

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.02 «Информационные системы и технологии»
 Отделение школы (НОЦ) Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Модели, алгоритмы и программное обеспечение для семантической сегментации хвойных деревьев на снимках с беспилотных летательных аппаратов
УДК <u>004.415.2:004.932.1:582.47</u>

Студент:

Группа	ФИО	Подпись	Дата
8ИМ92	Маслов Константин Андреевич		

Руководитель ВКР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	Д.Т.Н., профессор		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Верховская Марина Витальевна	К.Э.Н.		

По разделу «Социальная ответственность»:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Сечин Андрей Александрович	К.Т.Н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Шерстнев Владислав Станиславович	К.Т.Н., доцент		

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ
по направлению 09.04.02 «Информационные системы и технологии»

Код компетенции	Результат обучения (выпускник должен быть готов)
Универсальные компетенции	
УК(У)-1	Осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.
УК(У)-2	Управлять проектом на всех этапах его жизненного цикла.
УК(У)-3	Организовывать и руководить работой команды, вырабатывая стратегию для достижения поставленной цели.
УК(У)-4	Применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия.
УК(У)-5	Анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия.
УК(У)-6	Определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки.
Общепрофессиональные компетенции	
ОПК(У)-1	Самостоятельно приобретать, развивать и применять математические, естественно-научные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.
ОПК(У)-2	Разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач.
ОПК(У)-3	Анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.
ОПК(У)-4	Применять на практике новые научные принципы и методы исследований.
ОПК(У)-5	Разрабатывать и модернизировать программное и аппаратное обеспечение информационных систем.
ОПК(У)-6	Использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий.
ОПК(У)-7	Разрабатывать и применять математические модели процессов и объектов при решении задач анализа и синтеза распределенных информационных систем и систем поддержки принятия решений.
ОПК(У)-8	Осуществлять эффективное управление разработкой программных средств и проектов.
Профессиональные компетенции	
ПК(У)-1	Управлять программно-техническими, технологическими и человеческими ресурсами.
ПК(У)-2	Управлять развитием баз данных.
ПК(У)-3	Управлять работами по сопровождению и проектами создания (модификации) информационных систем, автоматизирующих задачи организационного управления и бизнес-процессы.
ПК(У)-4	Проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий.
ПК(У)-5	Осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ.

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.02 «Информационные системы и технологии»
 Отделение школы (НОЦ) Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП

 (Подпись) (Дата)

Шерстнев В.С.
 (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ИМ92	Маслов Константин Андреевич

Тема работы:

Модели, алгоритмы и программное обеспечение для семантической сегментации хвойных деревьев на снимках с беспилотных летательных аппаратов

Утверждена приказом директора	46-5/с от 15.02.2021
-------------------------------	----------------------

Срок сдачи студентом выполненной работы:	07.06.2021
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Данные ряда съемок территорий пихтовых насаждений, поврежденных уссурийским полиграфом, и кедровых насаждений, поврежденных союзным короедом, а также соответствующая экспертная разметка в виде точек различных цветов на кронах деревьев.
Перечень подлежащих исследованию, проектированию и разработке вопросов	<ul style="list-style-type: none"> – Анализ предметной области, связанной с автоматизированной интерпретацией данных съемки лесного покрова с БПЛА; – подготовка исходных данных; – проведение разведочного анализа данных; – программная реализация моделей машинного обучения и их обучение; – проектирование и программная реализация алгоритма подбора гиперпараметров; – сравнение и анализ полученных моделей машинного обучения; – финансовый менеджмент; – социальная ответственность; – раздел на иностранном языке.
Перечень графического материала	<ul style="list-style-type: none"> – Схемы реализованных моделей; – результирующие карты сегментации деревьев;

	– таблицы со сравнительным анализом результатов сегментации и характеристик моделей.
Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Верховская Марина Витальевна, доцент ОСГН ШБИП, к.э.н.
Социальная ответственность	Сечин Андрей Александрович, доцент ООД ШБИП, к.т.н.
Раздел на иностранном языке	Коротченко Татьяна Валериевна, доцент ОИЯ ШБИП, к.фил.н.
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Разделы на русском: мониторинг состояния хвойных деревьев; разведочный анализ данных; реализованные модели и алгоритмы; результаты исследования предлагаемых моделей; финансовый менеджмент, ресурсоэффективность и ресурсосбережение; социальная ответственность.	
Раздел на английском: Models and Algorithms.	

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику:	01.09.2020
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	Д.Т.Н., профессор		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ92	Маслов Константин Андреевич		

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.02 «Информационные системы и технологии»
 Уровень образования Магистратура
 Отделение школы (НОЦ) Отделение информационных технологий
 Период выполнения _____ 2020/2021 учебный год

Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	07.06.2021
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
05.10.2020	Анализ предметной области	10
23.11.2020	Подготовка исходных данных	5
23.11.2020	Проведение разведочного анализа данных	10
25.03.2021	Программная реализация моделей машинного обучения	15
25.03.2021	Проектирование и программная реализация алгоритма подбора гиперпараметров	15
23.05.2021	Сравнение и анализ полученных моделей	15
07.06.2021	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
07.06.2021	Социальная ответственность	10
07.06.2021	Раздел на иностранном языке	10

СОСТАВИЛ:

Руководитель ВКР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	Д.Т.Н., профессор		

СОГЛАСОВАНО:

Руководитель ООП:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Шерстнев Владислав Станиславович	К.Т.Н., доцент		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
8ИМ92	Маслов Константин Андреевич

Школа	Инженерная школа информационных технологий и робототехники	Отделение (НОЦ)	Отделение информационных технологий
Уровень образования	Магистратура	Направление/ специальность	09.04.02 «Информационные системы и технологии»

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НТИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад инженера – 22695 руб.; Оклад научного руководителя – 49129,5 руб.
2. Нормы и нормативы расходования ресурсов	Районный коэффициент – 30%; Коэффициент дополнительной заработной платы – 12%; Накладные расходы – 16%.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды – 27,1%.

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	Анализ конкурентных технических решений; SWOT-анализ.
2. Планирование и формирование бюджета научных исследований	Формирование плана и графика разработки: – определение структуры и трудоемкости работ; – создание диаграммы Гантта. Формирование бюджета затрат на разработку: – затраты на амортизацию оборудования; – заработная плата (основная и дополнительная); – отчисления во внебюджетные фонды; – накладные расходы.
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности разработки	Определение потенциального эффекта разработки.

Перечень графического материала:

<ul style="list-style-type: none"> – Оценочная карта конкурентных технических решений; – Матрица SWOT-анализа и интерактивная матрица проекта; – График проведения НТИ; – Диаграмма Гантта; – Бюджет НТИ; – Оценка ресурсной, финансовой и экономической эффективности НТИ.

Дата выдачи задания для раздела по линейному графику:

01.03.2021

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Верховская Марина Витальевна	К.Э.Н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ92	Маслов Константин Андреевич		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8ИМ92	Маслов Константин Андреевич

Школа	Инженерная школа информационных технологий и робототехники	Отделение (НОЦ)	Отделение информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.02 «Информационные системы и технологии»

Тема ВКР:

Модели, алгоритмы и программное обеспечение для семантической сегментации хвойных деревьев на снимках с беспилотных летательных аппаратов

Исходные данные к разделу «Социальная ответственность»:

<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Целью работы является разработка моделей и алгоритмов для семантической сегментации хвойных деревьев, поврежденных насекомыми-вредителями, на снимках с беспилотных летательных аппаратов с последующим анализом и сравнением этих моделей. Объекты исследования: цифровые снимки хвойных насаждений в видимом диапазоне, полученные с беспилотных летательных аппаратов.</p> <p>Область применения: экологический мониторинг.</p> <p>Оборудованием, на котором осуществляется работа является персональный компьютер. Рабочим местом является жилое помещение с площадью в 18 кв.м. Помещение оборудовано отоплением, вентиляцией и соответствует нормам освещенности. Имеется одностольный стол с настольной светодиодной лампой мощностью 11 Вт.</p>
---	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Правовые и организационные вопросы обеспечения безопасности:</p> <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	<ul style="list-style-type: none"> – ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования; – Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 27.12.2018); – СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95; – ГОСТ 12.0.003-2015 ССБТ. Опасные и вредные производственные факторы. Классификация; – СанПиН 2.2.4.548–96 Гигиенические требования к микроклимату производственных помещений; – СН 2.2.4/ 2.1.8.562-96 Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки. Санитарные нормы; – ГОСТ 12.1.038-82 ССБТ Электробезопасность; – ГОСТ 12.1.004-91 ССБТ Пожарная безопасность; – СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.
<p>2. Производственная безопасность:</p> <ul style="list-style-type: none"> – анализ выявленных вредных и опасных факторов; – обоснование мероприятий по снижению воздействия. 	<p>Анализ выявленных вредных факторов:</p> <ul style="list-style-type: none"> – недостаточная освещенность рабочей зоны; – отклонение параметров микроклимата; – повышенный уровень шума; – психофизиологические факторы (монотонность труда, нервно-психические перегрузки, перенапряжение зрительных

	анализаторов). Анализ выявленных опасных факторов: – поражение электрическим током; – статическое электричество; – короткое замыкание.
3. Экологическая безопасность:	Атмосфера: воздействие отсутствует; Гидросфера: воздействие отсутствует; Воздействие на литосферу происходит при утилизации: – компьютера и периферийных устройств; – люминесцентных ламп; – макулатуры.
4. Безопасность в чрезвычайных ситуациях:	Возможные ЧС: аварии на системах жизнеобеспечения населения; Наиболее типичная ЧС: возникновение пожара на рабочем месте.

Дата выдачи задания для раздела по линейному графику:	01.03.2021
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Сечин Андрей Александрович	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ92	Маслов Константин Андреевич		

РЕФЕРАТ

Выпускная квалификационная работа 111 с., 29 рис., 25 табл., 45 источников, 3 прил.

Ключевые слова: УССУРИЙСКИЙ ПОЛИГРАФ, ПИХТА СИБИРСКАЯ, СОЮЗНЫЙ КОРОЕД, СОСНА СИБИРСКАЯ КЕДРОВАЯ, БПЛА, СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ, МН-U-NET, МН-RES-U-NET

Целью данной работы является создание моделей и разработка алгоритмов машинного обучения для семантической сегментации хвойных деревьев, поврежденных насекомыми-вредителями, на снимках с беспилотных летательных аппаратов (БПЛА). Объектами исследования являются цифровые изображения хвойных насаждений, полученные в видимом диапазоне электромагнитного спектра.

Описана проблема анализа состояния хвойных деревьев по снимкам с БПЛА, дана характеристика объектов исследования и проведен разведочный анализ исходных снимков пихтовых и кедровых насаждений, поврежденных вредителями. Это позволило осуществить проектирование ряда моделей и алгоритмов машинного обучения. Описаны ансамблевая модель, обученная алгоритмом градиентного бустинга и использующая текстурные признаки Харалика, и модели полносверточных сетей, основанные на архитектурах U-Net, SegNet и ENet. Также предложены две существенные модификации U-Net: МН-U-Net и МН-Res-U-Net. С помощью описанных моделей решена задача семантической сегментации изображений. Результаты комплексных исследований показали высокую эффективность полученных моделей при классификации пикселей трех из четырех классов деревьев пихты, четырех из пяти классов деревьев кедра и фоновых пикселей. Модели МН-U-Net и МН-Res-U-Net показали меньшую чувствительность к выбору гиперпараметров.

Практическая ценность разработанных моделей и алгоритмов состоит в возможности их использования для проведения оперативного мониторинга поврежденных вредителями хвойных насаждений, что важно для лесной отрасли РФ. Научную новизну имеют предложенные модели МН-U-Net и МН-Res-U-Net.

Результаты магистерской диссертации использованы в ИМКЭС СО РАН.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В данной работе применены следующие обозначения и сокращения:

БПЛА – беспилотный летательный аппарат;

СНС – сверточная нейронная сеть;

GLCM – grey level co-occurrence matrix;

IoU – intersection over union;

mIoU – mean intersection over union;

CNN – convolutional neural network;

FCN – fully convolutional network;

EI – expected improvement.

СОДЕРЖАНИЕ

Введение.....	14
1 Мониторинг состояния хвойных деревьев	16
1.1 Задача анализа состояния пихты сибирской по снимкам с БПЛА	16
1.2 Задача анализа состояния сосны сибирской кедровой по снимкам с БПЛА	17
1.3 Методы анализа изображений дистанционного мониторинга.....	18
1.4 Постановка задач исследования	19
2 Разведочный анализ данных	23
2.1 Разведочный анализ изображений пихтовых насаждений	23
2.2 Разведочный анализ изображений кедровых насаждений	25
3 Реализованные модели и алгоритмы.....	28
3.1 Текстурные признаки Харалика	28
3.2 Градиентный бустинг	29
3.3 Полносверточные сети	30
3.3.1 Архитектура U-Net.....	30
3.3.2 Архитектура SegNet.....	32
3.3.3 Архитектура ENet	32
3.3.4 Архитектура МН-U-Net.....	35
3.3.5 Архитектура МН-Res-U-Net	36
3.4 Оценка качества сегментации.....	37
3.5 Обучение нейронных сетей.....	38
3.6 Подбор гиперпараметров	40
4 Результаты исследования предлагаемых моделей.....	43
4.1 Результаты сегментации тестовых участков.....	43
4.2 Сравнение производительности полносверточных сетей	46
4.3 Исследование чувствительности полносверточных сетей к выбору гиперпараметров.....	47
5 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	50
6 Социальная ответственность	63

Заключение	76
Список публикаций.....	77
Список источников	80
Приложение А Models and Algorithms	86
Приложение Б Дипломы, полученные на конференциях и конкурсах	101
Приложение В Акт об использовании результатов магистерской диссертации	110

ВВЕДЕНИЕ

Обширные очаги усыхания хвойных лесов, вызванные чужеродными насекомыми-вредителями, влекут за собой значимые экономические и экологические последствия [1, 2].

Так одним из основных негативных факторов, воздействующих на пихтовые леса ряда регионов Южной Сибири [1], Байкальского региона [3] и Предуралья [4], является инвазивный короед – уссурийский полиграф *Polygraphus proximus* Blandf. При заселении здоровых деревьев вредитель вносит под кору деревьев пропагулы фитопатогенных офиостомовых грибов, которые вызывают некрозы луба и более глубоких тканей ствола [5, 6]. В результате нарушения работы ассимиляционно-диссимиляционного аппарата происходит постепенная гибель дерева с характерной дехромацией хвои.

Другим негативным фактором является насекомое-вредитель союзный (многоходовый) короед *Ips amitinus* (Eichhoff), заселяющий деревья сосны сибирской кедровой, что приводит к массовому усыханию древостоя. В России вредитель встречается в западных, северо-западных и северных районах европейской части страны. В Западной Сибири вредитель обнаружен недавно в припоселковых кедровниках Томской и Кемеровской областей. Ареал союзного короеда имеет тенденцию к расширению [2].

Наличие таких факторов требует проведения непрерывного и оперативного мониторинга пихтовых насаждений с целью оценки их состояния, планирования санитарных и лесозащитных мероприятий.

Данные дистанционного зондирования Земли, полученные с использованием беспилотных летательных аппаратов (БПЛА), сегодня активно применяются при решении задач экологического мониторинга лесных насаждений, обеспечивая при этом высокое пространственное разрешение получаемых снимков и оперативность в принятии решений [7-9].

Данная работа посвящена разработке ряда моделей машинного обучения и их исследованию при решении задачи семантической сегментации деревьев пихты и кедра, поврежденных вредителями леса, на снимках с БПЛА.

В разделе 1 описаны задача анализа состояния деревьев пихты и кедра по снимкам с БПЛА и используемые в данной работе исходные данные, приведен краткий обзор используемых в литературе методов и поставлена задача исследования.

В разделе 2 приведен разведочный анализ исходных данных.

В разделе 3 дано описание моделей и алгоритмов машинного обучения: алгоритма градиентного бустинга, моделей полносверточных сетей U-Net, SegNet и ENet, а также новых архитектур МН-U-Net и МН-Res-U-Net; приведено описание алгоритмов их обучения, оценки качества и подбора гиперпараметров.

В разделе 4 представлены полученные результаты с использованием предлагаемых моделей: проведена оценка качества сегментации, сравнение моделей по качеству сегментации, времени вычислений и чувствительности к выбору гиперпараметров.

В разделе 5 изложены вопросы, связанные с финансовым менеджментом и ресурсоэффективностью разработки.

В разделе 6 рассмотрены экологическая безопасность и техника безопасности работников на рабочем месте.

Практическая ценность разработанных моделей и алгоритмов состоит в возможности их использования для проведения оперативного мониторинга поврежденных вредителями хвойных насаждений, что важно для лесной отрасли РФ. Научную новизну имеют предложенные модели МН-U-Net и МН-Res-U-Net.

Результаты работы опубликованы в ряде сборников (одна статья – в периодике, индексируемой базой данных Scopus), доложены на восьми конференциях международного и всероссийского уровня и отмечены дипломами (приложение Б). Практическая значимость работы подтверждается актом об использовании результатов магистерской диссертации в лаборатории мониторинга лесных экосистем Института мониторинга климатических и экологических систем СО РАН (приложение В).

1 Мониторинг состояния хвойных деревьев

Наземные лесопатологические обследования хвойных лесов позволяют провести прямую оценку состояния деревьев и численности популяции вредителей, однако существенно ограничены по площади исследуемых территорий и маловозможны для насаждений, труднодоступных для наземного транспорта. Осуществление лесопатологических обследований дистанционным и комбинированным методами с использованием космических аппаратов и БПЛА с соответствующей съемочной аппаратурой значительно расширяет возможности мониторинга хвойных лесов. Важно как можно раньше выявить наличие деревьев на начальных стадиях повреждения вредителями, чтобы спланировать и провести фитосанитарные мероприятия. При этом использование именно БПЛА может обеспечить большую оперативность наблюдений, так как они менее зависят от наличия облачности и более мобильны, и возможность фотографирования кроны деревьев в высоком пространственном разрешении для их более детального анализа.

1.1 Задача анализа состояния пихты сибирской по снимкам с БПЛА

У исследователей есть обширный опыт наземного наблюдения за хвойными насаждениями, в частности в очагах распространения вредителей. Так для оперативного мониторинга поврежденных уссурийском полиграфом деревьев пихты Институтом мониторинга климатических и экологических систем СО РАН разработана оценочная шкала их жизненного состояния [1]. На рисунке 1.1 приведены шесть категорий состояния деревьев пихты сибирской (*Abies sibirica*) в очагах массового размножения уссурийского полиграфа в соответствии с этой шкалой. Первая из них – здоровое дерево, хвоя зеленая, вторая категория – дерево атаковано полиграфом, но не заселено, крона его как у здорового дерева или изреженная, несколько ветвей имеют на концах хвою ярко-рыжего цвета. К третьей категории относятся атакованные полиграфом, но не заселенные, сильно ослабленные деревья. Их кроны могут быть как у здоровых деревьев, но более половины ветвей несут усыхающую хвою. Четвертая категория – заселенные полиграфом усыхающие деревья, хвоя

которых в верхней части кроны еще зеленая, но ниже на ветках ярко-рыжего цвета. Пятая категория – свежий сухостой, хвоя кроны мертвая, красная. Наконец, шестая категория – старый сухостой, крона мертвая, серая, хвоя полностью осыпалась.

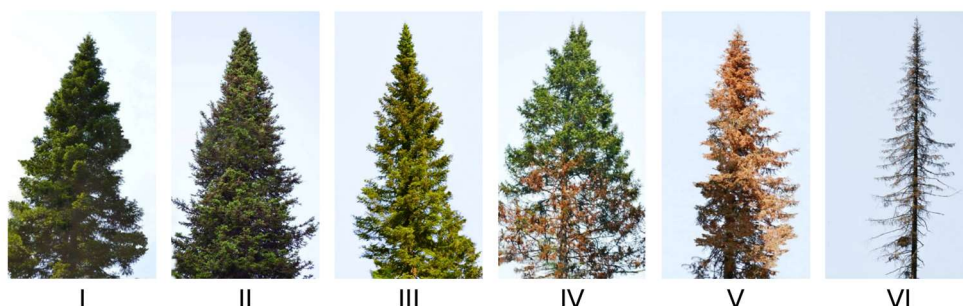


Рисунок 1.1 – Деревья пихты на различных стадиях повреждения полиграфом. Римскими цифрами обозначены категории деревьев [1]

1.2 Задача анализа состояния сосны сибирской кедровой по снимкам с БПЛА

Для анализа состояния кедр сибирского (*Pinus sibirica*) экспертом было выделено пять классов состояния дерева при глазомерной оценке:

– Класс I, «Условно здоровое»: дерево имеет ярко-зеленую крону без признаков дехромации или пожелтения, дерево условно здоровое;

– Класс II, «Свежезаселенное»: крона зеленая, но верхушечные ветви начинают приобретать бледный зелено-желтый оттенок. На этом этапе жуки под корой завершают свое 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 желто-бурый оттенок, нижняя часть кроны ярко- или бледно-зеленая. В весенний и осенний период жуки либо находятся на стадии зимовки в подстилке, либо весной после массового лета и летом осваивают нижний ярус ветвей, либо осваивают вершины соседних деревьев. Ствол этих деревьев может осваиваться сопутствующими видами;

– Класс IV, «Свежий сухостой»: вся крона буро-рыжего цвета 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 несет хвои или остатки неосыпавшейся рыже-бурой хвои на самых нижних ветках, союзный короед покинул дерево в предыдущем вегетационном сезоне, в древесине могут сохраняться личинки усачей.

На рисунке 1.2 представлены фотографии крон деревьев кедр различных классов с БПЛА.

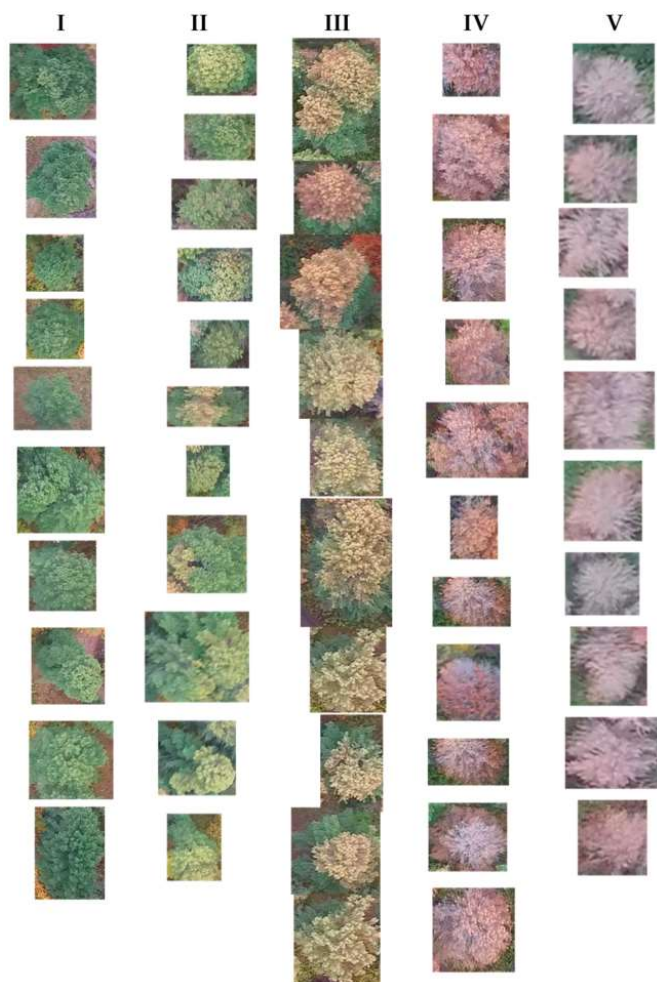


Рисунок 1.2 – Классы деревьев кедр, поврежденных союзным короедом

1.3 Методы анализа изображений дистанционного мониторинга

Для анализа изображений, полученных в результате съемки лесных массивов со спутников ресурсно-природного мониторинга или БПЛА, обычно используют различные традиционные методы классификации: , однако в таких случаях либо прибегают к использованию мультиспектральных данных, данных съемки LiDAR и пр. [10, 11], либо решают простые задачи обнаружения наличия нарушений леса по легко дешифрируемым признакам, например,

дехромации хвои [12].

В работе [13] предлагается классифицировать деревья различных пород с использованием методов глубокого обучения, однако эта работа также прибегает к данным LiDAR для построения цифровой модели поверхности с целью извлечения областей крон из изображения и их последующей классификации простыми сверточными нейронными сетями (СНС).

В работе [8] для классификации деревьев пихты также применены СНС, в [8] предлагается подход похожий на описанный в [13], но для выделения областей крон используется пороговая фильтрация изображения, что приводит к значительным ошибкам в разделении крон. При этом авторы никак не оценивают качество шага извлечения областей крон, что приводит к несоответствию заявленной точности модели с ее действительной точностью.

В данной работе для анализа этих снимков предлагается разработать, обучить и использовать модели машинного обучения, позволяющие решать задачу семантической сегментации обычных RGB изображений. В отличие от других моделей, используемых, например, в работе [8], эти модели позволяют выполнять попиксельную классификацию пораженных деревьев пихты за один прямой проход и, следовательно, более качественно производить разделение крон деревьев между собой и от фоновых объектов.

1.4 Постановка задач исследования

Целью данной работы является разработка моделей и алгоритмов для семантической сегментации хвойных деревьев, поврежденных насекомыми-вредителями, на снимках с беспилотных летательных аппаратов.

Для достижения цели необходимо:

- провести подготовку исходных данных;
- осуществить разведочный анализ данных;
- выбрать, спроектировать и реализовать модели машинного обучения;
- спроектировать и реализовать алгоритмы обучения и подбора гиперпараметров;

– исследовать и сравнить полученные модели.

В качестве исходных данных для анализа состояния деревьев пихты использованы два снимка участков пихтового леса в Томской области, заселенных уссурийским полиграфом, с пространственным разрешением около 0,1 м. Снимки получены с помощью БПЛА (квадрокоптер DJI Phantom 3 Standart) с установленной на нем фотокамерой, ведущей съемку в видимом диапазоне электромагнитного спектра (RGB) с высоты 365 и 388 м в период с 7 по 28 августа 2017 г.

