2018) показывает наличие двух различных кластеров: один преимущественно содержит сегменты крон деревьев классов «условно здоровое» и «свежезаселённое», другой — классов «свежий сухостой» и «старый сухостой». При этом сегменты класса «с усохшей вершиной» расположились в пространстве между выделенными кластерами. На UMAP-вложении, показывающем глобальные особенности (*рис. 3б*), можно видеть непрерывный переход между классами деревьев в порядке ухудшения их жизненного состояния. Сегментам соответствует низкое среднее значение коэффициентов силуэта (0,039) (Rousseeuw, 1987), при этом существенные доли отрицательных значений коэффициента наблюдаются для классов «свежезаселённое» и «с усохшей вершиной» (*рис. 3в*).



*Рис. 3.* UMAP-вложения для гистограмм (*а* и *б*) и диаграмма силуэтов для гистограмм (*в*) сегментов деревьев кедра; пунктирной линией показано среднее значение коэффициента силуэта

Таким образом, разведочный анализ исходных данных показал, что поставленная задача сегментации снимков поражённых деревьев *P. sibirica* с БПЛА не тривиальна, поскольку учёт только цветовых характеристик пикселей, составляющих сегменты, не позволяет достичь высокой разделимости классов. Более того, пиксели сегментов не сравнивались с фоновыми пикселями, чьи спектральные характеристики часто могут быть похожими, тем самым задача классификации пикселей усложнялась. Этим обосновывается необходимость использования моделей СНС, так как они способны учитывать текстуру объектов и их форму и автоматически извлекать соответствующие признаки. Например, можно применять модели U-Net, позволяющие решать задачи попиксельной классификации изображений.

### Предлагаемые модели свёрточных нейронных сетей

Для решения задачи сегментации изображений поражённых деревьев *P. sibirica* был разработан ряд моделей глубокого обучения: полносвёрточная сеть U-Net, а также две её модификации — МН-U-Net и МН-Res-U-Net.

**Модель U-Net.** Основывается на одной из наиболее известных моделей U-Net для решения задачи семантической сегментации. Эта модель U-Net была изначально разработана для решения задачи сегментации биомедицинских изображений, где продемонстрировала высокое качество результатов. Отличительная черта модели U-Net — наличие соединений пропуска (*англ.* skip-connections), соединяющих наборы карт признаков кодера с наборами карт признаков декодера с целью повышения детальности результирующей карты сегментации (Ronneberger et al., 2015). На *рис. 4* приведена архитектура нашей модели U-Net, разработанной на основе известной СНС U-Net. Прямоугольники соответствуют тензорам, или многомерным массивам, представляющим наборы карт признаков, числами указаны размеры соответствующих тензоров, стрелками обозначены следующие операции:

- свёртки ( $\text{Conv}3\times3$ ,  $\text{Conv}1\times1$ );
- вычисление функции активации ELU (*англ.* exponential linear unit — экспоненциальный выпрямитель);
- пакетная нормализация (*англ.* batch normalization — BN) (Ioffe, Szegedy, 2015);
- субдискретизация выбором максимального значения в окрестности  $2\times2$  пикселя (MaxPooling);
- увеличение масштаба методом ближайшего соседа (UpSampling);
- копирование тензора и его конкатенация с другим (Copying+Concatenation);
- исключение случайных сигналов слоя путём их приравнивания к нулю (Dropout) (Srivatsava et al., 2014).

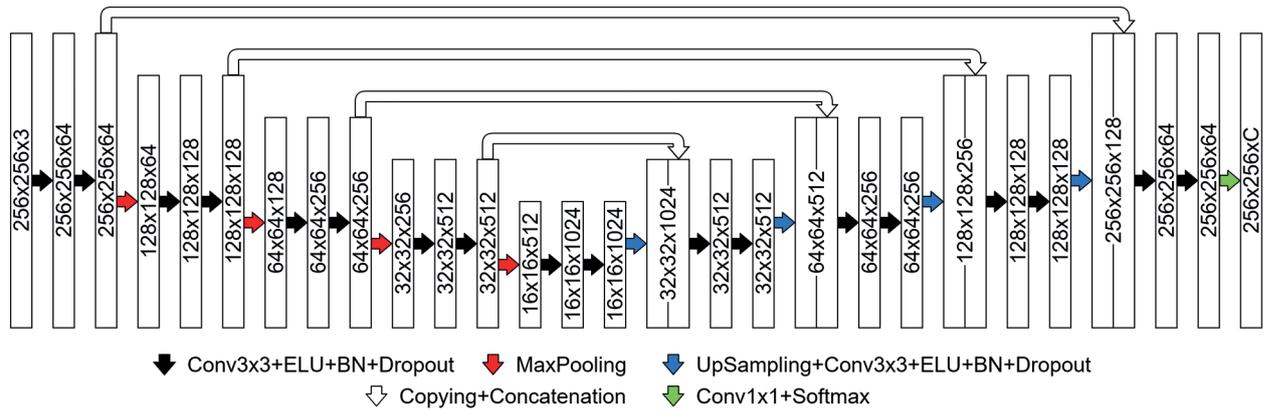


Рис. 4. Архитектура модели U-Net

Категориальное распределение на выходе декодера моделируется для каждого пикселя применением обобщения логистической функции для многомерного случая — Softmax.

