

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

Инженерная школа информационных технологий и робототехники
 Направление подготовки – 09.04.04 Программная инженерия
 Отделение школы (НОЦ) – Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Алгоритм локализации сельскохозяйственных растений в течение вегетативного периода на изображении

УДК: 004.415.2:004.421:631.5

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Сергеева Наталья Дмитриевна		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Аксенов С.В.	к.т.н.		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Меньшикова Е.В.	к.филос.н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Горбенко М.В.	к.т.н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Губин Е.И.	к.ф.-м.н.		

Томск – 2020 г.

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП

Код результата	Результат обучения
<i>Общие по направлению подготовки 09.04.04 «Программная инженерия»</i>	
P1	Способность проводить научные исследования, связанные с объектами профессиональной деятельности
P2	Способность разрабатывать новые и улучшать существующие методы и алгоритмы обработки данных в информационно-вычислительных системах
P3	Способность составлять отчеты о проведенной научно-исследовательской работе и публиковать научные результаты
P4	Способность проектировать системы с параллельной обработкой данных и высокопроизводительные системы
P5	Способность осуществлять программную реализацию информационно-вычислительных систем, в том числе распределенных
P6	Способность осуществлять программную реализацию систем с параллельной обработкой данных и высокопроизводительных систем
P7	Способность организовывать промышленное тестирование создаваемого программного обеспечения
<i>Профиль «Технологии больших данных»/ «Big data solutions»</i>	
P8	Способность исследовать и анализировать большие данные, создавать их модели и интерпретировать структуры данных в таких моделях
P9	Способность понимать принципы создания, хранения, управления, передачи и анализа больших данных с использованием новейших технологий, инструментов и систем обработки данных в высокопроизводительных сетях
P10	Способность применять теорию распределенной системы управления базами данных к традиционным распределенным системам реляционных баз данных, облачным базам данных, крупномасштабным системам машинного обучения и хранилищам данных

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

Инженерная школа информационных технологий и робототехники
 Направление подготовки – 09.04.04 Программная инженерия
 Отделение школы (НОЦ) – Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП

 (Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
8ПМ8И	Сергеева Наталья Дмитриевна

Тема работы:

Алгоритм локализации сельскохозяйственных растений в течение вегетативного периода на изображении	
Утверждена приказом директора	№59-62/с от 28.02.2020 г.

Срок сдачи студентом выполненной работы:	15.06.2020
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	3283 изображения растений сахарной свеклы на разных стадиях вегетативного роста
---------------------------------	---

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ol style="list-style-type: none"> 1. Аналитический обзор подходов к решению задач: <ul style="list-style-type: none"> – Сегментация изображений. – Детекция ключевых точек на изображении. 2. Разработка концепции алгоритма. 3. Выбор подходящих инструментов для реализации решения. 4. Реализация следующих частей алгоритма: <ul style="list-style-type: none"> – Нейронная сеть для сегментации изображений. – Функция потерь для задачи детекции ключевых точек. – Модуль определения относительных координат выделенных нейронной сетью областей. 5. Аннотация изображений и обучение модели. 6. Проведение оценки обученной модели. 7. Оценка потенциала разработки с точки зрения особенностей рынка, актуальности решаемой задачи и экономической эффективности 8. Рассмотрение условий труда исполнителей настоящего проекта. 9. Обсуждение результатов выполненной работы.
<p>Перечень графического материала</p>	<p>Рисунки и таблицы в описании результатов аналитического обзора предметной области, реализованного алгоритма, рассмотрения вопросов финансового менеджмента и социальной ответственности.</p>
<p>Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i></p>	
<p>Раздел</p>	<p>Консультант</p>
<p>Социальная ответственность</p>	<p>Горбенко Михаил Владимирович, доцент ООД ШБИП, к.т.н.</p>
<p>Финансовый менеджмент, ресурсоэффективность и ресурсосбережение</p>	<p>Меньшикова Екатерина Валентиновна, доцент ОСГН, к.филос.н.</p>
<p>Обязательное приложение на английском языке</p>	<p>Пичугова Инна Леонидовна, ст. преподаватель ОИЯ ШБИП</p>

<p>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</p>	<p>01.02.2020</p>
--	-------------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Аксенов С.В.	к.т.н		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Сергеева Наталья Дмитриевна		

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

Инженерная школа информационных технологий и робототехники
 Направление подготовки – 09.04.04 Программная инженерия
 Уровень образования магистратура
 Отделение школы (НОЦ) – Отделение информационных технологий
 Период выполнения: весенний семестр 2018 /2019 учебного года

Форма представления работы:

Магистерская диссертация

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:	15.06.2020
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
30.02.2020	<i>1 Обзор методов сегментации изображений и детекции ключевых точек</i>	20
18.03.2020	<i>2 Проектирование метода решения задачи локализации объектов на изображении</i>	30
25.04.2020	<i>3 Реализация алгоритма локализации объектов на данных с изображениями сахарной свеклы на разных стадиях вегетативного роста</i>	30
11.05.2020	<i>4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение</i>	10
27.05.2020	<i>5 Социальная ответственность</i>	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Аксенов С.В.	к.т.н.		