Кроны деревьев на снимках были размечены экспертом с учетом категории деревьев. В силу того, что при разметке снимков визуальным экспертом не удавалось отделить некоторые категории деревьев друг от друга, было решено объединить близкие категории деревьев в один класс. При этом в первый класс попали деревья первой и второй категорий, во второй класс – третьей и четвертой. В результате выделены следующие четыре класса деревьев: 1 – «Живые», 2 – «Отмирающие» (сильно ослабленные, усыхающие), 3 – «Свежий сухостой» и 4 – «Старый сухостой» (сухостой прошлых лет). Пример такой разметки на трех фрагментах снимков приведен на рисунке 1.3а. Далее в соответствии с первичной точечной разметкой эксперта на снимках были выделены эталонные сегменты (полигоны), представляющие кроны деревьев пихты разных классов и фон. Примеры эталонных карт сегментации приведены на рисунке 1.3б.

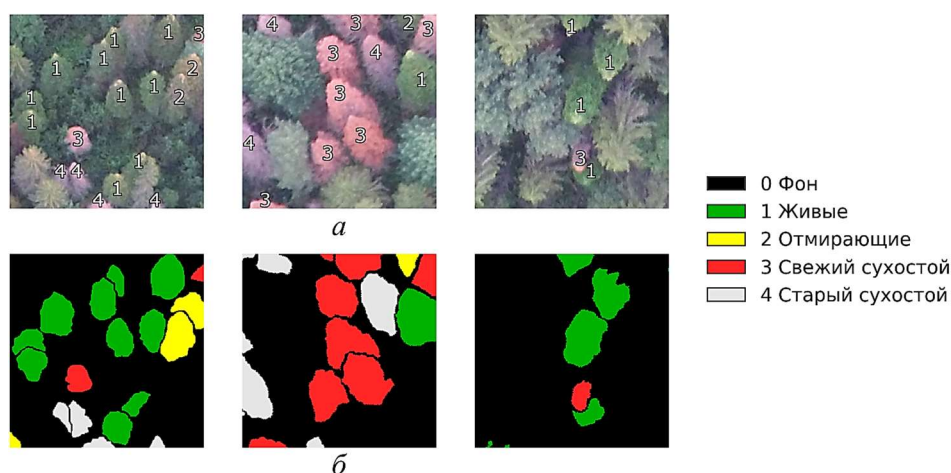


Рисунок 1.3 – Фрагменты исходных снимков пихтовых насаждений с разметкой по классам (а) и соответствующих им эталонных карт сегментации (б)

На одном из снимков был выбран тестовый участок. Он не участвовал в дальнейшем обучении моделей машинного обучения, а использовался исключительно для оценки эффективности предлагаемой модели. В итоге были сформированы обучающая и тестовая выборки – изображения и соответствующие им эталонные карты сегментации для разрабатываемых моделей.

В качестве исходных данных для задачи сегментации кедровых насаждений использовались снимки с того же БПЛА (квадрокоптер DJI Phantom 3 Standart), собранные в панорамы. Всего было получено четыре панорамы. Снимки для первой панорамы были сделаны 11 июля 2019 г. с высоты порядка 350 м, что обеспечило пространственное разрешение в примерно 10 см. Снимки для второй, третьей и четвертой панорамы были сделаны 23 июля 2019 г., высота порядка 150 м, что соответствует разрешению в около 3 см. Для нивелирования разницы в размерах деревьев в кадре из-за разной высоты съемки, панорамы были отмасштабированы. Четвертая панорама была использована в качестве тестовой.

Разметка данных проводилась аналогичным образом: эксперт ставил точечные метки на вершинах деревьев (рисунок 1.4а), которые в дальнейшем оконтуривались для получения масок сегментации (рисунок 1.4б).

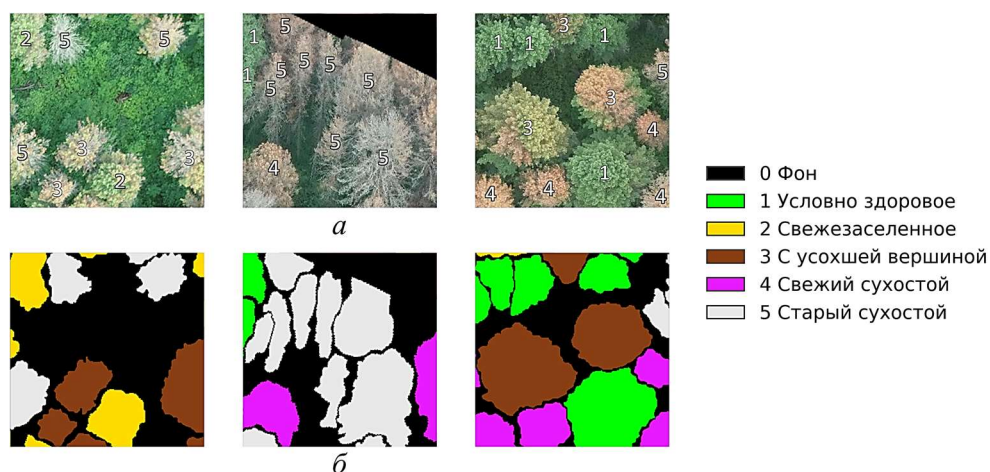


Рисунок 1.4 – Фрагменты исходных снимков с разметкой по классам (а) и соответствующих им эталонных карт сегментации (б)

Таким образом, был сформирован набор данных для решения двух поставленных задач классификации пикселей цифровых изображений с БПЛА:

– на четыре класса деревьев пихты («Живое», «Отмирающее», «Свежий сухостой» и «Старый сухостой») и класс «Фон»;

– на пять классов деревьев кедра («Условно здоровое», «Свежезаселенное», «С усохшей вершиной», «Свежий сухостой» и «Старый сухостой») и класс «Фон».

2 Разведочный анализ данных

Для разработки моделей машинного обучения, способных решать поставленные задачи, необходимо провести разведочный анализ исходных данных. Это позволит оценить сложность поставленных задач и, следовательно, необходимую сложность моделей машинного обучения, а также выявить существенные особенности данных, которые необходимо учитывать при обучении.

2.1 Разведочный анализ изображений пихтовых насаждений

На рисунке 2.1 представлены гистограммы, показывающие соотношения объектов различных классов в обучающей выборке для задачи сегментации крон пихты. Видно, что наблюдается значительная несбалансированность классов: преобладают пиксели класса «Фон», среди деревьев наименее представленными оказались деревья классов «Отмирающие» и «Свежий сухостой».

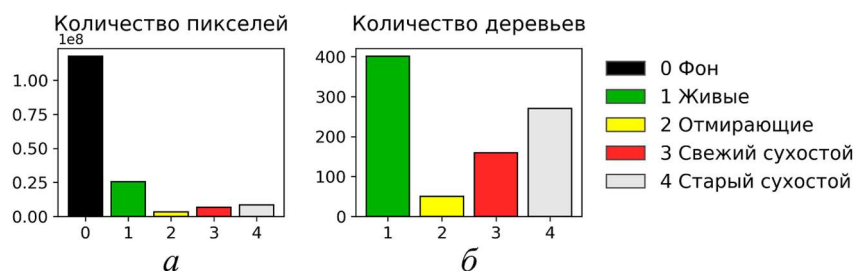


Рисунок 2.1 – Количество пикселей (а) и деревьев (б) пихты по классам в обучающей выборке

Для оценки разделимости классов был проведен анализ UMAP-вложений и коэффициентов силуэта (англ. silhouette coefficients), изображения которых приведены на рисунке 2.2. Поскольку стоит задача разделения крон деревьев пихты по классам, были рассмотрены спектральные характеристики сегментов, ограничивающих кроны. В качестве признаков сегментов использовались гистограммы яркостей в RGB каналах снимка: для каждого из трех каналов была построена гистограмма яркостей с 15 интервалами разбиения, таким образом вектор признаков для каждого сегмента представлял собой упорядоченный набор из 45 чисел.

UMAP – метод снижения размерности пространства признаков,

позволяющий анализировать характер распределения многомерных данных. Идея метода заключается в поиске такого представления данных в пространстве меньшей размерности, чтобы топологическая структура искомого представления была схожа с исходной [14]. При построении UMAP-вложений гистограмм сегментов (рисунки 2.2а и 2.2б) варьировались два параметра алгоритма: количество ближайших соседей в топологической структуре исходных данных (n_n) и минимальное расстояние между точками в результирующем представлении меньшей размерности (min_dist).

Коэффициенты силуэта позволяют оценивать разделимость классов, основываясь на вычислении относительных расстояний между векторами признаков. Коэффициент силуэта для конкретного экземпляра (сегмента, представляющего крону дерева) рассчитывается как разность среднего расстояния от этого экземпляра в пространстве признаков до всех экземпляров ближайшего соседнего класса и среднего расстояния от этого экземпляра до всех остальных экземпляров внутри своего класса, деленная на максимальное из этих расстояний. Близкие к единице значения коэффициента указывают на хорошо отделимые от ближайшего соседнего класса экземпляры, отрицательные – на наиболее вероятно относящиеся к ближайшему соседнему классу, близкие к нулю – на близость экземпляра к границе принятия решения [15]. Внутри каждого класса экземпляры сортируются по значению коэффициента силуэта и наносятся на диаграмму силуэтов, где вдоль вертикальной оси отображаются классы, а вдоль горизонтальной – значения коэффициентов силуэта для сегментов.

Анализируя локальные особенности распределения данных (рисунок 2.2а), можно сказать о наличии двух крупных кластеров, один из которых состоит преимущественно из точек, относящихся к классам «Живые» и «Отмирающие», а другой – из точек классов «Свежий сухостой» и «Старый сухостой», при этом сложно построить границы решений для разделения классов уже внутри этих кластеров. На рисунке 2.2б, показывающем глобальные особенности распределения, наблюдается переход между классами

в порядке улучшения состояния деревьев пихты. Анализ диаграммы силуэтов (рисунок 2.2в) позволяет сделать вывод о низкой разделимости классов в обучающей выборке в целом, так как среднее значение коэффициента силуэта составило 0,083. Значительной доле сегментов классов «Отмирающие», «Свежий сухостой» и «Старый сухостой» соответствуют отрицательные значения коэффициентов силуэта.

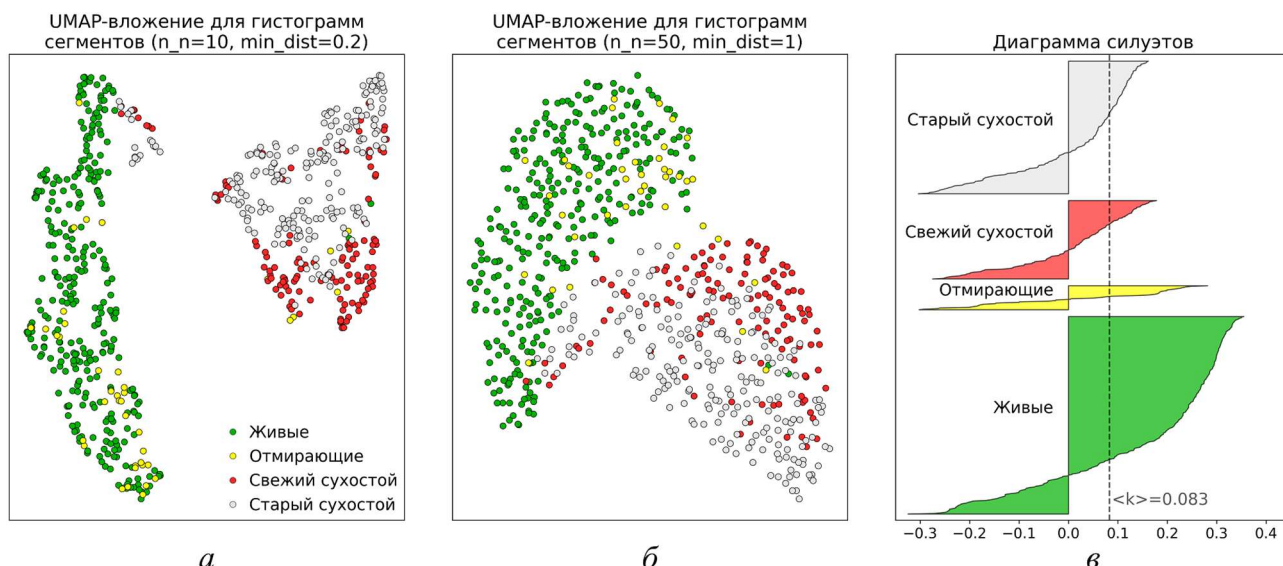


Рисунок 2.2 – УМАР-вложения для гистограмм (а и б) и диаграмма силуэтов для гистограмм (в) сегментов деревьев пихты, пунктирной линией показано среднее значение коэффициента силуэта

2.2 Разведочный анализ изображений кедровых насаждений

Аналогичный разведочный анализ данных был проведен для датасета для решения задачи сегментации кедровых деревьев.

На рисунке 2.3 представлены гистограммы, позволяющие оценить сбалансированность набора данных. Так, наблюдается похожая проблема: значительное преобладание пикселей класса «Фон» и слабая представленность таких классов деревьев, как «Свежезаселенное» и «Свежий сухостой».

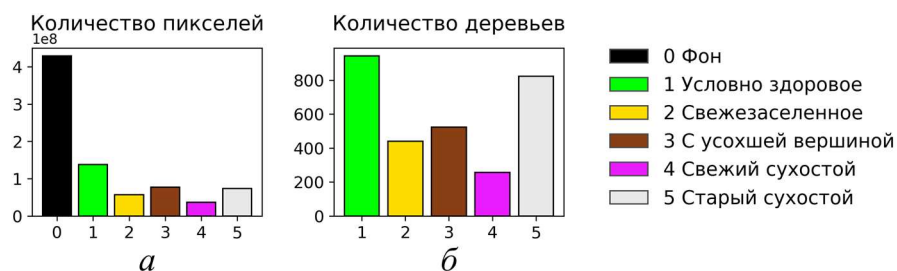


Рисунок 2.3 – Количество пикселей (а) и деревьев (б) кедра по классам в обучающей выборке

Анализ локальных особенностей распределения (рисунок 2.4а) показывает наличие двух различных кластеров: один преимущественно содержит сегменты крон деревьев классов «Условно здоровое» и «Свежезаселенное», другой – «Свежий сухостой» и «Старый сухостой». При этом сегменты класса «С усохшей вершиной» расположились в пространстве между выделенными кластерами. На UMAP-вложении, показывающем глобальные особенности (рисунок 2.4б), можно видеть непрерывный переход между классами деревьев в порядке ухудшения их состояния, как это наблюдалось и для сегментов деревьев пихты. Сегментам соответствует низкое среднее значение коэффициентов силуэта (0,039), при этом значительные доли отрицательных значений коэффициента наблюдаются для классов «Свежезаселенное» и «С усохшей вершиной» (рисунок 2.4в).

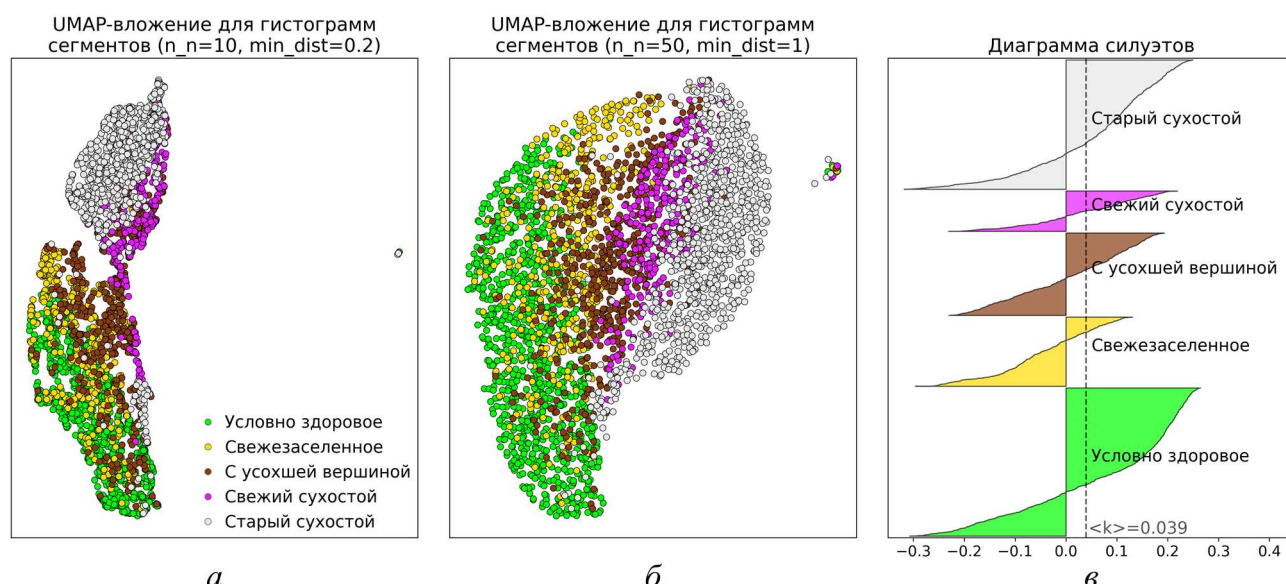


Рисунок 2.4 – UMAP-вложения для гистограмм (а и б) и диаграмма силуэтов для гистограмм (в) сегментов деревьев кедра, пунктирной линией показано среднее значение коэффициента силуэта

Таким образом, разведочный анализ исходных данных показал, что поставленные задачи сегментации не являются тривиальными, поскольку учет 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 усложняя задачу классификации пикселей. В связи с этим необходимо использовать достаточно

сложные модели, способные учитывать текстуры объектов и их формы, например, модели СНС, позволяющие решать задачи попиксельной классификации, или применять продвинутые техники извлечения признаков из изображений, например, текстурные признаки Харалика [16].

3 Реализованные модели и алгоритмы

Для решения двух поставленных задач сегментации изображений с БПЛА был рассмотрен ряд моделей машинного обучения: ансамблевая модель из деревьев принятия решений, использующая спектральные признаки и текстурные признаки Харалика и обученная с использованием градиентного бустинга, и полносверточные сети U-Net, SegNet и ENet, а также две модификации U-Net – МН-U-Net и МН-Res-U-Net.

3.1 Текстурные признаки Харалика

Одним из способов учета не только цветовых, но и текстурных характеристик пикселей вокруг них, является использование текстурных признаков Харалика [16].

Для вычисления текстурных признаков Харалика прежде всего необходимо получить матрицы совместной встречаемости яркостей (GLCM), считающие количества попарных вхождений яркостей пикселей в исходном изображении для определенного направления и расстояния. В данной работе использовались следующие четыре определения GLCM:

$$GLCM_{ij}(\theta_d, 0^\circ) = \left| \left\{ (k,l), (m,n) \in (L \times L) \times (L \times L) \mid k = m, |l - n| = \theta_d, I(k,l) = i, I(m,n) = j \right\} \right|, \quad (3.1)$$

$$GLCM_{ij}(\theta_d, 45^\circ) = \left| \left\{ \begin{array}{l} (k,l), (m,n) \in (L \times L) \times (L \times L) \mid \\ (k - m = \theta_d, l - n = -\theta_d) \text{ или } (k - m = -\theta_d, l - n = \theta_d), \\ I(k,l) = i, I(m,n) = j \end{array} \right\} \right|, \quad (3.2)$$

$$GLCM_{ij}(\theta_d, 90^\circ) = \left| \left\{ (k,l), (m,n) \in (L \times L) \times (L \times L) \mid k - m = \theta_d, l = n, I(k,l) = i, I(m,n) = j \right\} \right|, \quad (3.3)$$

$$GLCM_{ij}(\theta_d, 135^\circ) = \left| \left\{ \begin{array}{l} (k,l), (m,n) \in (L \times L) \times (L \times L) \mid \\ (k - m = \theta_d, l - n = \theta_d) \text{ или } (k - m = -\theta_d, l - n = -\theta_d), \\ I(k,l) = i, I(m,n) = j \end{array} \right\} \right|, \quad (3.4)$$

где $L = \{1, 2, \dots, \theta_{ps}\}$, θ_{ps} – размер окрестности пикселя, θ_d – расстояние между двумя пикселями.

Полученные матрицы могут быть использованы для вычисления дескрипторов изображений или определенных областей изображения. В данной работе использовалось шесть дескрипторов, реализованных в библиотеке

scikit-image [17]:

$$\text{Contrast} = \sum_{i,j}^{255} \text{GLCM}_{ij} (i-j)^2, \quad (3.5)$$

$$\text{Dissimilarity} = \sum_{i,j}^{255} \text{GLCM}_{ij} |i-j|, \quad (3.6)$$

$$\text{Homogeneity} = \sum_{i,j}^{255} \frac{\text{GLCM}_{ij}}{1+(i-j)^2}, \quad (3.7)$$

$$\text{ASM} = \sum_{i,j}^{255} \text{GLCM}_{ij}^2, \quad (3.8)$$

$$\text{Energy} = \sqrt{\text{ASM}}, \quad (3.9)$$

$$\text{Correlation} = \sum_{i,j}^{255} \text{GLCM}_{ij} \frac{(1-\mu_i)(1-\mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \quad (3.10)$$

При этом дескрипторы рассчитывались для каждой матрицы (3.1–3.4) и для всех трех цветовых каналов изображений в отдельности.

3.2 Градиентный бустинг

Для классификации пикселей на основе их спектральных характеристик и текстурных характеристик их окрестности использовался алгоритм градиентного бустинга. Градиентный бустинг – техника построения ансамблей простых моделей, в котором каждая последующая модель пытается скорректировать ошибки всех предыдущих [18]. Обычно в качестве простых моделей используются деревья принятия решений.

Для обучения ансамбля деревьев принятия решений из изображений обучающего набора извлекались пиксели с шагом 5, в качестве признаков брались спектральные характеристики этих пикселей и текстурные признаки Харалика, рассчитанные в их окрестности. Использовалась аугментация данных для генерирования синтезированных обучающих примеров путем добавления гауссова шума к окрестностям пикселя, при этом больше синтезированных примеров добавлялось наименее представленным в наборе данных классам. Во время обучения ансамблей минимизировалась перекрестная энтропия.

3.3 Полносверточные сети

Полносверточные сети образуют один из классов СНС. Типичная полносверточная сеть состоит из энкодера и декодера. Назначением энкодера является преобразование изображения в его скрытое представление с использованием слоев свертки и субдискретизации или уменьшения масштаба. Декодер преобразует полученное скрытое представление в итоговую карту сегментации, таким образом, позволяя классифицировать целые изображения или его фрагменты, а не отдельные пиксели. Обычно декодер включает в себя слои свертки и увеличения масштаба или обратной свертки. Чаще всего последним слоем полносверточной сети является сверточный слой с необходимым в зависимости от количества классов или других моделируемых переменных количеством фильтров размером 1 на 1 [19].

Следующие подразделы рассматривают пять архитектур полносверточных сетей: U-Net, SegNet, ENet, MH-U-Net и MH-Res-U-Net.

3.3.1 Архитектура U-Net

В данной работе одной из выбранных моделей для решения задачи семантической сегментации стала модель U-Net. Исходно она была спроектирована и разработана для решения задачи сегментации биомедицинских изображений, где продемонстрировала высокое качество результатов. U-Net до сих пор считается одной из лучших архитектур при сегментации биомедицинских изображений. Отличительной чертой архитектуры U-Net является наличие соединений проброса, соединяющих наборы карт признаков из энкодера с наборами карт признаков в декодере, в целях повышения детальности результирующей карты сегментации [20].

На рисунке 3.1 приведена архитектура модели, прямоугольники соответствуют тензорам, или многомерным массивам, представляющим наборы карт признаков, числами указаны размеры соответствующих тензоров, стрелками обозначены следующие операции:

– свертки (Conv3x3, Conv1x1);

- вычисление ELU;
- пакетная нормализация (BN) [21];
- субдискретизация путем выбора максимального значения в окрестности 2 на 2 (MaxPooling);
- увеличение масштаба методом ближайшего соседа (UpSampling);
- копирование тензора и его конкатенация с другим (Copying+Concatenation);
- исключение случайных сигналов слоя путем их приравнивания к нулю (Dropout).

Категориальное распределение на выходе декодера моделируется для каждого пикселя путем применение обобщения логистической функции для многомерного случая, Softmax.

По сравнению с исходной моделью U-Net, в используемую в данной работе архитектуру были добавлены следующие изменения:

- входное изображение сети представлено тензором $256 \times 256 \times 3$, соответствующим фрагменту RGB изображения;
- свертки не уменьшают размер карт признаков;
- обрезка карт признаков не используется для соединений проброса;
- пакетная нормализация применяется после каждой нелинейности;
- выходной тензор вычисляется C свертками с фильтрами размером 1 на 1, таким образом позволяя классифицировать пиксели сразу C классов.

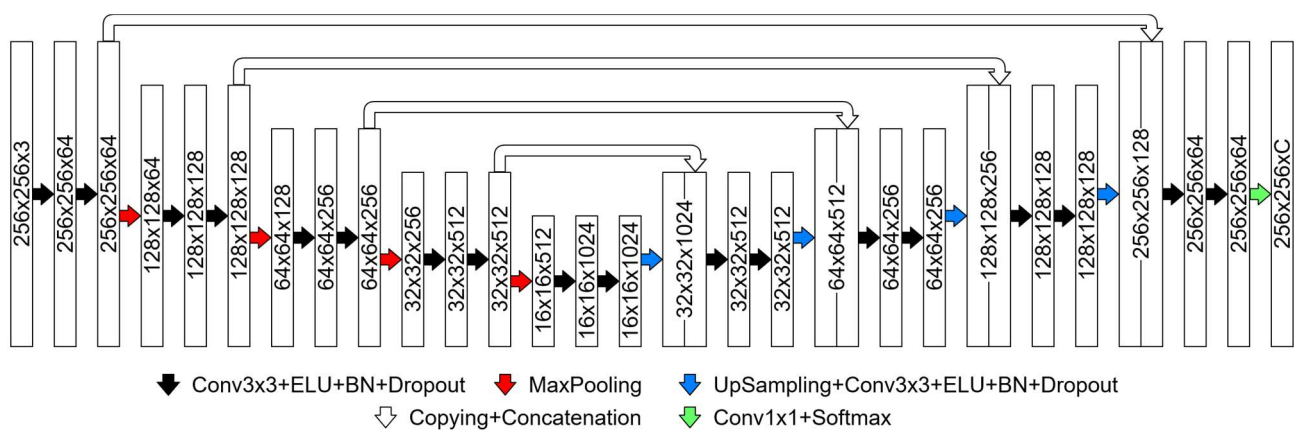


Рисунок 3.1 – Архитектура U-Net

3.3.2 Архитектура SegNet

Другой архитектурой полносверточной сети, используемой в данной работе, стала архитектура SegNet [22]. В нее были внесены аналогичные изменения, как это было сделано для U-Net. Также рассматривался менее глубокий вариант модели, чем в исходной статье. Архитектура сети SegNet, реализованная в данной работе, представлена на рисунке 3.2.

