

ОЦЕНКА СОСТОЯНИЯ ЗДОРОВЬЯ ПОРАЖЕННЫХ ДЕРЕВЬЕВ ПИХТЫ НА СНИМКАХ ВЫСОКОГО РАЗРЕШЕНИЯ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ С МЕХАНИЗМОМ ВНИМАНИЯ

Мачука М.К.Р.

Томский политехнический университет, ИШИТР, аспирант, гр. А2-36,
e-mail: kristianrodrigo1@tpu.ru

Аннотация

Работа посвящена оценке состояния здоровья деревьев пихты *Abies sibirica*, поврежденных уссурийским полиграфом, на снимках высокого разрешения с беспилотных летательных аппаратов с использованием нейронных сетей. Показано, что предложенная модель At-Mo-U-Net с механизмом внимания обеспечивает высокую точность классификации поврежденных деревьев пихты на снимках.

Ключевые слова: мониторинг хвойных лесов, семантическая сегментация снимков деревьев пихты, сверточная нейронная сеть с механизмом внимания

Введение

Современные методы дистанционного мониторинга лесов играют ключевую роль в своевременном выявлении патологических изменений древесных насаждений. Особую актуальность эта задача приобретает при массовом поражении хвойных деревьев насекомыми-вредителями. Например, уссурийским полиграфом *Polygraphus proximus* (далее – *P. Proximus*) поражаются деревья пихты *Abies sibirica* (далее – *A. Sibirica*), что наносит значительный ущерб пихтовым лесам [1].

Традиционные методы оценки состояния лесов требуют значительных временных и трудовых затрат, в то время как использование беспилотных летательных аппаратов (БПЛА) с высокоточными камерами позволяет оперативно получать детальные снимки лесных массивов. Однако интерпретация таких данных вручную остается трудоемкой задачей, что обуславливает необходимость разработки автоматизированных методов анализа таких снимков. В данном контексте особый интерес представляют сверточные нейронные сети (СНС), способные выделять значимые признаки крон деревьев на изображениях и весьма точно классифицировать степень повреждения деревьев.

В данной работе предложена гибридная модель At-Mo-U-Net, разработанная на основе классической модели полносверточной нейронной сети U-Net [2] с добавлением блока механизма внимания [3]. Целью исследования является оценка возможностей модели At-Mo-U-Net при классификации состояния здоровья (степени повреждения) деревьев пихты *A. sibirica* на снимках высокого разрешения, полученных с БПЛА.

Исходные данные для проведения исследований

Съемка проводилась в Томской области, преимущественно в районе деревни Парбиг Бакчарского района. Основным объектом изучения стали насаждения пихты *A. sibirica*, поврежденные уссурийским полиграфом *P. proximus*.

Для оценки состояния деревьев было собрано пять панорам с разрешением около 0.1 м, полученных с помощью БПЛА DJI Phantom 3 Standart. Съемка проводилась в видимом диапазоне (RGB) с высот 365-388 м в период с 7 по 28 августа 2017 года. Полученные панорамы имели следующие размеры (в пикселях):

- Панорама А: 1046×1912×3
- Панорама В: 1536×1048×3
- Панорама С: 1536×768×3
- Панорама D: 768×1792×3
- Панорама Е: 1046×1912×3 (использовалась для тестирования моделей)

Институт мониторинга климатических и экологических систем СО РАН разработал шкалу оценки состояния здоровья деревьев пихты *A. sibirica*, включающую 4 класса (рис. 1):

- Живое (рис. 1(а)) - Здоровые деревья, хвоя зеленая.
- Отмирающее (рис. 1(б)) - Кроны могут быть как у здоровых деревьев, но более половины ветвей несут усыхающую хвою или хвоя в верхней части кроны еще зеленая, а ниже на ветках хвоя рыжего цвета. Возможно наличие усыхающих ветвей в верхней части кроны.
- Свежий сухостой (рис. 1(в)) - Хвоя кроны мертвая, красная, частично осыпавшаяся. Густота кроны составляет только часть от кроны живого дерева, наличие усохших ветвей по всей кроне. Дерево погибло менее года назад.
- Старый сухостой (рис. 1(г)) - Крона мертвая, серая, часто хвоя полностью осыпалась. Дерево погибло несколько лет назад.
- Фон - другие виды деревьев и иные объекты поверхности Земли.