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Губин Е.И.	к.ф.-м.н.		

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

Студенту:

Группа	ФИО
8ПМ8И	Сергеева Наталья Дмитриевна

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	Программная инженерия

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

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	- Оклад инженера – 21760; - Оклад научного руководителя – 33664;
2. <i>Нормы и нормативы расходования ресурсов</i>	- Месячная норма амортизации – 2,8%
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	- Ставки налоговых отчислений во внебюджетные фонды (ст. 426 НК РФ) – 30% - Районный коэффициент по г. Томску (ст. 426 НК РФ, Постановление Правительства РФ от 13.05.92. №309) – 1,3

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. <i>Оценка коммерческого и инновационного потенциала НТИ</i>	- Анализ потенциальных потребителей - Анализ конкурентных технических решений - SWOT-анализ.
2. <i>Разработка устава научно-технического проекта</i>	Определение: - целей и ожидаемого результата проекта; - организационной структуры проекта; - ограничений и допущений проекта.
3. <i>Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок</i>	- Планирование структуры работ проекта. - Построение плана проекта - Формирование бюджета. - Анализ рисков проекта.
4. <i>Определение ресурсной, финансовой, экономической эффективности</i>	Расчет показателя финансовой эффективности.

Перечень графического материала (с точным указанием обязательных чертежей):

1. *Сегментация рынка*
2. *Результаты анализа конкурентных решений*
3. *Матрица SWOT разработки*
4. *Календарный план-график проекта*
5. *Бюджет затрат*
6. *Реестр рисков*

Дата выдачи задания для раздела по линейному графику

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Меньшикова Екатерина Валентиновна	к.филос.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Сергеева Наталья Дмитриевна		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8ПМ8И	Сергеева Наталья Дмитриевна

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	Программная инженерия

Тема ВКР:

Алгоритм локализации сельскохозяйственных растений в течение вегетативного периода на изображении	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Алгоритм локализации объектов на изображении, предназначенного для определения координат структурного центра объекта или нескольких объектов на изображении, на примере фотографий сельскохозяйственных растений в течение вегетативного периода.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: <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"> – специальные правовые нормы трудового законодательства при работе с компьютером и орг. техникой; – требования к организации рабочего места;
2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	<ul style="list-style-type: none"> – повышенная напряженность магнитного поля; – отклонение показателей микроклимата; – превышение уровня шума; – отсутствие или недостаток естественного света; – психофизиологический фактор.
3. Экологическая безопасность:	<ul style="list-style-type: none"> – анализ воздействия объекта на литосферу, гидросферу и атмосферу (отходы, связанные с утилизацией вышедшего из строя ПК, люминесцентных ламп и др.); – разработка решений по обеспечению экологической безопасности.
4. Безопасность в чрезвычайных ситуациях:	<ul style="list-style-type: none"> – возможная ЧС – пожар; – разработка мер по предупреждению пожара; – разработка действий при пожаре.

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

Задание выдал консультант:

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

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ8И	Сергеева Наталья Дмитриевна		

РЕФЕРАТ

Выпускная квалификационная работа 110 с., 41 рис., 24 табл., 36 источников.

Ключевые слова: компьютерное зрение, детекция ключевых точек, распознавание объектов, сегментация изображений, сверточные нейронные сети, взвешенное расстояние хаусдорфа.

Объект исследования: изображения сельскохозяйственных растений в течение вегетационного периода

Предмет исследования: методы компьютерного зрения для локализации объектов.

Цель работы: разработка алгоритма локализации сельскохозяйственных растений на разных стадиях роста в течение вегетативного периода.

В процессе исследования проводились: анализ задачи распознавания объектов; изучение, сравнение методов локализации объектов на изображениях; изучение способов устранения артефактов, возникающих при использовании методов; тестирование техник повышение качества обучаемой модели.

В результате исследования разработан алгоритм локализации объектов на изображениях, реализованный на изображениях сельскохозяйственных растений.

Область применения: разработанный алгоритм может быть применен в системах компьютерного зрения сельскохозяйственных полевых роботов, применяемых в земледелии.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ

Компьютерное зрение – междисциплинарная научная область, которая занимается компьютерной обработкой изображений и видеопоследовательностей с целью получения информации высокого уровня. С точки зрения инженерии, цель компьютерного зрения состоит в автоматизации задачи, которые может выполнять зрительная система человека.

Распознавание объектов – задача компьютерного зрения по идентификации объектов на изображении или видеоряде.

CNN (Convolutional Neural Network) – сверточная нейронная сеть, класс нейронных сетей для глубокого обучения, применяемая для анализа и создания изображений.

Машинное обучение – прикладная область искусственного интеллекта, занимающаяся алгоритмами, способными автоматически обучаться в процессе применения решений множества сходных задач.

Глубокое обучение – раздел методов машинного обучения, использующих нейронные сети, состоящие из множества слоев и способные обучаться на основе неструктурированных данных.