Принципиальным отличием архитектуры SegNet от архитектуры U-Net является способ восстановления исходного разрешения карты сегментации в декодере: вместо операций увеличения масштаба и полного копирования карт признаков, SegNet использует операции MaxUnPooling, выполняемые с использованием сохраненных позиций пикселей, выбранных операциями MaxPooling в энкодере.

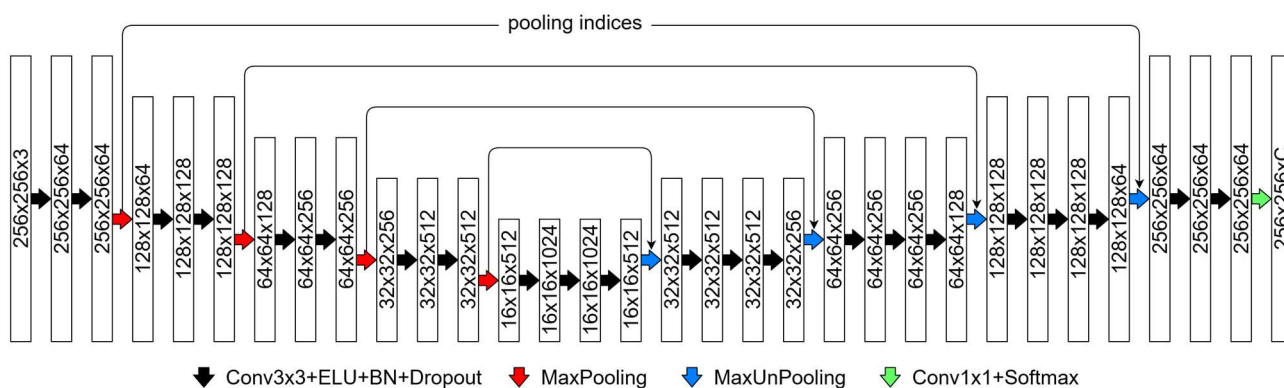


Рисунок 3.2 – Архитектура SegNet

3.3.3 Архитектура ENet

Также в работе использовалась полносверточная сеть ENet. В ее архитектуру не вносились никакие изменения, за исключением изменения размеров тензоров. Нейронная сеть была реализована согласно ее описанию, приведенному в [23].

Архитектура ENet представлена в таблице 3.1, ее блоки – на рисунке 3.3.

Таблица 3.1 – Архитектура ENet. Адаптировано из [23]

Секция	Название	Тип	Размер выхода
–	initial		128x128x16
1	bottleneck1.0	downsampling	64x64x64
	bottleneck1.1		64x64x64
	bottleneck1.2		64x64x64

	bottleneck1.3 bottleneck1.4		64x64x64 64x64x64
2	bottleneck2.0 bottleneck2.1 bottleneck2.2 bottleneck2.3 bottleneck2.4 bottleneck2.5 bottleneck2.6 bottleneck2.7 bottleneck2.8	downsampling dilated 2 asymmetric 5 dilated 4 dilated 8 asymmetric 5 dilated 10	32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128
3	Bottleneck3.1 bottleneck3.2 bottleneck3.3 bottleneck3.4 bottleneck3.5 bottleneck3.6 bottleneck3.7 bottleneck3.8	dilated 2 asymmetric 5 dilated 4 dilated 8 asymmetric 5 dilated 10	32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128
4	bottleneck4.0 bottleneck4.1 bottleneck4.2	upsampling	64x64x64 64x64x64 64x64x64
5	bottleneck5.0 bottleneck5.1	upsampling	128x128x16 128x128x16
–	deconvolution		256x256xC

Можно отметить следующие важнейшие особенности ENet:

- относительно неглубокий декодер в сравнении с энкодером;
- значительное уменьшение масштаба карт признаков на самых первых слоях сети;
- активное использование остаточных блоков [24];
- активное использование функции активации PReLU, позволяющей иметь обучаемые нелинейности [25];
- использование различных видов сверточных слоев (обычные, дилатационные/atrous и асимметричные) с целью увеличения области восприятия и уменьшения количества параметров сети.

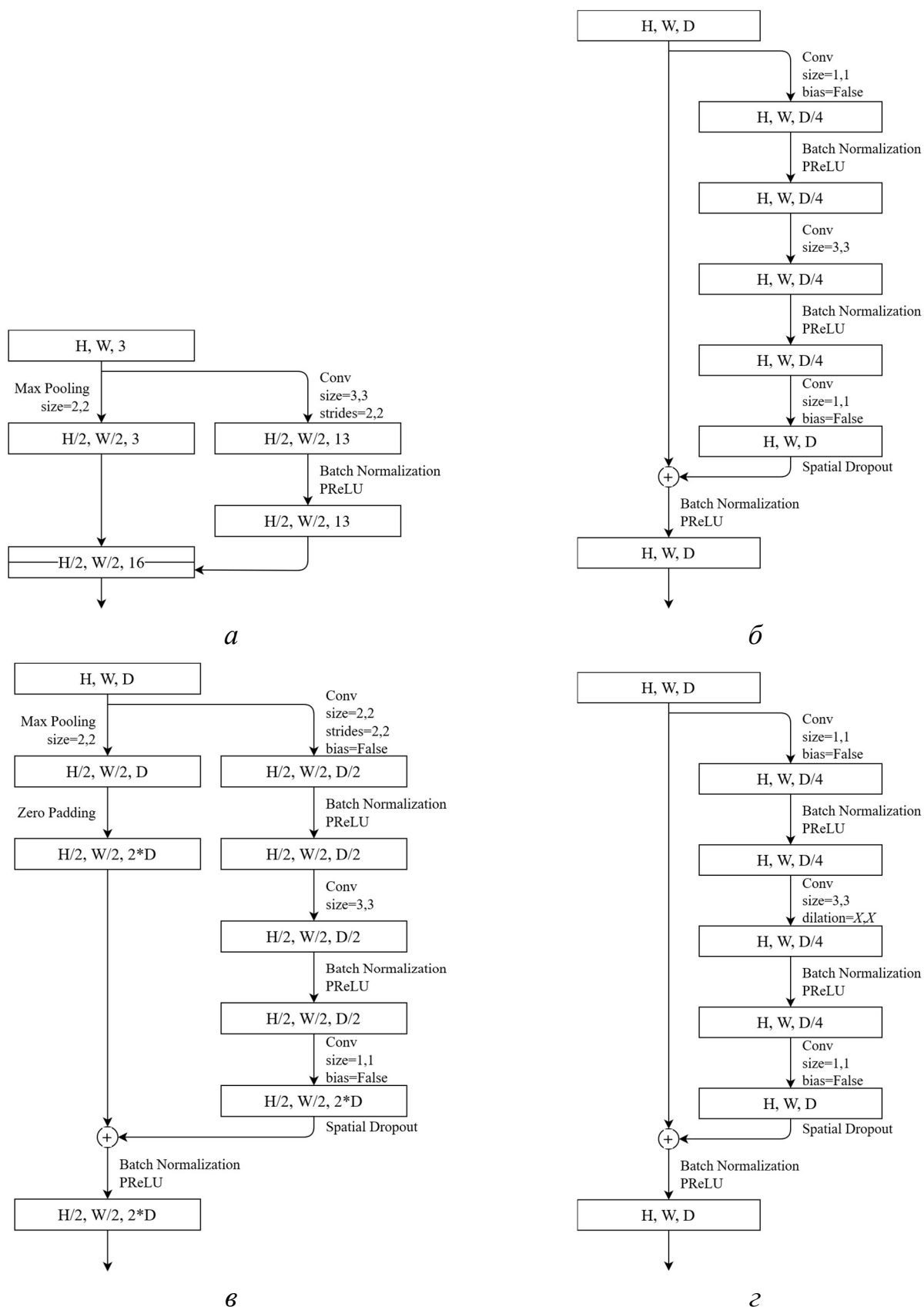


Рисунок 3.3 – Блоки ENet: а – initial, б – bottleneck, в – bottleneck downsampling, г – bottleneck dilated X , д – bottleneck asymmetric X , е – bottleneck upsampling. H , W и D обозначают соответственно высоту, ширину и количество карт признаков входного тензора. Адаптировано из [23]

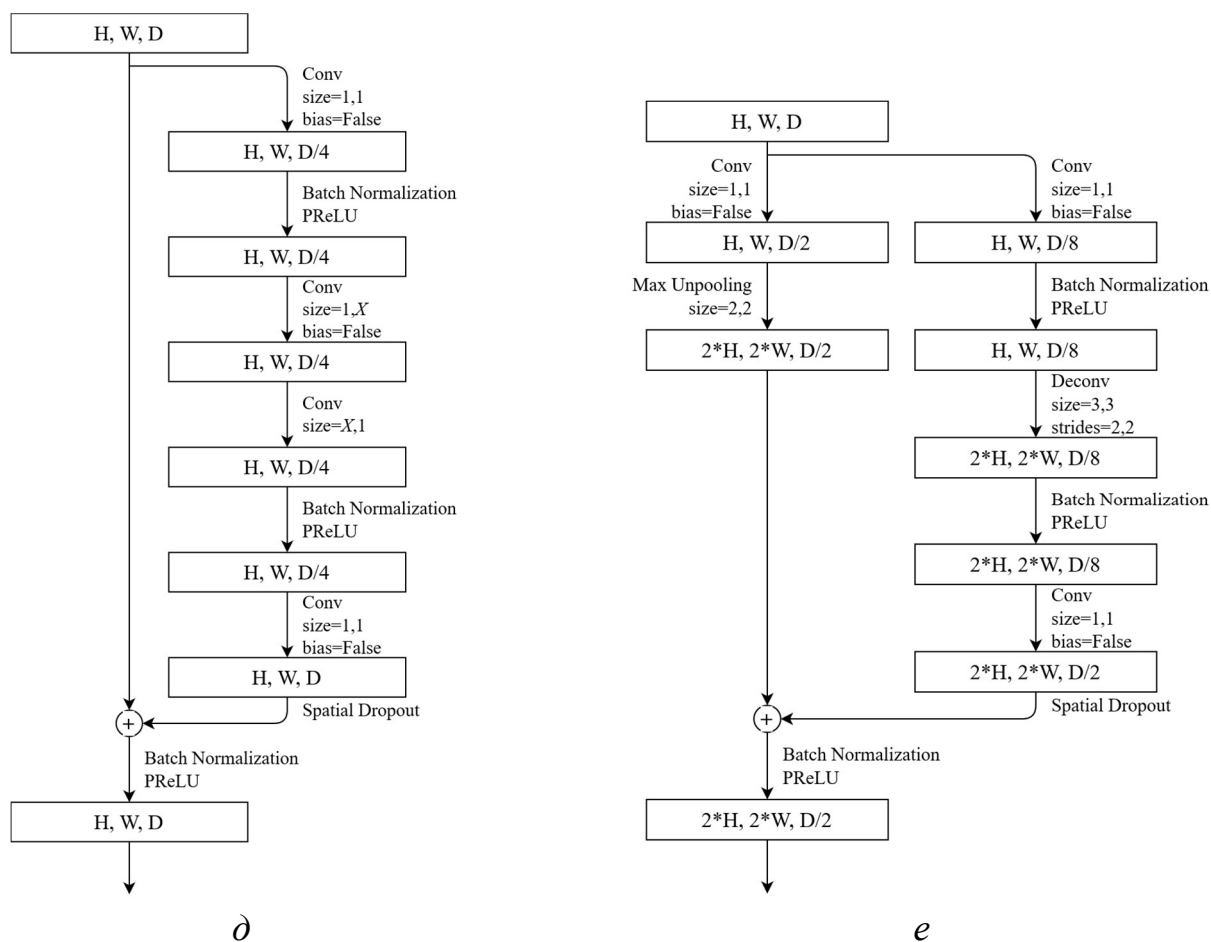


Рисунок 3.3 (продолжение)

Учитывая обозначенные особенности ENet, она позволяет осуществлять сегментацию изображений в реальном времени, что полезно при использовании на мобильных устройствах или при реализации систем компьютерного зрения на, например, программируемых логических интегральных схемах.

3.3.4 Архитектура MH-U-Net

В данной магистерской диссертации предлагаются две новые модели полносверточных сетей. Одна из них – multihead-U-Net (MH-U-Net). Данная модель основывается на архитектуре U-Net, но вводит следующие существенные изменения в ее топологию:

- вместо одного, используется три входа разных масштабов (256x256x3, 128x128x3 и 64x64x3);
- также имеется три выхода различных масштабов;
- дропаут заменен на пространственный дропаут, как в ENet.

На рисунке 3.4 показана архитектура предлагаемой нейронной сети.

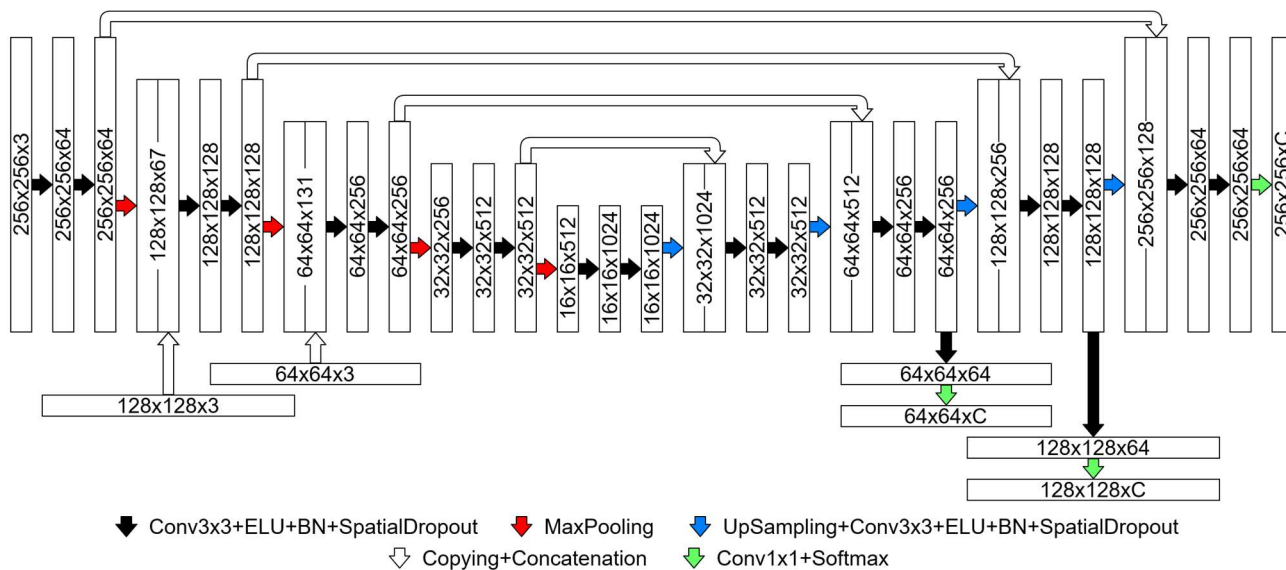


Рисунок 3.4 – Архитектура MH-U-Net

Во время обучения функция потерь рассчитывается на всех трех выходах сети, и алгоритм обучения минимизирует сумму всех трех значений. Во время использования сети выходы меньшего масштаба и соответствующие ветви вычислительного графа могут быть игнорированы.

Такая модель может рассматриваться как ансамбль из трех моделей U-Net различной глубины, где модели ансамбля делят между собой часть весов и пытаются одновременно анализировать изображение в различных масштабах. Таким образом, ожидается, что MH-U-Net способна обучаться более качественному скрытому представлению данных и это приводит к стабильно более качественным результатам сегментации. Данный вопрос требует соответствующих исследований, которые приведены далее в работе.

3.3.5 Архитектура MH-Res-U-Net

Идея архитектуры MH-U-Net была развита далее в архитектуру multihead-residual-U-Net (MH-Res-U-Net). MH-Res-U-Net обладает всеми особенностями MH-U-Net и дополнительно использует остаточные блоки [24].

Архитектура модели MH-Res-U-Net приведена на рисунке 3.5.

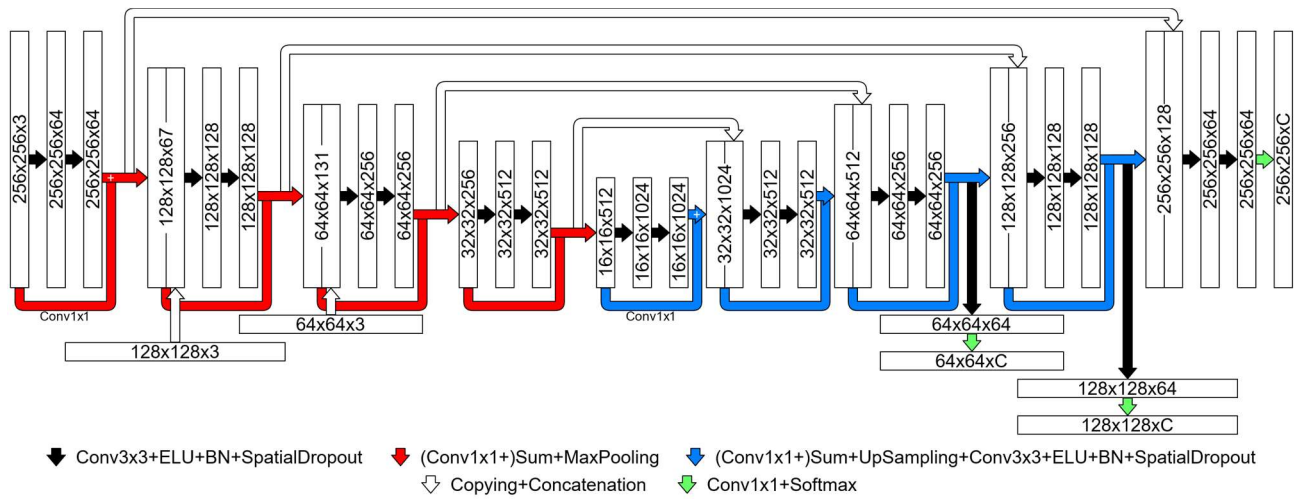


Рисунок 3.5 – Архитектура MH-Res-U-Net

Важно отметить, что операции проецирования (Conv1x1) в главных ветвях остаточных блоков не включают нелинейностей и смещений. Ожидается, что остаточные блоки в MH-Res-U-Net могут привести к более стабильному процессу обучения и упростить обратное распространение градиента.

Результаты анализа архитектур MH-U-Net и MH-Res-U-Net и их сравнение с другими архитектурами полносверточных сетей приведены в разделе 4.

3.4 Оценка качества сегментации

Для оценки качества предлагаемых моделей использовалась метрика intersection over union (IoU), или коэффициент Жаккара, считающаяся широко используемой при решении задач семантической сегментации. Метрика IoU может быть вычислена как [26, 27]:

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c}, \quad (3.11)$$

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

Значения IoU выше, чем 0,5, соответствуют приемлемо высокому качеству сегментации. Совокупным показателем качества моделей послужила усредненная метрика IoU (mean IoU, mIoU), или усредненный коэффициент Жаккара:

$$mIoU = \frac{1}{C} \sum_{c=1}^C IoU_c, \quad (3.12)$$

где C – количество классов в задаче сегментации.

3.5 Обучение нейронных сетей

Для уменьшения разницы в условиях освещенности различных снимков пихтовых насаждений была проведена процедура выравнивания гистограмм, при этом один из снимков из обучающей выборки был выбран в качестве опорного изображения. Для панорам кедровых насаждений процедура выравнивания гистограмм проведена не была из-за обилия черных пикселей на границах изображений.

Для формирования обучающей и валидационной выборок изображения были нарезаны на фрагменты размером 256 на 256 пикселей со сдвигом в 128 пикселей. Утечка данных из валидационной выборки устранялась путем наложения булевых масок на исходные изображения и блокирования возможности попадания соответствующих пикселей в обучающую выборку. Таким образом гарантировалось то, что признаки валидационных фрагментов не дублировались среди обучающих. Всего было получено 2004 обучающих и 672 валидационных фрагментов для задачи сегментации деревьев пихты и 12384 обучающих и 2400 валидационных фрагментов для задачи сегментации деревьев кедра.

В целях улучшения обобщающей способности обучаемых моделей полносверточных сетей использовалась техника аугментации «на лету», позволяющая расширять обучающую выборку синтезированными примерами непосредственно во время обучения. Аугментация «на лету» включала случайные:

- изменения масштаба изображений;
- изменения их яркости;
- изменения их контраста;
- отражения по вертикальной оси;

– эластические преобразования [28].

В данной работе для поиска оптимальных весовых коэффициентов полносверточных сетей максимизировалось значение mIoU с использованием алгоритма оптимизации Adam, основанного на вычислении градиента [29]. Оптимизация mIoU позволяет частично решить проблему несбалансированности обучающей выборки, обозначенную ранее. Чтобы свести задачу максимизации к задаче минимизации, можно решать следующую оптимизационную задачу:

$$1 - \text{mIoU} \rightarrow \min. \quad (3.13)$$

Во время обучения неудобно и вычислительно затратно вычислять матрицы ошибок для получения значений TP_c , FP_c и FN_c , поэтому вместо выражений (3.11) и (3.12) можно использовать аппроксимацию mIoU, или мягкую формулировку коэффициента Жаккара [26, 27], с дополнительными параметрами:

$$\text{mIoU}(\mathbf{T}, \mathbf{P}) = \frac{1}{C} \sum_{c=1}^C \left(\frac{\sum_{i=1}^H \sum_{j=1}^W \text{LS}(\mathbf{T})_{ijc} \cdot \mathbf{P}_{ijc} + \theta_s}{\sum_{i=1}^H \sum_{j=1}^W [\text{LS}(\mathbf{T})_{ijc} + (1 - \text{LS}(\mathbf{T})_{ijc}) \cdot \mathbf{P}_{ijc}] + \theta_s} \right), \quad (3.14)$$

$$\text{LS}(\mathbf{T}) = (1 - \theta_{ls}) \cdot \mathbf{T} + \frac{\theta_{ls}}{C}, \quad (3.15)$$

где \mathbf{T} – ожидаемый выход модели, \mathbf{P} – действительный выход полносверточной сети, H – высота изображения, W – ширина изображения, $\theta_s > 0$ – параметр сглаживания целевой функции, устраняющую точку сингулярности и большие по модулю значения градиента, $0 \leq \theta_{ls} < 1$ – коэффициент сглаживания меток, приводящий к менее уверенным предсказаниям сети, следовательно, помогающий избавиться от эффекта насыщения в последнем слое сети и сохраняющий весовые коэффициенты полносверточной сети более волатильными на большее количество эпох.

Важно отметить, что производная выражения (3.14) по \mathbf{P} может быть легко вычислена с использованием практически любого современного

фреймворка машинного обучения, что позволяет использовать методы оптимизации, основанные на вычислении градиента целевой функции, для поиска оптимальных и близких к оптимальным весовых коэффициентов.

3.6 Подбор гиперпараметров

Возможно добиться существенного улучшения качества моделей, если выявить множество гиперпараметров и решить задачу подбора оптимальных гиперпараметров или следующую задачу оптимизации:

$$mIoU(\theta) \rightarrow \max_{\theta} \quad (3.16)$$

Выявленные гиперпараметры для полносверточных сетей, градиентного бустинга с извлечением текстурных признаков Харалика и соответствующих алгоритмов аугментации данных и обучения перечислены в таблице 3.2.

Таблица 3.2 – Список гиперпараметров

Обозначение	Описание
Градиентный бустинг + текстурные признаки Харалика	
θ_{ne}	Максимальное количество деревьев в ансамбле
θ_{md}	Максимальная глубина дерева решений
θ_{lr}	Скорость обучения
θ_{ps}	Размер окна для расчета признаков
θ_d	Расстояние между двумя пикселями
θ_n	Стандартное отклонение гауссова шума для аугментации
Полносверточные сети	
θ_{lr}	Скорость обучения
θ_d	Коэффициент дропаута
θ_s	Коэффициент сглаживания функции потерь
θ_{ls}	Коэффициент сглаживания меток
θ_z	Коэффициент изменения масштаба
θ_b	Коэффициент изменения яркости
θ_c	Коэффициент изменения контраста
θ_{α}	Коэффициенты эластических преобразований [28]
θ_{σ}	
θ_{pf}	Уменьшение скорости обучения при выходе на «плато»
θ_{pp}	Количество эпох для подтверждения выхода на «плато»

Для поиска лучших наборов гиперпараметров для предложенных моделей использовался байесовский подход [31]. Идея алгоритма заключается в том, чтобы заменить действительную зависимость $mIoU(\theta)$, которая является

черным ящиком и для которой неизвестны выражения ни самой функции, ни ее производных, суррогатной моделью

$$\widehat{\text{mIoU}}(\theta) = \text{mIoU}(\theta) + \varepsilon, \varepsilon \sim N(0, s^2) \quad (3.17)$$

и решать задачу оптимизации (3.16) для этой суррогатной модели. В данной работе зависимость $\text{mIoU}(\theta)$ моделировалась гауссовским процессом, так как гауссовские процессы:

- являются непараметрическим методом регрессии, следовательно, позволяют воспроизводить зависимости сложного вида;
- позволяют проводить интервальную оценку моделируемого значения, что может быть использовано для балансирования степеней диверсификации и интенсификации.

На рисунке 3.6 приведена UML-диаграмма алгоритма подбора гиперпараметров.