По сравнению с исходной моделью U-Net в архитектуру предлагаемой модели были добавлены следующие изменения:

- входное изображение сети представлено тензором  $256\times256\times3$ , соответствующим фрагменту RGB-изображения (R — *англ.* red, красный; G — *англ.* green, зелёный; B — *англ.* blue, синий);
- свёртки не уменьшают размер карт признаков;
- обрезка карт признаков не используется для соединений пропуска;
- пакетная нормализация применяется после каждой нелинейности;
- выходной тензор вычисляется  $C$  свёртками с фильтрами размером  $1\times1$ , позволяя таким образом классифицировать пиксели сразу  $C$  классов (пять классов деревьев кедра и «фон»).

**Модель МН-U-Net.** Вторая из предлагаемых моделей полносвёрточных сетей — multi-head-U-Net, сокращённо МН-U-Net. Она основана на первой предложенной модели U-Net, но имеет следующие существенные изменения по сравнению с её архитектурой (*рис. 5*, см. с. 70):

- вместо одного входа разных масштабов используются три ( $256 \times 256 \times 3$ ,  $128 \times 128 \times 3$  и  $64 \times 64 \times 3$ );
- также имеется три выхода различных масштабов;
- операция Dropout заменена на операцию пространственного Dropout, как в СНС ENet (Paszke et al., 2016).

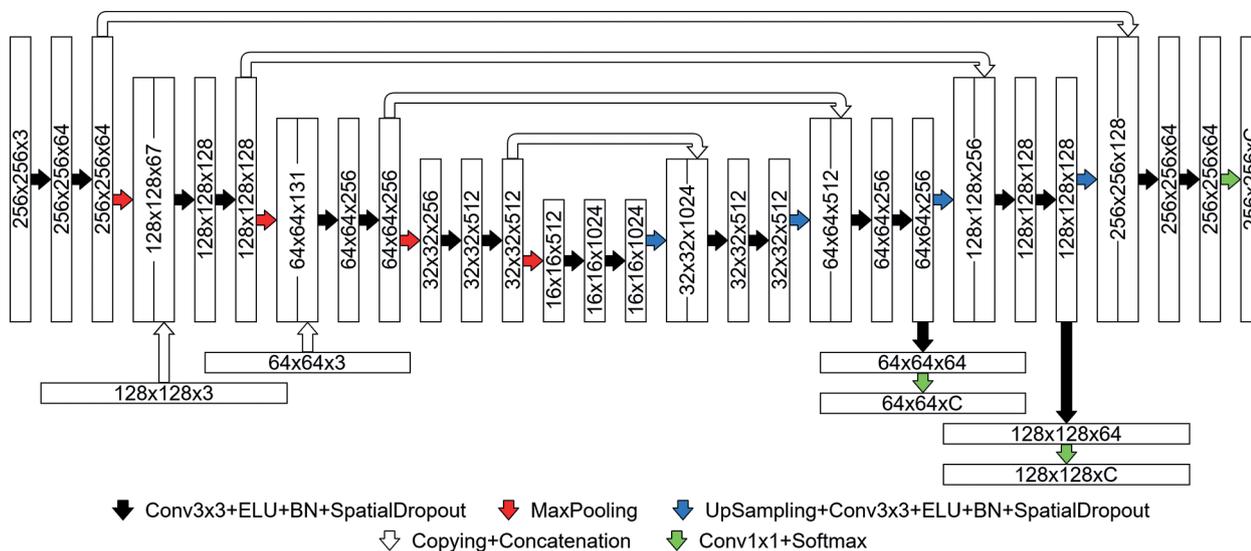


Рис. 5. Архитектура модели MH-U-Net

Во время обучения функция потерь рассчитывается на всех трёх выходах модели MH-U-Net, и алгоритм обучения минимизирует сумму всех трёх значений этой функции. При использовании обученной модели выходы меньшего масштаба и соответствующие ветви вычислительного графа могут быть проигнорированы. Такая модель может рассматриваться как ансамбль из трёх моделей U-Net различной глубины, где модели ансамбля делят между собой часть весов и пытаются одновременно анализировать изображение в различных масштабах. Это позволяет ожидать, что MH-U-Net способна обучаться более качественному скрытому представлению данных и может привести к стабильно более качественным результатам сегментации.

**Модель MH-Res-U-Net.** Третья предложенная модель — multihead-residual-U-Net (MH-Res-U-Net), обладает всеми особенностями модели MH-U-Net и дополнительно включает остаточные блоки (He et al., 2015). Архитектура модели MH-Res-U-Net приведена на рис. 6.