Содержание

Введение	12
1 Обзор литературы по теме распознавания объектов и методов локализации объектов на изображении	14
1.1 Общие понятия задачи распознавания объектов	14
1.2 Детекция ключевых точек.....	16
1.3 Методы распознавания объектов и детекции ключевых точек	18
1.4 Артефакты сегментации в задаче детекции ключевых точек	21
1.4.1 Причины возникновения артефактов	21
1.4.2 Методы устранения артефактов в сверточных нейронных сетях ...	24
2 Проектирование метода решения задачи локализации объектов на изображении	28
2.1 Описание подхода к решению задачи	28
2.2 Определение структуры алгоритма	30
2.3 Разработка модуля сегментации изображений	32
2.3.1 Функция ошибки для задачи детекции ключевых точек.....	32
2.3.2 Архитектура нейронной сети	35
2.3.3 Выбор метода оптимизации	38
2.4 Разработка модуля вычисления относительных координат объекта на изображении	40
2.5 Выбор метрик оценки обученной модели	42
Выводы	44
3 Реализация алгоритма локализации объектов на изображениях сахарной свеклы на разных стадиях вегетативного роста	45
3.1 Подготовка данных	45
3.1.1 Описание данных и аннотация изображений.....	45
3.1.2 Аугментация данных в обучающей выборке	46
3.1.3 Описание дополнительной выборки изображений для предтренировки модели	47
3.2 Описание стека технологий	49

3.3 Обсуждение результатов.....	50
3.3.1 Сравнение моделей.....	50
3.3.2 Результаты тренировки выбранной модели	50
Выводы	55
4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение...	56
4.1 Предпроектный анализ.....	57
4.1.1 Потенциальные потребители	57
4.1.2 Анализ конкурентных технических решений	58
4.1.3 SWOT-анализ	59
4.2 Инициация проекта	62
4.2.1 Цели и результат проекта.....	62
4.2.2 Организационная структура проекта.....	63
4.2.3 Ограничения и допущения проекта	63
4.3 Планирование управления проектом	64
4.3.1 Структура работ в рамках проекта	64
4.3.2 Определение трудоемкости выполнения работ	65
4.3.3 Бюджет проекта	67
4.3.4 Реестр рисков проекта	72
4.4 Определение экономической эффективности.....	74
Выводы	74
5 Социальная ответственность	76
5.1 Правовые и организационные вопросы обеспечения безопасности	76
5.2 Производственная безопасность.....	79
5.3 Экологическая безопасность.....	87
5.4 Безопасность в чрезвычайных ситуациях	89
Выводы	91
Заключение	92
Список использованных источников	94
Приложение I.....	98

Введение

Тенденция к роботизации в сельском хозяйстве обусловлена в числе прочих причин развитием точного земледелия, цель которого заключается в снижении негативного влияния современных методов обработки сельхозугодий на потребителей и фермерские хозяйства. Полевые роботы оснащаются системами компьютерного зрения, что позволяет им производить манипуляции с растениями точно и аккуратно. Актуальность этого направления робототехники объясняется стремлением к экологичному и безопасному производству продуктов и снижению экономических издержек. Например, система распознавания видов растений в роботизированном культиваторе делает возможным распыскивание гербицидов непосредственно на побеги сорняков, что отвечает обеим вышеупомянутым целям, так как такой метод не допускает попадания вредных веществ в возделываемую культуру и позволяет экономно использовать ресурсы.

Для роботов, применяемых в точном земледелии, требуется система компьютерного зрения, способная распознавать объекты на снимках или видеоряде, сделанных роботом во время обработки.

В данной работе рассматривается случай задачи распознавания, в котором необходимо определять положение культурных растений на изображении, неявно отделяя их от сорняков. Подобная система может быть применена в автоматических культиваторах, используемых для точечного внесения удобрений, механической прополки или рыхления почвы.

К техническому решению ставится требование определять на изображении относительные координаты структурных центров растений, которые определяются областью розеток побегов и в проекции соответствуют расположению корня растения. К задаче прилагается датасет, состоящий из снимков сахарной свеклы на различных стадиях вегетационного роста, сфотографированных камерой автоматического культиватора, направленной вертикально вниз (подробно датасет описан в параграфе 3.1).

Цель данной работы – разработка алгоритма локализации сельскохозяйственных растений на разных стадиях роста в течение вегетативного периода.

В работе разработка алгоритма проводится в два этапа: сначала проектируется общее решение для локализации объектов на изображении, затем реализуется решение на предоставленных данных с изображениями сахарной свеклы. Задача локализации интерпретирована как задача детекции ключевых точек объектов на изображении.

В рамках работы по разработке алгоритма локализации сельскохозяйственных растений выполнены следующие задачи:

- 1) Проведение обзора литературы по теме распознавания объектов и методов локализации объектов на изображении.
- 2) Выбор подхода к решению поставленной задачи и планирование структуры алгоритма.
- 3) Проектирование модуля сегментации изображений.
- 4) Разработка модуля вычисления относительных координат объекта на изображении.
- 5) Выбор методики оценки результатов обучения модели.
- 6) Реализация алгоритма локализации на имеющихся данных.
- 7) Обсуждение результатов.