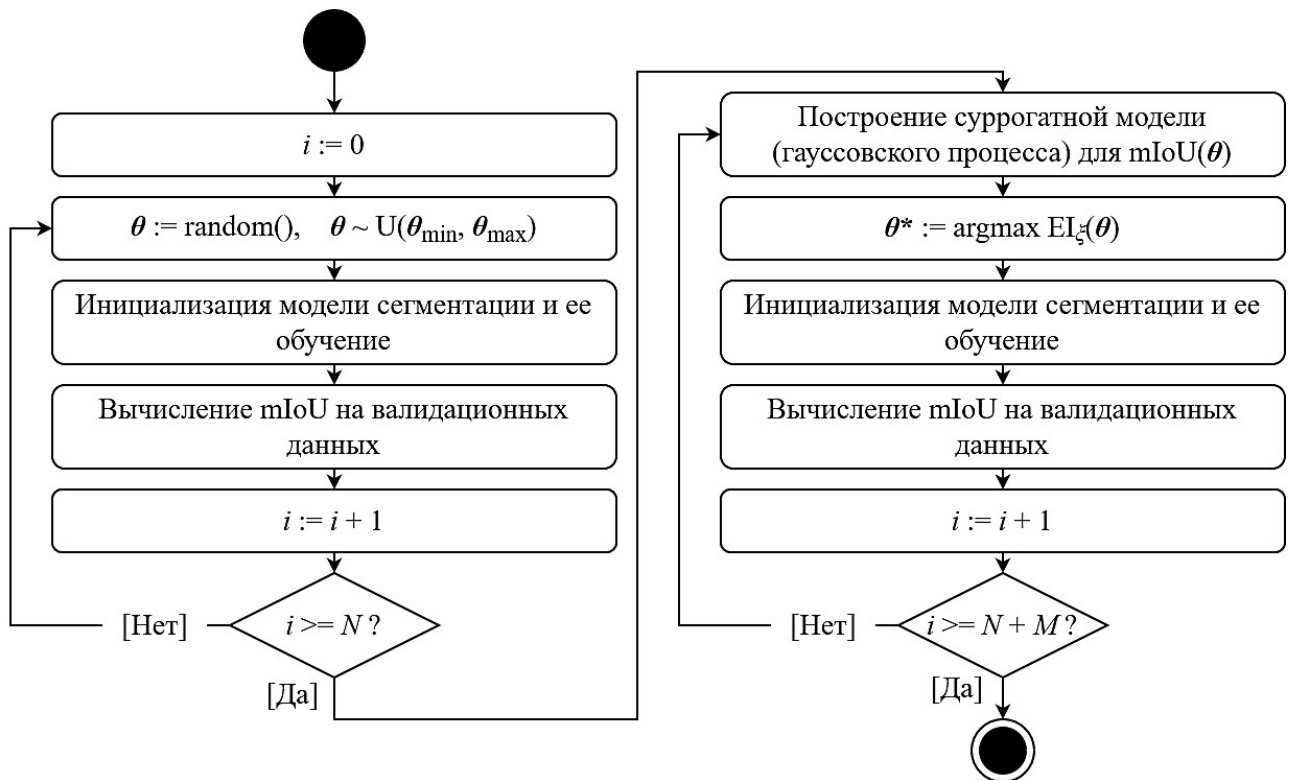


Рисунок 3.6 – Алгоритм подбора гиперпараметров

Алгоритм состоит из двух стадий. На первой стадии происходит выбор случайных N наборов гиперпараметров из заранее определенного равномерного распределения, инициализация соответствующих моделей, их обучение на

данных и оценка, т.е. вычисление mIoU на валидационной выборке. На второй стадии эти N пар из векторов гиперпараметров и соответствующих значений mIoU используются для построения суррогатной модели (3.17). Далее осуществляется максимизация функции выбора. В данной работе в качестве функции выбора использовалась функция expected improvement (EI) с введенным параметром ξ [30, 31]:

$$EI_{\xi}(\boldsymbol{\theta}) = E_{\xi}[\max(\widehat{\text{mIoU}}(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+), 0)] \quad (3.18)$$

или

$$EI_{\xi}(\boldsymbol{\theta}) = \begin{cases} \left(\mu(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+) - \xi \right) \Phi(Z) + \sigma(\boldsymbol{\theta}) \phi(Z) & , \text{ если } \sigma(\boldsymbol{\theta}) > 0 \\ 0 & , \text{ если } \sigma(\boldsymbol{\theta}) = 0 \end{cases}, \quad (3.19)$$

$$Z = \begin{cases} \frac{\mu(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+) - \xi}{\sigma(\boldsymbol{\theta})} & , \text{ если } \sigma(\boldsymbol{\theta}) > 0 \\ 0 & , \text{ если } \sigma(\boldsymbol{\theta}) = 0 \end{cases}, \quad (3.20)$$

где $\boldsymbol{\theta}^+$ – лучший из всех рассмотренных наборов гиперпараметров, $\mu(\boldsymbol{\theta})$ – среднее, $\sigma(\boldsymbol{\theta})$ – стандартное отклонение апостериорного распределения суррогатной модели в точке $\boldsymbol{\theta}$, Φ – интегральная функция распределения и ϕ – функция плотности вероятности стандартного нормального распределения. Параметр ξ позволяет регулировать степени интенсификации и диверсификации: более высокие значения ξ приводят к более частому сэмплингованию векторов гиперпараметров из областей с высокой неопределенностью. Чтобы получить новый потенциально лучший набор гиперпараметров, необходимо решить следующую оптимизационную задачу:

$$\boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta}} EI_{\xi}(\boldsymbol{\theta}). \quad (3.21)$$

Алгоритм решает (3.21) M раз и сохраняет результаты для обновления суррогатной модели на каждой итерации. В данной работе использовались значения $N = M = 30$ для ансамбля деревьев решений, обучаемого с использованием градиентного бустинга, $N = M = 40$ для полносверточных сетей и $\xi = 0.15$.

4 Результаты исследования предлагаемых моделей

4.1 Результаты сегментации тестовых участков

Обученные модели машинного обучения были применены для семантической сегментации изображения тестового участка пихтового насаждения (рисунок 4.1). Совместный визуальный анализ изображения тестового участка (рисунок 4.1а), эталонной карты сегментации (рисунок 4.1б) и результирующих выходных карт сегментации (рисунок 4.1в–з) для разных моделей показал, что все модели машинного обучения, предложенные в работе, способны воспроизводить границы между деревьями и успешно классифицировать значительную часть крон деревьев пихты. При этом важно заметить, что ансамблевая модель, полученная с использованием градиентного бустинга, не способна сегментировать кроны деревьев класса «Отмирающие».

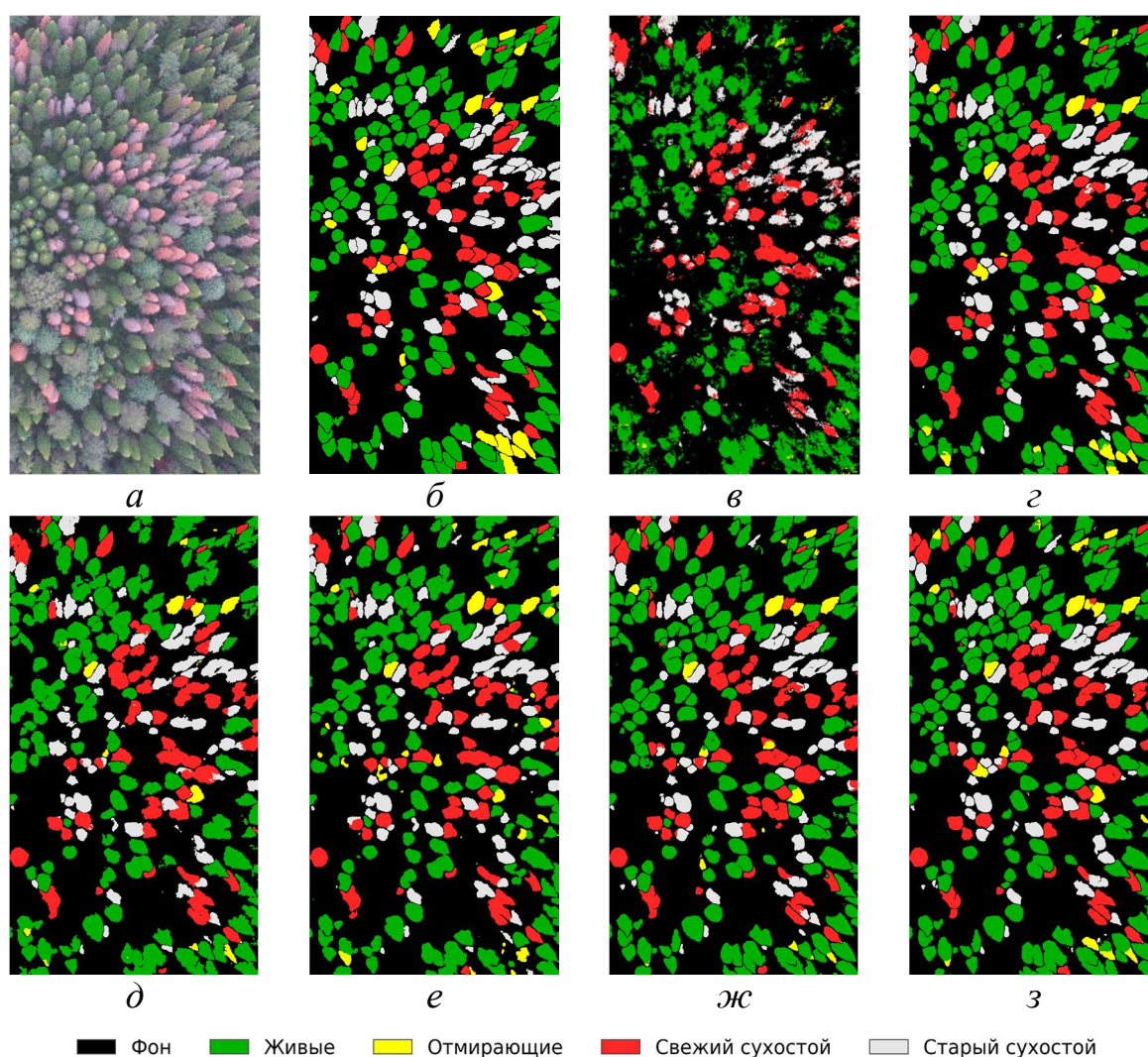


Рисунок 4.1 – Результаты сегментации деревьев пихты на изображении тестового участка: а – тестовый участок, б – эталонная карта сегментации,

в – результат сегментации ансамблевой моделью, обученной с использованием градиентного бустинга, г – моделью U-Net, д – моделью SegNet, е – моделью ENet, ж – моделью МН-U-Net, з – моделью МН-Res-U-Net

В таблице 4.1 приведены полученные значения IoU для различных моделей и классов, а также значения mIoU для оценки совокупного качества моделей машинного обучения при сегментации изображения тестового участка (столбец test) и на валидационной выборке (столбец val).

Таблица 4.1 – Метрики качества сегментации изображений пихтовых насаждений. Полужирным выделены лучшие значения метрик в столбце

Модель	IoU					mIoU	
	Фон	Живые	Отмирающие	Свежий сухостой	Старый сухостой	val	test
Градиентный бустинг	0,76	0,53	0,05	0,62	0,50	–	0,49
U-Net	0,86	0,76	0,46	0,76	0,63	0,65	0,69
SegNet	0,85	0,73	0,35	0,76	0,64	0,64	0,66
ENet	0,83	0,71	0,31	0,77	0,66	0,60	0,66
МН-U-Net	0,87	0,77	0,36	0,71	0,58	0,67	0,66
МН-Res-U-Net	0,87	0,77	0,48	0,75	0,65	0,66	0,71

Из таблицы 4.1 следует, что лучшей оказалась предложенная в данной работе модель полносверточной сети МН-Res-U-Net, однако на валидационных данных лучшее качество продемонстрировала МН-U-Net. При этом все модели показывают минимальную эффективность при классификации пикселей класса «Отмирающие». Невысокие результаты для этого класса объясняются малым количеством деревьев этого класса в обучающей выборке. Значения метрики IoU для классов «Фон», «Живые», «Свежий сухостой» и «Старый сухостой» превышают 0,5 для всех моделей, что говорит о высоком качестве сегментации снимков деревьев пихты этих классов и фона этими моделями.

Для сегментации деревьев кедра на снимках были выбраны модели U-Net, МН-U-Net и МН-Res-U-Net, как показавшие лучшие результаты на валидационной выборке в задаче сегментации пихты сибирской. На рисунке 4.2 представлены результаты сегментации тестового участка кедрового насаждения. Можно видеть, что СНС успешно воспроизводят границы между деревьями и сегментирует большую долю крон деревьев классов «Условно здоровое» и «Старый сухостой», однако в наличии значительный объем ошибок классификации для классов промежуточного состояния деревьев.

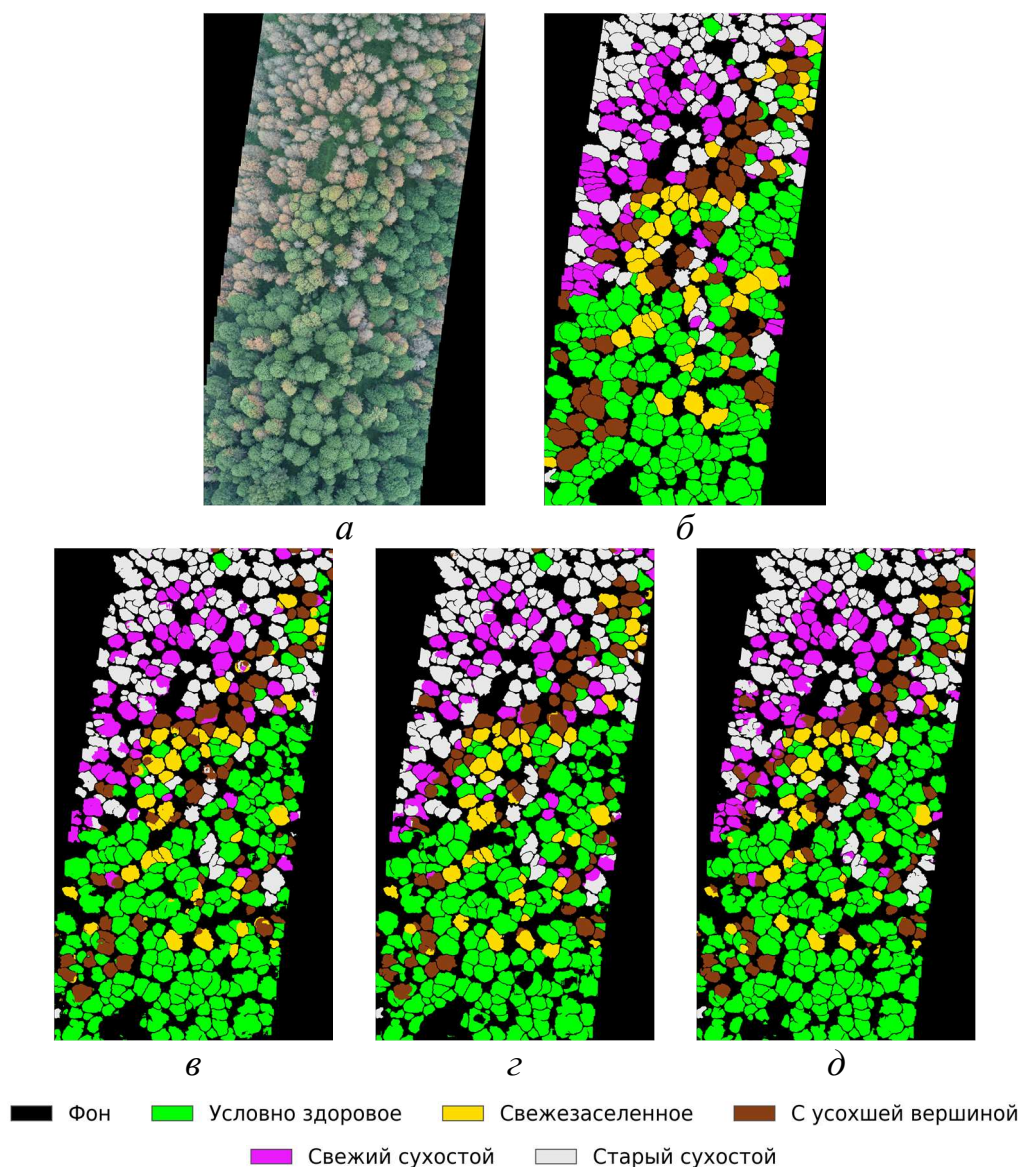


Рисунок 4.2 – Результаты сегментации деревьев кедра на изображении тестового участка: а – тестовый участок, б – эталонная карта сегментации, в – результат сегментации моделью U-Net, г – моделью MH-U-Net, д – моделью MH-Res-U-Net

В таблице 4.2 приведены значения IoU для класса «фон» и различных классов деревьев кедра, а также значения mIoU для валидационной выборки и всего тестового участка.

Таблица 4.2 – Метрики качества сегментации изображений кедровых насаждений. Полужирным выделены лучшие значения метрик в столбце

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

Все из приведенных моделей успешно классифицируют пиксели пяти

классов из шести: модели U-Net и MH-Res-U-Net успешно сегментируют все классы, кроме класса «Свежезаселенное», а модель MH-U-Net – все классы, кроме класса «Свежий сухостой». Как было отмечено выше, основную сложность для сегментации представили классы промежуточного состояния деревьев, при этом СНС MH-U-Net справляется с сегментацией промежуточного класса «Свежезаселенное», U-Net – класса «С усохшей вершиной», а MH-Res-U-Net – класса «Свежий сухостой». В связи с описанными особенностями одним из перспективных направлений дальнейших исследований является ансамблирование моделей в целях комбинации их лучших качеств.

Важно отметить, что представленные результаты для сегментации изображений кедровых насаждений являются промежуточными. Возможно, эти результаты могут быть значительно улучшены в процессе продолжающегося подбора оптимальных гиперпараметров.

4.2 Сравнение производительности полносверточных сетей

Одним из важнейших аспектов применения СНС на практике является объем вычислений и связанное с этим необходимое время для вычислений. В таблице 4.3 приведены медиана и медианное абсолютное отклонение (MAD) времени обработки одного фрагмента 256x256 для всех рассмотренных в работе полносверточных сетей, сравнение проводилось в среде Google Colaboratory [32], количество измерений – 1000; также приведено количество параметров, для MH-U-Net и MH-Res-U-Net в скобках указано количество параметров в режиме эксплуатации, в котором ветви вычислений для получения карт сегментации меньшего масштаба не используются.

Таблица 4.3 – Количество параметров и время обработки одного фрагмента размером 256x256 для предложенных моделей СНС

Модель СНС	Количество параметров	Время обработки фрагмента, мс	
		Медиана	MAD
U-Net	34540997	168,430	1,962
SegNet	31407557	131,696	1,817
ENet	372154	65,945	1,802
MH-U-Net	34794575 (34572101)	175,313	2,256
MH-Res-U-Net	36167823 (35945349)	180,156	2,052

Ожидается, лучшей по времени выполнения оказалась СНС ENet, практически вдвое обогнав SegNet и почти втрое – U-Net и ее модификации.

4.3 Исследование чувствительности полносверточных сетей к выбору гиперпараметров

Одним из наблюдений в ходе обучения моделей СНС стал стабильно более качественный результат сегментации архитектурами МН-U-Net и МН-Res-U-Net. Для исследования чувствительности СНС к выбору гиперпараметров были взяты первые $N = 40$ наборов случайных гиперпараметров и соответствующих значений mIoU для каждой из моделей СНС для сегментации деревьев пихты. На рисунке 4.3 приведена диаграмма размаха, на которой можно видеть, что медиана mIoU для МН-U-Net и МН-Res-U-Net больше, чем для других архитектур, а разброс значений меньше.

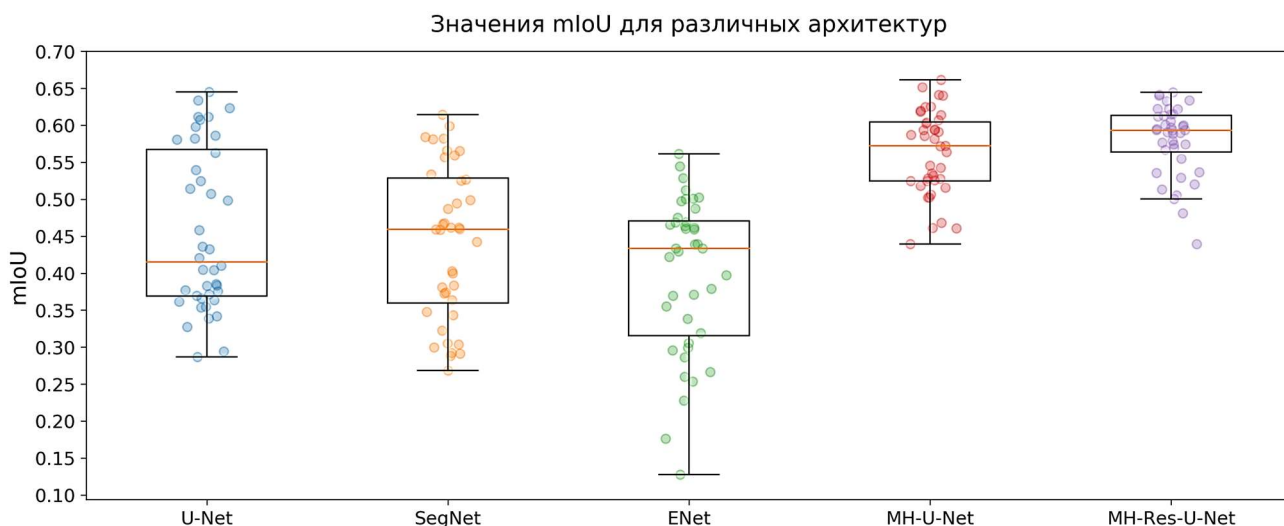


Рисунок 4.3 – Значения mIoU для различных архитектур СНС, полученные на этапе случайного поиска оптимальных гиперпараметров

Чтобы статистически подтвердить значимость отличия групп значений mIoU друг от друга для разных моделей, были проведены попарные тесты Краскела-Уоллиса, р-уровни представлены в таблице 4.4.

Таблица 4.4 – Значения p-уровней для теста Краскела-Уоллиса. Полужирным выделены значения, при которых принимается гипотеза о значимом статистическом различии совокупностей значений mIoU при $\alpha = 0,05$

	U-Net	SegNet	ENet	MH-U-Net	MH-Res-U-Net
U-Net	–	0,61	0,010	2,2e-5	1,8e-5
SegNet	0,61	–	0,11	2,6e-7	1,9e-9
ENet	0,010	0,11	–	1,7e-11	5,3e-13
MH-U-Net	2,2e-5	2,6e-7	1,7e-11	–	0,15
MH-Res-U-Net	1,8e-5	1,9e-9	5,3e-13	0,15	–

Можно утверждать, что результаты для конкретного набора данных, даваемые предложенными архитектурами MH-U-Net и MH-Res-U-Net, лучше, чем рассмотренными в работе U-Net, SegNet и ENet, и эти различия являются статистически значимыми. Таким образом, использование архитектур MH-U-Net и MH-Res-U-Net позволяет снизить количество обучаемых СНС для получения приемлемых результатов сегментации и, следовательно, значительно экономить время и вычислительные ресурсы при решении практических задач.

Чтобы выявить общий характер взаимосвязей между mIoU и гиперпараметрами, была проведена множественная линейная регрессия. В таблице 4.5 представлены коэффициенты регрессии для всех моделей СНС и выявленных гиперпараметров, для справки для каждой модели приведены значения коэффициента детерминации.

Таблица 4.5 – Значения коэффициентов множественной линейной регрессии для различных моделей нейронных сетей и гиперпараметров. Полужирным выделены коэффициенты, значимо отличающиеся от нуля при $\alpha = 0,05$

Переменные	Модели				
	U-Net ($R^2 = 0,93$)	SegNet ($R^2 = 0,95$)	ENet ($R^2 = 0,78$)	MH-U-Net ($R^2 = 0,93$)	MH-Res-U-Net ($R^2 = 0,85$)
const	0,54	0,54	0,24	0,52	0,54
θ_{lr}	-76	-94	180	-38	-27
θ_d	-0,56	-0,61	-0,33	-0,16	-0,12
θ_s	260	31	-190	-18	-46
θ_{ls}	0,040	-0,036	0,041	-8,9e-3	-0,065
θ_z	-0,051	-0,013	0,071	0,024	0,043
θ_b	0,17	0,11	-0,077	-0,069	-0,016
θ_c	0,058	9,8e-3	0,037	-0,042	-0,038
θ_α	-2,7e-5	-4,0e-5	-1,5e-4	-1,8e-5	1,5e-6
θ_σ	-6,6e-4	2,2e-3	6,2e-4	1,5e-3	2,1e-4
θ_{pf}	0,075	0,037	0,076	0,096	0,067
θ_{pp}	2,0e-3	7,0e-3	0,013	0,010	9,0e-3

Из таблицы 4.5 можно попытаться сделать вывод о том, что MH-U-Net и MH-Res-U-Net гораздо менее чувствительны к выбору коэффициентов скорости обучения и дропаута, для MH-Res-U-Net связь со скоростью обучения вообще

теряет статистическую значимость. Однако важно отметить, что такое утверждение нельзя считать однозначно доказанным, так как:

- количество рассмотренных значений mIoU ($N = 40$) не в значительное количество раз превышает количество переменных-предикторов, т.е. гиперпараметров;

- модель регрессии линейна, следовательно, не способна воспроизводить потенциально сложную функцию $mIoU(\theta)$, а лишь показывает общие тенденции;

- нельзя гарантировать, что такое поведение повторится для другого набора данных.

Таким образом, вопросы о том, к каким именно гиперпараметрам МН-U-Net и МН-Res-U-Net имеют более низкую чувствительность и почему для предложенных моделей наблюдается данный факт, требуют дополнительных объемных исследований, включающих большее количество обучаемых нейронных сетей и их тестирование на различных наборах данных.