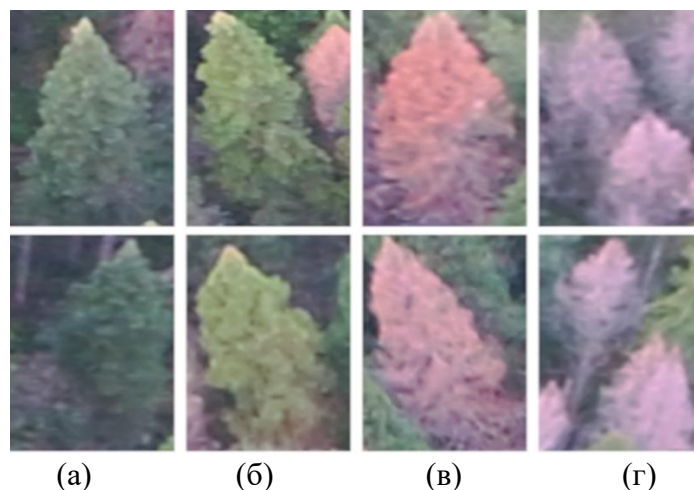


Рис. 1. Примеры состояния здоровья деревьев пихты *A. sibirica*: (а) Живое. (б) Отмирающее. (в) Свежий сухостой. (г) Старый сухостой.

Подготовка датасетов

Процесс разметки полученных панорам выполнялся с использованием графического редактора GIMP. На первом этапе эксперт присваивал числовые метки деревьям пихты *A. sibirica* на изображениях с БПЛА, отражающие их состояние здоровья. Затем другой специалист на основе этих меток создавал цветные маски сегментации, где каждый класс повреждений деревьев имел свое цветовое обозначение. Этот процесс позволил получить точные маски, разделяющие деревья на четыре категории состояния здоровья и фон [4].

Для формирования датасетов были отобраны четыре панорамных изображения для обучения и валидации, а также одна отдельная панорама для тестирования. Изображения и соответствующие маски сегментации были разделены на фрагменты размером $256 \times 256 \times 3$ пикселей с перекрытием в 128 пикселей. Чтобы исключить дублирование признаков между обучающими и валидационными выборками, применялись булевы маски к исходным изображениям. Дополнительно для увеличения объема данных использовалась аугментация, включающая повороты изображений под различными углами. В результате был создан первый датасет, содержащий 2004 обучающих фрагмента, 672 валидационных и 96 тестовых.

Параллельно был подготовлен второй датасет с фрагментами большего размера - $480 \times 480 \times 3$ пикселей. В этом случае перекрытие между соседними фрагментами составляло 240 пикселей. Данный датасет включал 502 обучающих образца, 180 валидационных и 24 тестовых фрагмента.

Анализ распределения деревьев по классам в исходных изображениях выявил существенный дисбаланс: классы «Отмирающее» (80 экземпляров) и «Свежий сухостой» (274 экземпляра) были представлены значительно меньше, чем «Живое» (574) и «Старый

сухостой» (491). Для компенсации этого дисбаланса в процессе обучения применялась аугментация данных «на лету», включавшая масштабирование, эластичные деформации, коррекцию яркости и контраста, а также вертикальные отражения изображений. Эти методы позволили увеличить разнообразие обучающей выборки и улучшить качество обучения моделей.

Предложенная гибридная модель At-Mo-U-Net

Гибридная модель, названная At-Mo-U-Net, – это усовершенствованная модель СНС на основе классической полносверточной модели U-Net [2]. Гибридная модель сочетает структурные модификации классической модели U-Net с механизмом пространственного внимания (рис. 2), что позволяет добиться более точного выделения значимых признаков объектов на снимках высокого разрешения. Гибридная модель разработана для решения задач семантической сегментации снимков хвойных деревьев, в частности, для оценки на снимках состояния здоровья пораженных уссурийским полиграфом *P. proximus* деревьев пихты *A. Sibirica*.

Модель At-Mo-U-Net имеет несколько принципиальных отличий от классической модели U-Net. Во-первых, входные данные обрабатываются в виде RGB-изображений размером $256 \times 256 \times 3$ или $480 \times 480 \times 3$ пикселей, тогда как модель U-Net разрабатывалась для работы с биомедицинскими изображениями в градациях серого. Во-вторых, в модели At-Mo-U-Net реализован принцип сохранения пространственной информации - в отличие от классической U-Net, где каждый сверточный блок уменьшает размер карт признаков, а операции пулинга вдвое снижают разрешение, здесь применяются свертки без уменьшения размерности, что позволяет сохранить больше деталей изображения.