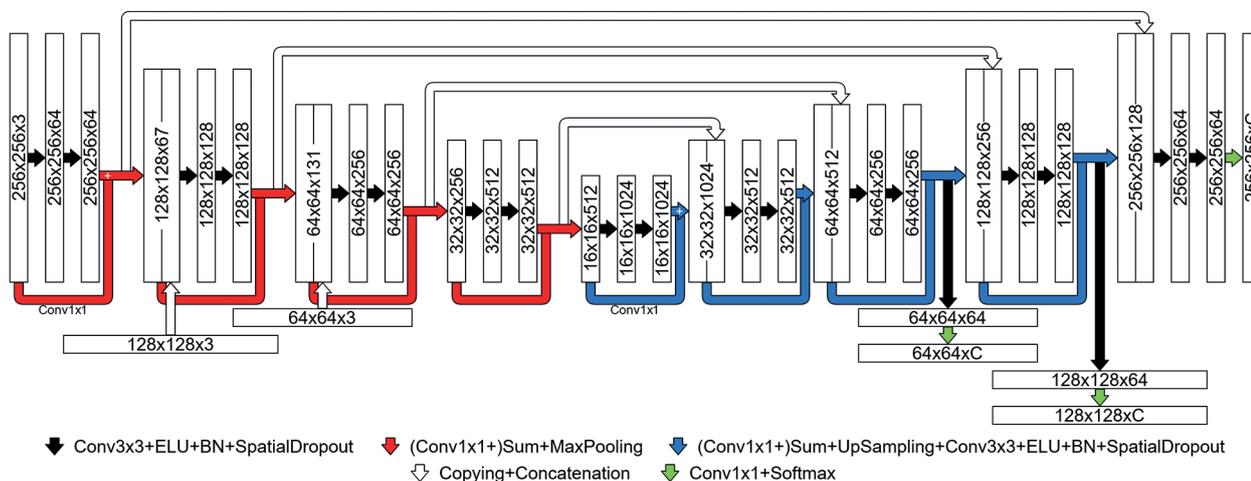


Рис. 6. Архитектура модели MH-Res-U-Net

Важно отметить, что операции проецирования ( $\text{Conv}1 \times 1$ ) в главных ветвях остаточных блоков модели МН-Res-U-Net не включают нелинейностей и смещений. Ожидается, что такие остаточные блоки могут привести к более стабильному процессу обучения и упростить обратное распространение градиента.

Для оценки эффективности предложенных моделей СНС использована метрика *intersection over union* (IoU), считающаяся общепринятой метрикой эффективности при решении задачи семантической сегментации цифровых изображений (Rahman, Wang, 2016). Значения IoU, превышающие 0,5, соответствуют приемлемому качеству сегментации. Использовалась также метрика *mean intersection over union* (mIoU), рассчитываемая как среднее значение IoU по всем классам.

**Процесс обучения моделей.** Для формирования обучающей и валидационной выборок первые три панорамы и соответствующие им эталонные карты сегментации в различных углах поворота были нарезаны на фрагменты. Утечка данных из валидационной выборки устранялась наложением на исходные изображения булевых масок — двумерных массивов, элементы которых принимают значения «ложь» или «истина». Сначала из панорам выбирались непересекающиеся фрагменты размером  $256 \times 256$  пикселей для валидационной выборки, а соответствующие по координатам значения булевой маски устанавливались как «истина». Затем все оставшиеся области панорамы, для которых значения булевой маски принимают только значения «ложь», нарезались на фрагменты размером  $256 \times 256$  пикселей со сдвигом в 128 пикселей. Это гарантировало, что признаки валидационных фрагментов не дублировались среди обучающих. Всего было получено 12 384 обучающих и 2400 валидационных фрагментов для решения задачи сегментации деревьев *P. sibirica*.

При обучении моделей для достижения их лучшей обобщающей способности обучающая выборка замещалась синтезированной, полученной случайными преобразованиями исходных примеров. Случайные преобразования включали в себя изменения масштаба и кадрирования, изменения яркости и контраста, отражения по вертикальной оси, а также эластические преобразования (Simard et al., 2003).

Для поиска оптимальных весовых коэффициентов моделей максимизировалась метрика mIoU в мягкой формулировке (Bertels et al., 2019) с использованием алгоритма оптимизации Adam, основанного на вычислении градиента целевой функции (Kingma, Ba, 2014). Оптимизация mIoU позволяет частично решить выявленную проблему несбалансированности выборки по классам деревьев.

Для определения оптимальных гиперпараметров предлагаемых моделей и алгоритма их обучения использовался метод случайного поиска. К гиперпараметрам были отнесены параметры алгоритма оптимизации весовых коэффициентов Adam и параметры степени случайных преобразований для формирования синтезированной обучающей выборки. Для итоговых моделей каждой из СНС были определены такие значения параметров, при которых целевая функция показала наибольшее значение, рассчитанное по валидационной выборке.

## Результаты исследования предлагаемых моделей свёрточных нейронных сетей

Для исследования были выбраны модели U-Net, МН-U-Net и МН-Res-U-Net с гиперпараметрами, позволившими получить лучшие результаты по качеству классификации на валидационной выборке. На *рис. 7* (см. с. 72) представлены результаты сегментации изображения тестового участка с деревьями *P. sibirica*. Можно видеть, что все исследуемые модели СНС успешно воспроизводят границы между деревьями и сегментируют большую долю крон деревьев классов «условно здоровое» и «старый сухостой», однако имеются ошибки сегментации для классов деревьев с промежуточными состояниями.

В *табл. 1* приведены значения метрики IoU для класса «фон» и пяти классов деревьев *P. sibirica*, а также значения mIoU для валидационной выборки (столбец val) и для изображения тестового участка (столбец test). Полужирным начертанием выделены лучшие значения метрик в столбце.

