

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

Школа ИШИТР
 Направление подготовки Информационные системы и технологии
 Отделение школы (НОЦ) ОИТ

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Детектирование с помощью сверточных нейронных сетей класса YOLO на изображениях деревьев пихты, пораженных уссурийским полиграфом
УДК 004.492.3..032.26:004.932:633.877

Студент

Группа	ФИО	Подпись	Дата
8ИМ81	Фоминский Виталий Сергеевич		

Руководитель ВКР

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

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Конотопский Владимир Юрьевич	К.Э.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД	Горбенко Михаил Владимирович	К.Т.Н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Савельев Алексей Олегович	К.Т.Н.		

Томск – 2020 г.

Планируемые результаты обучения

Код результатов	Результат обучения (выпускник должен быть готов)
Профессиональные и общепрофессиональные компетенции	
P1	Применять базовые и специальные естественнонаучные и математические знания для комплексной инженерной деятельности по созданию, внедрению и эксплуатации геоинформационных систем и технологий, а также информационных систем и технологий в бизнесе.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием геоинформационных систем и технологий, информационных систем в бизнесе, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Выполнять комплексные инженерные проекты по созданию информационных систем и технологий, а также средств их реализации (информационных, методических, математических, алгоритмических, технических и программных).
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания геоинформационных систем и технологий, а также информационных систем и технологий в бизнесе.
P6	Внедрять, эксплуатировать и обслуживать современные геоинформационные системы и технологии, информационные системы и технологии в бизнесе, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
Универсальные (общекультурные) компетенции	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Осуществлять коммуникации в профессиональной среде и в обществе в целом. Владеть иностранным языком (углублённый английский язык), позволяющим работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена команды, состоящей из специалистов различных направлений и квалификаций,
P10	Демонстрировать личную ответственность за результаты работы и готовность следовать профессиональной этике и нормам ведения комплексной инженерной деятельности.
P11	Демонстрировать знания правовых, социальных, экологических и культурных аспектов комплексной инженерной деятельности, а также готовность к достижению должного уровня физической подготовленности для обеспечения полноценной социальной и профессиональной деятельности.

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

Школа ИШИТР
 Направление подготовки (специальность) Информационные системы и технологии
 Отделение школы (НОЦ) ОИТ

УТВЕРЖДАЮ:
 Руководитель ООП

 (Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ИМ81	Фоминский Виталий Сергеевич

Тема работы:

Детектирование с помощью сверточных нейронных сетей класса YOLO на изображениях деревьев пихты, пораженных уссурийским полиграфом	
Утверждена приказом директора (дата, номер)	28.02.2020, 59 – 46/с

Срок сдачи студентом выполненной работы:

--	--

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

<p>Исходные данные к работе</p>	<p>Объектом исследования данной работы являются сверточные нейронные сети класса YOLO.</p> <p>В качестве исходных данных выступали снимки лесных массивов Томской области с беспилотных летательных аппаратов.</p> <p>Исследуемые сверточные нейронные сети должны соответствовать следующим требованиям:</p> <ul style="list-style-type: none"> • Сети должны детектировать пихты четырех классов повреждения; • Сети должны функционировать в режиме реального времени;
--	---

Перечень подлежащих исследованию, проектированию и разработке вопросов	<ul style="list-style-type: none"> • Анализ предметной области • Подготовка данных • Обучение СНС • Анализ полученных результатов • Экономическая часть • Социальная ответственность
Перечень графического материала	<ul style="list-style-type: none"> • Снимки с беспилотного летательного аппарата • Результаты детектирования пихт с помощью сверточных нейронных сетей
Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Конотопский Владимир Юрьевич
Социальная ответственность	Горбенко Михаил Владимирович
Раздел на иностранном языке	Маркова Наталья Александровна
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Обучение сверточных нейронных сетей	

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

Задание выдал руководитель:

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

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ81	Фоминский Виталий Сергеевич		

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

Школа ИШИТР

Направление подготовки (специальность) Информационные системы и технологии

Уровень образования Магистратура

Отделение школы (НОЦ) ОИТ

Период выполнения _____ (осенний / весенний семестр 2019 /2020 учебного года)

Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
02.03.2020	Анализ предметной области	15
09.03.2020	Подготовка данных	15
25.05.2020	Обучение СНС	30
31.05.2020	Анализ полученных результатов	20
31.05.2020	Экономическая часть	10
03.06.2020	Социальная ответственность	10

СОСТАВИЛ:

Руководитель ВКР

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

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Савельев Алексей Олегович	к.т.н.		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
8ИМ81	Фоминский Виталий Сергеевич

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	Информационные системы и технологии

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

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	Использовать действующие ценники и договорные цены на потребленные материальные и информационные ресурсы, а также указанную в МУ величину тарифа на эл. энергию
2. <i>Нормы и нормативы расходования ресурсов</i>	—
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	Действующие ставки единого социального налога и НДС, ставка дисконтирования = 0,1 (см. МУ)

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. <i>Организация и планирование работ</i>	Дать характеристику существующих и потенциальных потребителей (покупателей) результатов ВКР, ожидаемых масштабов их использования
2. <i>Расчет сметы на выполнение проекта</i>	Разработать проект такого устава в случае, если для реализации результатов ВКР необходимо создание отдельной организации или отдельного структурного подразделения (возможно временного) внутри существующей организации
3. <i>Оценка экономической эффективности проекта</i>	Построение плана-графика выполнения ВКР, составление соответствующей сметы затрат, расчет цены результата ВКР.

Перечень графического материала:

<ol style="list-style-type: none"> 1. <i>График проведения и бюджет НИИ</i> 2. <i>Оценка ресурсной, финансовой и экономической эффективности НИИ</i>
--

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Конотопский Владимир Юрьевич	к.э.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ81	Фоминский Виталий Сергеевич		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8ИМ81	Фоминский Виталий Сергеевич

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	Информационные системы и технологии

Тема ВКР:

Детектирование с помощью сверточных нейронных сетей класса YOLO на изображениях деревьев пихты, пораженных уссурийским полиграфом	
Исходные данные к разделу «Социальная ответственность»:	
<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>В ходе работы были проведены исследования эффективности алгоритма YOLO v2 для детектирования пораженных пихтовых деревьев на аэрофотоснимке с беспилотного летательного аппарата.</p> <p>Все работы проводились дистанционно, в общежитии Томского политехнического университета с помощью персонального компьютера.</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. 	<p>В ходе выполнения работы были использованы следующие документы:</p> <ul style="list-style-type: none"> – ГОСТ 12.0.003-74. ССБТ. Опасные и вредные производственные факторы. Классификация; – "Трудовой кодекс Российской Федерации" от 30.12.2001 N 197-ФЗ (ред. от 27.12.2018); – СанПиН 2.2.4.548–96. Гигиенические требования к микроклимату производственных помещений; – СанПиН 2.2.2/2.4.1340–03. Санитарно-эпидемиологические правила и нормативы «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы»; – ГОСТ 12.1.003–83 ССБТ. Шум. Общие требования безопасности;
<p>2. Производственная безопасность:</p> <p>2.1. Анализ выявленных вредных и опасных факторов</p> <p>2.2. Обоснование мероприятий по снижению воздействия</p>	<p>В ходе работы были выявлены следующие вредные факторы:</p> <ul style="list-style-type: none"> – Повышенный уровень шума; – Неоптимальные показатели микроклимата; – Электромагнитное излучение; – Недостаточная или избыточная

	<p>освещенность;</p> <ul style="list-style-type: none"> – Умственное и эмоциональное напряжение; – Монотонность труда; – Напряжение анализаторов. <p>К наиболее вероятным опасным факторам можно отнести:</p> <ul style="list-style-type: none"> – Поражение электрическим током; – Пожары; – Чрезвычайные ситуации;
3. Экологическая безопасность:	<ul style="list-style-type: none"> – При разработке опасность представляет утилизация компонентов ПЭВМ и макулатуры.
4. Безопасность в чрезвычайных ситуациях:	<p>Наиболее вероятные ЧС в процессе разработки и эксплуатации:</p> <ul style="list-style-type: none"> – указать наиболее типичную ЧС. – Ураганы; – Пожары; – Эпидемии.

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД	Горбенко Михаил Владимирович	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ИМ81	Фоминский Виталий Сергеевич		

Реферат

Выпускная квалификационная работа 74 с., 35 рис., 27 табл., 29 ист.

Ключевые слова: YOLO, сверточные нейронные сети, детектирование объектов, пихтовые деревья, уссурийский полиграф.

Объект исследования данной работы: Сверточная нейронная сеть YOLO v2 применительно к детектированию пихтовых деревьев разной степени поражения уссурийским полиграфом на аэрофотоснимке.

Цель работы: Исследовать эффективность алгоритмов класса YOLO для детектирования пораженных деревьев на аэрофотоснимке.

Область применения: Детектирование деревьев различной степени поражения на снимках, полученных с беспилотного летательного аппарата.

Результаты: Были подготовлены данные для обучения нейронных сетей, произведено обучение и подбор основных гиперпараметров сверточных нейронных сетей YOLO v2 и YOLO v2 tiny, выполнено детектирование на тестовых данных.

Использованные термины и определения

YOLO – you only look once, архитектура нейронной сети для детектирования объектов на изображении.

БПЛА – беспилотный летательный аппарат

Компьютерное зрение – теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов.

СНС – сверточная нейронная сеть, специальная архитектура нейронных сетей, предложенная Яном Лекуном, изначально нацеленная на эффективное распознавание изображений.

Уссурийский полиграф – вид жуков-короедов. Опасный инвазивный вредитель-дендрофаг пихты.

Содержание

1	Анализ предметной области.....	14
1.1	Мониторинг лесов.....	14
1.2	Постановка задачи исследования.....	15
2	Обзор архитектур нейронных сетей.....	17
2.1	Задачи компьютерного зрения.....	17
2.2	Сверточная нейронная сеть YOLO v2.....	17
2.3	Сверточная нейронная сеть YOLO v3.....	21
2.4	Сверточная нейронная сеть RetinaNet.....	23
2.5	Сверточная нейронная сеть Faster R-CNN.....	24
2.6	Сравнение СНС различных архитектур.....	25
2.7	Обзор средств реализации СНС.....	26
3	Подготовка данных.....	27
4	Обучение СНС.....	30
4.1	Выбор якорных ограничивающих прямоугольников.....	30
4.2	Обучение и исследование СНС YOLO v2.....	32
4.3	Обучение и исследование СНС YOLO v2 tiny.....	38
4.4	Исследование алгоритма NMS.....	40
5	Анализ полученных результатов.....	42
6	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	43
6.1	Организация и планирование работ.....	43
6.1.1	Структура работ в рамках научного исследования.....	43
6.1.2	Определение трудоемкости выполнения работ.....	44
6.2	Расчет сметы затрат на выполнение проекта.....	45
6.2.1	Расчет затрат на специальное оборудование для научных работ.....	45
6.2.2	Заработная плата исполнителей темы.....	46
6.2.3	Расчет затрат на электроэнергию.....	46
6.2.4	Расчет амортизационных расходов.....	47
6.2.5	Расчет общей себестоимости разработки.....	47
6.2.6	Расчет прибыли.....	47
6.2.7	Расчет НДС.....	48
6.2.8	Цена разработки НИР.....	48
6.3	Оценка экономической эффективности проекта.....	48
7	Социальная ответственность.....	49

7.1 Правовые и организационные вопросы обеспечения безопасности.....	49
7.1.1 Правовые нормы трудового законодательства для рабочей зоны оператора ПЭВМ	49
7.2 Производственная безопасность.....	50
7.2.1 Анализ вредных и опасных факторов	50
7.3 Экологическая безопасность.....	56
7.3.1 Анализ воздействия на окружающую среду	56
7.4 Безопасность в чрезвычайных ситуациях	56
Заключение	58
Список источников	59
Приложение А	62
1. Convolutional neural network training	63
1.1 Defining the size of anchor bounding boxes	63
1.2 YOLO v2 research and training	65
1.3 YOLO v2 tiny research and training	70
1.4 NMS algorithm researching	72
Приложение Б.....	74

Введение

Проблема своевременного обнаружения вредителей леса является актуальной и на сегодняшний день. Так в настоящее время уссурийский полиграф является одним из крупнейших вредителей деревьев пихты в таежных лесах Южной Сибири [1].

Одним из возможных способов решения данной проблемы является наблюдение за состоянием лесов с беспилотных летательных аппаратов (БПЛА). С помощью БПЛА можно не только находить поврежденные вредителями деревья, но и решать другие задачи, например, обнаружение возгораний.

В дальнейшем изображения, полученные с БПЛА, необходимо проанализировать. Это можно сделать как вручную, так и автоматизировать процесс с помощью алгоритмов компьютерного зрения и глубокого обучения на основе сверточных нейронных сетей (СНС).

Цель работы – исследовать эффективность СНС класса YOLO для детектирования пораженных деревьев пихты на аэрофотоснимках с БПЛА.

В первом разделе дано краткое описание предметной области. Данный раздел содержит информацию о различных видах состояния пихт, сформулирована постановка задачи.

Во втором разделе приведен обзор архитектур СНС и инструментов реализации СНС.

В третьем разделе представлены используемые для обучения данные, описаны способы подготовки данных для использования в процессе обучения нейронных сетей.

Четвертый раздел содержит описание процесса подготовки к обучению СНС, обучение СНС, а также подбор некоторых гиперпараметров.