5 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

5.1 Предпроектный анализ

5.1.1 Анализ конкурентных технических решений

Результатом разработки является набор моделей машинного обучения для решения задачи сегментации поврежденных вредителями хвойных деревьев на цифровых изображениях с БПЛА. Прямым конкурентом для решения тех же задач является работа Safonova et al. [8] (к1), где авторы используют обычные RGB-снимки пихтовых насаждений, проводят предварительную сегментацию крон деревьев пороговой фильтрацией по яркости и их дальнейшую классификацию с использованием СНС. Другие прямые конкуренты отсутствуют, поэтому сравнение проводилось дополнительно с методами картографирования лесного покрова, описанными в литературе:

– Lee et al. [12] (к2): в работе используется ряд простейших моделей машинного обучения для попиксельной классификации сцен в целях обнаружения нарушенности леса;

– Onishi and Ise [13] (к3): в работе используются данные аэрофотосъемки с большими перекрытиями, из которых восстанавливается цифровая модель поверхности для извлечения сегментов крон деревьев и их классификации СНС.

Были выбраны следующие критерии конкурентоспособности:

– точность модели машинного обучения (качество итоговой сегментации/классификации снимков с БПЛА);

– потребность в вычислительных ресурсах (необходимое количество операций процессора и графического ускорителя, а также объем оперативной памяти для обработки одного снимка с БПЛА);

– возможность дообучения модели (возможность использования дополнительных объемов данных для повышения качества сегментации);

– универсальность подхода (могут ли описанные в работах подходы применяться для решения похожих, но других задач картографирования

местности);

– требовательность к данным (требует ли подход наличия данных съемки в ближнем и среднем инфракрасном диапазонах, лазерной альтиметрии или больших площадей перекрытия двух соседних снимков).

Для каждого из критериев был оценен его вес, далее все конкурентные решения экспертным методом были оценены по десятибалльной шкале по приведенным критериям, на основании чего были получены итоговые оценки конкурентноспособности. В таблице 5.1 представлена оценочная карта конкурентных технических решений.

Таблица 5.1 – Оценочная карта конкурентных технических решений

Критерий оценки	Вес критерия	Баллы				Конкурентоспособность			
		Б _ф	Б _{к1}	Б _{к2}	Б _{к3}	К _ф	К _{к1}	К _{к2}	К _{к3}
Точность	0,30	8	7	5	8	2,40	2,10	1,50	2,40
Потребность в вычислительных ресурсах	0,10	3	5	8	4	0,30	0,50	0,80	0,40
Возможность дообучения	0,10	7	8	3	8	0,70	0,80	0,30	0,80
Универсальность	0,25	9	2	2	7	2,25	0,50	0,50	1,75
Требовательность к данным	0,25	9	9	9	5	2,25	2,25	2,25	1,25
Итого	1	36	31	27	32	7,90	6,15	5,35	6,60

Из таблицы 5.1 видно, что предлагаемый в работе подход имеет наибольшую конкурентоспособность, основным преимуществом является универсальность, сочетаемая с высокой точностью и невысокими требованиями к используемым исходным данным.

5.1.2 SWOT-анализ

Для определения потенциальных путей улучшения разработки был проведен SWOT-анализ, для этого были выделены сильные и слабые стороны разработки, потенциальные возможности и угрозы. Они были нанесены на матрицу SWOT-анализа (таблица 5.2), на пересечениях выделенных факторов приведены предложения по развитию проекта.

Таблица 5.2 – Матрица SWOT-анализа

		Внутренние факторы	
		Сильные стороны: С1. Высокая универсальность подхода. С2. Высокая точность сегментации. С3. Низкие требования к исходным данным.	Слабые стороны: Сл1. Высокие требования к вычислительным ресурсам. Сл2. Сложность эксплуатации для рядового пользователя.
Внешние факторы	Возможности: В1. Перенос подхода на более широкий спектр задач. В2. Создание расширения для геоинформационной системы.	В1В2С1С2С3. Создание набора готовых предобученных моделей для заранее определенного спектра задач. В1С1С2С3. Повышение точности моделей путем подготовки объемного набора данных для задачи семантической сегментации и предварительного обучения энкодера.	В1В2Сл2. Проектирование и реализация модуля для геоинформационной системы для решения персональных задач пользователя.
	Угрозы: У1. Конкуренция с аналогами. У2. Низкая степень интеграции в реальные решения.	У1С1С2. Отслеживание современных работ по тематике и улучшение предлагаемых моделей в соответствии с ними. У2С1С2. Публикация результатов в научных журналах и на профильных ресурсах.	У1Сл2. Реализация оболочки с графическим интерфейсом пользователя. У2Сл1. Применение новых подходов к проектированию и обучению нейронных сетей.

В таблице 5.3 приведена интерактивная матрица проекта, показывающая взаимосвязи выделенных областей матрицы SWOT-анализа.

Таблица 5.3 – Интерактивная матрица проекта

		Сильные стороны			Слабые стороны	
		С1	С2	С3	Сл1	Сл2
Возможности	В1	+	+	+		+
	В2	+	+	+		+
Угрозы	У1	+	+			+
	У2	+	+		+	

В целом, SWOT-анализ показал: перспективными выглядят улучшение предлагаемого в магистерской диссертации подхода путем изучения и применения современных техник искусственного интеллекта и понижение сложности использования разработки для рядового пользователя.

5.2 Планирование научно-исследовательских работ

5.2.1 Разработка графика проведения научного исследования

Для того чтобы определить трудоемкость работ, используются следующие показатели:

- ожидаемое значение трудоемкости,
- продолжительность каждой работы,
- продолжительность выполнения i -ой работы в календарных днях,
- коэффициент календарности.

Расчет ожидаемого значения продолжительности работ $t_{ож}$ осуществляется согласно формуле:

$$t_{ож} = \frac{3 \cdot t_{min} + 2 \cdot t_{max}}{5}, \quad (5.1)$$

где t_{min} – минимально возможная трудоемкость i -ой работы, чел.-дни; t_{max} – максимально возможная трудоемкость i -ой работы, чел.-дни.

Далее определяется продолжительность каждой работы в рабочих днях T_{pi} , которая учитывает параллельность выполнения работ несколькими исполнителями:

$$T_{pi} = \frac{t_{ожi}}{Ч_i}, \quad (5.2)$$

где $t_{ожi}$ – ожидаемая трудоемкость выполнения одной работы, чел.-дни; $Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Для построения календарного графика, осуществляется перевод длительности каждого из этапов работ из рабочих дней в календарные дни по следующей формуле:

$$T_{ki} = T_{pi} \cdot k_{кал}, \quad (5.3)$$

где T_{pi} – продолжительность выполнения i -ой работы в рабочих днях; $k_{кал}$ – коэффициент календарности.

Коэффициент календарности определяется по следующей формуле:

$$k_{кал} = \frac{T_{кал}}{T_{кал} - T_{вых} - T_{пр}}, \quad (5.4)$$

где $T_{кал}$ – количество календарных дней в году, $T_{вых}$ – количество выходных

дней в году, $T_{пр}$ – количество праздничных дней в году.

Согласно производственному календарю, в году 365 календарных дней, из них для шестидневной рабочей недели 66 выходных или праздничных дней, следовательно, для научного руководителя (НР) $k_{кал} = 1,22$, а для пятидневной – 118 выходных или праздничных дней, следовательно, для инженера (И) $k_{кал} = 1,48$.

Список работ, соответствующие им исполнители и результаты расчетов трудоемкости выполнения работ представлены в таблице 5.4.

Таблица 5.4 – Состав работ и их трудоемкость

№	Наименование работы	Исполнители	Трудоемкость, чел.-дни			Длительность, дни	
			t_{min}	t_{max}	$t_{ож}$	T_p	T_k
1	Выбор направления исследования	И	1	2	1,4	1	1
2	Составление и учреждение темы	И	2	3	2,4	1	1
		НР					1
3	Составление календарного плана	И	2	3	2,4	1	1
		НР					1
4	Анализ предметной области	И	18	22	19,6	20	30
5	Подготовка исходных данных	И	24	28	25,6	26	38
6	Проведение разведочного анализа данных	И	7	9	7,8	8	12
7	Анализ и обсуждение промежуточных результатов	И	1	2	1,4	1	1
		НР					1
8	Программная реализация моделей машинного обучения	И	44	50	46,4	46	68
9	Проектирование и программная реализация алгоритма подбора гиперпараметров	И	34	38	35,6	36	53
10	Сравнение и анализ полученных моделей	И	36	38	36,8	37	55
11	Согласование выполненной работы	И	5	6	5,4	3	4
		НР					4
12	Оформление пояснительной записки и сопутствующих документов	И	9	12	10,2	10	15
Итого		И			190		279
		НР			6		7

Диаграмма Гантта, построенная по рассчитанным показателям, представлена на рисунке 5.1.



■ Инженер ■ Научный руководитель

Рисунок 5.1 – Диаграмма Гантта

5.2.2 Бюджет научно-технической разработки

5.2.2.1 Расчет материальных затрат

Данная статья затрат включает в себя затраты на приобретение сырья, материалов, полуфабрикатов и комплектующих со стороны. Также в эту статью включаются транспортные расходы, равные 15 % от общей стоимости материальных затрат.

Общая сумма материальных затрат включает в себя только затраты на канцелярские принадлежности (250 руб.), для которых не учитываются транспортные расходы.

Таким образом, общая сумма материальных затрат составляет 250 руб.

5.2.2.2 Расчет затрат на специальное оборудование

В данную статью затрат входят суммы, необходимые на обеспечение амортизации используемого оборудования.

При подготовке и написании работы использовались персональный ноутбук стоимостью 60000 рублей и 3 персональных компьютера с установленными графическими ускорителями стоимостью 200000 рублей каждый. Срок полезного использования для офисных машин – 3 года, планируется использовать ноутбук и персональные компьютеры для постановки вычислительных экспериментов и написания ВКР в течение 8 месяцев. Тогда:

– норма амортизации:

$$A_n = \frac{1}{n} \cdot 100\% = \frac{1}{3} \cdot 100\% = 33,33\%, \quad (5.5)$$

– годовые амортизационные отчисления:

$$A_r = (60000 + 3 \cdot 200000) \cdot 0,33 = 220000 \text{ руб.}, \quad (5.6)$$

– ежемесячные амортизационные отчисления:

$$A_m = \frac{220000}{12} = 18333,33 \text{ руб.}, \quad (5.7)$$

– итоговая сумма амортизации основных средств:

$$A = 18333,33 \cdot 8 = 146666,67 \text{ руб.} \quad (5.8)$$

Итоговая сумма затрат на амортизацию составила 146666,67 руб.

5.2.2.3 Основная заработная плата исполнителей

Данная статья расходов включает основную заработную плату с учетом премий и доплат для исполнителей проекта: инженера и научного руководителя.

Месячный оклад руководителя с должностью профессора и степенью доктора наук, учитывая индексацию, составляет 49129,5 руб., для студента (инженера) был взят оклад ассистента без научной степени – 22695 руб.

Месячный должностной оклад исполнителя рассчитывается как:

$$Z_M = Z_{ок} \cdot k_p, \quad (5.9)$$

где $Z_{ок}$ – размер оклада, $k_p = 1,3$ – районный коэффициент (значение взято для г. Томска).

Среднедневная заработная плата рассчитывается как:

$$Z_{дн} = \frac{Z_M \cdot M}{F_d}, \quad (5.10)$$

где Z_M – месячный должностной оклад работника; M – количество месяцев работы без отпуска в течение года: для инженера (пятидневная рабочая неделя) $M = 11,2$ мес., для научного руководителя (шестидневная рабочая неделя) $M = 10,4$ мес.; F_d – действительный годовой фонд рабочего времени научно-технического персонала: для инженера $F_d = 223$ раб. дня, для научного руководителя $F_d = 243$ раб. дня.

Тогда основная заработная плата от предприятия может быть орасчитана следующим образом:

$$Z_{осн} = Z_{дн} \cdot T_p, \quad (5.11)$$

где $Z_{дн}$ – среднедневная заработная плата, T_p – продолжительность работ (см. таблицу 5.4).

Результаты расчетов заработной платы исполнителей приведены в таблице 5.5.

Таблица 5.5 – Расчет основной заработной платы

Исполнитель	З _{ок} , руб.	k _р	З _м , руб.	З _{дн} , руб.	T _р , раб. дни	З _{осн} , руб.
Инженер	22695,00	1,3	29503,50	1481,79	190	281540,13
Научный руководитель	49129,50	1,3	63868,35	2733,46	6	16400,76
Итого						297940,89

5.2.2.4 Дополнительная заработная плата исполнителей темы

В данную статью расходов входит заработная плата, начисленная рабочим и служащим не за фактически выполненные работы или проработанное время, а в соответствии с действующим законодательством, в том числе оплата очередных отпусков рабочих, времени, связанного с выполнением государственных и общественных обязанностей.

Зная основную заработную плату, можно рассчитать дополнительную заработную плату в размере 12 % от основной:

$$Z_{\text{доп}} = k_{\text{доп}} \cdot Z_{\text{осн}}, \quad (5.12)$$

где $k_{\text{доп}} = 0,12$ – коэффициент дополнительной заработной платы, $Z_{\text{осн}}$ – основная заработная плата.

Результаты расчетов дополнительной заработной платы приведены в таблице 5.6.

Таблица 5.6 – Расчет дополнительной заработной платы

Исполнитель	k _{доп}	З _{осн} , руб.	З _{доп} , руб.
Инженер	0,12	281540,13	33784,82
Научный руководитель		16400,76	1968,09
Итого			35752,91

5.2.2.5 Отчисления во внебюджетные фонды

В данную статью затрат включаются отчисления в органы социального страхования, пенсионные фонды и отчисления медицинского страхования.

Отчисления во внебюджетные фонды рассчитываются как:

$$Z_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}), \quad (5.13)$$

где $k_{\text{внеб}} = 0,271$ – коэффициент отчислений во внебюджетные фонды (принимается пониженная ставка для учреждений, осуществляющих образовательную и научную деятельность), $Z_{\text{осн}}$ – основная заработная плата,

$Z_{\text{доп}}$ – дополнительная заработная плата.

Результаты расчетов отчислений во внебюджетные фонды приведены в таблице 5.7.

Таблица 5.7 – Расчет отчислений во внебюджетные фонды

Исполнитель	$k_{\text{внеб}}$	$Z_{\text{доп}}$, руб.	$Z_{\text{осн}}$, руб.	$Z_{\text{внеб}}$, руб.
Инженер	0,271	33784,82	281540,13	85453,06
Научный руководитель		1968,09	16400,76	4977,96
Итого				90431,02

5.2.2.6 Накладные расходы

Накладные расходы учитывают прочие затраты организации, не попавшие в предыдущие статьи расходов: печать и ксерокопирование материалов, оплата услуг связи, электроэнергия и т.д. Их величина определяется согласно следующей формуле:

$$Z_{\text{накл}} = k_{\text{накл}} \cdot (\text{сумма статей расходов}), \quad (5.14)$$

где $k_{\text{накл}}$ – коэффициент накладных расходов, принятый за 16 %.

Результат расчета накладных расходов приведен в таблице 5.8.

Таблица 5.8 – Расчет накладных расходов

Статьи затрат	Сумма, руб.
Материальные затраты	250
Амортизация	146666,67
Основная заработная плата	297940,89
Дополнительная заработная плата	35752,91
Отчисления во внебюджетные фонды	90431,02
Накладные расходы	91366,64

5.2.2.7 Бюджет затрат

Рассчитанные величины затрат научно-исследовательской работы являются основой для формирования бюджета затрат проекта. Результаты составления итогового бюджета разработки представлены в таблице 5.9.

Таблица 5.9 – Бюджет затрат на разработку

Наименование	Сумма, руб.	Удельный вес, %
Материальные затраты	250	0,038
Амортизация	146666,67	22,14
Основная заработная плата	297940,89	44,98
Дополнительная заработная плата	35752,91	5,40
Отчисления во внебюджетные фонды	90431,02	13,65
Накладные расходы	91366,64	13,79
Общий бюджет	662408,13	100

5.3 Определение ресурсной, финансовой, бюджетной, социальной и экономической эффективности исследования

В данном подразделе сравнивается разработка, описанная в данной работе, со своим прямым конкурентом – работой Safonova et al. [8] (аналог 1), описанной в подразделе 5.1.1, так как данный аналог решает ту же самую задачу.

5.3.1 Определение финансовой эффективности исследования

Интегральный финансовый показатель разработки для i -го варианта исполнения рассчитывается как:

$$I_{\Phi}^{\text{исп.}i} = \frac{\Phi_i}{\Phi_{\text{max}}}, \quad (5.15)$$

где Φ_i – стоимость i -го варианта исполнения, Φ_{max} – максимальная стоимость исполнения из всех вариантов.

Реализация описанного в данной магистерской диссертации подхода займет большее время по причине его большей сложности (авторы [8] не рассматривают процесс подбора оптимальных гиперпараметров моделей, а также используют более простые модели), поэтому примем, что для разработки аналога потребуется на месяц меньше, или на 12,5 %.

Тогда:

$$I_{\Phi}^p = \frac{\Phi_p}{\Phi_{\text{max}}} = \frac{\Phi_p}{\Phi_p} = 1, \quad (5.16)$$

$$I_{\Phi}^1 = \frac{\Phi_1}{\Phi_{\text{max}}} = \frac{\Phi_1}{\Phi_p} = \frac{(1-0,125) \cdot \Phi_p}{\Phi_p} = 0,875. \quad (5.17)$$

5.3.2 Определение показателя ресурсоэффективности

Интегральный показатель ресурсоэффективности для i -го варианта исполнения рассчитывается как:

$$I_p^{\text{исп.}i} = \sum_{j=1}^n a_j \cdot b_j, \quad (5.18)$$

где a_j – весовой коэффициент j -го параметра сравнения, b_j – оценка j -го параметра сравнения по пятибалльной шкале, n – число параметров сравнения.

В качестве параметров сравнения были взяты критерии конкурентоспособности из подраздела 5.1.1 с теми же весовыми коэффициентами. Результаты определения показателя ресурсоэффективности для разработки и аналоги приведены в таблице 5.10.

Таблица 5.10 – Показатель ресурсоэффективности для разработки и аналога

Параметр	Весовой коэффициент параметра	Объект исследования	
		Разработка	Аналог 1
Точность	0,30	4	3
Потребность в вычислительных ресурсах	0,10	1	2
Возможность дообучения	0,10	3	4
Универсальность	0,25	4	1
Требовательность к данным	0,25	4	4
Итого	1	3,60	2,75

5.3.3 Интегральный показатель эффективности

Интегральный показатель эффективности i -го варианта исполнения вычисляется как:

$$I_{\text{исп.}i} = \frac{I_p^{\text{исп.}i}}{I_{\text{исп.}i}^{\text{ф}}}. \quad (5.19)$$

Для сравнения i -го и j -го вариантов исполнения разработки используется сравнительная эффективность проекта:

$$\mathcal{E}_{\text{ср}} = \frac{I_{\text{исп.}i}}{I_{\text{исп.}j}}. \quad (5.20)$$

Взяв за i -ый вариант разработку, описанную в данной работе, а за j -ый – прямого конкурента, можно сделать вывод о целесообразности предлагаемого варианта: если $\mathcal{E}_{\text{ср}} > 1$, то предлагаемая разработка более целесообразна. В таблице 5.11 приведены все рассчитанные значения интегральных показателей и сравнительная эффективность вариантов исполнения.

Таблица 5.11 – Сравнение интегральных показателей

Показатель	Разработка	Аналог 1
Интегральный финансовый показатель разработки	1	0,875
Интегральный показатель ресурсоэффективности разработки	3,60	2,75
Интегральный показатель эффективности	3,60	3,14
Сравнительная эффективность вариантов исполнения	1,15	

Таким образом, несмотря на то, что для предлагаемой разработки интегральный финансовый показатель выше, значение сравнительной эффективности показывает ее большую целесообразность.

5.4 Вывод по разделу

В ходе работы над разделом был проведен предпроектный анализ, включающий анализ конкурентных решений и SWOT-анализ. Анализ конкурентных решений показал преимущество разрабатываемого подхода по сравнению с описанными в литературе. SWOT-анализ позволил определить потенциальные пути улучшения разработки: перспективными можно считать улучшение предлагаемого в данной работе подхода путем применения новых методов машинного обучения и понижение порога вхождения для рядовых пользователей.

Также в ходе инициации проекта был составлен календарный план проекта, проведена оценка трудоемкости работ (таблица 5.4). График-план проекта был представлен в виде диаграммы Ганта (рисунок 5.1).

Был сформирован бюджет разработки, включающий материальные затраты, затраты на амортизацию, основную заработную плату исполнителям, дополнительную заработную плату, отчисления во внебюджетные фонды и накладные расходы (таблица 5.9). Общий бюджет разработки составил 662408,13 руб.

Было проведено сравнение эффективности выполнения для данной разработки и ее прямого аналога на основании сравнения значений интегральных показателей эффективности. Предлагаемый в данной работе подход оказался более целесообразным.

6 Социальная ответственность

Целью работы является разработка моделей и алгоритмов для семантической сегментации хвойных деревьев, поврежденных насекомыми-вредителями, на снимках с беспилотных летательных аппаратов с последующим анализом и сравнением этих моделей. Объектами исследования являются цифровые снимки хвойных насаждений в видимом диапазоне, полученные с БПЛА. Область применения разработки – экологический мониторинг лесного покрова.

В данном разделе исследованы вредные и опасные факторы среды, рассмотрены меры по защите работника (исполнителя) от возможного негативного воздействия среды, а также возможные чрезвычайные ситуации и необходимые действия, которые работник должен выполнить в случае возникновения чрезвычайной ситуации.

Были выделены и рассмотрены следующие факторы, воздействующие на разработчика программы:

- недостаточная освещенность рабочей зоны;
- отклонение параметров микроклимата;
- превышения уровня шума;
- нервно-психические перегрузки.

К опасным факторам при работе над разработкой (при пользовании персональным компьютером) относятся возможность короткого замыкания, влекущего за собой опасность поражения разработчика электрическим током и статическое электричество. Рассмотрены вопросы правового регулирования трудовых отношений, связанных с реализацией описанной в работе разработкой.

6.1 Правовые и организационные вопросы обеспечения безопасности

Работа в офисе относится ко второй категории тяжести труда – работы выполняются при оптимальных условиях внешней производственной среды и

при оптимальной величине физической, умственной и нервно-эмоциональной нагрузки. Согласно трудовому кодексу РФ продолжительность рабочего дня работников не должна превышать 40 часов в неделю. Возможно, сокращение рабочего времени. Для работников, возраст которых меньше 16 лет – не более 24 часа в неделю, от 16 до 18 лет – не более 35 часов, как и для инвалидов I и II группы [33]. Во время регламентированных перерывов целесообразно выполнять комплексы упражнений и осуществлять проветривание помещения.

Конструкция рабочего места и взаимное расположение всех его элементов должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы. Используемые средства отображения информации, требующие точного и быстрого считывания показаний, следует располагать в вертикальной плоскости под углом $\pm 15^\circ$ от нормальной линии взгляда и в горизонтальной плоскости под углом $\pm 15^\circ$ от сагиттальной плоскости. Клавиатура должна располагаться на поверхности стола на расстоянии 100-300 мм от края, обращенного к пользователю [34]. Соответствующие зоны зрительного наблюдения приведены на рисунке 6.1.

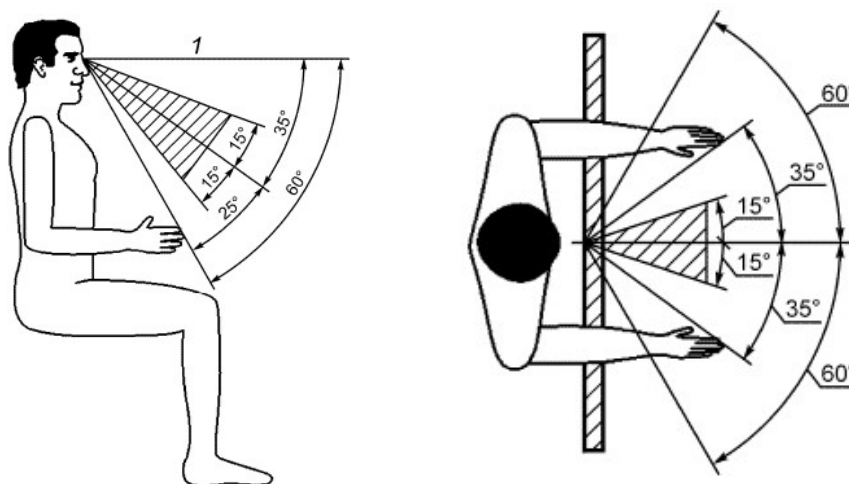


Рисунок 6.1 – Зоны зрительного наблюдения в горизонтальной плоскости и в вертикальной плоскости [35]

В соответствии с ГОСТ 12.2.032-78. «Система стандартов безопасности труда. Рабочее место, при выполнении работ сидя» рабочий стол может быть любой конструкции, отвечающей современным требованиям эргономики и

позволяющей удобно разместить на рабочей поверхности оборудование с учетом его количества, размеров и характера выполняемой работы [34]. Сравнение параметров действительного рабочего места с требуемыми приведены в таблице 6.1.

Таблица 6.1 – Требуемые и действительные параметры рабочего места

Параметр	Требуемое значение	Действительное значение
Высота рабочей поверхности стола	Регулируемая высота (680 – 800 мм), Нерегулируемая высота (725 мм)	Нерегулируемая высота (740 мм)
Рабочий стул	Подъемно-поворотный, регулируемый по высоте и углу наклона спинки	Подъемно-поворотный, регулируемый по высоте
Расположение монитора от глаз пользователя	600 – 700 мм	680 мм

6.2 Производственная безопасность

Вредные и опасные факторы, воздействующие на сотрудника, устанавливаются согласно ГОСТ 12.0.003-2015 «Опасные и вредные производственные факторы. Классификация» [36]. При разработке могут иметь место факторы, представленные в таблице 6.2.