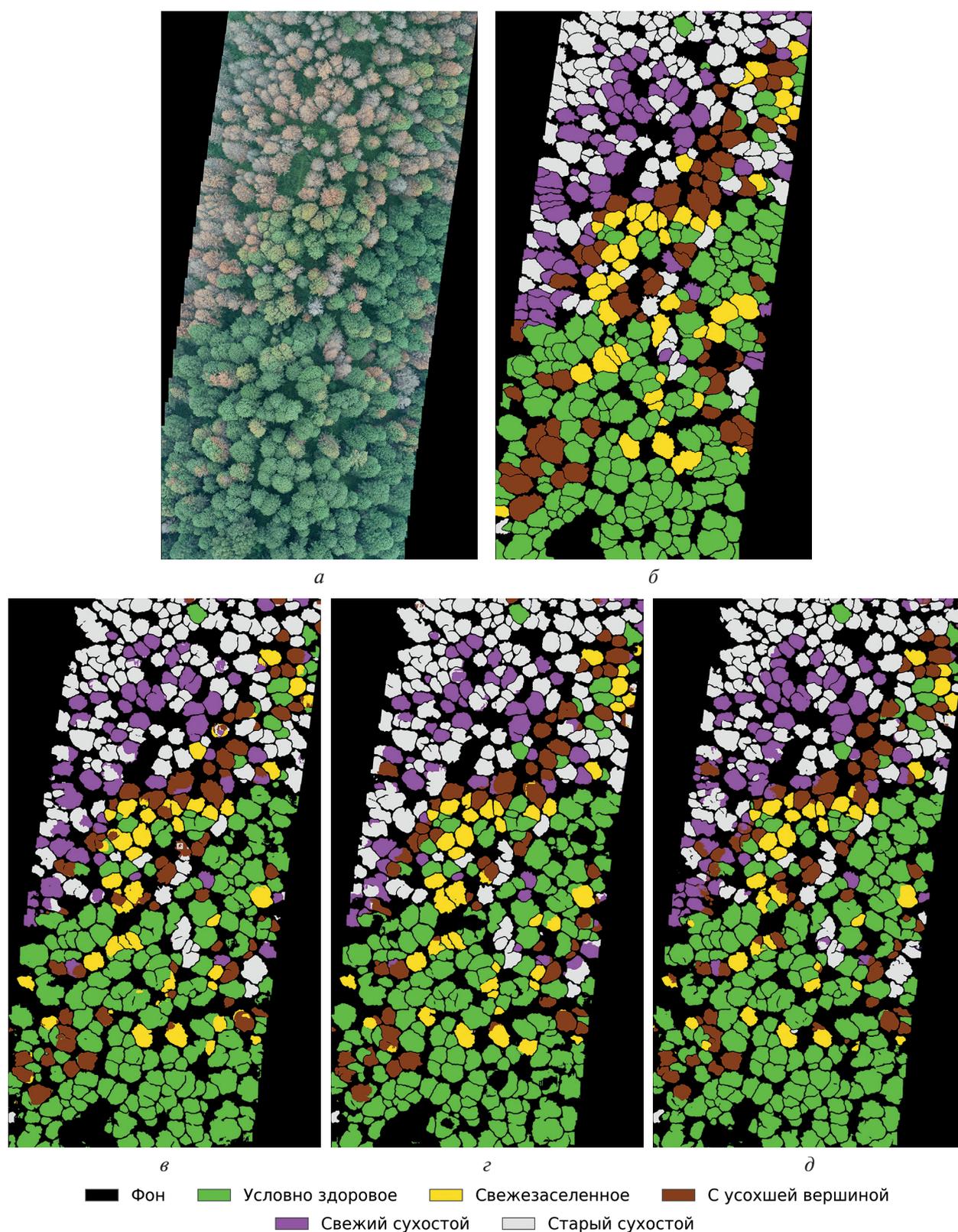


Рис. 7. Результаты сегментации деревьев *P. sibirica* на изображении тестового участка: *a* — изображение тестового участка; *б* — эталонная карта сегментации; *в* — результат сегментации моделью U-Net; *г* — моделью MH-U-Net; *д* — моделью MH-Res-U-Net

Все исследуемые модели СНС успешно классифицируют пиксели пяти классов из шести: модели U-Net и MH-Res-U-Net хорошо сегментируют все классы, кроме класса «свежезаселённое», а модель MH-U-Net — все классы, кроме класса «свежий сухостой». При этом

модель МН-U-Net справляется с сегментацией промежуточного класса «свежезаселённое», модель U-Net — класса «с усохшей вершиной», а модель МН-Res-U-Net — класса «свежий сухостой». В связи с этими особенностями разработанных моделей одним из перспективных направлений дальнейших исследований становится ансамблирование таких моделей в целях комбинации их лучших качеств.