В пятом разделе представлен анализ полученных результатов.

В шестом разделе рассмотрены вопросы финансового менеджмента.

Седьмой раздел исследования посвящен задаче социальной ответственности при выполнении ВКР.

1 Анализ предметной области

1.1 Мониторинг лесов

Уссурийский полиграф – жук из подсемейства короедов семейства долгоносиков. Первичный ареал обитания данного вредителя – южная часть Дальнего Востока, Северо-Восточный Китай, Япония и Корея. Спустя некоторое время уссурийский полиграф распространился в восточноевропейские и сибирские леса.

В настоящее время уссурийский полиграф является одним из крупнейших вредителей в таежных лесах Южной Сибири [1]. На рисунке 1 представлен ареал распространения уссурийского полиграфа в лесах Томской области.

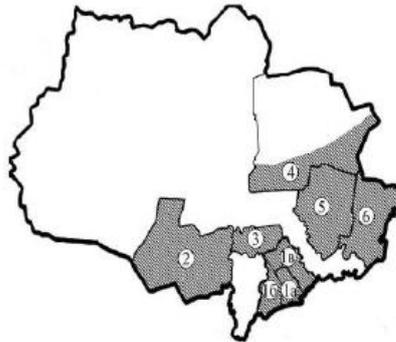


Рисунок 1. Распространение уссурийского полиграфа в районах Томской области [1]

Уссурийский полиграф встречается не только в эксплуатационных лесах, но и на особо охраняемых территориях.

Таким образом, вредитель наносит как экологический ущерб, так и экономический. Поэтому встает проблема скорейшего обнаружения вредителя, с целью снижения вредного воздействия на деревья.

Согласно шкале, по степени повреждения кроны пихт выделяют шесть классов пихт: здоровые, ослабленные, сильно ослабленные, усыхающие, свежие сухостои и старые сухостои (рис. 2).



Рисунок 2. Кроны деревьев пихты сибирской в очаге массового размножения уссурийского полиграфа [1]

Здоровые деревья имеют густую, ярко зеленую крону. Ослабленные и сильно ослабленные деревья атакованы полиграфом, но не заселены. Крона таких деревьев может быть как у здорового дерева, но чаще либо бледно зеленая, либо более половины ветвей имеют усыхающую хвою. Усыхающие деревья заселены полиграфом, их крона в нижней части ярко-рыжего цвета. Свежие сухостои – деревья, находящиеся в процессе обработки насекомыми, хвоя в кроне полностью мертвая, красная. Старые сухостои – деревья уже отработанные насекомыми, крона серая, хвоя осыпалась [1].

Для наблюдения за состоянием лесов применяются БПЛА, которые могут быть использованы, в том числе и для обнаружения пораженных уссурийским полиграфом пихтовых деревьев.

1.2 Постановка задачи исследования

Было решено разбить деревья на классы, сгруппировав при этом некоторые из них. Это было сделано в силу того, что при разметке данных визуально отделить некоторые классы друг от друга почти невозможно. Таким образом, итоговая классификация пихт по степени поражения выглядит следующим образом:

- Зеленые (здоровые, ослабленные);
- Желтые (сильно ослабленные, усыхающие);
- Красные (свежие сухостои);
- Серые (старые сухостои).

Целью работы является исследование эффективности СНС класса YOLO в задаче детектирования пораженных уссурийским полиграфом деревьев пихты на снимках с БПЛА.

2 Обзор архитектур нейронных сетей

2.1 Задачи компьютерного зрения

В настоящее время существуют множество алгоритмов для решения задач компьютерного зрения. Однако наиболее перспективными считаются нейронные сети.

СНС используются для решения таких задач компьютерного зрения, как классификация, детектирование и сегментация (рис. 3). Поскольку на снимках с БПЛА присутствует множество деревьев, то СНС необходимо решать либо задачу сегментации, либо детектирования.

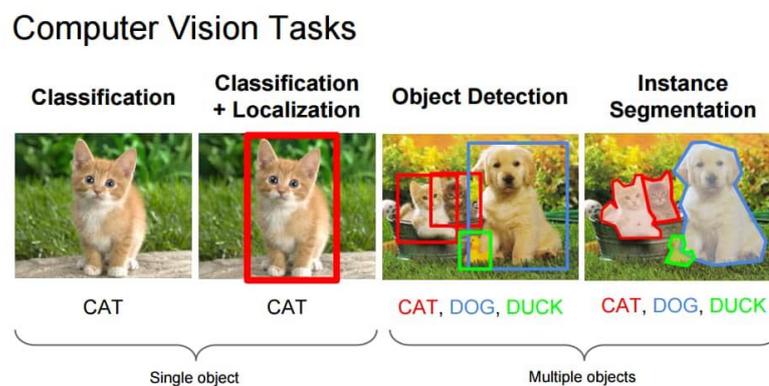


Рисунок 3. Задачи компьютерного зрения [2]

В данной работе СНС используются для решения задачи детектирования объектов на аэрофотоснимках. Для решения поставленной задачи существуют различные архитектуры СНС, такие как RetinaNet, Faster R-CNN, YOLO v2, YOLO v3 [3, 4, 5, 6].

Далее будет представлено краткое описание рассматриваемых архитектур, а также проведено сравнение точности и скорости детектирования.

2.2 Сверточная нейронная сеть YOLO v2

СНС YOLO v2 в качестве архитектуры для извлечения признаков использует Darknet-19. Архитектура СНС YOLO v2 представлена в таблице 1, где Convolutional – сверточный слой, Maxpool – слой подвыборки. Слой Reorg изменяет размерность тензора, чтобы можно было объединить

значения из нескольких слоев с помощью слоя Route. Размер входного изображения 416x416 пикселей [5].

Таблица 1. Архитектура СНС YOLO v2

Тип	Количество фильтров	Размер/Шаг	Выход
Darknet-19			
Convolutional	32	3x3	416x416
Maxpool		2x2/2	208x208
Convolutional	64	3x3	208x208
Maxpool		2x2/2	104x104
Convolutional	128	3x3	104x104
Convolutional	64	1x1	104x104
Convolutional	128	3x3	104x104
Maxpool		2x2/2	52x52
Convolutional	512	3x3	52x52
Convolutional	256	1x1	52x52
Convolutional	512	3x3	52x52
Maxpool		2x2/2	26x26
Convolutional	512	3x3	26x26
Convolutional	256	1x1	26x26
Convolutional	512	3x3	26x26
Convolutional	256	1x1	26x26
Convolutional	512	3x3	26x26
Maxpool		2x2/2	13x13
Convolutional	1024	3x3	13x13
Convolutional	512	1x1	13x13
Convolutional	1024	3x3	13x13
Convolutional	512	1x1	13x13
Convolutional	1024	3x3	13x13
Детектор			
Convolutional	1024	3x3	13x13
Route			13x13
Convolutional	1024	3x3	13x13
Reorg			13x13
Route			13x13

Convolutional	1024	3x3	13x13
Convolutional	anchors * (4 + 1 + classes)	1x1	13x13

Получая на вход изображение, сеть выдает предсказание в виде тензора размерностью [grid width, grid height, anchors * (4 + 1 + classes)], где grid width и grid height – количество ячеек по ширине и высоте, на которое разбивается изображение. Anchors – число якорных ограничивающих прямоугольников. Число 4 – координаты ограничивающих прямоугольников, а 1 – вероятность того, что в данном ограничивающем прямоугольнике находится какой-либо объект. Classes – условные вероятности классов. В данной работе количество ячеек равняется 13, поскольку размер входного изображения равнялся 416x416 пикселей, количество классов – 4. Число якорных ограничивающих прямоугольников один из гиперпараметров сети, который необходимо подобрать.

На рисунке 4 представлен пример сетки, каждая ячейка которой содержит по anchors * (4 + 1 + classes) значений.



Рисунок 4. Пример сетки на изображении

Так как YOLO v2 предсказывает больше объектов (рис. 5), чем нужно, то лишние ограничивающие прямоугольники отбрасываются с помощью алгоритма non-maximum suppression (NMS).



Рисунок 5. Возможные ограничивающие прямоугольники

У архитектуры YOLO v2 существует облегченная версия YOLO v2 tiny. СНС данной архитектуры использует меньшее количество слоев, следовательно скорость функционирования данной СНС выше, но точность меньше. Архитектура представлена в таблице 2.

Таблица 2. Архитектура СНС YOLO v2 tiny

Тип	Количество фильтров	Размер/Шаг	Выход
Darknet-19 tiny			
Convolutional	16	3x3	416x416
Maxpool		2x2/2	208x208
Convolutional	32	3x3	208x208
Maxpool		2x2/2	104x104
Convolutional	64	3x3	104x104
Maxpool		2x2/2	52x52
Convolutional	128	3x3	52x52
Maxpool		2x2/2	26x26
Convolutional	256	3x3	26x26
Maxpool		2x2/2	13x13
Convolutional	512	3x3	13x13
Maxpool		2x2/1	13x13
Convolutional	1024	3x3	13x13
Детектор			
Convolutional	1024	3x3	13x13
Route			13x13

Convolutional	1024	3x3	13x13
Reorg			13x13
Route			13x13
Convolutional	1024	3x3	13x13
Convolutional	anchors * (4 + 1 + classes)	1x1	13x13

2.3 Сверточная нейронная сеть YOLO v3

По принципу функционирования СНС архитектуры YOLO v3 схожа с СНС YOLO v2. Однако у данной сети есть несколько отличительных особенностей.

В качестве архитектуры для извлечения признаков используется Darknet-53. Как ясно из названия, сеть содержит 53 сверточных слоя. Помимо большего числа слоев, сеть данной архитектуры содержит остаточные блоки, как и сети архитектуры ResNet [6]. Кроме того, сеть данной архитектуры напоминает Feature Pyramid Network (FPN). Это позволяет улучшить точность детектирования объектов разного масштаба. В таблице 3 представлена архитектура данной СНС, где Residual – остаточные слои.

Таблица 3. Архитектура СНС YOLO v3 [6]

Количество	Тип	Количество фильтров	Размер/Шаг	Выход
Darknet-53				
	Convolutional	32	3x3	416
	Convolutional	64	3x3/2	208
1	Convolutional	32	1x1	
	Convolutional	64	3x3	
	Residual			208
	Convolutional	128	3x3/2	104
2	Convolutional	64	1x1	
	Convolutional	128	3x3	
	Residual			104
	Convolutional	256	3x3/2	52
8	Convolutional	128	1x1	

	Convolutional	256	3x3	
	Residual			52
	Convolutional	512	3x3/2	26
8	Convolutional	256	1x1	
	Convolutional	512	3x3	
	Residual			26
	Convolutional	1024	3x3/2	13
4	Convolutional	512	1x1	
	Convolutional	1024	3x3	
	Residual			13
Детектор (выход 1)				
	Convolutional	512	1x1	13
	Convolutional	1024	3x3	13
	Convolutional	512	1x1	13
	Convolutional	1024	3x3	13
	Convolutional	512	1x1	13
	Convolutional	1024	3x3	13
	Convolutional	anchors * (4 + 1 + classes)	1x1	13
Детектор (выход 2)				
	Route			13
	Convolutional	256	1x1	13
	UpSample			26
	Route			26
	Convolutional	256	1x1	26
	Convolutional	512	3x3	26
	Convolutional	256	1x1	26
	Convolutional	512	3x3	26
	Convolutional	256	1x1	26
	Convolutional	512	3x3	26
	Convolutional	anchors * (4 + 1 + classes)	1x1	26
Детектор (выход 3)				
	Route			26

	Convolutional	128	1x1	26
	UpSample			52
	Route			52
	Convolutional	128	1x1	52
	Convolutional	256	3x3	52
	Convolutional	128	1x1	52
	Convolutional	256	3x3	52
	Convolutional	128	1x1	52
	Convolutional	256	3x3	52
	Convolutional	anchors * (4 + 1 + classes)	1x1	52

Как видно из таблицы 3 сеть данной архитектуры содержит несколько выходов. Таким образом, получая на вход изображение размером 416x416 пикселей, получаем 3 тензора, размерности которых равны [grid width, grid height, anchors * (4 + 1 + classes)]. В первом случае размер сетки 13x13 ячеек, во втором – 26x26 ячеек, 52x52 ячеек – в третьем. Это позволяет улучшить детектирование объектов разного масштаба.

После получения результатов детектирования лишние ограничивающие прямоугольники отбрасываются с помощью алгоритма NMS.

2.4 Сверточная нейронная сеть RetinaNet

СНС архитектуры RetinaNet также как СНС типа YOLO является одноэтапным детектором. В качестве архитектуры для извлечения признаков из изображений используется адаптированная СНС типа FPN, что улучшает детектирование объектов разных масштабов. В то же время, СНС FPN имеет архитектуру подобную либо ResNet-50 либо ResNet-101 [3].

В качестве выхода FPN получает 5 уровней пирамиды [P₃; P₇], где размер каждого уровня в 2ⁱ раз меньше входного изображения [3].

Кроме того, сеть использует на каждом уровне пирамиды по 9 якорных ограничивающих прямоугольников, поскольку учитывается как отношение размеров прямоугольников $\{2^0, 2^{1/3}, 2^{2/3}\}$, так и их сторон $\{2:1, 1:1, 1:2\}$. Каждый такой якорный ограничивающий прямоугольник содержит вероятности каждого класса и 4 координаты [3].