Таблица 6.2 – Возможные опасные и вредные факторы

Факторы	Нормативные документы
Отклонение показателей микроклимата	СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату производственных помещений.
Недостаточная освещенность рабочей зоны	СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95.
Психофизиологические факторы	Трудовой кодекс Российской Федерации от 30.12.2001 №197-ФЗ (ред. от 27.12.2018).
Зрительное напряжение	СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.
Превышение уровня шума	СН 2.2.4/2.1.8.562-96. Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки.
Поражение электрическим током	ГОСТ 12.1.038-82 ССБТ Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов.
Статическое электричество	ГОСТ 12.1.006-84 ССБТ Электромагнитные поля радиочастот. Допустимые уровни на рабочих местах и требования к проведению контроля.
Короткое замыкание	ГОСТ 12.1.030-81 ССБТ Электробезопасность. Защитное заземление. Зануление. ГОСТ 26522-85 Короткие замыкания в электроустановках.

6.2.1 Отклонение показателей микроклимата

В производственных помещениях, в которых работа с использованием ПЭВМ является основной (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.) и связана с нервно-эмоциональным напряжением, должны обеспечиваться оптимальные параметры микроклимата для категории работ 1а и 1б в соответствии с действующими санитарно-эпидемиологическими нормативами микроклимата производственных помещений. Содержание вредных химических веществ в производственных помещениях, в которых работа с использованием ПЭВМ является основной, не должно превышать предельно допустимых концентраций загрязняющих веществ в атмосферном воздухе населенных мест в соответствии с действующими гигиеническими нормативами. Также указывается, что в помещениях с ПЭВМ должна ежедневно проводиться влажная уборка [37].

Уставленные гигиенические нормативы для помещений с ПЭВМ для категории работы 1б приведены в таблице 6.3.

Таблица 6.3 – Оптимальные и допустимые величины показателей микроклимата на рабочих местах производственных помещений [37]

Оптимальные значения характеристик микроклимата				
Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	22 – 24	21 – 25	40 – 60	0,1
Теплый	23 – 25	22 – 26	40 – 60	0,1
Допустимые значения характеристик микроклимата				
Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	20 – 25	19 – 26	15 – 75	0,1
Теплый	21 – 28	20 – 29	15 – 75	0,1 – 0,2

6.2.2 Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны помещения, оборудованной ПЭВМ, также является одной из причин нарушения зрительной функции, а также влияет на общее самочувствие и эффективность труда.

Искусственное освещение в помещениях для эксплуатации

персональный вычислительных устройств должно осуществляться системой общего равномерного освещения. В случаях преимущественной работы с документами, следует применять системы комбинированного освещения (к общему освещению дополнительно устанавливаются светильники местного освещения, предназначенные для освещения зоны расположения документов). Окна в помещениях, где эксплуатируется вычислительная техника, преимущественно должны быть ориентированы на север и северо-восток. Нормируемые показатели естественного, искусственного и совмещенного освещения в соответствии с СП 52.13330.2016 указаны в таблице 6.4 [38].

Таблица 6.4 – Нормируемые показатели естественного, искусственного и совмещенного освещения

Помещения	Рабочая поверхность и плоскость нормирования КЕО и освещенности (Г – горизонтальная, В – вертикальная) и высота плоскости над полом, м	Естественное освещение		Совмещенное освещение	
		КЕО e_n , %		КЕО e_n , %	
		При верхнем или комбинированном	При боковом	При верхнем или комбинированном	При боковом
Кабинеты, рабочие комнаты, офисы, представительства	Г – 0,8	3,0	1,0	1,8	0,6
Помещения	Искусственное освещение				
	Освещенность, лк			Показатель дискомфорта, М, не более	Коэффициент пульсации освещенности, Кп, %, не более
	При комбинированном освещении		При общем освещении		
Всего	От общего				
Кабинеты, рабочие комнаты, офисы, представительства	400	200	300	40	15

Проведем расчет освещенности рабочего места в помещении, в котором проводилась подготовка данной работы. Параметры помещения: длина $a = 6$ м, ширина $b = 3$ м, высота $H = 2,4$ м. Высота рабочей поверхности $h_{рп} = 0,74$ м. Требуемая освещенность $E = 300$ лк. Коэффициент отражения стен $R_c = 0,30$, потолка – $R_{п} = 0,50$. Коэффициент запаса $k = 1,5$, коэффициент неравномерности $Z = 1,1$. Для светильника типа ОД $\lambda = 1,4$.

Приняв $h_c = 0,5$ м, определяем расчетную высоту:

$$h = H - h_c - h_{рп} = 2,4 \text{ м} - 0,5 \text{ м} - 0,74 \text{ м} = 1,16 \text{ м}. \quad (6.1)$$

Индекс помещения:

$$i = \frac{S}{(a+b)h} = \frac{18}{9 \cdot 1,16} = 1,72. \quad (6.2)$$

Определяем коэффициент использования светового потока:

$$\eta = 0,55. \quad (6.3)$$

Расстояние между светильниками L определяется как:

$$L = \lambda * h = 1,4 * 1,16 \text{ м} = 1,624 \text{ м}. \quad (6.4)$$

Количество рядов светильников с люминесцентными лампами определяется по формуле:

$$n_{\text{ряд}} = \frac{b - \frac{2}{3}L}{L} + 1 = \frac{3 - \frac{2}{3} \cdot 1,624}{1,624} + 1 = 2,18 \approx 2. \quad (6.5)$$

Однако ширина помещения, с учетом расстояния между светильниками не позволяет разметить более одного ряда, поэтому $n_{\text{ряд}} = 1$.

Количество светильников определяем как:

$$n_{\text{св}} = \frac{a - \frac{2}{3}L}{l_{\text{св}} + 0,5} = \frac{5 - \frac{2}{3} \cdot 1,624}{1,23 + 0,5} = 2,26 \approx 2. \quad (6.6)$$

Таким образом, размещаем светильники в один ряд. В ряду можно установить два светильника типа ОД мощностью 40 Вт. Учитывая, что в каждом светильнике установлено две лампы, общее число ламп в помещении $N = 4$.

Определяем потребный световой поток лампы:

$$\Phi = \frac{E \cdot S \cdot k \cdot Z}{N \cdot \eta} = \frac{30 \cdot 18 \cdot 1,5 \cdot 1,1}{4 \cdot 0,55} \text{ лм} = 5400,75 \text{ лм}. \quad (6.7)$$

Выбираем ближайшую стандартную лампу – ЛТБ 80 Вт с потоком 5200 лм. Делаем проверку выполнения условий:

$$-10\% \leq \frac{\Phi_{\text{ст}} - \Phi_{\text{расч}}}{\Phi_{\text{ст}}} \cdot 100\% \leq +20\%, \quad (6.8)$$

$$-10\% \leq -3,7\% \leq +20\%. \quad (6.9)$$

Условие выполнено. При этом общая электрическая мощность осветительной установки:

$$P = 4 * 40 \text{ Вт} = 160 \text{ Вт}. \quad (6.10)$$

6.2.3 Зрительное напряжение

Работа на ПК сопровождается постоянным и значительным напряжением функций зрительного анализатора. Одной из основных особенностей является иной принцип чтения информации, чем при обычном чтении. Чтобы снизить зрительное напряжение нужно соблюдать визуальные параметры экрана (таблица 6.5).

Таблица 6.5 – Допустимые визуальные параметры устройств отображения информации [39]

Параметры	Допустимые значения
Яркость белого поля	Не менее 35 кд/кв.м
Неравномерность яркости рабочего поля	Не более $\pm 20\%$
Контрастность (для монохромного режима)	3:1
Пространственная нестабильность изображения (непреднамеренное изменение положения фрагментов изображения экрана)	Не более $2 \cdot 10L - 4L$, где L – расстояние наблюдения

6.2.4 Психофизические факторы

Работа с ПК сопряжена с воздействием вредных психофизиологических факторов, в частности, нервно-психических перегрузок. Нервно-психические перегрузки – совокупность таких сдвигов в психофизиологическом состоянии организма человека, которые развиваются после совершения работы и приводят к временному снижению эффективности труда. Состояние утомления (усталость) характеризуется определенными объективными показателями и субъективными ощущениями.

Нервно-психические перегрузки подразделяются на следующие:

- умственное перенапряжение;
- перенапряжение анализаторов;
- монотонность труда;
- эмоциональные перегрузки.

При первых симптомах психического перенапряжения необходимо:

- дать нервной системе расслабиться;
- рационально чередовать периоды отдыха и работы;
- начать заниматься спортом;

- ложиться спать в одно и то же время;
- в тяжелых случаях обратиться к врачу.

Естественно, что полностью исключить провоцирующие факторы практически невозможно, но можно уменьшить их негативное воздействие, давая нервной системе необходимый отдых. Суммарное время регламентированных перерывов при работе с ПК зависит от категории трудовой деятельности и уровня нагрузки за рабочую смену. В таблице 6.6 приведено суммарное время отдыха для каждой категории работ.

Таблица 6.6 – Суммарное время перерывов в зависимости от категории работы и нагрузки

Категория работы с ПЭВМ	Уровень нагрузки за рабочую смену при видах работ с ПЭВМ			Суммарное время регламентированных перерывов при 8-часовой смене, мин
	Группа А, количество знаков	Группа Б, количество знаков	Группа В, часов	
I	До 20000	До 15000	До 2	50
II	До 40000	До 30000	До 4	70
III	До 60000	До 40000	До 6	90

При проведении исследований уровень нагрузки относился к группе В, категория работы III. Согласно таблице, суммарное время перерывов необходимо установить не менее 90 минут. Для данной категории работ требуется установить перерывы по 15 минут каждый трудовой час [39].

6.2.5 Превышение уровня шума

В производственных помещениях при выполнении основных или вспомогательных работ с использованием ПЭВМ уровни шума на рабочих местах не должны превышать предельно допустимых значений, установленных для данных видов работ в соответствии с действующими санитарно-эпидемиологическими нормативами СН 2.2.4/2.1.8.562-96 [40]. Допустимые значения уровней звукового давления, создаваемого ПЭВМ, приведены в таблице 6.7.

Таблица 6.7 – Допустимые значения уровней звукового давления в октавных полосах частот и уровня звука, создаваемого ПЭВМ

Уровни звукового давления, дБ в октавных полосах со среднегеометрическими частотами, Гц.									Уровни звука в дБА
31,5 Гц	63 Гц	125 Гц	250 Гц	500 Гц	1000 Гц	2000 Гц	4000 Гц	8000 Гц	
6	71	61	54	49	45	42	40	38	50

6.2.6 Повышенное значение напряжения в электрической цепи

При разработке широко используется электричество для питания компьютерной техники, которая может являться источником опасности. Несоблюдение правил ГОСТ 12.1.038-82 ССБТ «Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов» может привести к тяжелым последствиям [41]. Поражение электрическим током может произойти при прикосновении к токоведущим частям, находящимся под напряжением, на которых остался заряд или появилось напряжение.

Электрический ток оказывает на человека термическое, электролитическое, биологическое и механическое воздействие. Действие электрического тока на человека приводит к травмам или гибели людей. Для переменного тока частотой 50 Гц допустимое значение напряжения прикосновения составляет 2 В, а силы тока – 0,3 мА, для тока частотой 400 Гц соответственно 2 В и 0,4 мА, для постоянного тока – 8 В и 1 мА.

Мерами защиты от воздействия электрического тока являются оградительные устройства, устройства автоматического контроля и сигнализации, изолирующие устройства и покрытия, устройства защитного заземления, устройства автоматического отключения, предохранительные устройства.

6.2.7 Статическое электричество

В помещениях, оборудованных ПЭВМ, токи статического электричества чаще всего возникают при прикосновении персонала к любому из элементов ПЭВМ. Такие разряды опасности для человека не представляют, однако кроме неприятных ощущений могут привести к выходу оборудования из строя.

Для предотвращения образования и защиты от статического электричества в помещении используются нейтрализаторы и увлажнители, а полы имеют антистатическое покрытие в виде поливинилхлоридного антистатического линолеума. В ГОСТ 12.1.006-84 ССБТ «Электромагнитные поля радиочастот. Допустимые уровни на рабочих местах и требования к проведению контроля» [42] установлен максимальный допустимый электростатический потенциал экрана видеомонитора – 500 В.

В качестве мер уменьшения влияния вредных факторов на пользователя используются защитные фильтры для мониторов, увлажнители воздуха. Должны использоваться розетки с заземлением. Требуется проводить регулярную влажную уборку.

6.3 Экологическая безопасность

Загрязнение окружающей среды выражается в виде воздействия на литосферу в результате образования отходов при поломке предметов вычислительной техники и оргтехники.

В ходе разработке возможно повышение температуры ПК, при этом он является источником электромагнитного и ионизирующего излучения, а также шума. Для защиты здоровья сотрудников, работающих с программным обеспечением, рекомендуется соблюдать необходимую дистанцию при работе с компьютером (от 0,5 до 1 м), а также использовать эффективные охлаждающие системы и наиболее современное и эффективное оборудование.

Федеральный закон № 89 от 1998 г. «Об отходах производства и потребления» [43] запрещает юридическим лицам самовольно избавляться от опасных отходов. Этим видом деятельности, согласно постановлению Правительства РФ № 340 от 2002 г., могут заниматься только специализированные структуры. В их число входят и фирмы, которые занимаются утилизацией электронных отходов. Обращение с отходами регламентируется ГОСТ Р 53692-2009 «Ресурсосбережение. Обращение с отходами» [44].

Поэтому при необходимости утилизировать вышедшую из употребления

электронику наиболее безопасным для окружающей среды способом необходимо обращаться в специализированную компанию по утилизации. Такие компании действуют на всей территории Российской Федерации, в том числе и в Томской области.

6.4 Безопасность в чрезвычайных ситуациях

Наиболее вероятной чрезвычайной ситуацией при разработке алгоритма является пожар на рабочем месте. В качестве противопожарных мероприятий должны быть применены следующие меры:

- в помещении должны находиться средства тушения пожара, средства связи;
- электрическая проводка электрооборудования и осветительных приборов должна быть исправна;
- все сотрудники должны знать место нахождения средств пожаротушения и уметь ими воспользоваться, средств связи и номера экстренных служб.

Согласно ГОСТ 12.1.004-91 «Пожарная безопасность. Общие требования», при работе с компьютером необходимо соблюдать следующие нормы пожарной безопасности [45]:

- для предохранения сети от перегрузок запрещается одновременно подключать к сети количество потребителей, превышающих допустимую нагрузку;
- работы за компьютером проводить только при исправном состоянии оборудования, электропроводки;
- иметь средства для тушения пожара (огнетушитель);
- установить количество, размеры и соответствующее конструктивное исполнение эвакуационных путей и выходов;
- обеспечить возможность беспрепятственного движения людей по эвакуационным путям.

В связи с возможностью возникновения пожара разработан следующий план действий:

1. в случае возникновения пожара сообщить о нем руководителю, постараться устранить очаг возгорания имеющимися силами при помощи первичных средств пожаротушения (огнетушитель порошковый, углекислотный О-1П0 (з)-АВСЕ);

2. привести в действие ручной пожарной извещатель, если очаг возгорания потушить не удастся;

3. сообщить о возгорании в службу пожарной охраны по телефону 01 или 101, сообщить адрес, место и причину возникновения пожара;

4. принять меры по эвакуации людей и материальных ценностей;

5. встретить пожарную охрану, при необходимости сообщить всю необходимую информацию и оказать помощь при выборе наилучшего подхода к очагу возгорания.

Рабочее помещение оборудовано в соответствии с требованиями пожарной безопасности. Имеется порошковый огнетушитель, а также пожарная сигнализация и средства связи. В случае возникновения пожара в здании необходимо немедленно эвакуироваться из здания и направиться на выход в соответствии с планом эвакуации при пожарах и других ЧС (рисунок 6.2).



Рисунок 6.2 – План эвакуации при пожарах и других ЧС

6.5 Вывод по разделу

В ходе работы над разделом были выявлены опасные и вредные факторы, воздействию которых может подвергнуться человек, проводящий описанное в магистерской диссертации исследование, а также проведен анализ нормативной документации.

Установлено, что рабочее место удовлетворяет требованиям безопасности. Выполняемая работа не сопряжена с высоким риском травматизма. Освещение на рабочем месте соответствует нормам. Уровни шума находятся в допустимых пределах – источником шума при эксплуатации ПК являются системы охлаждения. Микроклиматические условия соблюдаются за счет использования систем отопления и кондиционирования. В виду того, что корпус ПК является токонепроводящим, необходимость в мерах против статического электричества отсутствует. Во время работы осуществляются перерывы для снижения нагрузки и предотвращения нервно-психических перегрузок. Рабочее помещение оборудовано согласно требованиям электро- и пожарной безопасности.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена задача анализа полученных с БПЛА снимков хвойных лесов, поврежденных насекомыми-вредителями. Для решения этой задачи реализован ряд моделей машинного обучения. Разработанные модели позволяют осуществлять попиксельную классификацию пораженных деревьев пихты и кедра в зависимости от степени их повреждения вредителями. Модели показали достаточно высокую эффективность при классификации пикселей классов «Фон», «Живые», «Свежий сухостой», «Старый сухостой» и низкую – для класса «Отмирающие» в задаче сегментации пихтовых деревьев, поврежденных уссурийским полиграфом. В задаче сегментации кедровых деревьев, поврежденных союзным короедом, модели U-Net и МН-Res-U-Net успешно сегментируют все классы, кроме класса «Свежезаселенное», а модель МН-U-Net – все классы, кроме класса «Свежий сухостой», ожидается, что продолжающийся процесс подбора гиперпараметров позволит добиться более высокого результата.

Предложенные в работе архитектуры полносверточных сетей МН-U-Net и МН-Res-U-Net продемонстрировали меньшую чувствительность к выбору гиперпараметров, чем модели U-Net, SegNet и ENet, однако вопросы о причинах такого поведения и деталях зависимости от значений конкретных гиперпараметров требуют сбора большего количества данных и их дальнейшего изучения.

Эффективность предложенных моделей может быть улучшена путем их обучения на большем количестве снимков разных участков хвойных лесов. Также она может быть повышена при использовании дополнительных данных, например, снимков в ближнем инфракрасном диапазоне спектра, карт глубины и т. п. Также перспективным может быть использование ансамблей полученных моделей. Возможно, имеет смысл рассматривать задачу нечеткой классификации крон деревьев, так как процесс ухудшения состояния деревьев является непрерывным и часто сложно отнести дерево однозначно к какой-либо из категорий, однако это является объектом дальнейших исследований.

СПИСОК ПУБЛИКАЦИЙ

Публикации по тематике магистерской диссертации:

1. Керчев И. А., **Маслов К. А.**, Марков Н. Г., Токарева О. С. Семантическая сегментация поврежденных деревьев пихты на снимках с беспилотных летательных аппаратов // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. №. 1. С. 116–126. (Scopus, WoS)
2. Маслов К. А. Байесовский подход к подбору гиперпараметров нейронных сетей на примере задачи семантической сегментации // Научная сессия ТУСУР – 2021. Материалы международной научно-технической конференции студентов, аспирантов и молодых ученых. Томск, 19-21 мая 2021. Томск: В-Спектр. 2021. В печати.
3. Маслов К. А. Автоматизированные методы выявления поврежденных деревьев на снимках с беспилотных летательных аппаратов // Труды XXV Международного научного симпозиума студентов и молодых ученых имени академика М. А. Усова «Проблемы геологии и освоения недр». Томск, 5-9 апреля 2021. Томск: ТПУ. 2021. В печати.
4. Маслов К. А. Применение методов машинного и глубокого обучения в задачах семантической сегментации изображений лесного покрова // Сборник трудов XVIII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии. Томск, 22-26 марта 2021. Томск: ТПУ. 2021. В печати.
5. **Маслов К. А.**, Токарева О. С., Керчев И. А., Марков Н. Г. Методы машинного обучения в задачах семантической сегментации снимков с беспилотных летательных аппаратов // Материалы 18-й Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Москва, 16-20 ноября 2020. Москва: Институт космических исследований РАН. 2020. С. 31.
6. Марков Н. Г., **Маслов К. А.**, Керчев И. А., Токарева О. С. Сверточная нейронная сеть для сегментации пораженных деревьев пихты на снимках с

беспилотных летательных аппаратов // Региональные проблемы дистанционного зондирования Земли – 2020. Материалы VII Международной научной конференции. Красноярск, 29 сентября-2 октября 2020. Красноярск: СФУ. 2020. С. 102–105.

7. Маслов К. А. Сегментация крон поврежденных деревьев на цифровых снимках с БПЛА с использованием нейросетевых моделей // Научная сессия ТУСУР – 2020. Материалы XXV Международной научно-технической конференции студентов, аспирантов и молодых учёных: в 3 частях. Томск, 25-27 мая 2020. Часть 2. Томск: В-Спектр. 2020. С. 135–138.

8. Маслов К. А. Классификация поврежденной уссурийским полиграфом пихты на снимках с БПЛА методами компьютерного зрения // Труды XXIV Международного научного симпозиума студентов и молодых ученых имени академика М. А. Усова «Проблемы геологии и освоения недр». Томск, 6-10 апреля 2020. Том 1. Томск: ТПУ. 2020. С. 526–527.

9. Маслов К. А. Классификация пихты, поврежденной уссурийским полиграфом, по снимкам с БПЛА с использованием нейронной сети U-Net // Сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». Томск, 17-20 февраля 2020. Томск: ТПУ. 2020. С. 109–110.

10. Маслов К. А. Детерминированный алгоритм классификации пихты, поврежденной уссурийским полиграфом, по снимкам с БПЛА // Сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». Томск, 17-20 февраля 2020. Томск: ТПУ. 2020. С. 107–108.

Другие публикации:

11. **Maslov K. A.** and Tokareva O. S. System for Automated Environmental Monitoring using Remote Sensing Data of the Earth from Open Data Sources // Proceedings of the 14th International Forum on Strategic Technology (IFOST 2019). October 2019. IOP Publishing. DOI:10.1088/1757-899X/1019/1/01210. (Scopus)

12. **Маслов К. А.**, Токарева О. С. Мониторинг изменений состояния хвойных лесов, поврежденных насекомыми-вредителями, с использованием космических снимков // Тринадцатое Сибирское совещание и школа молодых ученых по климато-экологическому мониторингу. Тезисы докладов российской конференции. Томск, 15-19 октября 2019 / Под редакцией М. В. Кабанова. Томск: Аграф-Пресс. 2019. С. 210–211.

13. **Маслов К. А.**, Токарева О. С. Система автоматизированного экологического мониторинга с использованием данных дистанционного зондирования Земли // Материалы 17-й Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Москва, 11-15 ноября 2019. Москва: Институт космических исследований РАН. 2019. С. 107.

14. Маслов К. А. Система автоматизированного мониторинга изменений земной поверхности с использованием данных космической съемки // Труды XXIV Международного научного симпозиума студентов и молодых ученых имени академика М. А. Усова «Проблемы геологии и освоения недр». Томск, 6-10 апреля 2020. Том 1. Томск: ТПУ. 2020. С. 524–525.

СПИСОК ИСТОЧНИКОВ

1. Кривец С. А., Бисирова Э. М., Волкова Е. С., Дебков Н. М., Керчев И. А., Мельник М. А., Никифоров А. Н., Чернова Н. А. Технология мониторинга пихтовых лесов в зоне инвазии уссурийского полиграфа в Сибири. Методическое пособие. Томск: УМИУМ, 2018. 74 с.
2. Керчев И. А., Мандельштам М. Ю., Кривец С. А., Илинский Ю. Ю. Союзный короед *Ips Amitinus* (Eichhoff, 1872) (Coleoptera, Curculionidae: Scolytinae) – новый чужеродный вид в Западной Сибири // Энтомологическое обозрение. 2019. Т. 98. № 3. С. 592–599.
3. Bystrov S. O., Antonov I. A. First record of the four-eyed fir bark beetle *Polygraphus proximus* Blandford, 1894 (Coleoptera, Curculionidae: Scolytinae) from Irkutsk province Russia // Entomological Review. 2019. V. 99. № 1. P. 54–55. DOI: 10.1134/S001387381901007X.
4. Проведение санитарно-оздоровительных мероприятий в лесных насаждениях, повреждённых Уссурийским полиграфом, на территории Удмуртской Республики. – URL: <http://perm.rcfh.ru/news/22860.html> (дата обращения: 01.07.2020).
5. Пашенова Н. В., Петько В. М., Керчев И. А., Бабичев Н. С. Перенос офиостомовых грибов уссурийским полиграфом *Polygraphus proximus* Blandford (Coleoptera, Scolytidae) в Сибири // Известия Санкт-Петербургской лесотехнической академии. 2012. № 200. С. 114–120.
6. Pashenova N. V., Pertsovaya A. A., Baranchikov Y. N., Kononov A. V., Ustyantsev K. V., Blinov A. G. Ophiostomatoid fungi associated with the four-eyed fir bark beetle on the territory of Russia // Russian Journal of Biological Invasions. 2018. V. 9. № 1. P. 63–74. DOI: 10.1134/S2075111718010137.
7. Санников П. Ю., Андреев Д. Н., Бузмаков С. А. Выявление и анализ сухостоя при помощи беспилотного летательного аппарата // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15. № 3. С. 103–113. DOI: 10.21046/2070-7401-2018-15-3-103-113.
8. Safonova A., Tabik S., Alcaraz-Segura D., Rubtsov A., Maglinets Y., and 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. № 6. P. 643. DOI: 10.3390/rs11060643.
9. Berie H. T., and Burud I. Application of unmanned aerial vehicles in earth resources monitoring: focus on evaluating potentials for forest monitoring in Ethiopia // *European Journal of Remote Sensing*. 2018. V. 51. Issue 1. P. 326–335. DOI: 10.1080/22797254.2018.1432993.
10. Chenari A., Erfanifard Y., Dehghani M., and Pourghasemi H. R. Woodland Mapping at Single-Tree Levels Using Object-Oriented Classification of UAV Images // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. V. XLII-4/W4. 2017. P. 43–49. DOI: 10.5194/isprs-archives-XLII-4-W4-43-2017.
11. Gini R., Sona G., Ronchetti G., Passoni D., and Pinto L. Improving Tree Species Classification Using UAS Multispectral Images and Texture Measures // *International Journal of Geo-Informatics*. 2018. V. 7. Issue 8. P. 315. DOI: 10.3390/ijgi7080315.
12. Lee S., Park S., Baek G., Kim H., and Lee C. Detection of Damaged Pine Tree by the Pine Wilt Disease Using UAV Image // *Korean Journal of Remote Sensing*. 2019. V. 35. № 3. P. 359–373. DOI:10.7780/kjrs.2019.35.3.2.
13. Onishi M., and 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> (дата обращения: 01.07.2020).
14. McInnes L., Healy J., and 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> (дата обращения: 01.07.2020).
15. Rousseeuw P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis // *Journal of Computational and Applied Mathematics*. 1987. V. 30. P. 33–65. DOI: 10.1016/0377-0427(87)90125-7.
16. Haralick R. M., Shanmugam K., and Dinstein I. Textural Features for Image Classification // *IEEE Transactions on Systems, Man, and Cybernetics*. 1973. Vol. SMC-3. No. 6. pp. 610–621.