Таблица 1. Метрики качества сегментации изображений деревьев *P. sibirica*

Модель	IoU						mIoU	
	Фон	Условно здоровое	Свежезаселённое	С усохшей вершиной	Свежий сухостой	Старый сухостой	val	test
U-Net	<b>0,80</b>	0,74	0,49	<b>0,53</b>	0,56	0,73	0,51	0,64
МН-U-Net	<b>0,80</b>	0,72	<b>0,54</b>	0,52	0,46	0,70	0,55	0,62
МН-Res-U-Net	<b>0,80</b>	<b>0,75</b>	0,47	0,50	<b>0,63</b>	<b>0,75</b>	<b>0,57</b>	<b>0,65</b>

Основную сложность для сегментации представляют классы промежуточного состояния деревьев *P. sibirica* («свежезаселённое», «с усохшей вершиной» и «свежий сухостой»). Относительно невысокое качество сегментации этих классов обосновывается их недостаточной представленностью в обучающей выборке, что было отмечено при проведении разведочного анализа исходных данных. Важно отметить, что деревья смежных классов часто очень похожи друг на друга визуально. Это также объясняет более низкое качество сегментации для деревьев промежуточных классов, так как для классов «условно здоровое» и «старый сухостой» моделям СНС проще извлечь характерные признаки деревьев из-за потенциально меньшего количества границ принятия решений (что видно, например, при анализе спектральных характеристик на *рис. 2б*).

**Исследование чувствительности моделей класса U-Net к выбору гиперпараметров.** Одним из результатов в ходе обучения предложенных моделей U-Net стала стабильно более качественная сегментация изображений крон деревьев с помощью моделей МН-U-Net и МН-Res-U-Net. Для исследования чувствительности всех трёх моделей к выбору гиперпараметров для каждой из них были взяты все наборы случайных гиперпараметров и соответствующих значений метрики mIoU. На *рис. 8* приведена диаграмма размаха, на которой можно видеть, что медиана mIoU (отмечена оранжевым) для моделей МН-U-Net и МН-Res-U-Net больше, чем для U-Net.

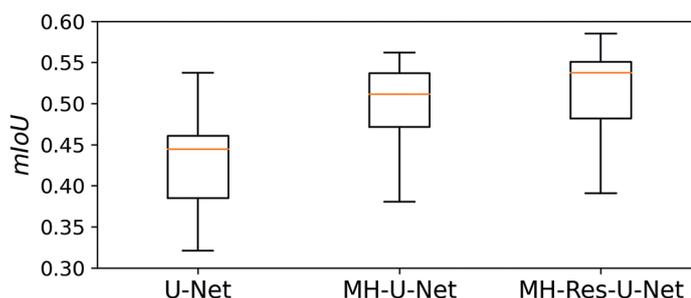


Рис. 8. Значения метрики mIoU для различных моделей СНС, полученные на этапе случайного поиска оптимальных гиперпараметров

Чтобы статистически подтвердить значимость отличия групп значений mIoU друг от друга для разных моделей, были проведены попарные тесты Краскела–Уоллиса. Полученные *p*-уровни представлены в *табл. 2*. Полужирным начертанием выделены значения, при которых принимается гипотеза о значимом статистическом различии совокупностей значений mIoU при  $\alpha = 0,05$ .

Таблица 2. Значения  $p$ -уровней для теста Краскела – Уоллиса

	U-Net	MH-U-Net	MH-Res-U-Net
U-Net	–	<b>1,7e-3</b>	<b>1,2e-4</b>
MH-U-Net	<b>1,7e-3</b>	–	6,5e-2
MH-Res-U-Net	<b>1,2e-4</b>	6,5e-2	–

Проанализировав значения  $p$ -уровней, можно утверждать, что результаты сегментации для конкретного набора исходных снимков с БПЛА, получаемые с помощью моделей MH-U-Net и MH-Res-U-Net, лучше, чем в случае модели U-Net, и эти различия статистически значимы. Таким образом, использование моделей MH-U-Net и MH-Res-U-Net позволяет снизить количество обучаемых моделей для получения приемлемых результатов сегментации изображений поражённых деревьев *P. sibirica* и, следовательно, значительно экономить время и вычислительные ресурсы при решении этой практически важной задачи.

### Заключение

Рассмотрена актуальная задача классификации деревьев *P. sibirica*, повреждённых союзным короедом, на снимках с БПЛА. Показано, что для её решения перспективными становятся современные модели СНС, например U-Net. Разработанные модели U-Net позволяют осуществлять попиксельную классификацию на снимках поражённых деревьев *P. sibirica* в зависимости от степени их повреждения союзным короедом. В результате исследований показано, что модели U-Net и MH-Res-U-Net успешно сегментируют все классы деревьев *P. sibirica*, кроме класса «свежезаселённое», а модель MH-U-Net — все классы, кроме класса «свежий сухостой». Для успешной сегментации всех классов деревьев *P. sibirica*, возможно, перспективным направлением исследований окажется создание ансамблей из полученных моделей. Другим направлением исследований может быть решение задачи регрессии, так как процесс ухудшения состояния деревьев при поражении союзным короедом непрерывный и часто даже экспертам весьма сложно однозначно отнести дерево к тому или иному классу.

Исследования также показали, что модели MH-U-Net и MH-Res-U-Net имеют меньшую чувствительность к выбору гиперпараметров, чем модель U-Net. Однако задача выявления зависимостей чувствительности моделей от значений конкретных гиперпараметров требует дальнейшего рассмотрения.

Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований и администрации Томской обл. (научный проект № 18-47-700010 p\_a) и в рамках госбюджетной темы № 121031300226-5 «Динамические и эволюционные процессы в природных экосистемах Сибири: индикаторы, мониторинг, прогноз».