Особое внимание уделено архитектуре пропускных соединений (skip connections). Если в оригинальной U-Net для согласования размеров используются обрезанные (cropped) карты признаков, то в модели At-Mo-U-Net обрезка не применяется - вместо этого для выравнивания размеров используется комбинация интерполяции и конкатенации.

Существенные изменения внесены в процесс нормализации и активации. После каждого слоя активации добавлена пакетная нормализация, что значительно ускоряет процесс обучения модели. Кроме того, традиционная функция активации ReLU заменена на экспоненциальную линейную единицу (ELU) [5], которая благодаря плавной обработке отрицательных значений улучшает сходимость модели. На выходном слое применяются свертки 1×1 с количеством фильтров, соответствующим числу классов классификации (включая фоновый класс).

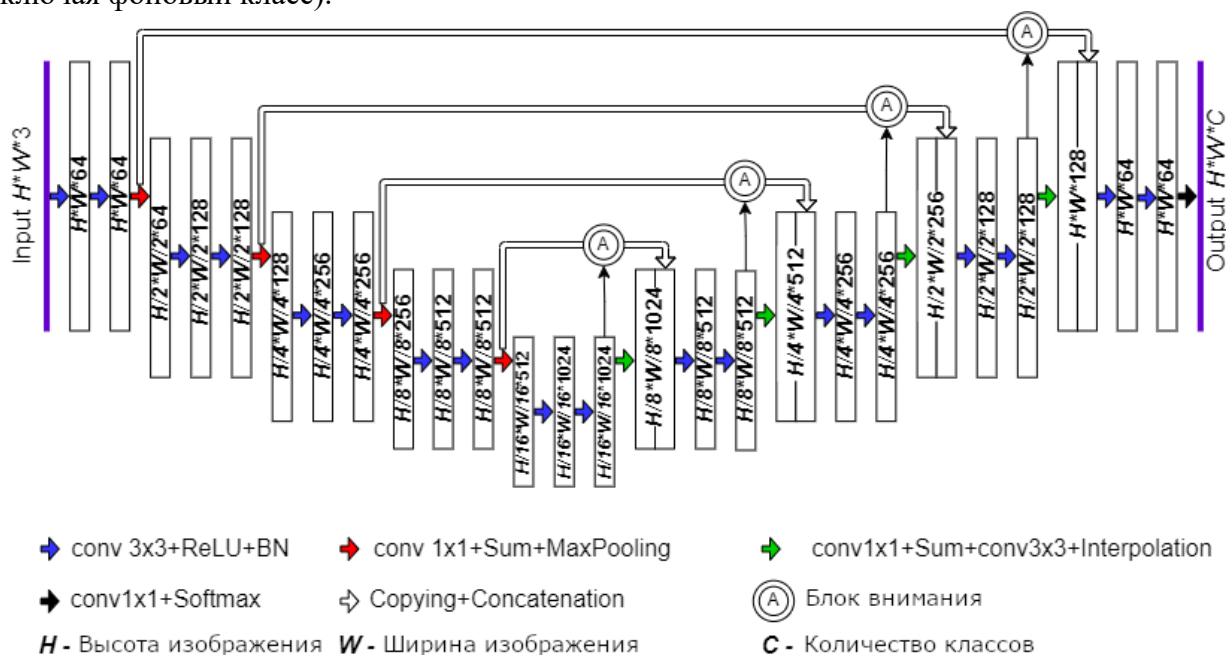


Рис. 2. Архитектура модели At-Mo-U-Net.

Главным инновационным элементом гибридной модели являются блоки внимания, интегрированные как в энкодер, так и в декодер базовой модели U-Net. Эти блоки позволяют сети автоматически выделять наиболее значимые области изображения (например, поврежденные участки крон деревьев пихты *A. sibirica*) и эффективно игнорировать шумовые компоненты. Механизм внимания работает по принципу динамического перевзвешивания значимости различных участков изображения, что существенно повышает точность сегментации по сравнению с точностью в случае применения модели U-Net.