У каждого слоя пирамиды есть две подсети: первая предсказывает класс объекта, вторая – координаты ограничивающего прямоугольника.

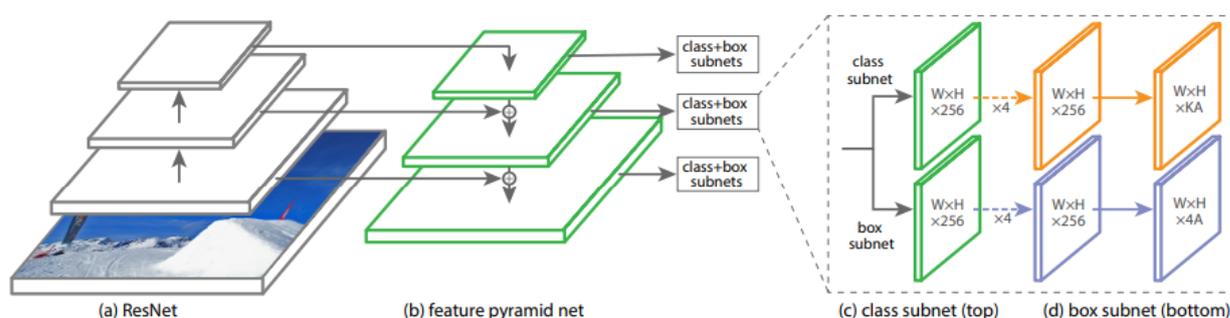


Рисунок 6. СНС архитектуры RetinaNet [3]

2.5 Сверточная нейронная сеть Faster R-CNN

Одной из наиболее популярных архитектур СНС является Faster R-CNN.

Сначала из входного изображения с помощью СНС извлекаются признаки. В качестве сети для извлечения признаков авторы статьи использовали VGG-16. Однако возможно также использование сетей других архитектур, например, ResNet-50 или ResNet-101 [4].

Далее полученные признаки поступают в Region Proposal Network (RPN). Из RPN признаки поступают в две ветви: первая предсказывает класс объекта, вторая описывает положение ограничивающего прямоугольника. Схема сети представлена на рисунке 7 [4].

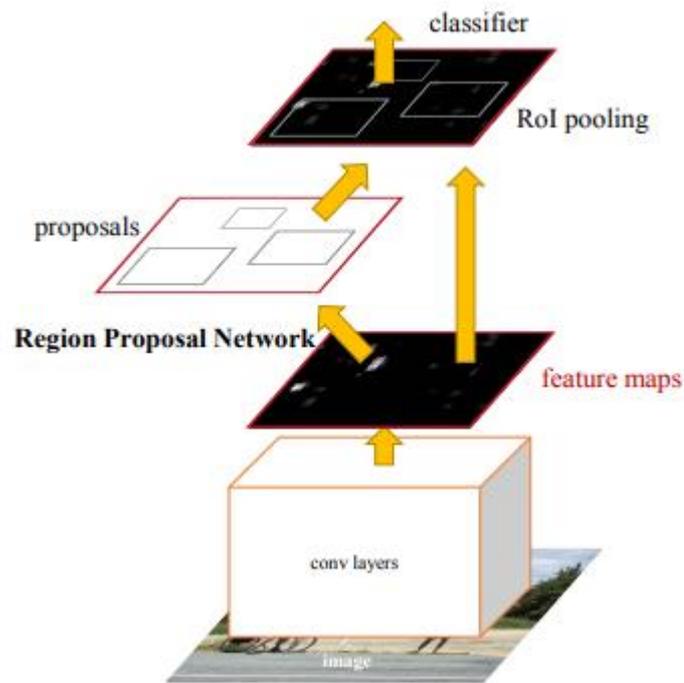


Рисунок 7. СНС архитектуры Faster R-CNN [4]

Для получения предсказаний сеть использует якорные ограничивающие прямоугольники. Авторы предлагают использовать 9 ограничивающих прямоугольников, 3 из которых отвечают за масштаб прямоугольников, 3 – за отношение сторон [4].

2.6 Сравнение СНС различных архитектур

Описанные СНС отличаются как по скорости функционирования, так и по точности детектирования. Поэтому необходимо провести их сравнение по заданным критериям.

В качестве метрики оценки точности детектирования использовалась mean average precision (mAP). Для вычисления данной метрики обычно задается порог IOU, который чаще всего равен либо 0,5, либо 0,75, либо вычисляется среднее значение в диапазоне от 0,05 до 0,95 с шагом 0,05 [8]. Решено было остановиться на $IOU = 0,5$. Оценка производительности была оценена с помощью метрики frames per second (FPS).

Поскольку разброс значения скорости детектирования выше, то решено было выставить оценку по пятибалльной шкале. Кроме того, для

расчета итоговых значений были использованы коэффициенты, отвечающие за важность показателей. Так для скорости детектирования коэффициент равен 0,6, а для точности детектирования 0,4. Результаты представлены в таблице 4.

Таблица 4. Сравнение СНС различных архитектур [9]

Архитектура СНС	Точность детектирования		Скорость детектирования		Итого, балл
	mAP	Балл	FPS	Балл	
YOLO v2	48,1	3	40	5	4,2
YOLO v3	57,9	5	20	3	3,8
RetinaNet-101	53,1	4	11	3	3,4

Таким образом, в качестве исследуемой архитектуры СНС была выбрана YOLO v2, поскольку она лучше соответствует оценочным критериям.

2.7 Обзор средств реализации СНС

Для реализации СНС доступны различные библиотеки под разные языки программирования. Так для языка C++ наибольшей популярностью пользуется библиотека Caffe, для Java – DL4J [10, 11]. Для языка Python существует множество различных библиотек, таких как, Tensorflow, Keras, Pytorch, MXNet [12, 13, 14, 15]. Кроме того, разработчики YOLO предлагают собственную разработку, написанную на языке C – Darknet [16].

Проанализировав все возможные варианты было решено реализовывать СНС с помощью библиотеки Tensorflow, поскольку данная библиотека является наиболее популярной из перечисленных, содержит подробную документацию, а также идет в связке с фреймворком Keras, который упрощает процесс написания нейронных сетей.

3 Подготовка данных

Для обучения и тестирования СНС были использованы два больших снимка с БПЛА пихтовых лесов. Сначала снимки были размечены с помощью инструментария CVAT [17]. В результате были получены два больших снимка, содержащих ограничивающие прямоугольники для пихтовых деревьев четырех классов: зеленые, желтые, красные, серые (рис. 8).

В процессе разметки возникла сложность в разделении красного и серого классов крон, поскольку встречались почти полностью осыпавшиеся пихтовые деревья с красной хвоей. Их решено было выделить в класс красных крон. Кроме того, сложность состояла в отделении пихтовых деревьев от деревьев других видов, таких как береза и ель. Кроме того, часть деревьев по краям исходного изображения были размыты и имели непонятную окраску. Решено было ограничить присутствие таких деревьев как в обучающей, так и тестовой выборках.



Рисунок 8. Размеченный снимок

Затем размеченные исходные снимки были разделены на фрагменты, поскольку в качестве входных данных для СНС YOLO v2 обычно используются изображения размером 416x416 пикселей [3]. Для получения обучающей выборки первое изображение было разделено на фрагменты, сдвигаемые по изображению с шагом 104 пикселя. Для получения тестовой и

валидационной выборки второе изображение было разделено на такие же фрагменты, но сдвигаемые с шагом 208 пикселей. Полученная выборка была разбита на 2 части, 80% изображений составили тестовую выборку, 20% - валидационную. Таким образом, первоначальный размер обучающей выборки составил 766 фрагментов (изображений на входе СНС для обучения), тестовой выборки – 161 фрагмент, а валидационной – 41 фрагмент. На рис. 9 представлен пример фрагмента из тестовой выборки.

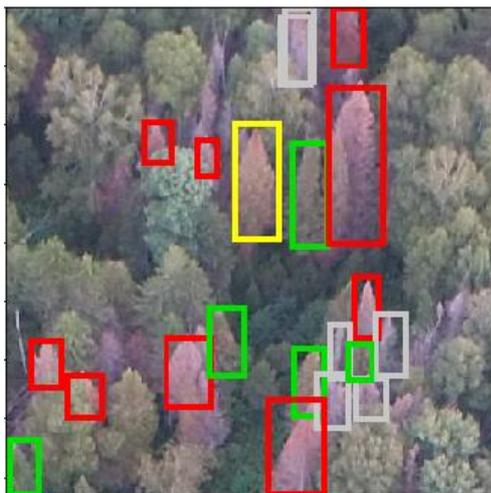


Рисунок 9. Фрагмент изображения из тестовой выборки

Предварительный анализ обучающей и тестовой выборок показал, что на изображениях преобладают красный и серый классы крон пихт, в то время как желтых крон на изображениях меньше всего (рис. 10, 11). Данный факт стоит учитывать при обучении СНС.

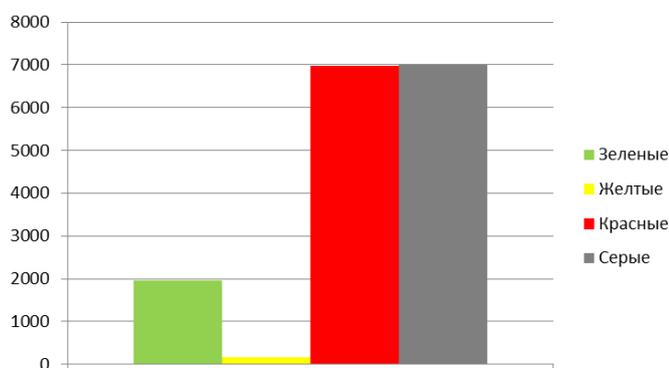


Рисунок 10. Распределение классов на обучающей выборке

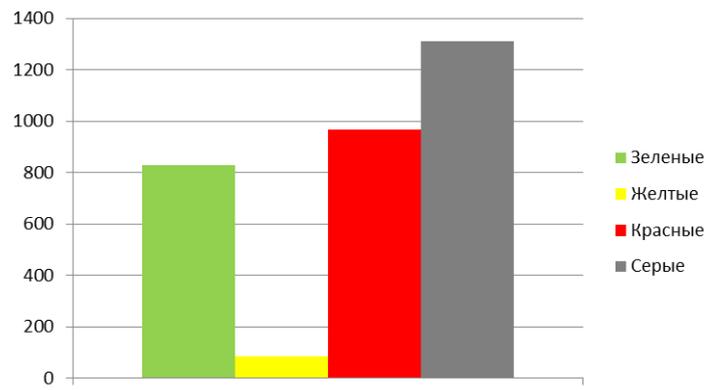


Рисунок 11. Распределение классов на тестовой выборке

4 Обучение СНС

4.1 Подбор якорных ограничивающих прямоугольников

Прежде чем приступить к обучению СНС YOLO v2, необходимо подобрать основные гиперпараметры. Одним из таких параметров являлись размеры якорных ограничивающих прямоугольников, которые были подобраны с помощью алгоритма k-means, как это описано в статье [5]. Метрика расстояния высчитывалась по следующей формуле:

$$d(box, centroid) = 1 - IOU(box, centroid),$$

где d – расстояние между текущим ограничивающим прямоугольником (box) и ограничивающим прямоугольником, выбранным в качестве центра кластера ($centroid$). В качестве метрики использовалось intersection over union (IOU) – отношение площади пересечения к площади объединения двух прямоугольников.

На рис. 12, 13, 14, представлены результаты работы алгоритма k-means. На рис. 15 представлена зависимость точности детектирования от количества кластеров, на которые разбивается выборка. Видно, что чем больше кластеров, тем точнее будет детектирование. Однако с увеличением количества кластеров уменьшается скорость работы СНС. Поэтому решено было ограничиться 5 кластерами, поскольку при этом достигается баланс между скоростью работы СНС и точностью детектирования.