### Литература

1. Керчев И. А., Кривец С. А., Бисирова Э. М., Смирнов Н. А. (2021a) Распространение союзного короеда *Ips amitinus* (Eichhoff, 1872) в Западной Сибири // Российский журн. биолог. инвазий. 2021. № 4. С. 77–84. DOI: 10.35885/1996-1499-2021-14-4.
2. Керчев И. А., Маслов К. А., Марков Н. Г., Токарева О. С. (2021b) Семантическая сегментация повреждённых деревьев пихты на снимках с беспилотных летательных аппаратов // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 1. С. 116–126. DOI: 10.21046/2070-7401-2021-18-1-116-126.
3. Bertels J., Eelbode T., Berman M., Vandermeulen D., Maes F., Bisschops R., Blaschko M. Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice // arXiv preprint. arXiv:1911.01685. 2019. 9 p. URL: <https://arxiv.org/abs/1911.01685> (accessed 18.11.2021).
4. Bradshaw C. J. A., Leroy B., Bellard C., Roiz D., Albert C., Fournier A., Barbet-Massin M., Salles J.-M., Simard F., Courchamp F. Massive yet grossly underestimated global costs of invasive insects // Nature Communications. 2016. V. 7. Art. No. 12986. DOI: 10.1038/ncomms12986.

5. *Chenari A., Erfanjifard Y., Dehghani M., Pourghasemi H.R.* Woodland Mapping at Single-Tree Levels Using Object-Oriented Classification of UAV Images // The Intern. Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences. 2017. V. XLII-4/W4. P. 43–49. DOI: 10.5194/isprs-archives-XLII-4-W4-43-2017.
6. *Gini R., Sona G., Ronchetti G., Passoni D., Pinto L.* Improving Tree Species Classification Using UAS Multispectral Images and Texture Measures // Intern. J. Geo-Informatics. 2018. V. 7. Iss. 8. Art. No. 325. DOI: 10.3390/ijgi7080315.
7. *He K., Zhang X., Ren S., Sun J.* Deep Residual Learning for Image Recognition // arXiv preprint. arXiv:1512.03385. 2015. 12 p. URL: <https://arxiv.org/abs/1512.03385> (accessed 21.11.2021).
8. *Ioffe S., Szegedy C.* Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift // arXiv preprint. arXiv:1502.03167. 2015. 11 p. URL: <https://arxiv.org/abs/1502.03167> (accessed 07.11.2021).
9. *Kerchev I.A., Mandelshtam M.Yu., Krivets S.A., Ilinsky Yu. Yu.* Small spruce bark beetle *Ips amitinus* (Eichhoff, 1872) (Coleoptera, Curculionidae: Scolytinae): a new alien species in West Siberia // Entomological Review. 2019. V. 99. No. 5. P. 639–644.
10. *Kingma D.P., Ba J.* Adam: A Method for Stochastic Optimization // arXiv preprint. arXiv:1412.6980. 2014. 15 p. URL: <https://arxiv.org/abs/1412.6980> (accessed 07.11.2021).
11. *Lee S., Park S., Baek G., Kim H., Lee C.* Detection of Damaged Pine Tree by the Pine Wilt Disease Using UAV Image // Korean J. Remote Sensing. 2019. V. 35. No. 3. P. 359–373. DOI: 10.7780/kjrs.2019.35.3.2.
12. *Liu C., Diagne C., Angulo E., Banerjee A.-K., Chen Y., Cuthbert R.N., Haubrock P.J., Kirichenko N., Pattison Z., Watari Y., Xiong W., Courchamp F.* Economic costs of biological invasions in Asia // The economic costs of biological invasions around the world / eds. Zenni R. D., McDermott S., Garcia-Berthou E., Essl F. *NeoBiota*, 2021. V. 67. P. 53–78. URL: <https://neobiota.pensoft.net/article/58147/>.
13. *McInnes L., Healy J., Melville J.* UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction // arXiv preprint. arXiv:1802.03426. 2018. 51 p. URL: <https://arxiv.org/abs/1802.03426> (accessed 01.11.2021).
14. *Økland B., Flø D., Schroeder M., Zach P., Cocos D., Martikainen P., Siitonen J., Mandelshtam M. Y., Musolin D. L., Neuvonen S., Vakula J., Nikolov C., Lindelöw Å., Voolma K.* Range expansion of the small spruce bark beetle *Ips amitinus*: a newcomer in northern Europe // *Agricultural and Forest Entomology*. 2019. V. 21. No. 3. P. 286–298.
15. *Onishi M., Ise T.* Automatic Classification of Trees Using a UAV Onboard Camera and Deep Learning // arXiv preprint. arXiv:1804.10390. 2018. 9 p. URL: <https://arxiv.org/abs/1804.10390> (accessed 01.11.2021).
16. *Paszke A., Chaurasia A., Kim S., Culurciello E.* ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation // arXiv preprint. arXiv:1606.02147. 2016. 10 p. URL: <https://arxiv.org/abs/1606.02147> (accessed 13.11.2021).
17. *Rahman M.A., Wang Y.* Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation // *Advances in Visual Computing: Intern. Symp. Visual Computing*. Cham: Springer, 2016. V. 10072. P. 234–244. DOI: 10.1007/978-3-319-50835-1\_22.
18. *Ronneberger O., Fischer P., Brox T.* U-Net: Convolutional Networks for Biomedical Image Segmentation // arXiv preprint. arXiv:1505.04597. 2015. 8 p. URL: <https://arxiv.org/abs/1505.04597> (accessed 12.11.2021).
19. *Rousseeuw P.J.* Silhouettes: A graphical aid to the interpretation and validation of cluster analysis // *J. Computational and Applied Mathematics*. 1987. V. 30. P. 33–65. DOI: 10.1016/0377-0427(87)90125-7.
20. *Safonova A., Tabik S., Alcaraz-Segura D., Rubtsov A., Maglinets Y., Herrera F.* Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning // *Remote Sensing*. 2019. V. 11. No. 6. Art. No. 643. DOI: 10.3390/rs11060643.
21. *Simard P.Y., Steinkraus D., Platt J.C.* Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis // *Proc. Intern. Conf. Document Analysis and Recognition*. 2003. pp. 958–963.
22. *Srivatsava N., Hinton J., Krizhevsky A., Sutskever I., Salakhutdinov R.* Dropout: A Simple Way to Prevent Neural Networks from Overfitting // *J. Machine Learning Research*. 2014. V. 15. P. 1929–1958.