Механизм внимания в модели At-Mo-U-Net реализован через три ключевых этапа обработки данных. На первом этапе – этапе вычисления весов внимания – на вход специального слоя поступают карты признаков, полученные либо из энкодера, либо из предыдущих слоев модели. Этот слой, представляющий собой комбинацию сверточной операции и сигмоидной функции активации, генерирует матрицу весовых коэффициентов в диапазоне от 0 до 1, где значения, близкие к 1, указывают на наиболее значимые области изображения, требующие повышенного внимания модели.

На втором этапе происходит перевзвешивание признаков: исходные карты признаков поэлементно умножаются на полученную матрицу весов. Эта операция позволяет усилить релевантные участки изображения и одновременно подавить малозначимые или шумовые компоненты. В результате формируются уточненные карты признаков, где важные для решения задачи области визуально выделены.

Третий этап – слияние с картами декодера – обеспечивает точное восстановление пространственной структуры объектов. Уточненные карты признаков объединяются (конкатенируются) с соответствующими картами из декодера, что особенно важно для корректного восстановления границ сегментируемых объектов и сохранения мелких деталей изображения.

Обучение модели At-Mo-U-Net

В процессе обучения применялась стратегия, направленная на достижение максимальной точности при работе с несбалансированными классами степени повреждения (состояния здоровья) деревьев пихты *A. sibirica*. В качестве функции потерь использовалась *Focal Loss*, которая эффективно справляется с неравномерным распределением классов, а для оптимизации весов модели был выбран алгоритм Adam.

Для предотвращения переобучения в процесс обучения был встроен механизм ранней остановки – обучение прекращалось, если в течение 11 последовательных эпох не наблюдалось улучшения показателей на валидационных выборках.

Эффективность модели оценивалась с помощью метрики Intersection over Union (IoUc) для каждого класса повреждений и средней величины mIoU, вычисляемой как усредненное значение IoUc по всем классам. Значения IoUc и mIoU, превышающие 0.5, свидетельствуют о высоком качестве сегментации снимков, достаточном для практического применения при решении лесопатологических задач в лесной отрасли.

Результаты исследования модели At-Mo-U-Net и их обсуждение

На рисунке 3 представлены визуальные результаты обученной модели At-Mo-U-Net при решении задачи семантической сегментации снимков пораженных деревьев пихты *A. sibirica*. Исследование проводилось на тестовой выборке с размерами фрагментов 256×256×3 пикселей из первого датасета. Показано, что модель успешно определяет границы крон деревьев и правильно классифицирует значительную часть крон, даже где кроны деревьев частично перекрываются или имеют неоднородную структуру.

Аналогичные визуальные результаты были получены при тестировании этой модели на тестовой выборке второго датасета с увеличенным размером фрагментов 480×480×3 пикселей (рис. 4). В этом случае модель At-Mo-U-Net также показала высокую точность классификации степени повреждения (состояния здоровья) деревьев пихты.

Количественные результаты этих экспериментов показаны в таблице 1 и таблице 2. Сюда же внесены результаты исследований базовой модели U-Net, полученные с использованием этих же двух датасетов. Эти результаты для разных классов деревьев демонстрируют преимущества модели At-Mo-U-Net по сравнению с моделью U-Net при оценке состояния здоровья пораженных деревьев. Анализ результатов для фрагментов размером $256 \times 256 \times 3$ пикселей в таблице 1 показывает, что модель At-Mo-U-Net достигает более высоких показателей точности по большинству классов деревьев. Особенно заметно преимущество ее при сегментации деревьев класса «Отмирающее», когда разница по метрике IoUc составляет 0.51 против 0.45 у модели U-Net.

Из таблицы 2 следует, что при увеличении размера фрагментов до $480 \times 480 \times 3$ пикселей преимущества модели At-Mo-U-Net сохраняются и даже усиливаются. Модель демонстрирует по сравнению с моделью U-Net улучшенные показатели для всех классов, особенно для класса «Отмирающее» (0.52 против 0.46) и класса «Живое» (0.76 против 0.73).