1 Обзор литературы по теме распознавания объектов и методов локализации объектов на изображении

1.1 Общие понятия задачи распознавания объектов

Распознавание объектов – общее название класса задач компьютерного зрения. Распознавание объектов подразумевает идентификацию объектов на изображениях или видеопоследовательности с помощью алгоритмов машинного обучения и глубокого обучения.

Выделяют несколько основных типов задач компьютерного зрения, относящихся к распознаванию объектов:

- Классификация изображений – предсказание класса одного объекта на изображении. Отдельно выделяют задачу теггинга, в которой определяется наличие нескольких классов объектов на изображении.
- Локализация объектов – идентификация положения одного или нескольких объектов на изображении. Чаще всего подразумевает также рисование ограничивающей рамки (в англоязычной литературе – bounding box) вокруг объекта.
- Детекция объектов – задача, объединяющая две предыдущие и подразумевает определение положения и классификацию одного или нескольких объектов на изображении. Часто под распознаванием объектов имеют в виду именно задачу детекции.
- Сегментация экземпляров объектов – расширенная задача детекции объектов, в которой экземпляры распознанных объектов указываются не ограничивающей рамкой, а выделением всех пикселей экземпляров объектов.
- Семантическая сегментация – попиксельная классификация изображения. В данной задаче каждому пикселю изображения присваивается один из нескольких классов. Если в задаче определяется один класс, то пиксели определяются по принадлежности либо к этому классу, либо к фону.

Экземпляры объектов одного класса не разделяются, как в предыдущей задаче.

Рисунки 1.1 и 1.2 иллюстрируют перечисленные типы задач компьютерного зрения и разницу между двумя видами задач сегментации.

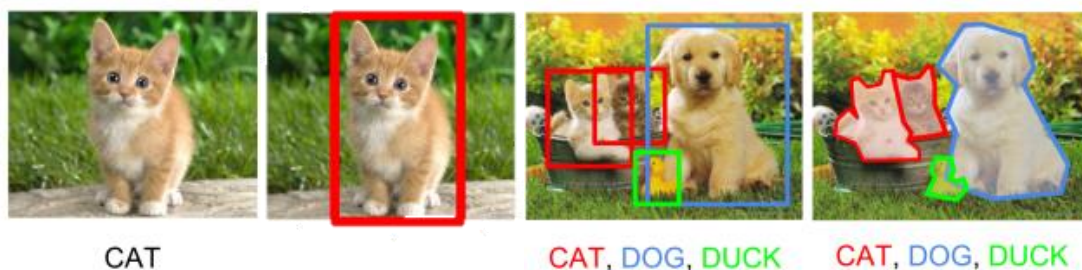


Рисунок 1.1 – Задачи компьютерного зрения: а) классификация, б) локализация, в) детекция, г) сегментация экземпляров объектов.

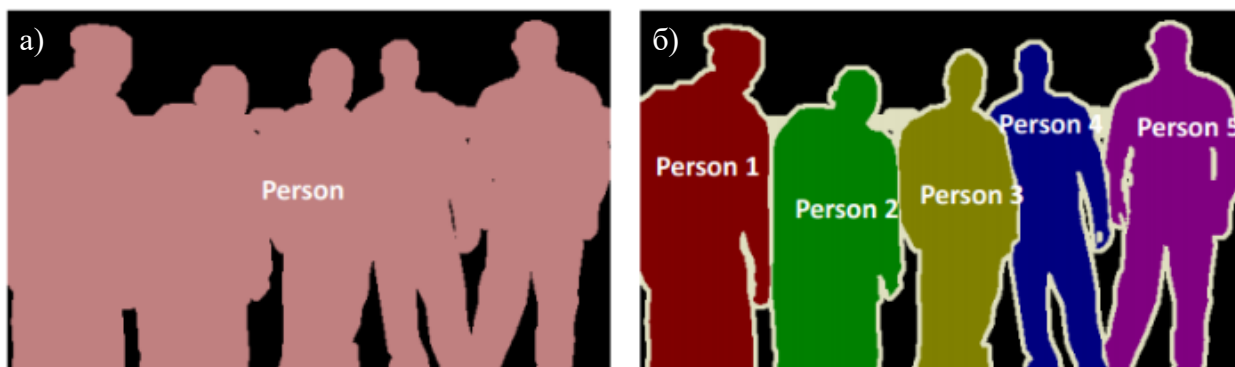


Рисунок 1.2 – Разница между семантической сегментацией (а) и сегментацией экземпляров объектов (б).

Иногда бывает трудно определить, к какому виду относится рассматриваемая задача, так как она может обладать чертами различных типов задач компьютерного зрения или особенности подходов к решению этих задач могут накладывать ограничения. Такого рода задачи могут быть решены с помощью комбинаций подходов или иных модификаций методов.

Задача, рассматриваемая в данной работе, в общем виде является задачей локализации объектов, так как требует определить положение объекта на изображении, не предсказывая его класс. Однако, цель задачи состоит в том, чтобы определить положение розетки побега, а не всего растения, что заставляет отказаться от методов решения задач локализации и детекции. С

этой точки зрения задача определения координат структурного центра побега должна быть рассмотрена как задача детекции ключевых точек.