## U-Net models for semantic segmentation of damaged *Pinus sibirica* trees in UAV imagery

N. G. Markov<sup>1</sup>, K. A. Maslov<sup>1</sup>, I. A. Kerchev<sup>2</sup>, O. S. Tokareva<sup>1</sup>

<sup>1</sup> National Research Tomsk Polytechnic University, Tomsk 634050, Russia  
E-mails: markovng@tpu.ru, kam20@tpu.ru, ost@tpu.ru

<sup>2</sup> Institute of Monitoring of Climatic and Ecological Systems SB RAS  
Tomsk 634055, Russia  
E-mail: ivankerchev@gmail.com

Since 2019, there have been rapid increases in mortality rates of *Pinus sibirica* Du Tour stands caused by a new alien bark beetle — *Ips amitinus* Eichh. — in three regions of Western Siberia: Tomsk, Kemerovo and Novosibirsk regions. The success of pest management directly depends on the timeliness of identifying the colonised trees. Dried out treetops, which are hardly noticeable during ground surveys, is a distinct feature of the damaged trees. The use of unmanned aerial vehicles (UAVs) ensures time efficiency of observations and provides an ultra-high spatial resolution of tree crown images. This paper is devoted to the development of U-Net models and their testing when solving the problem of semantic segmentation of *P. sibirica* trees damaged by the pest in UAV imagery. To assess the condition of the trees, experts identify five classes of them: healthy, recently colonised, with a dried out treetop, current year deadwood and old deadwood. Trees of other species and the remaining objects in the images belong to the background. Images acquired in July 2019 with the DJI Phantom 3 Standard drone were used to train, validate and test the models. To solve the addressed problem several fully convolutional networks were proposed: a minor modification of U-Net and two major modifications — multihead-U-Net (MH-U-Net) and multihead-residual-U-Net (MH-Res-U-Net). MH-U-Net is an ensemble of three U-Nets of different depths. The models in the ensemble share part of their weights and simultaneously analyse an image at three different scales. MH-Res-U-Net has all the properties of MH-U-Net and additionally introduces residual blocks. The research has shown that all the models successfully classify pixels of five classes out of six: U-Net and MH-Res-U-Net successfully classify all the classes except the recently colonised trees, and MH-U-Net — all the classes except the current year deadwood trees. Intermediate classes of the tree condition represent the main difficulty for segmentation. However, MH-U-Net copes with the segmentation of the intermediate class of the recently colonised trees, U-Net — of the trees with a dried out treetop, and MH-Res-U-Net — of the current year deadwood trees.

**Keywords:** unmanned aerial vehicle, deep learning, fully convolutional network, U-Net, semantic segmentation, *Pinus sibirica* Du Tour, *Ips amitinus* Eichh.

Accepted: 13.01.2022

DOI: 10.21046/2070-7401-2022-19-1-65-77

### References

1. Kerchev I. A., Krivets S. A., Bisirova E. M., Smirnov N. A. (2021a), Distribution of the small spruce bark beetle *Ips amitinus* (Eichhoff, 1872) in Western Siberia, *Rossiiskii zhurnal biologicheskikh invazii*, 2021, No. 4 (in Russian), DOI: 10.35885/1996-1499-2021-14-4.
2. Kerchev I. A., Maslov K. A., Markov N. G., Tokareva O. S. (2021a), Semantic Segmentation of Damaged Fir Trees in Unmanned Aerial Vehicle Images, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2021, Vol. 18, No. 1, pp. 116–126 (in Russian), DOI: 10.21046/2070-7401-2021-18-1-116-126.
3. Bertels J., Eelbode T., Berman M., Vandermeulen D., Maes F., Bisschops R., Blaschko M., Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice, *arXiv preprint, arXiv:1911.01685*, 2019, 9 p., available at: <https://arxiv.org/abs/1911.01685> (accessed 18.11.2021).
4. Bradshaw C. J. A., Leroy B., Bellard C., Roiz D., Albert C., Fournier A., Barbet-Massin M., Salles J.-M., Simard F., Courchamp F., Massive yet grossly underestimated global costs of invasive insects, *Nature Communications*, 2016, Vol. 7, Art. No. 12986, DOI: 10.1038/ncomms12986.
5. Chenari A., Erfanfard Y., Dehghani M., Pourghasemi H. R., Woodland Mapping at Single-Tree Levels Using Object-Oriented Classification of UAV Images, *The Intern. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, Vol. XLII-4/W4, pp. 43–49, DOI: 10.5194/isprs-archives-XLII-4-W4-43-2017.