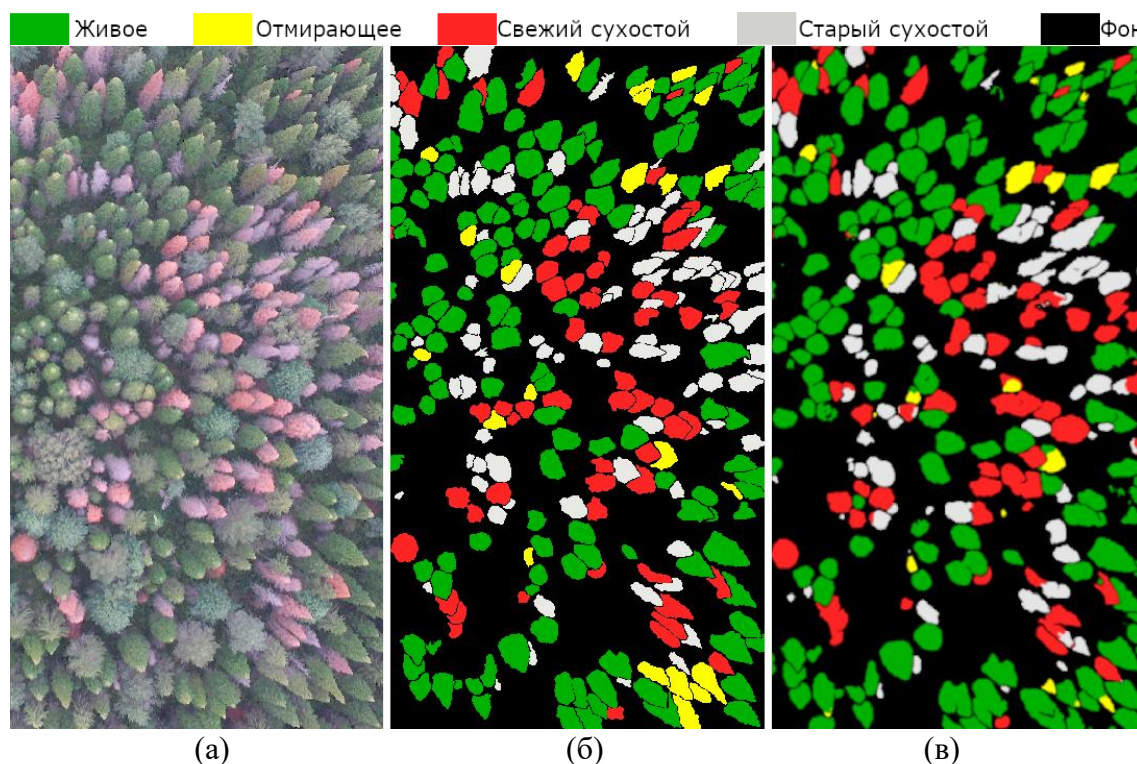
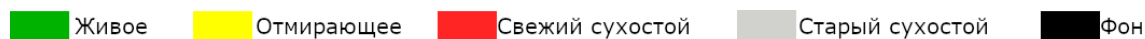


Рис. 3. Результаты семантической сегментации снимка тестового участка *A. sibirica* с использованием фрагментов размером $256 \times 256 \times 3$ пикселей: а – снимок тестового участка, б – эталонная маска сегментации, с – использование модели At-Mo-U-Net.