1.2 Детекция ключевых точек

Детекция ключевых точек включает одновременное обнаружение объектов на изображении и локализацию их ключевых точек. Ключевые точки, или точки интереса, это пространственные локации или точки на изображении, которые определяют значимые области на изображении. Они инвариантны к повороту изображения, сжатию, сдвигу, искажению и так далее.

Детекция ключевых точек применяется во многих задачах компьютерного зрения, таких как SLAM (одновременная локализация и картирование), SfM (структура из движения), калибровка камеры и многие другие.

В основном, детекция ключевых точек обсуждается в контексте задачи распознавания человеческого тела или распознавания ключевых точек лица (в англоязычной литературе – *human pose estimation, facial keypoint detection*). Примеры приведены на рисунках 1.3 и 1.4. В данном случае детекция точек часто применяется для захвата структуры объектов и далее используется для различных манипуляций с изображением [1], анимации [2], в дополненной реальности [3].

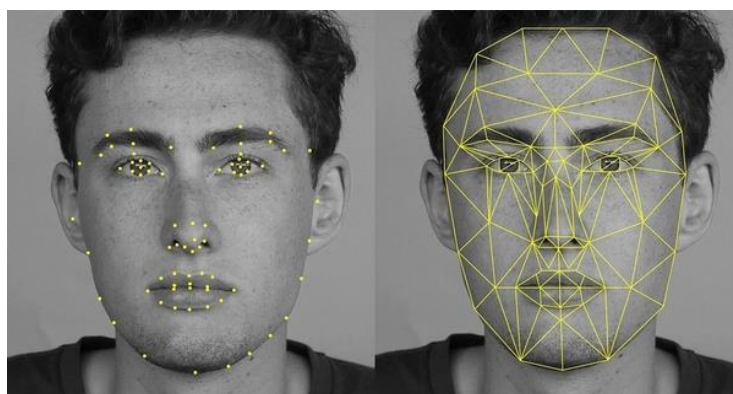


Рисунок 1.3 – Пример детекции ключевых точек лица.



Рисунок 1.4 – Пример детекции человеческой позы

В более широком контексте применения детекция ключевых точек помогает в распознавании объектов в случаях, когда объекты перекрывают друг друга и плохо разделимы с помощью ограничивающих рамок [4] – подобный случай проиллюстрирован на рисунке 1.5, – или в качестве вспомогательного средства при анализе более комплексных параметров объектов [5].



Рисунок 1.5 – Пример изображения с перекрывающимися объектами, ограничивающие рамки которых совпадают.

В рассматриваемой задаче наблюдается подобная сложность обнаружения интересующей области объекта, так как искомая часть растения

может быть закрыта листьями, а побеги – перекрывать друг друга. На рисунке 1.6 приведены примеры изображений из датасета, на которых видна данная проблема.

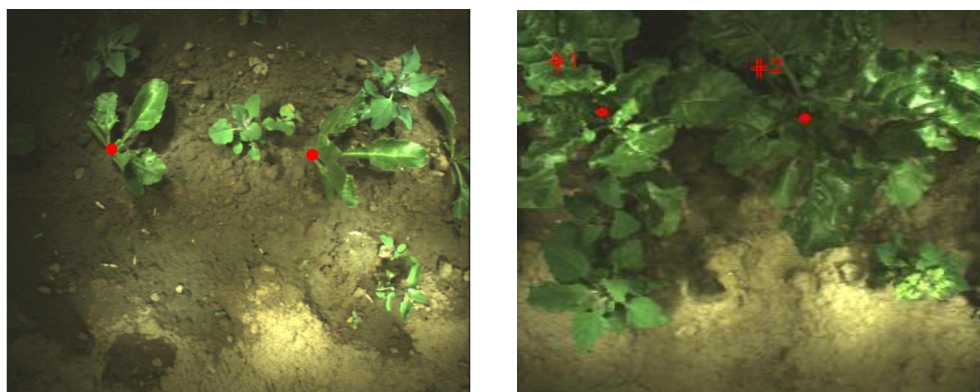


Рисунок 1.6 – Изображения датасета с перекрытиями побегов.

1.3 Методы распознавания объектов и детекции ключевых точек

Для распознавания объектов существуют различные подходы. Наиболее популярными методами являются семейства моделей R-CNN и YOLO.

Создание архитектуры R-CNN (“Regions with CNN Features”, “Region-Based Convolutional Neural Network”) связано с высокой эффективностью сверточных нейронных сетей в классификации изображений. Принцип подхода заключается в выделении регионов-кандидатов, на которых предполагается наличие каких-либо объектов, и обработке их сверточной нейронной сетью отдельно от целого изображения.