6. Gini R., Sona G., Ronchetti G., Passoni D., Pinto L., Improving Tree Species Classification Using UAS Multispectral Images and Texture Measures, *Intern. J. Geo-Informatics*, 2018, Vol. 7, Issue 8, Art. No. 325, DOI: 10.3390/ijgi7080315.
7. He K., Zhang X., Ren S., Sun J., Deep Residual Learning for Image Recognition, *arXiv preprint, arXiv:1512.03385*, 2015, 12 p., available at: <https://arxiv.org/abs/1512.03385> (accessed 21.11.2021).
8. Ioffe S., Szegedy C., Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, *arXiv preprint, arXiv:1502.03167*, 2015, 11 p., available at: <https://arxiv.org/abs/1502.03167> (accessed 07.11.2021).
9. Kerchev I. A., Mandelshtam M. Yu., Krivets S. A., Ilinsky Yu. Yu., Small spruce bark beetle *Ips amitinus* (Eichhoff, 1872) (Coleoptera, Curculionidae: Scolytinae): a new alien species in West Siberia, *Entomological Review*, 2019, Vol. 99, No. 5, pp. 639–644.
10. Kingma D. P., Ba J., Adam: A Method for Stochastic Optimization, *arXiv preprint, arXiv:1412.6980*, 2014, 15 p., available at: <https://arxiv.org/abs/1412.6980> (accessed 07.11.2021).
11. Lee S., Park S., Baek G., Kim H., Lee C., Detection of Damaged Pine Tree by the Pine Wilt Disease Using UAV Image, *Korean J. Remote Sensing*, 2019, Vol. 35, No. 3, pp. 359–373, DOI: 10.7780/kjrs.2019.35.3.2.
12. Liu C., Diagne C., Angulo E., Banerjee A.-K., Chen Y., Cuthbert R. N., Haubrock P. J., Kirichenko N., Pattison Z., Watari Y., Xiong W., Courchamp F., Economic costs of biological invasions in Asia, In: *The economic costs of biological invasions around the world*, Zenni R. D., McDermott S., García-Berthou E., Essl F. (eds), NeoBiota, 2021, Vol. 67, pp. 53–78, available at: <https://neobiota.pensoft.net/article/58147/>.
13. McInnes L., Healy J., Melville J., UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, *arXiv preprint, arXiv:1802.03426*, 2018, 51 p., available at: <https://arxiv.org/abs/1802.03426> (accessed 01.11.2021).
14. Økland B., Flø D., Schroeder M., Zach P., Cocos D., Martikainen P., Siitonen J., Mandelshtam M. Y., Musolin D. L., Neuvonen S., Vakula J., Nikolov C., Lindelöw Å., Voolma K., Range expansion of the small spruce bark beetle *Ips amitinus*: a newcomer in northern Europe, *Agricultural and Forest Entomology*, 2019, Vol. 21, No. 3, pp. 286–298.
15. Onishi M., Ise T., Automatic Classification of Trees Using a UAV Onboard Camera and Deep Learning, *arXiv preprint, arXiv:1804.10390*, 2018, 9 p., available at: <https://arxiv.org/abs/1804.10390> (accessed 01.11.2021).
16. Paszke A., Chaurasia A., Kim S., Culurciello E., ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation, *arXiv preprint, arXiv:1606.02147*, 2016, 10 p., available at: <https://arxiv.org/abs/1606.02147> (accessed 13.11.2021).
17. Rahman M. A., Wang Y., Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation, In: *Advances in Visual Computing: Intern. Symp. Visual Computing*, Cham: Springer, 2016, Vol. 10072, pp. 234–244, DOI: 10.1007/978-3-319-50835-1\_22.
18. Ronneberger O., Fischer P., Brox T., U-Net: Convolutional Networks for Biomedical Image Segmentation, *arXiv preprint, arXiv:1505.04597*, 2015, 8 p., available at: <https://arxiv.org/abs/1505.04597> (accessed 12.11.2021).
19. Rousseeuw P. J., Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Computational and Applied Mathematics*, 1987, Vol. 30, pp. 33–65, DOI: 10.1016/0377-0427(87)90125-7.
20. Safonova A., Tabik S., Alcaraz-Segura D., Rubtsov A., Maglinets Y., Herrera F., Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning, *Remote Sensing*, 2019, Vol. 11, No. 6, Art. No. 643, DOI: 10.3390/rs11060643.
21. Simard P. Y., Steinkraus D., Platt J. C., Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis, *Proc. Intern. Conf. Document Analysis and Recognition*, 2003, pp. 958–963.
22. Srivatsava N., Hinton J., Krizhevsky A., Sutskever I., Salakhutdinov R., Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *J. Machine Learning Research*, 2014, Vol. 15, pp. 1929–1958.