17. Документация к функции `greycorprobs` [Электронный ресурс] / `scikit-image docs`. – URL: <https://scikit-image.org/docs/dev/api/skimimage.feature.html#skimimage.feature.greycorprobs> (дата обращения: 13.03.2021).
18. Friedman J. H. Greedy Function Approximation: A Gradient Boosting Machine // *The Annals of Statistics*. 2001. Vol. 29. No. 5. pp. 1189–1232.
19. Shelhamer E., Long J., and Darrell T. Fully Convolutional Networks for Semantic Segmentation // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017. V. 39. No. 4. pp. 640–651.
20. Ronneberger O., Fischer P., and 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> (дата обращения: 14.03.2021).
21. Ioffe S., and 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> (дата обращения: 07.07.2020).
22. Badrinarayanan V., Kendall A., and Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation: arXiv preprint arXiv:1511.00561. 2015. 14 p. – URL: <https://arxiv.org/abs/1511.00561> (дата обращения: 13.03.2021).
23. Paszke A., Chaurasia A., Kim S., and 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> (дата обращения: 13.03.2021).
24. He K., Zhang X., Ren S., and Sun J. Deep Residual Learning for Image Recognition: arXiv preprint arXiv:1512.03385. 2015. 12 p. – URL: <https://arxiv.org/abs/1512.03385> (дата обращения: 21.05.2021).
25. He K., Zhang X., Ren S., and Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification: arXiv preprint arXiv:1502.01852. 2015. 11 p. – URL: <https://arxiv.org/abs/1502.01852> (дата обращения: 22.05.2021).
26. Rahman M. A., and Wang Y. (2016) Optimizing Intersection-Over-Union in

- Deep Neural Networks for Image Segmentation. In: Bebis G. et al. (eds) Advances in Visual Computing. ISVC 2016. Lecture Notes in Computer Science. Springer, Cham. 2016. V. 10072. P. 234–244. DOI: 10.1007/978-3-319-50835-1_22.
27. Bertels J., Eelbode T., Berman M., Vandermeulen D., Maes F., Bisschops R., and Blaschko M. Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory & Practice: arXiv preprint arXiv:1911.01685. 2019. 9 p. – URL: <https://arxiv.org/abs/1911.01685> (дата обращения: 18.05.2021).
28. Simard P.Y., Steinkraus D., and Platt J.C. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis // Proc. of the International Conference on Document Analysis and Recognition. 2003. pp. 958–963.
29. Kingma D. P., and Ba J. Adam: A Method for Stochastic Optimization: arXiv preprint arXiv:1412.6980. 2015. 15 p. – URL: <https://arxiv.org/abs/1412.6980> (дата обращения: 01.07.2020).
30. Jones D.R., Schonlau M., and Welch W.J. Efficient Global Optimization of Expensive Black-Box Functions // Journal of Global Optimization. 1998. Vol. 13. pp. 455–492.
31. Krasser M. Bayesian optimization [Электронный ресурс] / github.io. – URL: <http://krasserm.github.io/> (дата обращения: 18.05.2021).
32. Google Colaboratory [Электронный ресурс] / Google research. – URL: <https://research.google.com/colaboratory/> (дата обращения: 23.05.2021).
33. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 27.12.2018) [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/12125268> (дата обращения: 23.05.2021).
34. ГОСТ 12.2.032-78 Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя. Общие эргономические требования [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/3924216> (дата обращения: 01.06.2021).
35. СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/400274954/> (дата

обращения: 01.06.2021).

36. ГОСТ 12.0.003-2015 Опасные и вредные производственные факторы. Классификация [Электронный ресурс] / Электронный фонд. – URL: <https://docs.cntd.ru/document/120013607> (дата обращения: 03.06.2021).

37. СанПиН 2.2.4.548–96 Гигиенические требования к микроклимату производственных помещений [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/4173106/> (дата обращения: 03.06.2021).

38. СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95 [Электронный ресурс] / КонсультантПлюс. – URL: http://www.consultant.ru/document/cons_doc_LAW_237487/ (дата обращения: 04.06.2021).

39. ТОИ Р-45-084-01 Типовая инструкция по охране труда при работе на персональном компьютере [Электронный ресурс] / Электронный фонд. – URL: <https://docs.cntd.ru/document/1200030047> (дата обращения: 03.06.2021).

40. СН 2.2.4/2.1.8.562-96 Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/4174553/> (дата обращения: 04.06.2021).

41. ГОСТ 12.1.038-82 ССБТ Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/3924202/> (дата обращения: 01.06.2021).

42. ГОСТ 12.1.006-84 ССБТ Электромагнитные поля радиочастот. Допустимые уровни на рабочих местах и требования к проведению контроля [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/6305025/> (дата обращения: 03.06.2021).

43. Федеральный закон № 89 от 1998 г. Об отходах производства и потребления [Электронный ресурс] / КонсультантПлюс. – URL: http://www.consultant.ru/document/cons_doc_LAW_19109/ (дата обращения: 05.06.2021).

44. ГОСТ Р 53692-2009 Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов [Электронный ресурс] / Гарант. – URL:

<https://base.garant.ru/2175160/> (дата обращения: 01.06.2021).

45. ГОСТ 12.1.004-91 Пожарная безопасность. Общие требования [Электронный ресурс] / Гарант. – URL: <https://base.garant.ru/3922226/> (дата обращения: 01.05.2021).

Приложение А
(обязательное)
Models and Algorithms

Студент:

Группа	ФИО	Подпись	Дата
8ИМ92	Маслов Константин Андреевич		

Руководитель ВКР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	Д.т.н., профессор		

Консультант-лингвист отделения иностранных языков ШБИП:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИЯ ШБИП	Коротченко Татьяна Валериевна	к.фил.н.		

A Models and Algorithms

A.1 Haralick Texture Features

To take into account not only the spectral features of the pixels, but also the texture features of the image area around them, it is possible to use the Haralick texture features [16].

To calculate the Haralick texture features, it is first necessary to calculate so-called gray level co-occurrence matrices (GLCMs), which count the number of pairwise occurrences of brightness in the original image for a certain direction and distance offset. In this work, four GLCMs were calculated for each channel of the original images for four predefined directions:

$$\text{GLCM}_{ij}(\theta_d, 0^\circ) = \left| \left\{ (k, l), (m, n) \in (L \times L) \times (L \times L) \mid k = m, |l - n| = \theta_d, I(k, l) = i, I(m, n) = j \right\} \right|, \quad (\text{A.1})$$

$$\text{GLCM}_{ij}(\theta_d, 45^\circ) = \left| \left\{ \begin{array}{l} (k, l), (m, n) \in (L \times L) \times (L \times L) \\ (k - m = \theta_d, l - n = -\theta_d) \text{ or } (k - m = -\theta_d, l - n = \theta_d), \\ I(k, l) = i, I(m, n) = j \end{array} \right\} \right|, \quad (\text{A.2})$$

$$\text{GLCM}_{ij}(\theta_d, 90^\circ) = \left| \left\{ (k, l), (m, n) \in (L \times L) \times (L \times L) \mid k - m = \theta_d, l = n, I(k, l) = i, I(m, n) = j \right\} \right|, \quad (\text{A.3})$$

$$\text{GLCM}_{ij}(\theta_d, 135^\circ) = \left| \left\{ \begin{array}{l} (k, l), (m, n) \in (L \times L) \times (L \times L) \\ (k - m = \theta_d, l - n = \theta_d) \text{ or } (k - m = -\theta_d, l - n = -\theta_d), \\ I(k, l) = i, I(m, n) = j \end{array} \right\} \right|, \quad (\text{A.4})$$

where $L = \{1, 2, \dots, \theta_{ps}\}$, θ_{ps} is a patch size around the central pixel (a window size), θ_d is the pixel pair distance offset.

The obtained GLCMs are used to calculate the descriptors of the original area around the central pixel. In this work, six texture features implemented in the scikit-image library [17] were used:

$$\text{Contrast} = \sum_{i,j}^{255} \text{GLCM}_{ij} (i - j)^2, \quad (\text{A.5})$$

$$\text{Dissimilarity} = \sum_{i,j}^{255} \text{GLCM}_{ij} |i - j|, \quad (\text{A.6})$$

$$\text{Homogeneity} = \sum_{i,j}^{255} \frac{\text{GLCM}_{ij}}{1 + (i - j)^2}, \quad (\text{A.7})$$

$$\text{ASM} = \sum_{i,j}^{255} \text{GLCM}_{ij}^2, \quad (\text{A.8})$$

$$\text{Energy} = \sqrt{\text{ASM}}, \quad (\text{A.9})$$

$$\text{Correlation} = \sum_{i,j}^{255} \text{GLCM}_{ij} \frac{(1-\mu_i)(1-\mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \quad (\text{A.10})$$

A.2 Gradient Boosting

To classify pixels according to their spectral and texture features in their vicinity, the gradient boosting algorithm was used. Gradient boosting is a technique for constructing ensembles of weak models, in which each next classifier in the ensemble tries to correct the mistakes of the previous ones [18]. Usually, decision trees are used as weak classifiers for combining into an ensemble, the ensemble is then called gradient-boosted trees.

To train the gradient-boosted trees, pixels were extracted from the training images with a step of 5. The spectral features of the pixels and the calculated Haralick texture features in its vicinity were taken as training features. Data augmentation was used for the training samples, more samples were synthesized by adding Gaussian noise to the whole patches of the images. More synthesized samples were added to the less represented classes in the dataset. During the training, crossentropy loss function was minimized.

A.3 Fully Convolutional Networks

Fully convolutional networks (FCNs) are a class of convolutional neural networks (CNNs). A typical FCN consists of an encoder and a decoder. The purpose of the encoder is the same as the purpose of the backbone in CNNs; it transforms the image into its hidden representation with convolutional and pooling or other downsampling layers. The decoder transforms the hidden representation into the resulting segmentation map, therefore, allowing classifying pixels in batches, not in one-by-one manner. The decoder usually includes convolutions and upsampling or deconvolution layers. Most often, the last layer of an FCN is convolutional with 1 by 1 kernels to obtain needed amount of outputs [19].

The following subsections consider five FCN architectures: U-Net, SegNet, ENet, MH-U-Net and MH-Res-U-Net.

A.3.1 U-Net

One of the selected FCN architectures for solving the problem of tree segmentation was U-Net. It was originally designed and developed to solve the problem of biomedical image segmentation, where its extremely high efficiency has been shown. It is still considered among state-of-the-art models. A distinctive feature of U-Net is the presence of concatenation operations connecting feature maps in the encoder with feature maps in the decoder in order to increase the detail of the resulting segmentation maps [20].

Figure A.1 shows the architecture of the model. Every rectangle depicts a tensor, or a multidimensional array representing a set of feature maps. The numbers correspond to the sizes of the tensors. The arrows depict the following operations:

- convolution (Conv3x3, Conv1x1);
- ELU calculation;
- batch normalization (BN) [21];
- downsampling by choosing the maximum value in the 2 by 2 patches of the features maps (MaxPooling);
- upsampling with the nearest neighbor interpolation (UpSampling);
- copying of the tensor and its concatenation with another one (Copying+Concatenation);
- dropping out values of the tensor by setting them to zero at random positions during the training (Dropout).

The categorical distribution at the output of the decoder is modeled for each pixel by applying the multivariate logistic function, Softmax.

In contrast to the original U-Net, the described model has the following minor modifications:

- the input image of the network is represented by a 256x256x3 tensor, which corresponds to an ordinary RGB image;

- convolutions do not reduce the size of the feature maps;
- cropping is not used;
- batch normalization is applied after every non-linearity;
- the output tensor is calculated using C convolutions with 1×1 kernels, thereby allowing to classify the pixels of C classes.

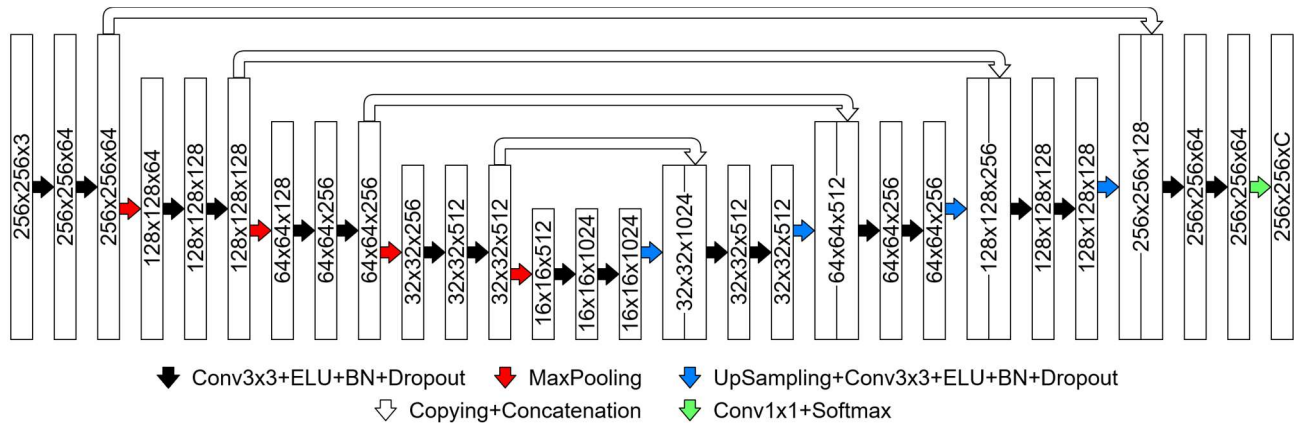


Figure A.1 – U-Net architecture

A.3.2 SegNet

Another fully convolutional network for solving the problems was the SegNet model [22]. Similar minor modifications were made to it, as it was done for the U-Net model. Also, a less deep version of the neural network was considered than given in the original paper. The architecture of the SegNet network implemented in this work is shown in figure A.2.

The fundamental difference between SegNet and U-Net architectures is the way to restore the original image resolution in the decoder: instead of completely copying feature maps and upsampling, SegNet utilizes MaxUnPooling, performed using the saved pixel positions by the MaxPooling operations in the encoder.

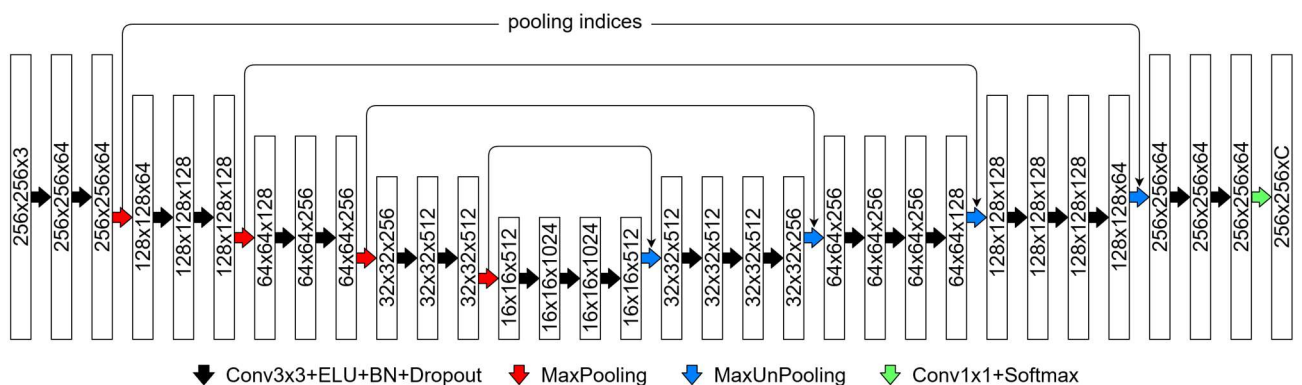


Figure A.2 – SegNet architecture

A.3.3 ENet

Also in the work, fully convolutional network ENet was used. No changes were made to it, except for changing the sizes of the tensors. The network was implemented according to the description given in [23].

The ENet architecture is presented in table A.1; its blocks are described in figure A.3.

Table A.1 –ENet architecture. Adapted from [23]

Section	Name	Type	Output size
–	initial		128x128x16
1	bottleneck1.0 bottleneck1.1 bottleneck1.2 bottleneck1.3 bottleneck1.4	downsampling	64x64x64 64x64x64 64x64x64 64x64x64 64x64x64
2	bottleneck2.0 bottleneck2.1 bottleneck2.2 bottleneck2.3 bottleneck2.4 bottleneck2.5 bottleneck2.6 bottleneck2.7 bottleneck2.8	downsampling dilated 2 asymmetric 5 dilated 4 dilated 8 asymmetric 5 dilated 10	32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128
3	Bottleneck3.1 bottleneck3.2 bottleneck3.3 bottleneck3.4 bottleneck3.5 bottleneck3.6 bottleneck3.7 bottleneck3.8	 dilated 2 asymmetric 5 dilated 4 dilated 8 asymmetric 5 dilated 10	32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128 32x32x128
4	bottleneck4.0 bottleneck4.1 bottleneck4.2	upsampling	64x64x64 64x64x64 64x64x64
5	bottleneck5.0 bottleneck5.1	upsampling	128x128x16 128x128x16
–	deconvolution		256x256xC

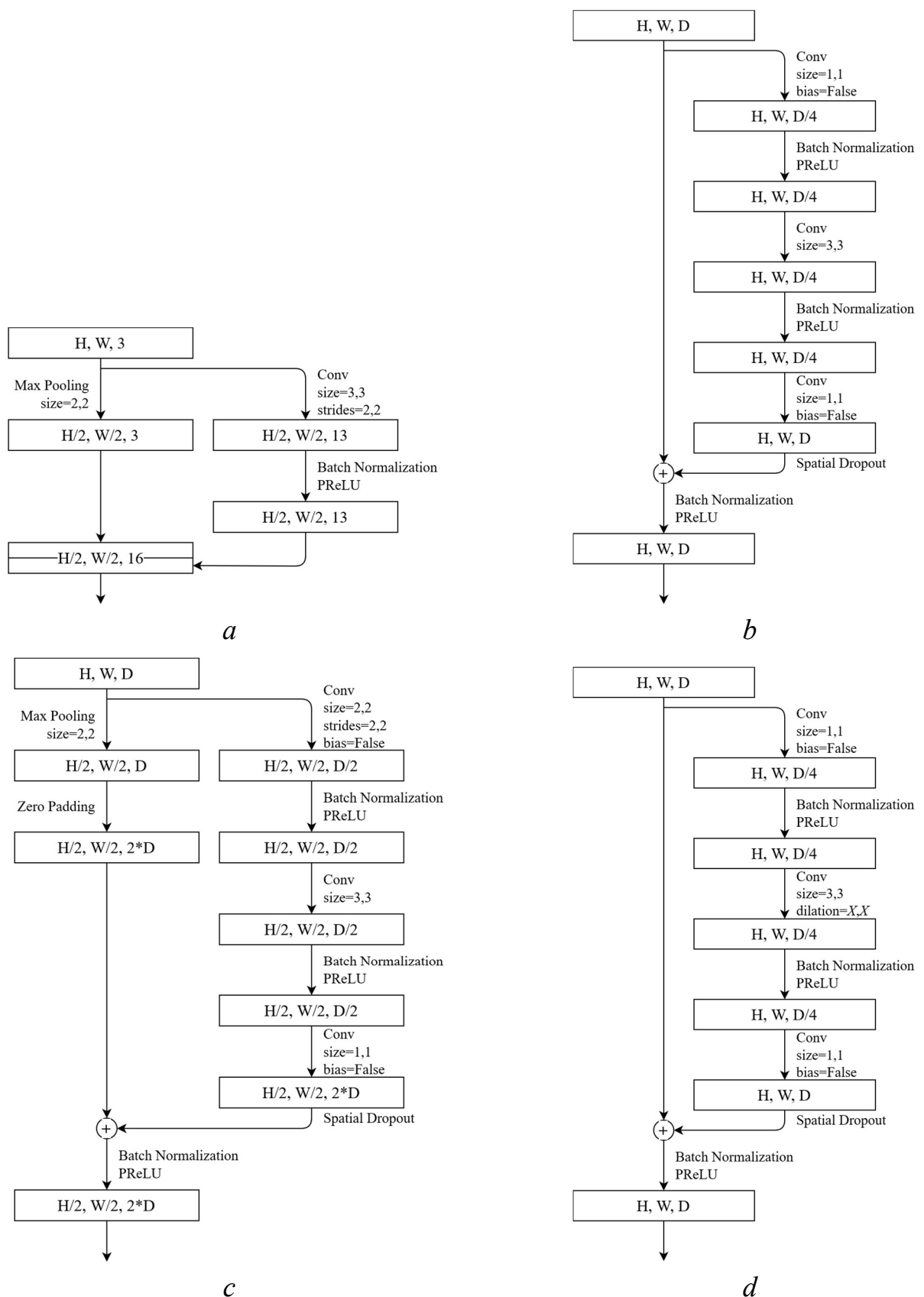


Figure A.3 – E-Net blocks: a – initial, b – bottleneck, c – bottleneck downsampling, d – bottleneck dilated X , e – bottleneck asymmetric X , f – bottleneck upsampling. H , W and D denote, respectively, the height, the width and the amount of feature maps of the input tensor. Adapted from [23]

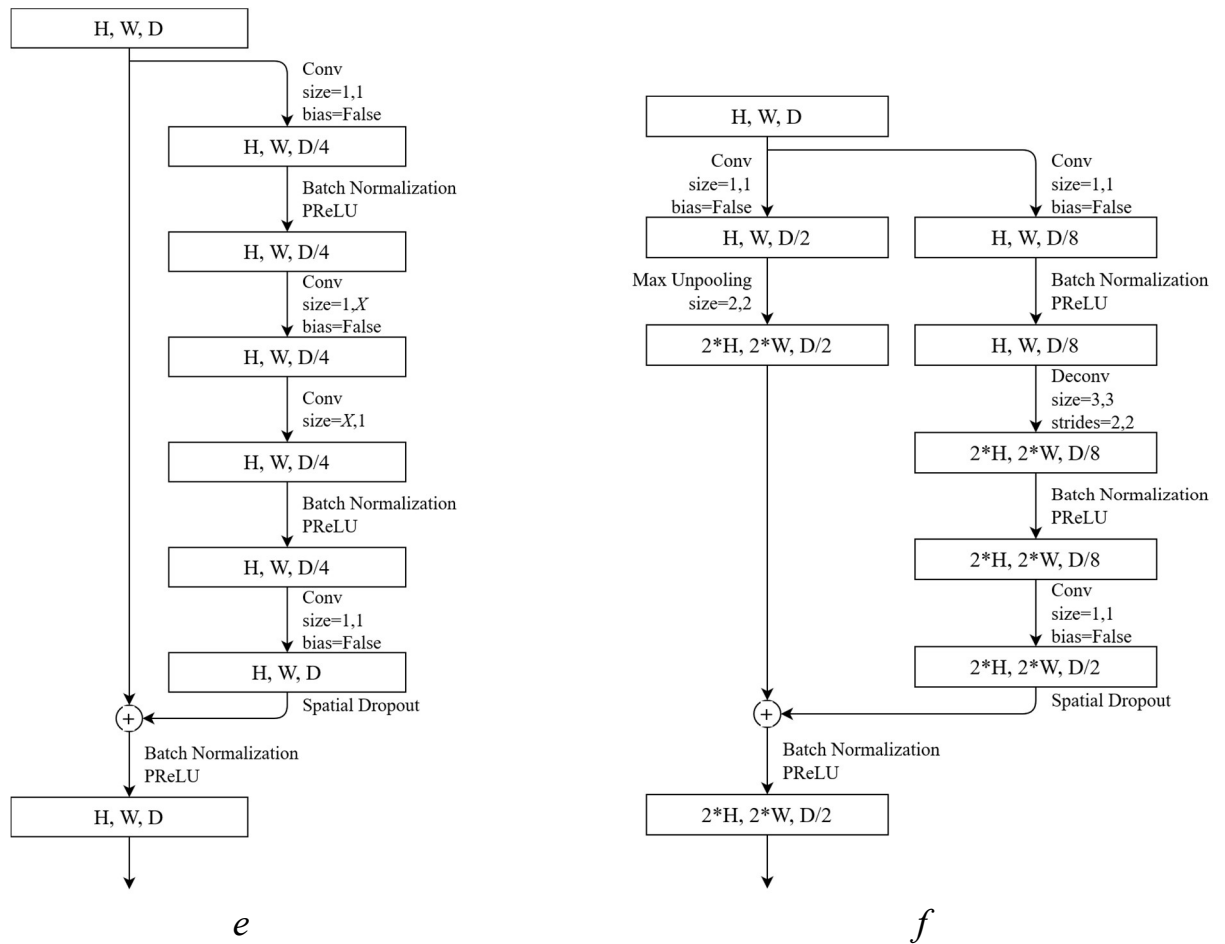


Figure A.3 (continued)

The following features of the ENet architecture are essential:

- the relatively shallow decoder compared to the encoder;
- significant downscaling of the feature maps in the first layers;
- active use of residual-blocks [24];
- active use of the PReLU activation function, which allows training non-linearities [25];
- application of the various types of convolutions (ordinary, dilated/atrous and asymmetric) in order to increase the perception area and reduce the number of the model weights.

In connection with the indicated features of ENet, it allows real-time semantic segmentation of images, which is useful when used on mobile devices or when implementing computer vision systems on, for instance, field-programmable gate arrays.

A.3.4 MH-U-Net

This thesis proposes two new FCN models for semantic image segmentation. One of them is called multihead-U-Net (MH-U-Net). The model is based on the U-Net architecture but introduces major modifications of its topology:

- instead of one, it utilizes three inputs of different scales (256x256x3, 128x128x3 and 64x64x3);
- it also has three outputs of different scales;
- dropout is replaced with spatial dropout, as it is in ENet.