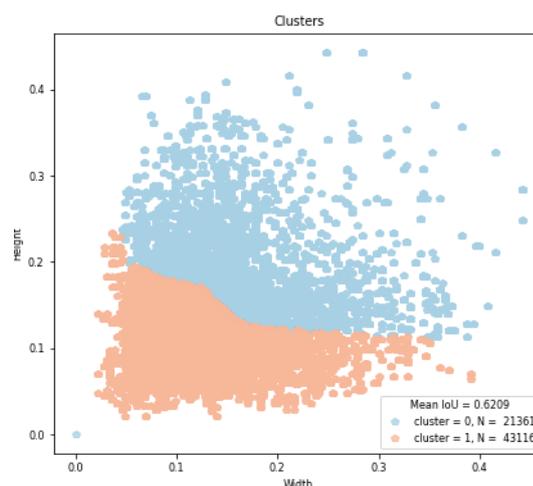


Рисунок 12. Результат разбиение на 2 кластера

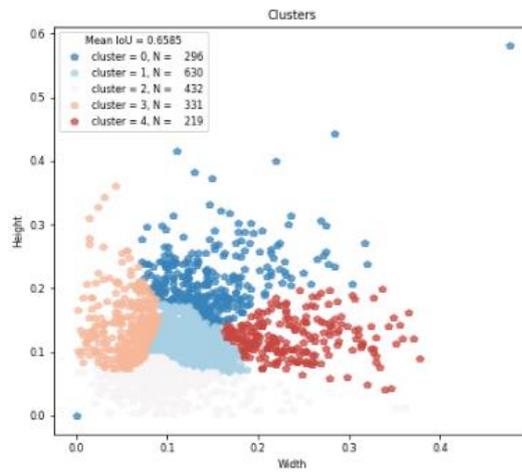


Рисунок 13. Результат разбиения на 5 кластеров

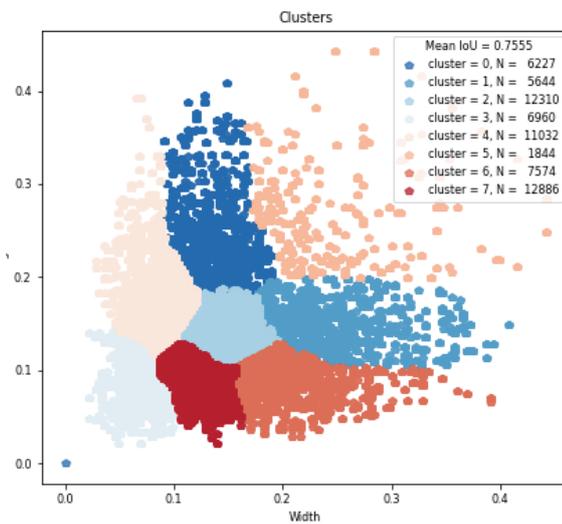


Рисунок 14. Результат разбиения на 8 кластеров

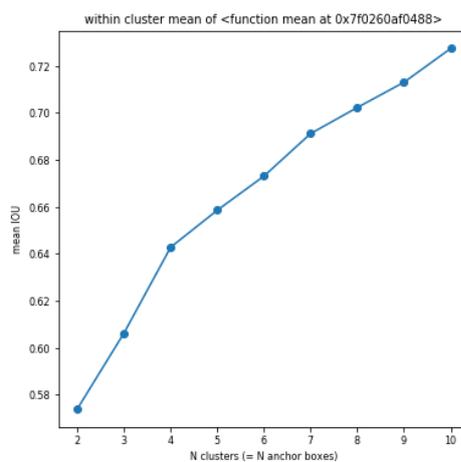


Рисунок 15. Зависимость IOU от количества кластеров

4.2 Обучение и исследование СНС YOLO v2

В ходе экспериментов были исследованы зависимость качества детектирования от числа эпох и от различных методов аугментации. Во время обучения гиперпараметры сети не менялись.

Так как СНС YOLO v2 предсказывает множество ограничивающих прямоугольников, то необходимо оставить наиболее вероятные прямоугольники. Среди оставшихся прямоугольников нужно выбрать наиболее вероятные, и оставить те прямоугольники, у которых значение IOU с наиболее вероятными не превышает пороговое. Среди оставшихся прямоугольников нужно отбросить те прямоугольники, вероятность которых ниже порогового значения. Пороговое значение для IOU равняется 0,45, для вероятности 0,35. В дальнейшем, данные значения необходимо скорректировать.

Обучение проводилось с помощью оптимизатора Adam со скоростью обучения $3 \cdot 10^{-4}$. Во время обучения веса сети сохранялись при уменьшении ошибки на валидационной выборке. Тестирование проводилось с использованием полученных весов.

В качестве метрики была использована mAP с порогом IOU, равным 0.5.

В качестве функции потерь использовалась функция, описанная в статье [5]. Ниже представлена формула используемой функции потерь СНС:

$$loss_{coord} = \lambda_{coord} \sum_i^{S^2} \sum_j^B 1^{obj} [(x_{ij} - \hat{x}_{ij})^2 + (y_{ij} - \hat{y}_{ij})^2 + (\sqrt{w_{ij}} - \sqrt{\hat{w}_{ij}})^2 + (\sqrt{h_{ij}} - \sqrt{\hat{h}_{ij}})^2]$$

$$loss_{conf} = \lambda_{obj} \sum_i^{S^2} \sum_{c \in C} (1^{obj} + \lambda_{no\ obj} (1 - 1_{ij}^{obj})) (C_{ij} - \hat{C}_{ij})^2$$

$$loss_{class} = \lambda_{class} \sum_i^{S^2} \sum_{c \in C} 1^{obj} p_i(c) \log(\hat{p}_i(c))$$

$$loss = loss_{coord} + loss_{conf} + loss_{class}$$

где итоговая функция потерь $loss$ вычисляется как сумма функции потерь для координат $loss_{coord}$, функции потерь вероятности нахождения объекта в данном прямоугольнике $loss_{conf}$ и функции потерь вероятности принадлежности к какому-либо классу $loss_{class}$. x, y – координаты центров ограничивающих прямоугольников, w, h – ширина и высота ограничивающих прямоугольников. C – вероятность нахождения объекта в данной ячейке сетки для данного ограничивающего прямоугольника, $p(c)$ – вероятность принадлежности к классу. $\lambda_{coord}, \lambda_{obj}, \lambda_{no obj}, \lambda_{class}$ – коэффициенты, равные 1, 5, 1, 1 соответственно. Данные коэффициенты можно считать гиперпараметрами сети, но в данной работе значения коэффициентов были взяты из репозитория [16].

Сначала обучение было проведено на выборке, содержащей 766 изображений. Затем для увеличения количества изображений каждое исходное изображение обучающей выборки было повернуто на $90^\circ, 180^\circ$ и 270° . В итоге расширенная обучающая выборка составила 3064 изображения. Следующим способом увеличения обучающей стало случайное изменение яркости и контраста входных изображений на этапе обучения модели. Результаты исследований представлены в таблице 5.

Таблица 5. Зависимость точности детектирования от числа эпох и методов аугментации

Методы аугментации	Число эпох		
	5	10	25
Отсутствует	0,032	0,036	0,029
Поворот	0,072	0,080	0,078
Поворот, изменение яркости и контраста	0,101	0,116	0,078

Видно, что с добавлением методов аугментации точность детектирования увеличивается. Однако увеличение числа эпох не оказывает влияния на точность детектирования.

Далее был исследован подход [19], когда на первых этапах обучения СНС используется большой набор аугментаций. Обучение происходило в два этапа: сначала СНС обучалась на данных, где для увеличения выборки применялись поворот изображений, а также случайное изменение яркости и

контраста. На втором этапе яркость и контраст не менялись. Результаты представлены в таблице 6.

Таблица 6. Зависимость точности детектирования от длительности этапов обучения

Число эпох без изменения яркости и контраста	Число эпох с изменением яркости и контраста		
	5	10	25
1	0,134	0,102	0,111
3	0,175	0,131	0,126
5	0,148	0,108	0,131
10	0,158	0,156	0,119

Наилучший результат был достигнут, когда количество эпох с аугментацией равно 5, а количество эпох без аугментации равнялось 3.

Для сравнения было выбрано три фрагмента из тестовой выборки. На первом фрагменте деревья расположены под углом к камере БПЛА, на втором – напротив камеры, на третьем – под камерой. На рис. 16, 17, 18 представлены сравнение работы сетей, обученных разными способами, с исходной разметкой. Слева – исходная разметка, в центре сеть, обученная на 5 эпохах с применением всех аугментаций, справа результаты обучения сети сначала на 5 эпохах с применением аугментаций, а затем на 3 эпохах без аугментаций.



Рисунок 16. Результаты детектирования для снимка, расположенного под углом к камере

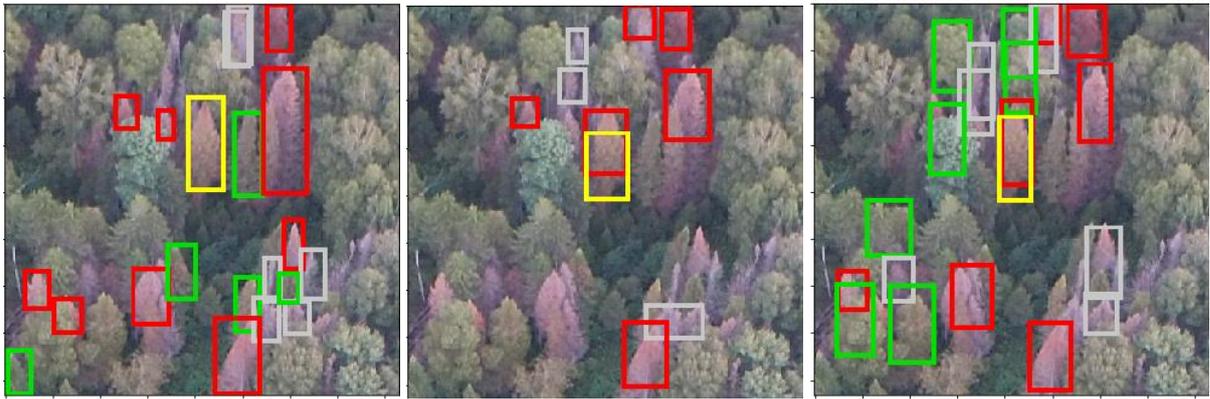


Рисунок 17. Результаты детектирования для снимка, расположенного напротив камеры



Рисунок 18. Результаты детектирования для снимка, расположенного под камерой

Для всех трех снимков видно, что при двухэтапном обучении сеть выдает больше ограничивающих прямоугольников. Из этого можно сделать несколько выводов. Во-первых, каждый ограничивающий прямоугольник для сети, обученной в два подхода, имеет более высокую вероятность для каждого класса, то есть сеть выдает более надежные результаты. Во-вторых, можно заметить, что для сети, обученной в два этапа, для некоторых деревьев сеть выдает несколько ограничивающих прямоугольников, что свидетельствует о необходимости настройки пороговых значений сети.

После того, как сеть была обучена, необходимо подобрать пороговые значения для алгоритма NMS. Как видно из таблицы 7, что с увеличением порога вероятности уменьшается точность детектирования. Однако зависимость точности от значения порога IOU нелинейная, поскольку наибольшее значение точности достигается при пороговом значении 0.45.

Таблица 7 – Зависимость метрики от пороговых значений

Порог вероятности	0,1	0,2	0,3	0,35	0,4	0,5	0,6
Порог IOU							
1	0,182	0,175	0,160	0,154	0,139	0,113	0,080
0,9	0,182	0,175	0,161	0,154	0,139	0,113	0,080
0,8	0,184	0,177	0,162	0,155	0,140	0,114	0,081
0,7	0,194	0,186	0,170	0,163	0,147	0,119	0,083
0,6	0,204	0,196	0,179	0,171	0,153	0,124	0,085
0,5	0,208	0,200	0,182	0,173	0,155	0,124	0,086
0,45	0,212	0,203	0,184	0,174	0,158	0,126	0,086
0,4	0,211	0,202	0,183	0,174	0,158	0,126	0,087
0,3	0,207	0,199	0,180	0,172	0,156	0,125	0,086

На рисунке 19 представлены результаты детектирования при фиксированном значении порога IOU, равным 0,45, и различных значениях порога вероятности. В верхнем левом углу – размеченное изображение. Остальные изображения отражают результаты детектирования сети, порог вероятности изменялся на отрезке [0,2; 0,5]. Видно, что при пороге вероятности 0,2 сеть предсказывает достаточно много лишних объектов, а при пороге 0,5 часть ограничивающих прямоугольников отбрасывается.



Рисунок 19. Результаты детектирования при различных значениях порога вероятности

Далее представлены результаты детектирования при фиксированном значении порога вероятности, равного 0,3. Порог IOU изменялся в интервале [0,4; 0,7]. Результаты представлены на рисунке 20. В верхнем левом углу находится размеченное изображение, остальные изображения – результаты детектирования при различных значениях порога IOU. Видно, что чем больше порог, тем больше ограничивающих прямоугольников на изображении.

Исходя из полученных результатов, можно сделать следующие выводы. Наилучшими пороговыми значениями вероятности как с точки зрения значения метрики mAP, так и с визуальной точки зрения, является 0,3 и 0,35. Значение порога IOU в данном случае равняется 0,45. Таким образом, оставить исходные значение.



Рисунок 20. Результаты детектирования при различных значениях порога IOU

4.3 Обучение и исследование СНС YOLO v2 tiny

Помимо YOLO v2, было произведено обучение СНС YOLO v2 tiny. В ходе экспериментов были исследованы зависимость качества детектирования от числа эпох. Исследование методов аугментации для YOLO v2 tiny не проводилось.

Обучение проводилось с помощью оптимизатора Adam со скоростью обучения $3 \cdot 10^{-4}$. Во время обучения веса сети сохранялись при уменьшении ошибки на валидационной выборке. Тестирование проводилось с использованием полученных весов.

Как и для YOLO v2, для YOLO v2 tiny было проведено исследование двухэтапного обучения СНС, когда на первом этапе используется больший набор аугментаций, чем во втором. Результат представлен в таблице 8. Результат были получены при пороговом значении вероятности, равным 0,35, и пороговом значении IOU, равным 0,45.

Таблица 8. Зависимость точности детектирования от длительности этапов обучения

Число эпох без изменения яркости и контраста	Число эпох с изменением яркости и контраста		
	5	10	25
1	0,068	0,073	0,074
3	0,084	0,099	0,075
5	0,071	0,079	0,058
10	0,076	0,074	0,067

Наилучший результат был получен, когда количество эпох с аугментацией равно 10, а количество эпох без аугментации равнялось 3. Результаты представлены на рисунке 21, где изображение слева – исходный снимок с разметкой.