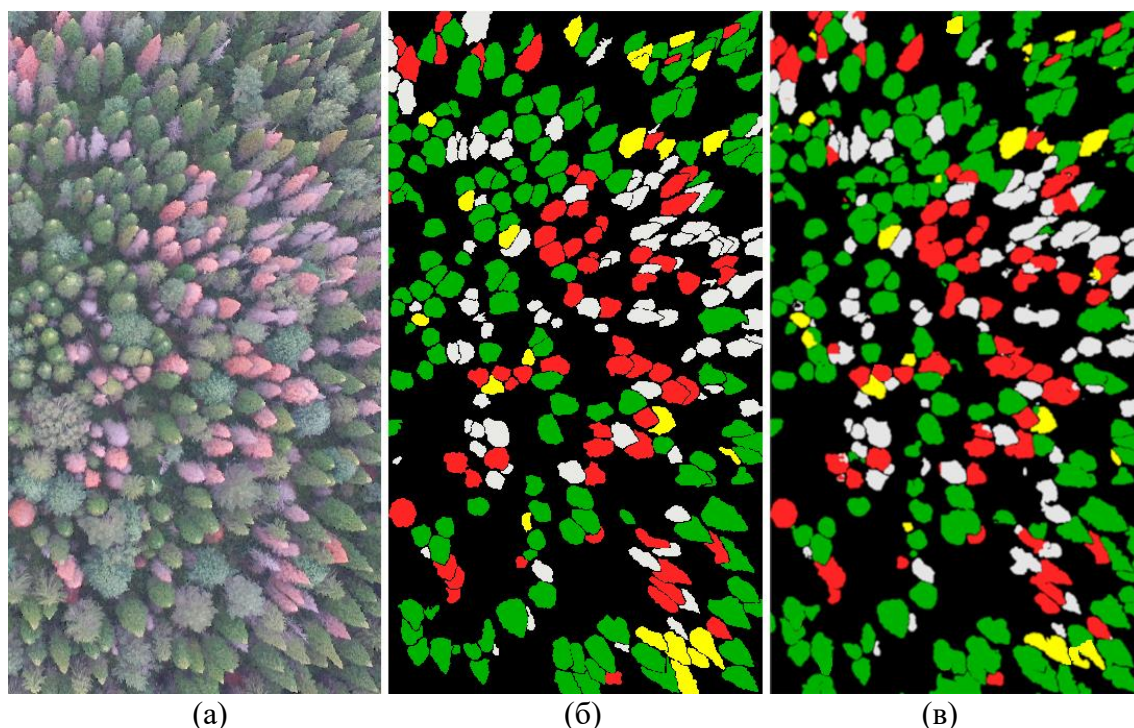


Рис. 3. Результаты семантической сегментации снимка тестового участка *A. sibirica* с использованием фрагментов размером 480x480x3 пикселей: (a) Снимок тестового участка. (b) Эталонная маска сегментации. (c) Использование модели At-Mo-U-Net.

Таблица 1. Качество семантической сегментации моделей, обученных и протестированных с использованием фрагментов снимков деревьев пихты *A. sibirica* размером 256x256x3 пикселей.

Модель	IoUc					mIoU
	Фон	Живое	Отмирающее	Свежий сухостой	Старый сухостой	
U-Net	0.87	0.73	0.45	0.79	0.68	0.70
At-Mo-U-Net	0.86	0.74	0.51	0.79	0.68	0.72

Таблица 2. Качество семантической сегментации моделей, обученных и протестированных с использованием фрагментов снимков деревьев пихты *A. sibirica* размером 480x480x3 пикселей.

Модель	IoUc					mIoU
	Фон	Живое	Отмирающее	Свежий сухостой	Старый сухостой	
U-Net	0.87	0.73	0.46	0.79	0.69	0.71
At-Mo-U-Net	0.86	0.76	0.52	0.80	0.70	0.73

Заключение

Результаты исследования, проведенные с использованием созданных двух датасетов, убедительно доказывают эффективность разработанной гибридной модели At-Mo-U-Net. Введение механизма внимания в базовую модель U-Net позволило добиться стабильного улучшения качества сегментации на 2-3% по метрике mIoU по сравнению с моделью U-Net для обоих размеров фрагментов. Наибольший выигрыш наблюдается для наиболее сложного промежуточного класса деревьев пихты, такого как «Отмирающее», когда разница по метрике IoUc достигает 13%. При этом модель сохраняет все преимущества базовой модели U-Net при классификации простых классов деревьев и фона. Полученные результаты особенно важны для практического применения при решении задач лесопатологического мониторинга, когда

точная идентификация состояния здоровья каждого дерева имеет критическое значение на пораженном участке пихтового леса. Преимущества модели At-Mo-U-Net открывают перспективы для ее использования при анализе снимков высокого и сверхвысокого разрешения.

Список использованных источников

1. Chang W.Y., Lantz V.A., Hennigar C.R., MacLean D.A. Economic impacts of forest pests: a case study of spruce budworm outbreaks and control in New Brunswick, Canada // Canadian Journal of Forest Research. – 2012. – Т. 42. – С. 490–505. DOI: 10.1139/x11-190.
2. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention. – 2015. – Т. 9351. DOI: 10.1007/978-3-319-24574-4_28.
3. Vaswani A. Attention is All you Need / A. Vaswani, N.M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin // Neural Information Processing Systems. – 2017.
4. Керчев И.А. Семантическая сегментация поврежденных деревьев пихты на снимках с беспилотных летательных аппаратов / И.А. Керчев, К.А. Маслов, Н.Г. Марков, О.С. Токарева // Современные проблемы дистанционного зондирования Земли из космоса. – 2021. – Т. 18. – № 1. – С. 116–126. DOI: 10.21046/2070-7401-2021-18-1-116-126.
5. Clevert D.A., Unterthiner T., Hochreiter S. Fast and accurate deep network learning by Exponential Linear Units (ELUs) // arXiv preprint arXiv:1511.07289. – 2015.