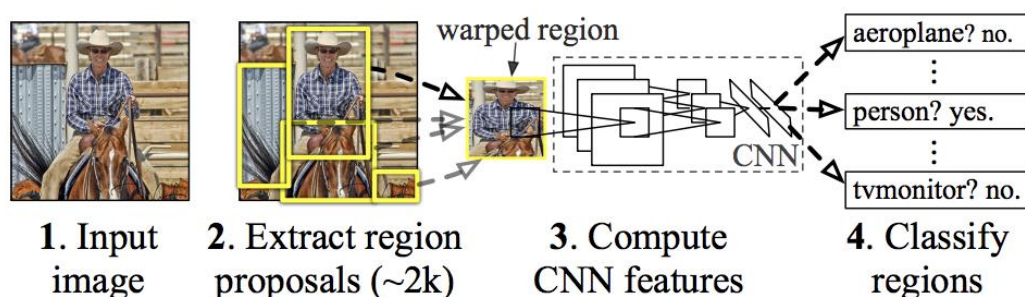


Рисунок 1.7 – Структура архитектуры R-CNN.

Архитектура зарекомендовала себя в решении задач локализации, детекции объектов и сегментации. Семейство включает архитектуры R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN [6-9].

Модели семейства YOLO (“You Only Look Once”) уступают R-CNN в качестве, однако их основное преимущество заключается в скорости обработки. Модели YOLO спроектированы для распознавания объектов в реальном времени [10]. Подход включает в себя единую обученную нейронную сеть, которая берет фотографию в качестве входных данных и прогнозирует ограничивающие рамки и метки классов для каждой ограничивающей рамки напрямую. Этот метод обеспечивает более низкую точность прогнозирования (например, больше ошибок локализации), но работает со скоростью 45 кадров в секунду (до 155 кадров в секунду для оптимизированной по скорости версии модели).

Модель работает, сначала разбивая входное изображение на сетку ячеек, где каждая ячейка отвечает за предсказание ограничивающей рамки, если центр ограничивающей рамки попадает в нее. Каждая ячейка сетки предсказывает ограничивающий прямоугольник, включающий координаты x , y , ширину, высоту и достоверность. Предсказание класса также основано на каждой ячейке.

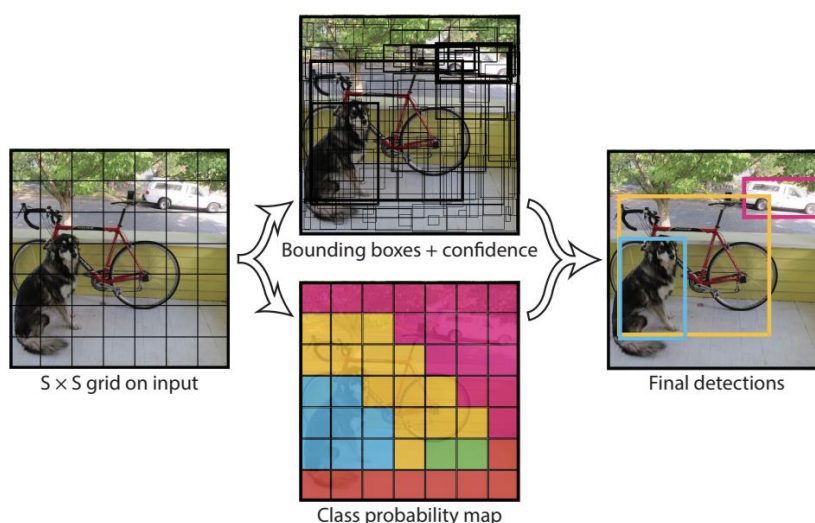


Рисунок 1.8 – Структура архитектуры YOLO.

Для решения задачи детекции точек с помощью CNN требуется добавление в архитектуру полносвязных слоев для того, чтобы модель была способна генерировать на выход координаты найденных ключевых точек. Однако в этом случае количество ключевых точек на выходе неявно задается числом выходов нейронной сети. Этот подход является удобным, если каждое изображение датасета содержит один экземпляр объекта, для которого нужно локализовать заданное количество ключевых точек.

Для варианта задачи, когда изображения содержат несколько экземпляров объектов, требуется более сложный подход. Наиболее успешные решения для задачи с несколькими объектами на изображении приводятся в работах [11-14].

В работе [15] приводится решение для извлечения ключевых точек для случая задачи, когда каждому объекту соответствует одна ключевая точка, и на изображении находится несколько объектов, количество которых заранее неизвестно. Авторы статьи модифицируют сверточную нейронную сеть дополнительной веткой для оценки количества объектов на изображении. На двух выходах нейронной сети генерируются маска сегментации и число, соответствующее предсказанному количеству объектов. Далее для обработки маски применяется алгоритм кластеризации с известным числом кластеров. Центры найденных кластеров интерпретируются как координаты ключевых точек объектов.

Подход с использованием техники сегментации для выделения областей интереса взят за основу при разработке алгоритма в данной работе. Однако прием с дополнительной веткой для оценки количества объектов не подходит для рассматриваемой задачи, так как ошибка в предсказании с меньшим количеством объектов, чем имеется на изображении, приведет к нежелательным последствиям при использовании алгоритма в конечной системе. Так как особенностью датасета является крупный масштаб снимков, то расстояние между точками интереса (розетками побегов) считается достаточным для того, чтобы модель выделила их в виде отдельных, четко

разделенных областей. В связи с этим решено извлекать координаты сегментированных зон посредством обработки маски методами классического компьютерного зрения.