Рисунок 21. Результат детектирования для снимка, расположенного напротив камеры

После того, как сеть обучена, необходимо подобрать пороги для алгоритма NMS. Учитывая ранее полученные результаты, диапазон поиска можно уменьшить. Так значения порога IOU были взяты из интервала [0,3; 0,6], а значения порога вероятности из интервала [0,2; 0,5].

Как видно из таблицы 9, что при постоянном значении порога вероятности, наибольшее значение метрики достигается при значениях порога IOU 0.5 или 0.45. Таким образом, в качестве значения порога IOU можно выбрать 0,45.

С увеличением порога вероятности уменьшается и точность детектирования, поскольку сеть предсказывает меньше объектов.

Таблица 9. Зависимость метрики от пороговых значений

Порог вероятности	0,2	0,3	0,35	0,4	0,5
Порог IOU					
0,6	0,119	0,104	0,097	0,088	0,072
0,5	0,123	0,107	0,099	0,090	0,074
0,45	0,122	0,106	0,099	0,090	0,073
0,4	0,117	0,103	0,097	0,088	0,072
0,3	0,114	0,101	0,094	0,086	0,072

На рисунке 22 представлены результаты детектирования при постоянном значении порога IOU. Левое изображение содержит исходную разметку. Два других изображения – результаты детектирования при наименьшем и наибольшем значениях порога вероятности соответственно.



Рисунок 22. Результаты детектирования при различных значениях порога вероятности

Видно, что в первом случае у модели много ложных срабатываний, в то время как во втором случае некоторые из ограничивающих прямоугольников отсутствуют. Таким образом, значение порога IOU равняется 0,45, а значение порога вероятности 0,35.

4.4 Исследование алгоритма NMS

После завершения обучения СНС было проведено исследование алгоритма NMS. Поскольку NMS алгоритм обнуляет вероятности ограничивающих прямоугольников, пересекающихся с наиболее вероятными, то часть ограничивающих прямоугольников теряется.

Для решения данной проблемы было решено исследовать алгоритм soft-NMS [20], который, в отличие от предыдущего, не обнуляет вероятность, а уменьшает на величину, пропорциональную IOU данных прямоугольников. Используемая формула представлена ниже

$$pr = \begin{cases} pr, & IOU < IOU_{threshold} \\ pr * (1 - IOU), & IOU \geq IOU_{threshold} \end{cases}$$

где pr – вероятность принадлежности к заданному классу. Кроме того, была исследована Гауссовская функция штрафов, представленная ниже. В отличие от предыдущей функции, она непрерывная [20].

$$pr = pr * e^{-\frac{IOU^2}{IOU_{threshold}}}$$

Согласно статье [20], алгоритм позволил увеличить точность детектирования YOLO v2 на 0,5%.

Таким образом, было решено провести исследование данного алгоритма как для СНС YOLO v2, так и для YOLO v2 tiny. В ходе исследования значение порога вероятности было константным. Порог IOU менялся в диапазоне [0,3; 0,6]. Результаты представлены в таблице 10, где Gaussian soft-NMS – алгоритм soft-NMS, использующий Гауссовскую функцию штрафов.

Таблица 10. Зависимость точности детектирования от порога IOU для разных алгоритмов

Метод	Порог IOU	0,3	0,4	0,45	0,5	0,6
	Архитектура СНС					
NMS	YOLO v2	0,172	0,174	0,174	0,173	0,171
	YOLO v2 tiny	0,094	0,097	0,099	0,099	0,097
Soft-NMS	YOLO v2	0,173	0,174	0,174	0,172	0,170
	YOLO v2 tiny	0,095	0,096	0,098	0,099	0,097
Gaussian Soft-NMS	YOLO v2	0,171	0,172	0,173	0,173	0,174
	YOLO v2 tiny	0,093	0,094	0,095	0,096	0,097

Как видно из таблицы 10, при различном пороге IOU результаты и используемом алгоритме результаты отличаются как в большую, так и в меньшую сторону. Однако ни в одном из экспериментов не удалось улучшить значение метрики.

5 Анализ полученных результатов

В ходе проделанной работы была исследована точность детектирования пихтовых деревьев на изображениях с помощью архитектур СНС класса YOLO. Помимо точности детектирования необходимо также учесть и скорость функционирования данных нейронных сетей, поскольку в дальнейшем СНС данных архитектур планируется реализовать для применения на БПЛА. Результаты представлены в таблице 11. Результаты представлены с использованием облачного сервиса Google Colaboratory, использующего видеокарту NVIDIA Tesla K80.

Таблица 11. Сравнение скорости и точности детектирования

Архитектура СНС	Скорость, FPS	Точность, mAP
YOLO v2	13 FPS	0,174
YOLO v2 tiny	33 FPS	0,099

Видно, что точность детектирования YOLO v2 выше, чем точность детектирования YOLO v2 tiny. Однако скорость детектирования YOLO v2 меньше.

В таблице 12 приведено сравнение точности детектирования СНС для разных классов пихтовых деревьев.

Таблица 12. Сравнение точности детектирования различных классов

Архитектура СНС	Зеленые, AP	Желтые, AP	Красные, AP	Серые, AP
YOLO v2	0,010	0,161	0,299	0,229
YOLO v2 tiny	0,001	0,026	0,232	0,138

Из таблицы 12 ясно, что наименьшая точность детектирования характерна для объектов, относящихся к классу зеленых деревьев, то есть для здоровых пихт. Это можно объяснить как наличием большого количества зеленых деревьев других видов, так и недостаточно точной разметкой исходных изображений.

6 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

6.1 Организация и планирование работ

6.1.1 Структура работ в рамках научного исследования

В проекте задействовано два исполнителя: руководитель и исследователь. Задачи руководителя – выдача задания, а также консультация исследователя по различным разделам работы. Задачи исследователя – реализация функционала и формирование отчетности о проделанной работе. В таблице 13 представлен подробный перечень этапов работ.

Таблица 13. Перечень этапов работ и распределение исполнителей

Этапы	Содержание работ	Загрузка исполнителей	
Разработка технического задания	Разработка технического задания	Руководитель	100
Анализ предметной области	Обзор предметной области	Исследователь	100
	Обзор аналогов	Исследователь	100
	Предъявление требований	Руководитель	100
		Исследователь	40
Подготовка данных	Получение снимков с БПЛА	Руководитель	100
	Разметка исходных снимков	Исследователь	100
	Нарезка снимков на фрагменты	Исследователь	100
	Аугментация данных	Исследователь	100
Исследование	Обзор литературы	Руководитель	50
		Исследователь	100
	Проведение экспериментов	Руководитель	10
		Исследователь	100
Написание отчета	Написание отчета	Исследователь	100

6.1.2 Определение трудоемкости выполнения работ

После определения структуры работ была произведена оценка трудоемкости выполнения работ, поскольку трудовые затраты при разработке программного компонента могут составлять большую часть от итоговой стоимости. Ожидаемая трудоемкость измеряется в человеко-днях и вычисляется по следующей формуле.

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5}$$

Затем вычисляется продолжительность каждой работы в рабочих днях, на основании полученной ожидаемой трудоемкости по формуле:

$$T_{рд} = \frac{t_{ож}}{K_{вн}} * K_{д}$$

Значение коэффициента выполнения работ и коэффициента дополнительного времени соответственно равны $K_{вн} = 1$, $K_{д} = 1$. Длительность работ в календарных днях вычисляется по формуле:

$$T_{кд} = T_{рд} * T_{к}$$

Коэффициент календарности $T_{к} = 1,2$. Результаты представлены в таблице 14.

Таблица 14. Временные показатели проведения разработки

Название работ	Продолжительность работ, дни			Загрузка исполнителей		Длительность, рабочий день		Длительность, календарный день	
	t_{min}	t_{max}	$t_{ож}$	НР	И	НР	И	НР	И
Разработка технического задания	1	3	1,8	0	1	0	1,8	0	2,16
Обзор предметной области	1	3	1,8	0	1	0	1,8	0	2,16
Обзор аналогов	1	2	1,4	0	1	0	1,4	0	1,68
Предъявление требований	1	2	1,4	1	0,4	1,4	0,56	1,68	0,672

Получение снимков с БПЛА	2	5	3,2	1	0	3,2	0	3,84	0
Разметка исходных снимков	3	14	7,4	0	1	0	7,4	0	8,88
Нарезка снимков на фрагменты	1	3	1,8	0	1	0	1,8	0	2,16
Аугментация данных	1	2	1,4	0	1	0	1,4	0	1,68
Обзор литературы	1	2	1,4	0,5	1	0,7	1,4	0,84	1,68
Проведение экспериментов	60	120	84	0,1	1	8,4	84	10,08	100,8
Написание отчета	2	7	4	0	1	0	4	0	4,8
Итого			109,6			13,7	105,56	16,44	126,672

После проведения оценки трудоемкости работ была построена диаграмма Ганта (см. приложение А).

6.2 Расчет сметы затрат на выполнение проекта

После определения календарного плана, необходимо рассчитать бюджет научно-технического исследования (НТИ), который представлен в дальнейших пунктах работы.

6.2.1 Расчет затрат на специальное оборудование для научных работ

Основными инструментами при выполнении научного исследования являются ноутбук и установленное программное обеспечение.

Поскольку коммерческие продукты являются очень дорогими, то решено было использовать либо продукты с открытым исходным кодом, которые распространяются бесплатно, либо ограниченные по функциональности версии коммерческих продуктов. Кроме того, некоторые из коммерческих продуктов студентам можно получить бесплатно, например, по программе Dream Spark.

Что касается ПК, то его стоимость рассчитывалась на основе амортизационных исчислений, поскольку он не был приобретен специально для проведения исследований.

Однако стоит учесть, что в дальнейшем стоимость разработки может возрасти, если будут использованы полные лицензии коммерческих продуктов или будет приобретен новый ПК.

6.2.2 Заработная плата исполнителей темы

Далее была рассчитана заработная плата научного руководителя и исследователя. Результаты представлены в таблице 15. При этом социальный налог $C_{\text{соц.}} = 109640,37 * 0,3 = 32892,11$ руб.

Таблица 15. Расчет заработной платы

Исполнитель	Оклад, руб/месяц	Среднедневная ставка, руб/рд	Затраты времени, рд	Коэффициент	Фонд зп, руб
НР	47104	1878,14	14	1,699	44670,11
И	9489	378,35	106	1,62	64970,26
Итого					109640,37

6.2.3 Расчет затрат на электроэнергию

Рассчитаем затраты на электроэнергию в ходе выполнения проекта. Во время работы были задействованы настольная лампа и ноутбук. Вычисления производились по формуле: $C_{\text{эл.об.}} = P_{\text{об.}} \cdot t_{\text{об.}} \cdot C_{\text{Э}}$

$$C_{\text{эл}} = P_{\text{об}} * t_{\text{об}} * C_{\text{об}}$$

Цена электроэнергии за квт*ч равна 6,59 руб. Значение коэффициентов равны 1. Результаты представлены в таблице 16.

Таблица 16. Расчет затрат на электроэнергию

Наименование оборудования	Время работы оборудования $t_{\text{об}}$, час	Потребляемая мощность $P_{\text{об}}$, кВт	Затраты $\text{Э}_{\text{об}}$, руб.
Ноутбук	848	0,06	335,3
Итого:			335,3

6.2.4 Расчет амортизационных расходов

Далее были рассчитаны амортизационные расходы на оборудование.

В ходе работ было использовано такое оборудование как ноутбук. Вычисление производилось по формуле:

$$C_{AM} = \frac{N_A * C_{Об} * t_{рф} * n}{F_D}$$

Годовая норма амортизации оборудования равна 0.4, цена ноутбука 25000 рублей, время работы 838 часов, действительный годовой фонд работы оборудования равен 2384 часов. Поэтому полученная амортизационная стоимость равна 3557 рубля.

6.2.5 Расчет общей себестоимости разработки

Таким образом, общая себестоимость разработки составила 163526,33 рублей (таблица 17).

Таблица 17. Общая себестоимость разработки

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{мат}$	5600
Основная заработная плата	$C_{зп}$	109640,37
Отчисления в социальные фонды	$C_{соц}$	32892,11
Расходы на электроэнергию	$C_{эл.}$	335,3
Амортизационные отчисления	$C_{ам}$	3557
Прочие расходы	$C_{проч}$	14866,03
Итого:		163526,33

6.2.6 Расчет прибыли

Рассчитаем прибыль разработки. Поскольку сфера применения проекта узкоспециализированная, то будем считать, что прибыль равна 7,5% от суммы, потраченной на разработку, то есть 12542,04 рубля.

6.2.7 Расчет НДС

Поскольку НДС составляет 20% от затрат на разработку и прибыли, то значение НДС равно 33445,45 рублей.

6.2.8 Цена разработки НИР

Цена разработки НИР равна сумме полной себестоимости, прибыли и НДС, то есть 213 214,75 рублей.

6.3 Оценка экономической эффективности проекта

Данная работа не предназначена для коммерческого использования, так как получение прямого экономического эффекта не предполагается, поскольку работа носит исследовательский характер.

Однако данный алгоритм можно реализовать в составе прикладного программного обеспечения, которое в дальнейшем можно использовать на предприятиях лесного хозяйства для проведения мониторинга состояния леса. Своевременное обнаружение зараженных деревьев позволяет избежать заражения здоровых деревьев.

Таким образом, факторы получения экономического эффекта носят косвенный характер.