Figure A.4 demonstrates the architecture of the proposed neural network.

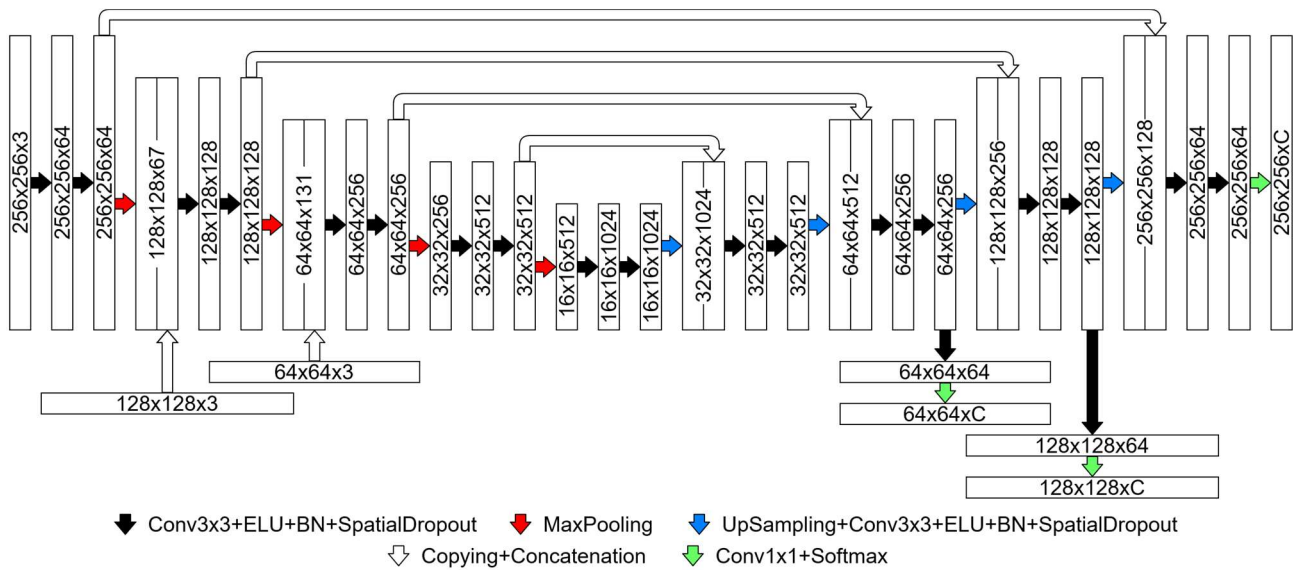


Figure A.4 – MH-U-Net architecture

During the training, loss function is calculated for every output, and the training algorithm minimizes the sum of all three values. During the testing, the two outputs of smaller scales and the corresponding branches of the network can be neglected.

The model can be considered an ensemble of three U-Nets of different depths, the models share common weights and try to analyse an image in its different scales, therefore, it is expected that MH-U-Net learns a more sensible hidden representation of the data. Furthermore, it can be speculated that such an approach strongly affects the process of training, thus, giving better results and reducing the sensitivity to hyperparameters, as it has been shown (see subsection 4.3).

A.3.5 MH-Res-U-Net

The idea of MH-U-Net was elaborated resulting in the multihead-residual-U-Net (MH-Res-U-Net) architecture. MH-Res-U-Net shares all the features of the MH-U-Net and makes use of residual blocks [24].

The architecture of the MH-Res-U-Net model is presented in figure A.5.

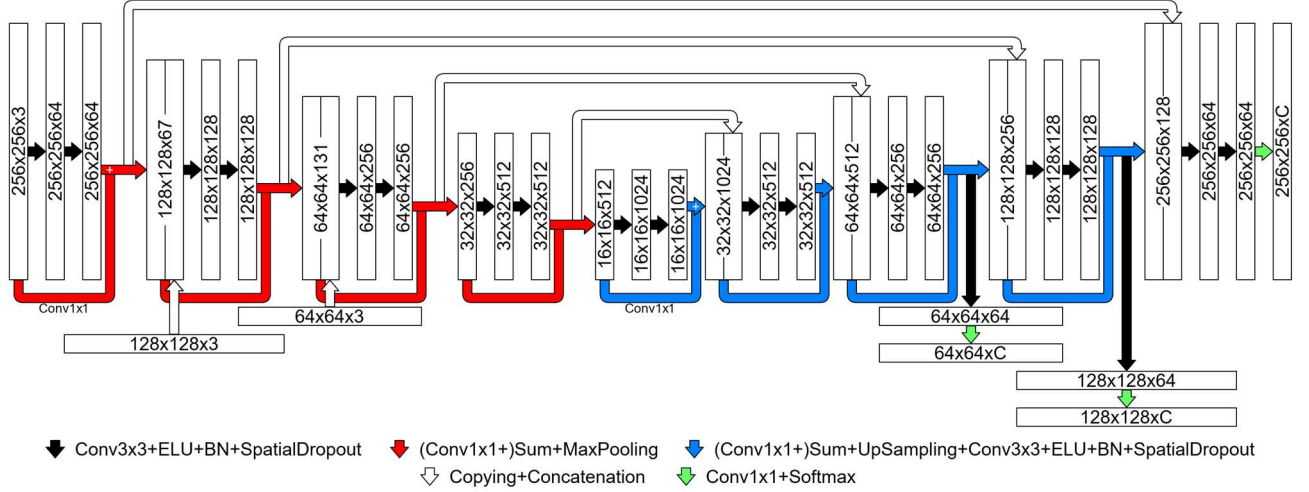


Figure A.5 – MH-Res-U-Net architecture

Note that projection (Conv1x1) operations in the main paths of the residual blocks do not involve any non-linearities and bias coefficients. It is expected that the residual blocks in MH-Res-U-Net will lead to a more stable learning and better gradient propagation.

The results of the analysis of the MH-U-Net and MH-Res-U-Net models are presented in section 4.

A.4 Model Evaluation

In order to evaluate the quality of the models, intersection over union (IoU), or Jaccard coefficient, was used. It is considered a widely recognized metric in the problems of semantic image segmentation. IoU can be calculated as [26, 27]:

$$\text{IoU}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c + \text{FN}_c}, \quad (\text{A.11})$$

where TP_c , FP_c and FN_c are, respectively, true positive, false positive and false negative results for class c .

IoU values greater than 0.5 are considered to correspond to acceptably high segmentation quality. An aggregate indicator of the quality of the models could be

mean intersection over union (mIoU), or mean Jaccard coefficient, which is average of IoU by class:

$$\text{mIoU} = \frac{1}{C} \sum_{c=1}^C \text{IoU}_c, \quad (\text{A.12})$$

where C is the number of classes in the segmentation problem.

3.5 Neural Network Training

The procedure of histograms matching was carried out in order to decrease the difference in the illumination conditions of different scenes in the dataset for *Abies Sibirica* trees, while one image from the training set was selected as a reference image. For *Pinus Sibirica*, though, the same procedure was not conducted due to the abundance of the black pixels at the boundary regions of the panoramas.

For the formation of training and validation sets for the FCNs, the images from the dataset were cut into fragments of 256 by 256 pixels with an offset of 128 pixels. Data leakage from the validation set was eliminated by imposing Boolean masks on the original images and blocking the possibility of moving the corresponding pixels into the training set. Thus, it was ensured that the features of the validation samples are not repeated among the training samples. In total, 2004 training and 672 validation fragments for the problem of segmentation of *Abies Sibirica* trees and 12384 training and 2400 validation fragments for the problem of segmentation of *Pinus Sibirica* trees were obtained.

Online augmentation technique was applied in order to improve the generalizing ability of the FCN models. Online augmentation allows expanding the training set during the model fitting with randomly synthesized data. The augmentation included:

- changes of the image scale;
- changes of its brightness;
- changes of its contrast;
- vertical flips;
- elastic transformations [28].

In this work, to find optimal FCN weights, mIoU was maximized and the gradient-based optimization algorithm Adam was used [29]. Optimization of mIoU allows to partially solve the issue of the imbalanced datasets shown earlier. To get rid of the maximization problem and get the minimization problem, the optimization problem can be defined as:

$$1 - \text{mIoU} \rightarrow \min. \quad (\text{A.13})$$

During the training, it is inconvenient and computationally expensive to calculate confusion matrices to get the values of TP_c , FP_c and FN_c , so instead of (A.11) and (A.12) one can use an approximation of mIoU, or soft Jaccard coefficient [26, 27], with some parametrizations:

$$\text{mIoU}(\mathbf{T}, \mathbf{P}) = \frac{1}{C} \sum_{c=1}^C \left(\frac{\sum_{i=1}^H \sum_{j=1}^W \text{LS}(\mathbf{T})_{ijc} \cdot \mathbf{P}_{ijc} + \theta_s}{\sum_{i=1}^H \sum_{j=1}^W [\text{LS}(\mathbf{T})_{ijc} + (1 - \text{LS}(\mathbf{T})_{ijc}) \cdot \mathbf{P}_{ijc}] + \theta_s} \right), \quad (\text{A.14})$$

$$\text{LS}(\mathbf{T}) = (1 - \theta_{ls}) \cdot \mathbf{T} + \frac{\theta_{ls}}{C}, \quad (\text{A.15})$$

where \mathbf{T} is an expected output, \mathbf{P} is a predicted output of the FCN, H is the image height, W is the image width, $\theta_s > 0$ is a loss function smoothing parameter eliminating the singularity point and large gradient values of the objective function, $0 \leq \theta_{ls} < 1$ is a label smoothing coefficient making the resulting models less confident in their predictions, therefore, assisting to avoid saturation effect in the final layer and keeping the FCN weights volatile for a longer time.

It is worth noting that the derivative of (A.14) with respect to \mathbf{P} is easy to calculate with almost any modern deep learning framework. It allows using gradient-based optimization methods to find optimal or nearly optimal model weights.

3.6 Hyperparameter Optimization

To gain a significant increase in model quality, one should identify a list of hyperparameters and perform hyperparameter optimization:

$$\text{mIoU}(\boldsymbol{\theta}) \rightarrow \max_{\boldsymbol{\theta}} \quad (\text{A.16})$$

Identified hyperparameters for the FCNs, gradient boosting learning algorithm coupled with the extraction of Haralick texture features and the corresponding algorithms of data augmentation and training are listed in table A.2.

Table A.2 – The list of the hyperparameters

Name	Description
Gradient boosting + Haralick texture features	
θ_{ne}	Maximum number of decision trees in an ensemble
θ_{md}	Maximum depth of a decision tree
θ_{lr}	Learning rate
θ_{ps}	Patch size
θ_d	Pixel pair distance offset
θ_n	Standard deviation of Gaussian noise for augmentation
Fully convolutional networks	
θ_{lr}	Learning rate
θ_d	Dropout coefficient
θ_s	Loss function smoothing coefficient
θ_{ls}	Label smoothing coefficient
θ_z	Coefficient of scale change
θ_b	Coefficient of brightness change
θ_c	Coefficient of contrast change
θ_α	Coefficients of elastic transformations [28]
θ_σ	
θ_{pf}	Learning rate drop factor on reaching a “plateau”
θ_{pp}	Number of epochs to confirm reaching a “plateau”

To find the best sets of the hyperparameters for the proposed models, Bayesian approach to optimization was utilized [31]. The idea behind the algorithm is to replace the real dependency $mIoU(\theta)$, which is a black-box function with no analytical expressions for itself or its derivatives, with a surrogate model

$$\widehat{mIoU}(\theta) = mIoU(\theta) + \varepsilon, \quad \varepsilon \sim N(0, s^2) \quad (A.17)$$

and solve the optimization problem (A.16) for the surrogate model. In this thesis, $mIoU(\theta)$ is modelled as a Gaussian process as it provides the following advantages:

- it allows modelling arbitrary complex functions, as it is a non-parametric method;
- possibility of uncertainty estimation, which Gaussian processes provide, can be utilized to balance between exploitation and exploration.

Figure A.6 depicts a UML-diagram of the algorithm for hyperparameter tuning.

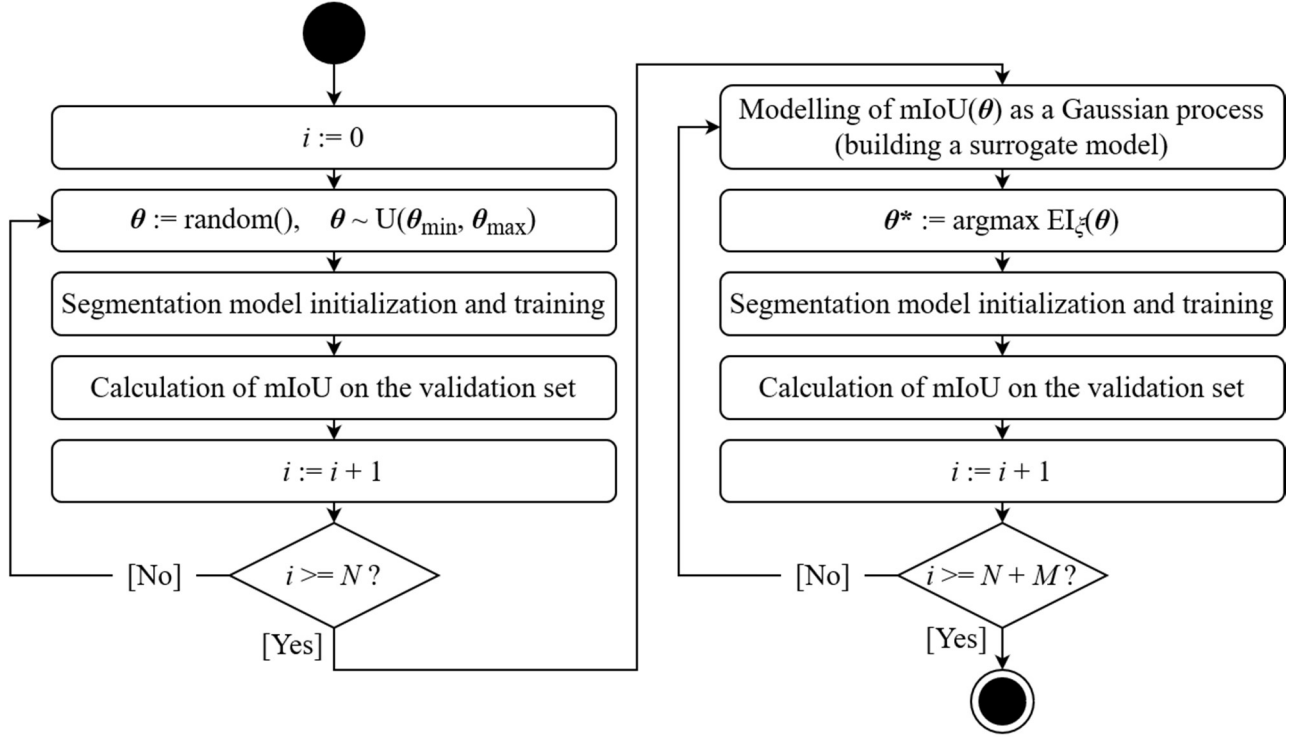


Figure A.6 – Hyperparameter tuning algorithm

The algorithm consists of two stages. First, it samples N hyperparameter vectors from a predefined uniform distribution, initializes the corresponding models, fits them to the data and evaluates them, i.e. calculates mIoUs on the validation set. Second, it uses these N pairs of the hyperparameter vectors and the corresponding mIoU values to build a surrogate model (A.17). Then it maximizes an acquisition function. In this thesis, expected improvement (EI) parametrized with ξ [30, 31] was used as an acquisition function:

$$\text{EI}_{\xi}(\boldsymbol{\theta}) = \mathbb{E}_{\xi}[\max(\widehat{\text{mIoU}}(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+), 0)] \quad (\text{A.18})$$

or

$$\text{EI}_{\xi}(\boldsymbol{\theta}) = \begin{cases} \left(\mu(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+) - \xi \right) \Phi(Z) + \sigma(\boldsymbol{\theta}) \phi(Z) & \text{if } \sigma(\boldsymbol{\theta}) > 0 \\ 0 & \text{if } \sigma(\boldsymbol{\theta}) = 0 \end{cases}, \quad (\text{A.19})$$

$$Z = \begin{cases} \frac{\mu(\boldsymbol{\theta}) - \widehat{\text{mIoU}}(\boldsymbol{\theta}^+) - \xi}{\sigma(\boldsymbol{\theta})} & \text{if } \sigma(\boldsymbol{\theta}) > 0 \\ 0 & \text{if } \sigma(\boldsymbol{\theta}) = 0 \end{cases}, \quad (\text{A.20})$$

where θ^+ is the best set of the hyperparameters sampled, $\mu(\theta)$ is the mean, $\sigma(\theta)$ is the standard deviation of the surrogate model posterior predictive at θ , Φ is the cumulative distribution function and ϕ is the probability density function of $N(0, 1)$. Parameter ξ balances exploitation and exploration, higher values of ξ lead to more frequent sampling of the hyperparameters from the regions of high uncertainty. To find a new candidate for the best set of the hyperparameters, the following optimization problem is to be solved:

$$\theta^* = \arg \max_{\theta} \text{EI}_{\xi}(\theta). \quad (\text{A.21})$$

The algorithm solves (A.21) M times and saves the results to rebuild the surrogate model at each iteration. The thesis uses values of $N = M = 30$ for gradient-boosted trees, $N = M = 40$ for neural networks and $\xi = 0.15$.

Приложение Б
(справочное)

Дипломы, полученные на конференциях и конкурсах



TUSUR | TUSUR UNIVERSITY

ДИПЛОМ I СТЕПЕНИ

награждается

Маслов

Константин

Андреевич

за лучший доклад на
Международной научно-технической конференции
студентов, аспирантов и молодых ученых
«**НАУЧНАЯ СЕССИЯ ТУСУР**»
Подсекция 3.4 «Вычислительный интеллект»

19-21 мая 2021 г., Томск

Председатель конференции
Ректор ТУСУРа



В.М. Рулевский

Рисунок Б.1 – Диплом первой степени, «Научная сессия ТУСУР – 2021»

ДИПЛОМ

I степени

НАГРАЖДАЕТСЯ

МАСЛОВ КОНСТАНТИН АНДРЕЕВИЧ

магистрант Томского политехнического университета

за доклад, представленный на

**XVIII Международной научно-практической конференции
студентов, аспирантов и молодых ученых**

«Молодёжь и современные информационные технологии»

**«ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО И ГЛУБОКОГО ОБУЧЕНИЯ
В ЗАДАЧАХ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ
ЛЕСНОГО ПОКРОВА»**

22-26 марта 2021 г.

г. Томск

**Председатель организационного
комитета МСИТ-2021**

 **Н.Г. Марков**

**Проректор по науке
Томского политехнического университета**

 **М.С. Юсубов**



Рисунок Б.2 – Диплом первой степени, «Молодежь и современные информационные технологии 2021»



ТУСУР | TUSUR
UNIVERSITY

ДИПЛОМ I СТЕПЕНИ

награждается

Маслов Константин Андреевич

за лучший доклад на подсекции
3.4 «Вычислительный интеллект»
Международной научно-технической конференции
студентов, аспирантов и молодых ученых
«**НАУЧНАЯ СЕССИЯ ТУСУР**»
25-27 мая 2020 г., Томск

Председатель конференции
Ректор ТУСУРа



В.М. Рулевский

Рисунок Б.3 – Диплом первой степени, «Научная сессия ТУСУР – 2020»

Министерство науки и высшего образования РФ
Национальный исследовательский Томский политехнический университет
Инженерная школа природных ресурсов



ДИПЛОМ ИШПР

МАСЛОВУ КОНСТАНТИНУ АНДРЕЕВИЧУ

магистранту

Национального исследовательского Томского политехнического университета
за доклад

«КЛАССИФИКАЦИЯ ПОВРЕЖДЕННОЙ УССУРИЙСКИМ ПОЛИГРАФОМ
ПИХТЫ НА СНИМКАХ С БПЛА МЕТОДАМИ КОМПЬЮТЕРНОГО ЗРЕНИЯ»,

представленный на XXIV Международном научном симпозиуме

студентов и молодых ученых

имени академика М.А. Усова

«ПРОБЛЕМЫ ГЕОЛОГИИ И ОСВОЕНИЯ НЕДР»

6 – 10 апреля 2020 года

г. Томск

Председатель оргкомитета,
директор ИШПР ТПУ,
К.Х.Н.



Рисунок Б.4 – Диплом ИШПР, симпозиум имени академика М. А. Усова, 2020 г.

ДИПЛОМ

I степени

НАГРАЖДАЕТСЯ

МАСЛОВ КОНСТАНТИН АНДРЕЕВИЧ

магистрант Томского политехнического университета

*за доклады, представленные на
XVII Международной научно-практической конференции
студентов, аспирантов и молодых ученых
«Молодёжь и современные информационные технологии»*

1. «ДЕТЕРМИНИРОВАННЫЙ АЛГОРИТМ КЛАССИФИКАЦИИ ПИХТЫ, ПОВРЕЖДЕННОЙ УССУРИЙСКИМ ПОЛИГРАФОМ, ПО СНИМКАМ С БПЛА»
2. «КЛАССИФИКАЦИЯ ПИХТЫ, ПОВРЕЖДЕННОЙ УССУРИЙСКИМ ПОЛИГРАФОМ, ПО СНИМКАМ С БПЛА С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ U-NET»

17-20 февраля 2020 г.

г. Томск

Зам. председателя организационного
комитета МСИТ-2020



В.Г. Спицын

Проректор по научной
работе и инновациям ППУ



М.С. Юсубов

Рисунок Б.5 – Диплом первой степени, «Молодежь и современные информационные технологии 2020»

XIII

СИБИРСКОЕ СОВЕЩАНИЕ И ШКОЛА МОЛОДЫХ УЧЕНЫХ ПО
КЛИМАТО-ЭКОЛОГИЧЕСКОМУ МОНИТОРИНГУ
ТОМСК, 15-19 ОКТЯБРЯ 2019

ДИПЛОМ

ПОБЕДИТЕЛЯ КОНКУРСА ДОКЛАДОВ МОЛОДЫХ УЧЕНЫХ

НАГРАЖДАЕТСЯ

Маслов Константин Андреевич

ПРЕДСЕДАТЕЛЬ ОРГАНИЗАЦИОННОГО КОМИТЕТА
ЧЛЕН-КОРРЕСПОНДЕНТ РАН



Югорский
государственный
университет



М.В. КАБАНОВ

Рисунок Б.6 – Диплом победителя, Сибирское совещание по климато-экологическому мониторингу, 2019 г.

TOMSK
POLYTECHNIC
UNIVERSITY



ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ



**ЛУЧШИЙ
СТУДЕНТ
ТПУ 2021**

ДИПЛОМ

Лауреата III СТЕПЕНИ

XXIII Университетского конкурса
на соискание звания «Лучший студент ТПУ»
в номинации «Лучший в науке»
в 2021 году

НАГРАЖДАЕТСЯ

Маслов Константин Андреевич,

магистрант группы 8ИМ92

Инженерной школы информационных технологий и робототехники

Научный руководитель – Марков Николай Григорьевич,

д.т.н., профессор Отделения информационных технологий

Инженерной школы информационных технологий и робототехники

Врио ректора



А.А. Яковлев

Томск 2021

Рисунок Б.7 – Лауреат третьей степени, «Лучший студент ТПУ 2021»

ДИПЛОМ

Лауреата III СТЕПЕНИ

XXII Университетского конкурса
на соискание звания «Лучший студент ТПУ»
в 2020 году

НАГРАЖДАЕТСЯ

Маслов Константин Андреевич,
магистрант группы 8ИМ92

Инженерной школы информационных технологий и
робототехники,

Научный руководитель – Токарева Ольга Сергеевна,
к.т.н., доцент Отделения информационных технологий
Инженерной школы информационных технологий и робототехники



Врио ректора

В.В. Дёмин

Томск 2020

Рисунок Б.8 – Лауреат третьей степени, «Лучший студент ТПУ 2020»

Приложение В
(справочное)

Акт об использовании результатов магистерской диссертации

Минобрнауки России

Федеральное государственное бюджетное учреждение науки
Институт мониторинга климатических и экологических систем
Сибирского отделения Российской академии наук
(ИМКЭС СО РАН)

Академический пр., 10/3, г.Томск, 634055, Россия
Тел. (382-2) 492-265. Факс (382-2) 491-950 e-mail: post@imces.ru http://www.imces.ru
ОКПО 03534200, ОГРН 1027000880170, ИНН/КПП 7021001400 / 701701001

08.06.2021 № 15323/07-181

На № _____ от _____

АКТ

об использовании результатов магистерской диссертации
«Модели, алгоритмы и программное обеспечение для семантической сегментации хвойных
деревьев на снимках с беспилотных летательных аппаратов»
Маслова Константина Андреевича

Комиссия в составе:

Председатель: Керчев Иван Андреевич, к.б.н. старший научный сотрудник ЛМЛЭС,

Члены комиссии:

– Бисирова Эльвина Михайловна, научный сотрудник ЛМЛЭС,

– Пац Елена Николаевна, к.б.н. научный сотрудник ЛМЛЭС

составила настоящий акт о том, что результаты магистерской диссертации «Модели, алгоритмы и программное обеспечение для семантической сегментации хвойных деревьев на снимках с беспилотных летательных аппаратов» студента Национального исследовательского Томского политехнического университета Маслова К. А. апробированы в лаборатории мониторинга лесных экосистем, а именно полученные карты сегментации изображений поврежденных пихтовых и кедровых насаждений.

Использование указанных результатов позволяет значительно сократить время на проведение оценки состояния хвойных лесов, проводить пространственный анализ характера повреждения древостоев насекомыми – лесными вредителями, а также упрощает процесс планирования санитарных и лесозащитных мероприятий.

Председатель комиссии




Керчев И.А.

Члены комиссии



Бисирова Э.М.



Пац Е.Н.

МП

Подписи Керчева И.А., Бисировой Э.М., Пац Е.Н. заверяю.

Ученый секретарь
Института мониторинга климатических, к.т.н.
и экологических систем СО РАН



О.В. Яблокова

Рисунок В.1 – Акт об использовании результатов магистерской диссертации