1.4 Артефакты сегментации в задаче детекции ключевых точек

1.4.1 Причины возникновения артефактов

На изображениях, сгенерированных сверточными нейронными сетями, часто наблюдаются артефакты в форме шахматной доски. Данный недостаток – общая проблема сверточных нейронных сетей и встречается при решении задач различного вида. Например, на рисунке 1.9 иллюстрируется возникновение артефактов в задаче генерации изображений при использовании генеративно-сопоставительных сетей [16, 17].

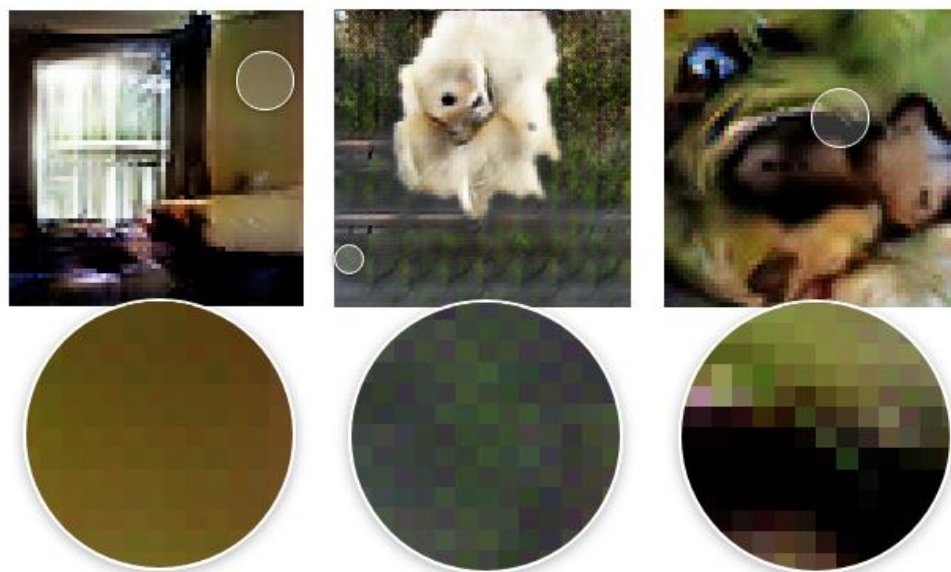


Рисунок 1.9 – Артефакты на изображениях в задаче генерации изображений

Можно заметить, что артефакты в виде шахматной доски более заметны на тех изображениях, где имеются более контрастные цвета. Поэтому артефакты сильно выражены на картах вероятностей в рассматриваемой задаче сегментации растений (рисунок 1.10), потому что они представляют из себя изображения, состоящие из пикселей со значениями либо близкими к нулю, либо к единице.

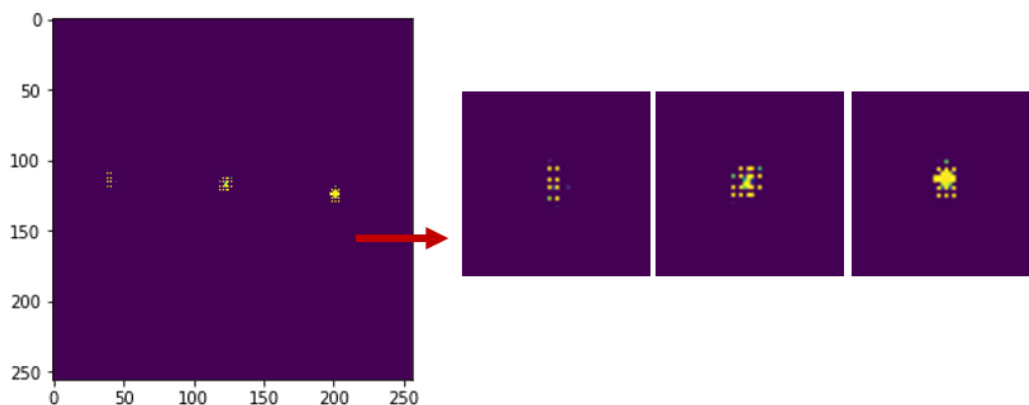


Рисунок 1.10 – Карта вероятностей, созданная моделью

Причина возникновения данного вида артефактов заключается в особенностях работы декодировщика сверточной нейронной сети. Нейронная сеть, применяемая в данной работе, имеет U-net-подобную архитектуру, также известную как архитектура типа encoder-decoder [18]. Схема архитектуры U-Net изображена на рисунке 1.11.