7 Социальная ответственность

7.1 Правовые и организационные вопросы обеспечения безопасности

7.1.1 Правовые нормы трудового законодательства для рабочей зоны оператора ПЭВМ

Согласно трудовому кодексу, продолжительность рабочего дня не должна превышать 24 часов в неделю для работников до 16 лет, 35 часов для работников в возрасте от 16 до 18 лет или являющихся инвалидами I или II групп. В остальных случаях рабочая неделя должна длиться не более 40 часов.

Так как работа программиста относится к категории работ требующей постоянного взаимодействия с ПЭВМ, то рекомендуется организовывать перерывы длительностью 10 – 15 минут через каждые 40–60 минут работы. Кроме того, продолжительность непрерывной работы с ЭВМ не должна превышать 60 минут. При работе в ночную смену, с 22:00 до 6:00 необходимо увеличивать длительность перерывов на 30%. Во время таких перерывов рекомендуется выполнять комплекс упражнений для снижения нервно-эмоционального напряжения, утомления зрительного анализатора, устранения влияния гиподинамии и гипокинезии, предотвращения развития позотонического утомления [21]. Помимо этого, организацией должен быть предоставлен перерыв длиной не менее 30 минут для приема пищи.

Также организация обязана предоставлять ежегодный отпуск продолжительностью 28 календарных дней. Дополнительные отпуска предоставляются работникам, занятым на работах с вредными или опасными условиями труда, работникам имеющими особый характер работы, работникам с ненормированным рабочим днем и работающим в условиях Крайнего Севера и приравненных к нему местностях [22].

7.2 Производственная безопасность

7.2.1 Анализ вредных и опасных факторов

Согласно ГОСТ [21] опасные и вредные факторы согласно природе воздействия выделяют в следующие группы:

- Физические;
- Химические;
- Биологические;
- Психофизиологические.

При работе с персональными электронно-вычислительными машинами (ПЭВМ) возможны воздействия факторов каждой из этой группы. Ниже представлены опасные и вредные факторы, которые могут возникнуть при работе программиста с ПЭВМ.

Таблица 18. Опасные и вредные факторы

Источник фактора, наименование видов работ	Факторы (по ГОСТ 12.0.003-74)		Нормативные документы
	Вредные	Опасные	
Работа за ПЭВМ	<ul style="list-style-type: none">• Повышенный уровень шума на рабочем месте;• Повышенная или пониженная температура воздуха рабочей зоны;• Повышенная или пониженная влажность воздуха;• Повышенный уровень электромагнитных излучений;• Отсутствие или недостаток естественного света;• Недостаточная освещенность рабочей зоны;• Повышенная яркость света;• Пониженная контрастность;	<ul style="list-style-type: none">• Повышенное значение напряжения;	<ul style="list-style-type: none">• СанПиН 2.2.2/2.4.1340–03. Санитарно-эпидемиологические правила и нормативы «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы»;• СанПиН 2.2.4.548–96. Гигиенические требования к микроклимату

	<ul style="list-style-type: none"> • Прямая и отраженная блескость; • Повышенная пульсация светового потока; • Умственное перенапряжение; • Перенапряжение анализаторов; • Монотонность труда; • Эмоциональные перегрузки. 		<p>производственных помещений;</p> <ul style="list-style-type: none"> • ГОСТ 12.1.003–83 ССБТ. Шум. Общие требования безопасности; • "Трудовой кодекс Российской Федерации" от 30.12.2001 N 197-ФЗ (ред. от 05.02.2018).
--	--	--	--

7.2.1.1 Монотонность труда

Монотонность труда – один из вредных факторов при работе с ПЭВМ. Во время работы происходит постоянное статическое напряжение мышц шеи, спины, рук и ног, что негативно сказывается как на здоровье человека, так и на его трудоспособность и эмоциональное состояние.

7.2.1.2 Микроклимат

Для уменьшения негативного влияния микроклимата помещения рекомендуется придерживаться следующих показателей температуры и влажности помещения, в зависимости от времени года. Ниже представлена таблица 19 оптимальных показателей микроклимата для категории работ 1a [23].

Таблица 19. Оптимальные величины показателей микроклимата на рабочих местах производственных помещений

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	22–24	21–25	60–40	0,1
Теплый	23–25	22–26	60–40	0,1

Таблица 20. Допустимые величины показателей микроклимата на рабочих местах производственных помещений

Период года	Температура воздуха, °С		Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с	
	Диапазон ниже оптимальных величин	Диапазон выше оптимальных величин			Ниже оптимальных величин, не более	Выше оптимальных величин, не более
Холодный	20,0–21,9	24,1–25,0	19–26	15–75	0,1	0,1
Теплый	21,0–22,9	25,1–28,0	20–29	15–75	0,1	0,2

В результате измерений была получена температура воздуха, равная 26 °С в теплое время года, что превышает оптимальное значение. Однако, согласно СанПиН 2.2.4.548-96, данный показатель находится в допустимом пределе 25,1–28,0 °С (табл. 20).

7.2.1.3 Освещение

Для снижения зрительной нагрузки при работе за компьютером важно подобрать комфортное освещение.

Источники освещения делятся на два основных типа: естественные и искусственные. Как правило, естественное освещение используется в дневное время суток, а искусственное преимущественно в темное время суток. Однако при недостаточном освещении оба этих типа могут комбинироваться.

Рабочее место в данной работе организовано таким образом, что тыльная сторона монитора обращена к окну. Размер рабочего места 1,0x0,7 м., размер окна 2,5x1,7 м.

Индекс помещения рассчитывается по следующей формуле [25]:

$$I_{\text{п}} = \frac{ab}{(h_1 - h_2)(a + b)}$$

Где $I_{\text{п}}$ – индекс помещения, h_1 – высота помещения, h_2 – высота рабочего стола, a – длина помещения, b – ширина помещения. Так длина помещения составила 5,12 м, ширина – 3,5 м, высота помещения – 2,5 м, а высота стола 0,75 м. Таким образом, индекс помещения равен:

$$I_{\text{п}} = \frac{3,5 * 5,12}{(2,5 - 0,75)(3,5 + 5,12)} = 1,188$$

Далее была рассчитана освещенность по формуле [26]:

$$E = \frac{K_{\text{св}} * K_{\text{л}} * \text{СП}_{\text{л}} * U}{S * k_3 * 100}$$

где $K_{\text{св}}$ – количество светильников, $K_{\text{л}}$ – количество ламп в светильнике, $\text{СП}_{\text{л}}$ – световой поток лампочки, U – коэффициент использования, S – площадь, k_3 – коэффициент запаса.

Поскольку потолок белого цвета, стены и пол – светлые, то коэффициент отражения потолка и стен равен 0,7, для пола – 0,3 [25]. Согласно таблице, коэффициент использования помещения U равен 55 [25].

В качестве основных источников освещения используются 3 светильника, каждый из которых содержит по 2 светодиодные лампы.

Мощность лампы 13 W, цветовая температура 4500 К. Коэффициент запаса равен 1,2 [26]. Таким образом, освещенность равна:

$$E = \frac{3 * 2 * 1200 * 55}{17,92 * 1,2 * 100} = 184 \text{ лк}$$

Нормативные значения для операций с высоким уровнем зрительной (объекты от 0,3 до 0,5 мм) работы составляют от 200 до 400 лк [27]. Таким образом, для достаточного освещения рабочего места необходимо либо работать в светлое время суток, либо использовать настольную лампу. В процессе работы в качестве дополнительного точечного источника света выступала настольная лампа, оснащенная источником света с цветовой температурой 3000 К и мощностью 9 W, расположенная слева от монитора.

7.2.1.4 Шум

Важное значение имеет показатель уровня шума в помещении. Повышенный уровень шума негативно воздействует на нервную и слуховую системы человека, приводя к различным заболеваниям, а также снижая работоспособность. К основным источникам шума при работе с ПЭВМ можно отнести шум систем охлаждения ПЭВМ, шум работающего жесткого диска. К другим источникам шума относятся уличный шум, бытовой шум.

Согласно ГОСТ [28] уровень шума не должен превышать 50 дБ. Для снижения уровня шума рекомендуется проводить мероприятия по техническому обслуживанию ПЭВМ.

7.2.1.5 Электробезопасность

В связи с тем, что работа выполняется с помощью ПЭВМ, то соблюдение правил электробезопасности имеет непосредственное отношение к работе. К основным источникам электрического воздействия, находящимся непосредственно рядом с рабочим местом, относятся ПЭВМ, настольная лампа, электрические розетки.

На территории России электроприборы включены в сеть под напряжением 220 В, с частотой 50 Гц, что является опасным фактором воздействия на организм человека. Поражающими факторами электрического тока являются термическое, биологическое и электролитическое воздействия.

7.2.1.6 Электромагнитные излучения

Компьютер, как и многие электроприборы, является источником электромагнитного излучения. Воздействие электромагнитного излучения на организм определяется различными параметрами, такими как напряженность поля, поток энергии, частота колебаний. Ниже приведены временные допустимые уровни электромагнитного поля (ЭМП) [21].

Таблица 21. Временные допустимые уровни ЭМП, создаваемых ПЭВМ на рабочих местах

Наименование параметров		Временные допустимые уровни электромагнитного поля
Напряженность электрического поля	5 Гц–2 кГц	25 В/м
	2 кГц–400 кГц	2,5 В/м
Плотность магнитного потока	5 Гц–2 кГц	250 нТл
	2 кГц–400 кГц	25 нТл
Напряженность электростатического тока		15 кВ/м

7.2.1.7 Пожарная безопасность

Пожары являются опасным фактором, который может привести к потере информации, хранящейся на ПЭВМ, а также, что немало важно, причинить вред здоровью человека. Поэтому меры противопожарной безопасности помогут избежать негативных последствий.

Главными вероятными источниками пожара могут стать неисправная электропроводка, поврежденные электроприборы и легковоспламеняющиеся вещества, например, бумага.

Для предотвращения возможных пожаров используются исправные электроприборы, а также сетевые фильтры с плавким предохранителем. В качестве мер быстрого реагирования используются дымовые датчики, а также пожарная сигнализация. Помимо этого, проводятся учения по эвакуации.

7.3 Экологическая безопасность

7.3.1 Анализ воздействия на окружающую среду

В процессе выполнения работы возможны такие отходы, как бумага, неисправные детали ПЭВМ, неработающие электролампы.

В состав компонентов ПЭВМ входят такие загрязняющие вещества, как ртуть, входящая в состав жидкокристаллических экранов, мышьяк и бериллий, используемые при производстве плат, свинец, применяемый для пайки, поливинилхлорид, используемый для изготовления изоляции кабелей. Сейчас некоторые из данных веществ запрещены для использования, например свинец, но остальные используются до сих пор. Добыча данных материалов уже сама по себе наносит вред окружающей среде [29].

Поэтому ПЭВМ и ее компоненты по окончании срока службы необходимо утилизировать соответствующим образом, чтобы избежать дальнейшего негативного влияния на окружающую среду. Кроме того, необходимо сдавать макулатуру в специальные пункты приема.

7.4 Безопасность в чрезвычайных ситуациях

В ходе выполнения ВКР могут возникнуть чрезвычайные ситуации (ЧС) техногенного, экологического, стихийного и биолого-социального характера. Ниже представлены наиболее вероятные и опасные ЧС по каждой категории:

- Техногенные:
 - Пожары;
- Стихийные:

- Ураганы;
- Биолого-социальные:
 - Эпидемии;

В случае пожара необходимо покинуть помещение согласно плану эвакуации, представленному на рисунке 23.



Рисунок 23. План эвакуации

Как видно из плана эвакуации, место проведения работ находится приблизительно в 5 м от ближайшего лестничного проема, что позволит оперативно эвакуироваться в случае ЧС.

Заключение

В последние годы для мониторинга лесов используются БПЛА. Именно с их помощью можно оперативно сделать снимок лесов, и по снимкам обнаружить вредителей лесов, и оценить степень повреждения деревьев.

В работе предложено автоматизировать процесс интерпретации снимков с БПЛА с помощью сверточных нейронных сетей. Для этого решается задача детектирования на снимках деревьев пихты на снимках, пораженных уссурийским полиграфом. Исследуются возможности СНС класса YOLO при решении этой задачи детектирования. При этом получаются следующие результаты.

Была проведена разметка исходного изображения, которое затем было нарезано на фрагменты.

Перед началом обучения СНС была произведена подборка размеров якорных ограничивающих прямоугольников с помощью алгоритма k-means.

В ходе обучения СНС было проведено исследование зависимости точности детектирования от числа эпох и используемых методов аугментации данных. В ходе исследования было получено, что при использовании двухэтапного подхода при обучении СНС достигаются наилучшие результаты детектирования на тестовой выборке.

Кроме того, для обученных СНС были подобраны значения пороговых параметров. В результате было выявлено, что при двухэтапном обучении СНС с большей уверенностью детектирует пихты, поскольку был использован порог с большим значением для вероятности классов.

После завершения обучения было проведено исследование различных модификаций алгоритма NMS. Однако для данного набора данных первоначальный алгоритм NMS продемонстрировал наилучшие по точности результаты.

СПИСОК ИСТОЧНИКОВ

1. С.А.Кривец, И.А.Керчев, Э.М.Бисирова, Н.В.Пашенова, Д.А.Демидко, В.М.Петько, Ю.Н.Баранчиков. Уссурийский полиграф в лесах Сибири (распространение, биология, экология, выявление и обследование поврежденных насаждений). Методическое пособие. Томск-Красноярск, 2015. 48 с.
2. Spatial localization and detection [Электронный ресурс] / Stanford University CS231n: Convolutional Neural Networks for Visual Recognition – URL: http://cs231n.stanford.edu/slides/2016/winter1516_lecture8.pdf (Дата обращения 08.09.2019);
3. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollar. Focal Loss for Dense Object Detection;
4. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks;
5. Joseph Redmon, Ali Farhadi. YOLO9000: Better, Faster, Stronger;
6. Joseph Redmon Ali Farhadi. YOLOv3: An Incremental Improvement;
7. YOLO v2 in Keras and applications [Электронный ресурс] / The world's leading software development platform – GitHub – URL: <https://github.com/experiencor/keras-yolo2> (Дата обращения 08.09.2019);
8. Detection Evaluation [Электронный ресурс] / COCO – Common objects in context – URL: <http://cocodataset.org/#detection-eval> (Дата обращения 09.12.2019);
9. YOLO: Real time object detection [Электронный ресурс] – URL: <https://pjreddie.com/darknet/yolo/> (Дата обращения: 29.05.2020);
10. Caffe [Электронный ресурс] – URL: <https://caffe.berkeleyvision.org/> (Дата обращения 08.09.2019);
11. DL4J [Электронный ресурс] – URL: <https://deeplearning4j.org/> (Дата обращения 08.09.2019);
12. Tensorflow [Электронный ресурс] – URL: <https://www.tensorflow.org/> (Дата обращения 08.09.2019);

13. Keras [Электронный ресурс] – URL: <https://keras.io/> (Дата обращения 08.09.2019);
14. Pytorch [Электронный ресурс] – URL: <https://pytorch.org/> (Дата обращения 08.09.2019);
15. MXNet [Электронный ресурс] – URL: <https://mxnet.apache.org/> (Дата обращения 08.09.2019);
16. Darknet [Электронный ресурс] – URL: <https://pjreddie.com/darknet/> (Дата обращения 08.09.2019);
17. Computer Vision Annotation Tool (CVAT) [Электронный ресурс] / github – URL: <https://github.com/opencv/cvat> (Дата обращения 11.02.2020);
18. Darknet: Convolutional neural networks [Электронный ресурс] / github – URL: <https://github.com/pjreddie/darknet> (Дата обращения 06.03.2020);
19. Kaggle: Amazon from Space – трюки и хаки при обучении нейросетей. [Электронный ресурс] / habr – URL: <https://habr.com/ru/company/ods/blog/413667/> (Дата обращения 25.02.2020);
20. Navaneeth Vodla, Bharat Singh, Rama Chellappa, Larry S. Davis: Improving Object Detection With One Line of Code;
21. ГОСТ 12.0.003-74. ССБТ. Опасные и вредные производственные факторы. Классификация;
22. "Трудовой кодекс Российской Федерации" от 30.12.2001 N 197-ФЗ (ред. от 27.12.2018).
23. СанПиН 2.2.4.548–96. Гигиенические требования к микроклимату производственных помещений;
24. СанПиН 2.2.2/2.4.1340–03. Санитарно-эпидемиологические правила и нормативы «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы»;
25. Таблица «Коэффициент использования светового потока» [Электронный ресурс] – URL: <https://www.websor.ru/metodkoeffi.html> (дата обращения 28.05.2020);

26. Как самостоятельно выполнить расчет освещенности помещения [Электронный ресурс] / Электрика своими руками – URL: <https://elektrika-svoimi-rykami.com/raschet-osveshheniya/raschet-osveshheniya> (дата обращения 02.06.2020);

27. СНиП 23-05-95. Естественное и искусственное освещение;

28. ГОСТ 12.1.003–83 ССБТ. Шум. Общие требования безопасности;

29. Грязная и опасная сторона технологий [Электронный ресурс] / Мир ПК – URL: <https://www.osp.ru/pcworld/2013/06/13035804> (дата обращения 19.05.2020).

Приложение А
(справочное)

Convolutional neural network training

Студент

Группа	ФИО	Подпись	Дата
8ИМ81	Фоминский Виталий Сергеевич		

Руководитель ВКР

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

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Маркова Наталия Александровна			

1. Convolutional neural network training

1.1 Defining the size of anchor bounding boxes

It is necessary to get the hypoparameters of the network before starting the training. The anchor bounding boxes' size is one of them. To define it the k-means algorithm was used [5]. The distance metric was calculated with following formula:

$$d(box, centroid) = 1 - IOU(box, centroid),$$

where d is distance between current bounding box and bounding boxes selected as center of a cluster. Intersection over union (IOU) was used as metric. The results of k-means algorithm are shown in Figures 1- 3. The IOU dependence on different number of clusters is shown in Figure 4. The more the number of clusters, the more accurate the results. But the speed of detection increases with decreasing number of clusters. 5 clusters were chosen to achieve the balance between speed and detection quality.

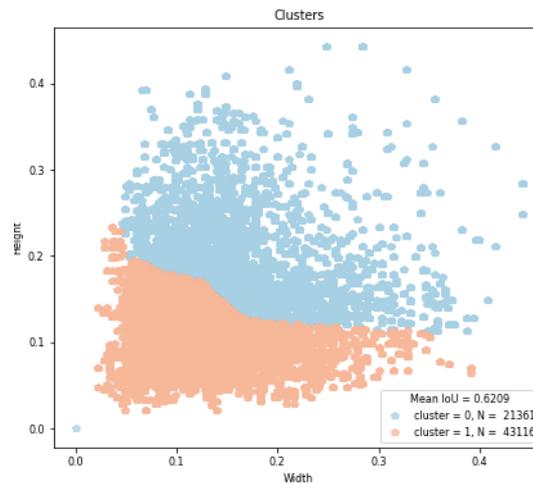


Figure 1. The result of splitting into 2 clusters

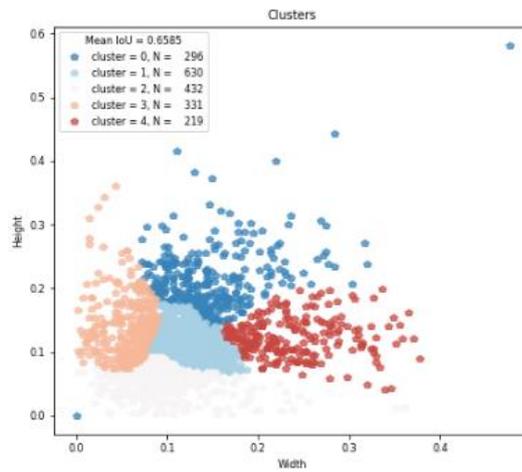


Figure 2. The result of splitting into 5 clusters

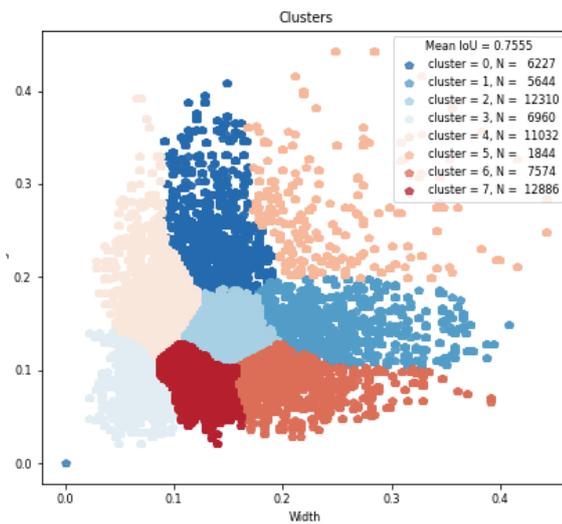


Figure 3. The result of splitting into 8 clusters

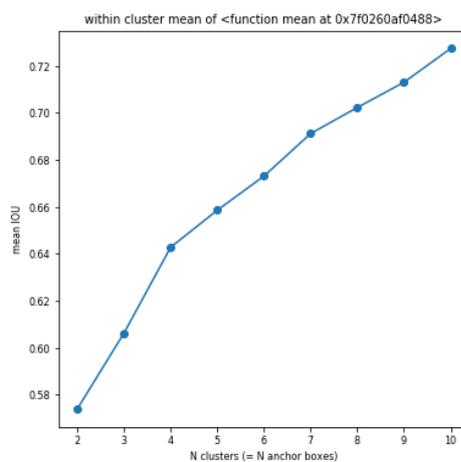


Figure 4. IOU dependence on different number of clusters

1.2 YOLO v2 research and training

During the experiments quality dependence in data augmentation methods and number of epochs was researched. Network hyperparameters were constant during the training process.

Convolutional neural network (CNN) YOLO v2 predicts many bounding boxes. It is necessary to remove some of them. At first, the more probable bounding boxes are chosen. Among the remaining bounding boxes, bounding boxes with IOU value keep less than IOY threshold. And finally, it is necessary to get rid of bounding boxes with low probability. The IOU threshold value equals 0,45 and the probability threshold value equals 0,35. These values may be changed later.

Adam optimization algorithm with learning rate $3 \cdot 10^{-4}$ was used during the training. The weights of the network were saving during the training process if the error on validation set was reduced. These weights were used later during the testing.

To estimate the results mean average precision (mAP) metric was used. To calculate it, it is necessary to choose intersection over union (IOU) threshold. In this work IOU threshold equals 0.5.

Loss function was chosen from the paper [5], it is shown below:

$$\begin{aligned}
 loss_{coord} &= \lambda_{coord} \sum_i^{S^2} \sum_j^B 1^{obj} [(x_{ij} - \hat{x}_{ij})^2 + (y_{ij} - \hat{y}_{ij})^2 + (\sqrt{w_{ij}} - \sqrt{\hat{w}_{ij}})^2 \\
 &\quad + (\sqrt{h_{ij}} - \sqrt{\hat{h}_{ij}})^2] \\
 loss_{conf} &= \lambda_{obj} \sum_i^{S^2} \sum_{c \in C} (1^{obj} + \lambda_{no\ obj} (1 - 1_{ij}^{obj})) (C_{ij} - \hat{C}_{ij})^2 \\
 loss_{class} &= \lambda_{class} \sum_i^{S^2} \sum_{c \in C} 1^{obj} p_i(c) \log(\hat{p}_i(c)) \\
 loss &= loss_{coord} + loss_{conf} + loss_{class}
 \end{aligned}$$

where total loss equals sum of coordinates loss function $loss_{coord}$, object confidence loss function $loss_{conf}$ and class probability loss function $loss_{class}$. x, y are centers of

bounding boxes, w , h are their weights and heights. C is confidence of object existence in the grid cell for each anchor box, $p(c)$ is class probability. λ_{coord} , λ_{obj} , $\lambda_{\text{no obj}}$, λ_{class} are coefficients, equal to 1, 5, 1, 1 respectively. They were taken from repository [18] and were not changed.

The first training set had 766 images. Then, number of images in training set was increased to 3064 with 90° , 180° and 270° rotation. The next method of increasing number of images was random brightness and contrast changes. The results are shown in Table 1.

Table 1. Detection quality dependence on different data augmentation methods and number of epochs

Augmentation method	Epochs number		
	5	10	25
No	0.032	0.036	0.029
Rotation	0.072	0.080	0.078
Rotation, brightness and contrast changes	0.101	0.116	0.078

It is clear that adding of augmentation methods increases quality of detection. But increment number of epochs does not influence on detection quality.

After that two-stage approach was explored [19]. On the first stage Convolutional neural network (CNN) used more augmentation techniques. On the second stage brightness and contrast changes were not used. The results are shown in Table 2.

Table 2. Detection quality dependence on duration of training stages

Number of epochs without brightness and contrast changes	Number of epochs with brightness and contrast changes		
	5	10	25
1	0.134	0.102	0.111
3	0.175	0.131	0.126
5	0.148	0.108	0.131
10	0.158	0.156	0.119

The best result was achieved when number of epochs with augmentation was 5 and number of epochs without augmentation was 3.

To demonstrate result 3 fragments from test set is chosen. Fir trees are at the angle to the camera of unmanned aerial vehicle (UAV) in the first image. In the second image trees are in front of camera and in the third image trees are under the camera.

The results of CNN evaluation are shown in Figures 5- 7. Left image is image with ground true bounding boxes. Center image shows detection result after 5 epochs training with all augmentation methods and the right image demonstrates detection results after 5 epochs training with augmentation and 3 epochs without augmentation.

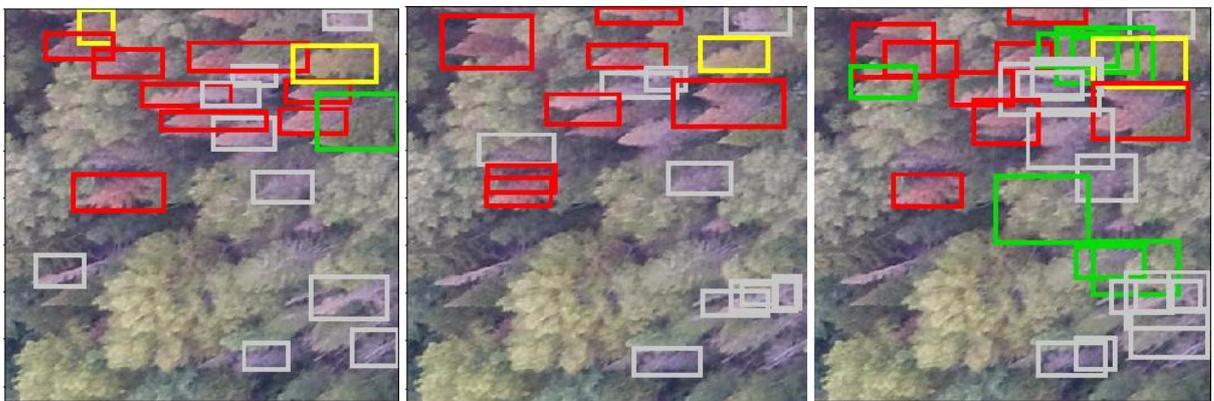


Figure 5. Detection results for trees at the angle to the camera

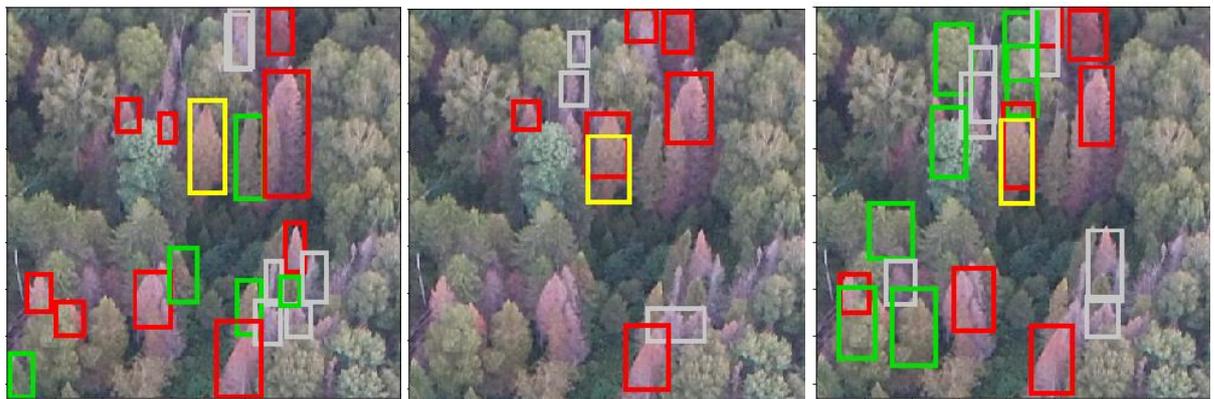


Figure 6. Detection results for trees in front of camera



Figure 7. Detection results for trees in under the camera

It is obvious that after two-stage training process network produces more bounding boxes. Therefore, each bounding box of network that was trained in two stages has greater probability for all classes, so these results are more confident. Secondly, two-stage trained network predicts some bounding boxes for one tree. This demonstrates the need to tune the threshold values.

After network has trained it is necessary to find threshold values for non-maximum suppression algorithm (NMS).

As shown in Table 3, detection quality decreases with increasing of the probability threshold. Nevertheless, dependence on IOU threshold is non-linear. Detection quality reaches the peak at 0.45 IOU threshold value.

Table 3. Metric dependence on threshold values

Probability threshold	0.1	0.2	0.3	0.35	0.4	0.5	0.6
IOU threshold							
1	0.182	0.175	0.160	0.154	0.139	0.113	0.080
0.9	0.182	0.175	0.161	0.154	0.139	0.113	0.080
0.8	0.184	0.177	0.162	0.155	0.140	0.114	0.081
0.7	0.194	0.186	0.170	0.163	0.147	0.119	0.083
0.6	0.204	0.196	0.179	0.171	0.153	0.124	0.085
0.5	0.208	0.200	0.182	0.173	0.155	0.124	0.086
0.45	0.212	0.203	0.184	0.174	0.158	0.126	0.086
0.4	0.211	0.202	0.183	0.174	0.158	0.126	0.087
0.3	0.207	0.199	0.180	0.172	0.156	0.125	0.086

The detection results at the 0.45 IOU threshold value and different probability threshold values are shown in Figure 8. The original image with ground true bounding boxes is located on the left-up corner. The other images show

network detection results, when probability threshold was changing in interval [0.2; 0.5]. It is seen that at the 0.2 probability threshold network predicts extra bounding boxes, but at the 0.5 probability threshold many bounding boxes are left.

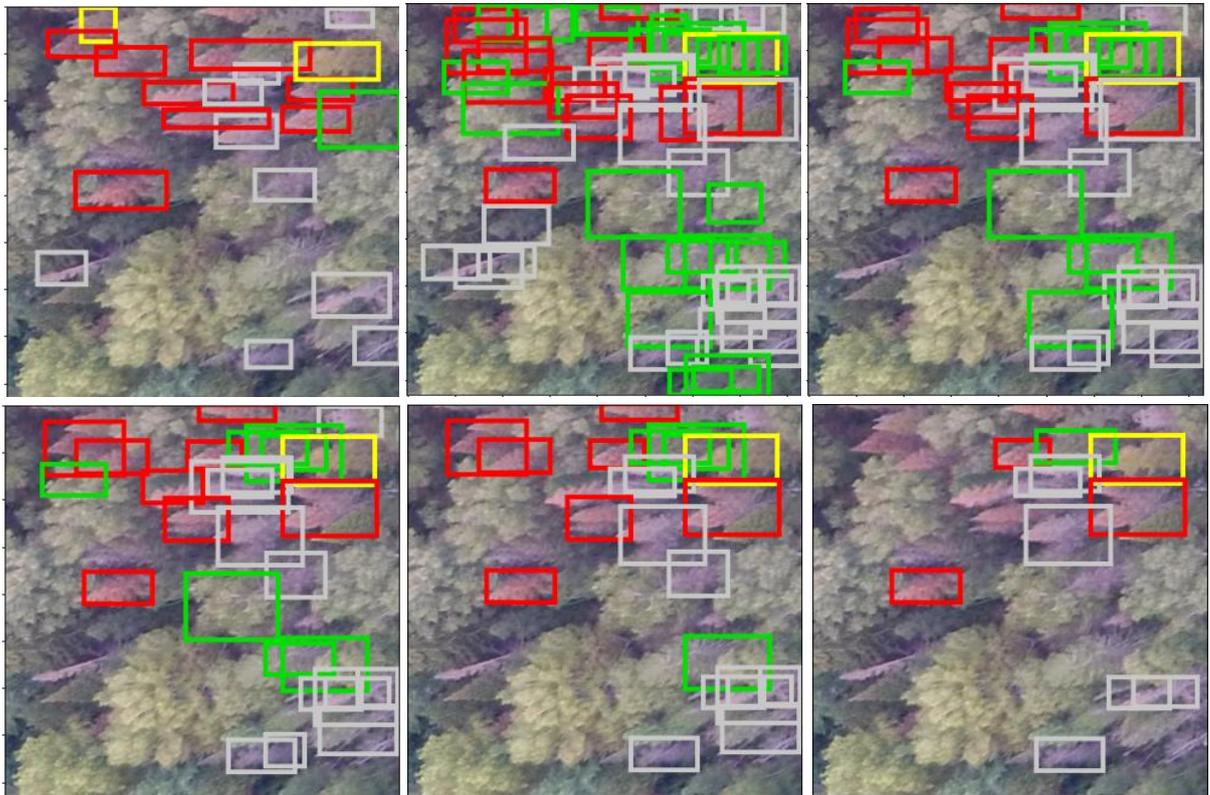


Figure 8. Detection results at different values of probability threshold

Detection results at fixed value of 0.3 probability threshold are described next. The IOU threshold was changing in interval [0.4; 0.7]. The results are shown in Figure 9. The original image with ground true bounding boxes is located on the left-up corner. The more IOU threshold is, the more bounding boxes are on the image. So, the best threshold according to the mAP metric and visual similarity of true and predicted images is 0.3 or 0.35. The IOU threshold equals 0.45. So, default values can be used.



Figure 9. Detection results at different values of IOU threshold

1.3 YOLO v2 tiny research and training

In addition to YOLO v2, YOLO v2 tiny training was run. Dependence on number of epochs was explored. Research of dependence on data augmentation methods was not conducted.

Adam optimization algorithm with $3 \cdot 10^{-4}$ learning rate was used as well. The best weights of the model were saving during the training process. They were used in testing process.

As for YOLO v2, two-stage training approach was researched. The results were received with 0.35 probability threshold and 0.45 IOU threshold (Table 4).

Table 4. Detection quality dependence on duration of training stages

Number of epochs without brightness and contrast changes	Number of epochs with brightness and contrast changes		
	5	10	25
1	0,068	0,073	0,074
3	0,084	0,099	0,075
5	0,071	0,079	0,058
10	0,076	0,074	0,067

The best result was achieved, when the number of epochs with augmentation 10 and number of epochs without augmentation was 3. Results are shown in Figure 10.



Figure 10. Detection results for trees at the angle to the camera

After the network was trained, it was necessary to find threshold values for NMS algorithm. According to previous results, the range of values was reduced. So, the range of IOU threshold was [0,3; 0,6] and range of probability threshold was [0,2; 0,5].

As shown in Table 5, the maximum value of mAP was achieved at 0,5 or 0,45 IOU threshold. So, IOU threshold equals 0,5. The quality of detection decreases with increasing of probability threshold.

Table 5. Metric dependence on threshold values

Probability threshold	0,2	0,3	0,35	0,4	0,5
IOU threshold					
0,6	0.119	0.104	0.097	0.088	0.072
0,5	0.123	0.107	0.099	0.090	0.074
0,45	0.122	0.106	0.099	0.090	0.073
0,4	0.117	0.103	0.097	0.088	0.072
0,3	0.114	0.101	0.094	0.086	0.072

Detection results at different values of probability threshold are shown in Figure 11. The left image is original image with ground true bounding boxes. Detection results with the least and the largest probability threshold values are shown in the other images.



Figure 11. Detection results at different values of probability threshold

It is clear that first model predicts many false objects, while in the second case some of bounding boxes are left. Thereby, IOU threshold value equals 0,45 and probability threshold value equals 0,35.

1.4 NMS algorithm researching

After CNN training, researching of NMS algorithm was conducted. The original NMS algorithm sets probability of intersected bounding boxes to zero. It means that some of bounding boxes are left.

To solve this problem soft-NMS algorithm was used [20]. Unlike original NMS, it reduces the probability of bounding boxes by value, proportional to IOU. The formula is shown bellow

$$pr = \begin{cases} pr, & IOU < IOU_{threshold} \\ pr * (1 - IOU), & IOU \geq IOU_{threshold} \end{cases}$$

where pr is a class probability. Moreover, the Gaussian penalty function was researched. It is shown below. Unlike the last function, it is continuous [20].

$$pr = pr * e^{-\frac{IOU^2}{IOU_{threshold}}}$$

According to the paper [20], this algorithm increases quality detection of YOLO v2 by 0,5%.

So, the research was conducted for both YOLO v2 and YOLO v2 tiny. The probability threshold was constant during the research. The IOU threshold was changed in range [0,3; 0,6]. The results are shown in Table 6, where Gaussian soft-NMS is soft-NMS algorithm with Gaussian penalty function.

Table 6. Quality detection dependence on IOU threshold and different NMS algorithms

Algorithm	IOU threshold	0,3	0,4	0,45	0,5	0,6
	CNN architecture					
NMS	YOLO v2	0,172	0,174	0,174	0,173	0,171
	YOLO v2 tiny	0,094	0,097	0,099	0,099	0,097
Soft-NMS	YOLO v2	0,173	0,174	0,174	0,172	0,170
	YOLO v2 tiny	0,095	0,096	0,098	0,099	0,097
Gaussian	YOLO v2	0,171	0,172	0,173	0,173	0,174
Soft-NMS	YOLO v2 tiny	0,093	0,094	0,095	0,096	0,097

It is seen from Table 6 that detection metric was not improved.

Приложение Б

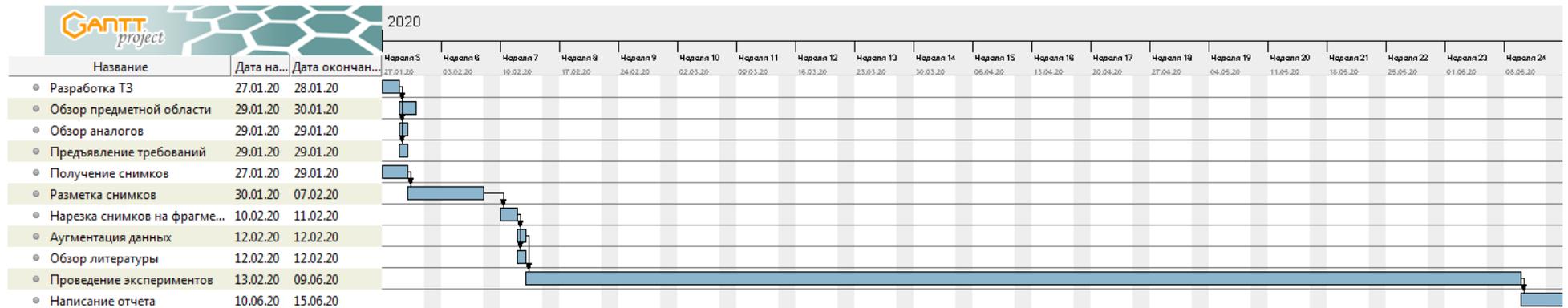


Рисунок 24. Диаграмма Ганта