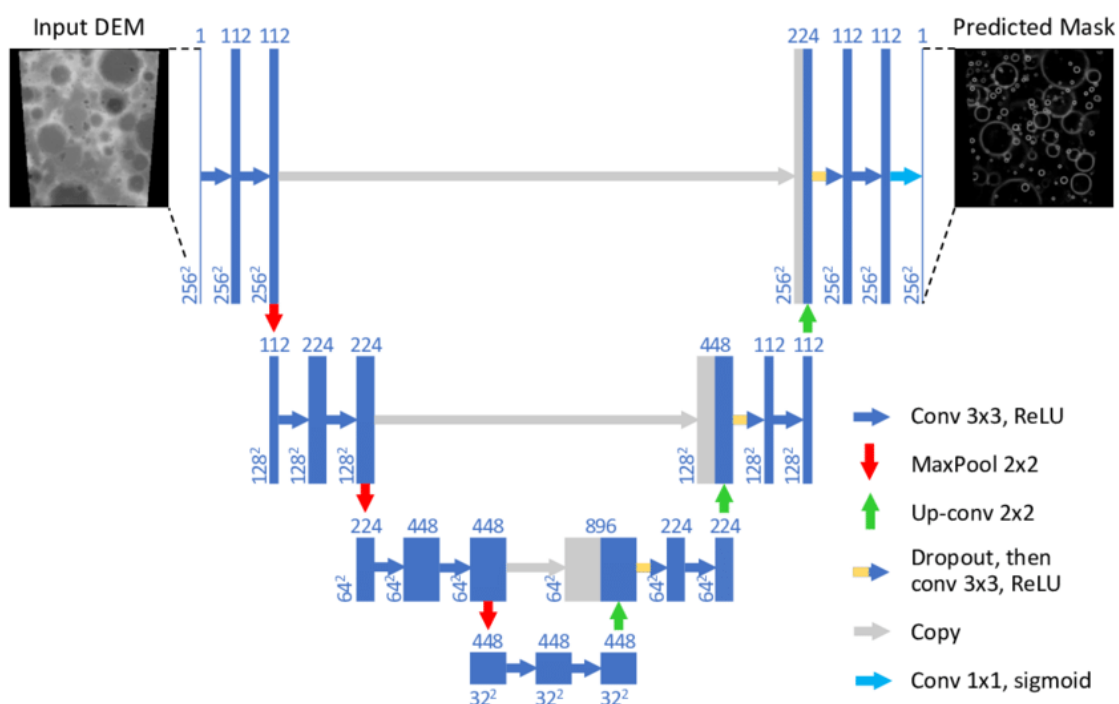


Рисунок 1.11 – Архитектура encoder-decoder.

Архитектура состоит из кодировщика и декодировщика – суживающегося пути для захвата контекста и симметрично расширяющегося пути, который создает выходное изображение на основе карты признаков, полученной кодировщиком.

Суживающийся путь состоит из чередующихся операций свертки и пуллинга. Свертка применяется для извлечения признаков и не меняет размера изображения, но увеличивает количество каналов признаков. Пуллинг уменьшает размер изображения в 4 раза, не изменяя количества каналов.

Расширяющийся путь осуществляет обратные преобразования, применяя к карте признаков операции деконволюции. В результате генерируется конечное изображение, многократно превышающее по разрешению карту признаков, ставшую его основой.

Таким образом, декодировщик формирует изображение высокого разрешения из изображения с низким разрешением, используя несколько шагов деконволюции. Так как именно на этом этапе появляются артефакты, рассмотрим подробнее операцию деконволюции, которая позволяет модели использовать информацию каждого пикселя меньшего изображения для того, чтобы заполнить значениями некоторый квадратный участок большего.

Деконволюция имеет несколько интерпретаций и названий, в числе которых «транспонированная свертка». Деконволюция аналогична операции свертки на изображении, дополненном паддингами (нулевыми пикселями) вокруг пикселей или групп пикселей. На рисунке 1.12 приведены два варианта транспонированной свертки для различных размеров ядра свертки и шага.

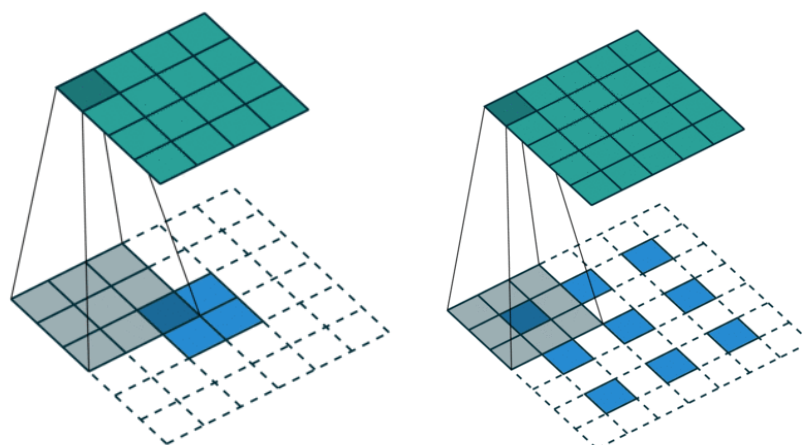


Рисунок 1.13 – Транспонированная свёртка (синий слой – исходный, зеленый – результат деконволюции)

Деконволюция может иметь неравномерное покрытие слоя, создавая наложения результатов свертки от соседних пикселей [19]. В частности, такое может случаться, если размер ядра свертки (то есть размер окна вывода) не делится на размер шага. На рисунке 1.14 проиллюстрирована проблема неравномерного покрытия в одномерном случае.

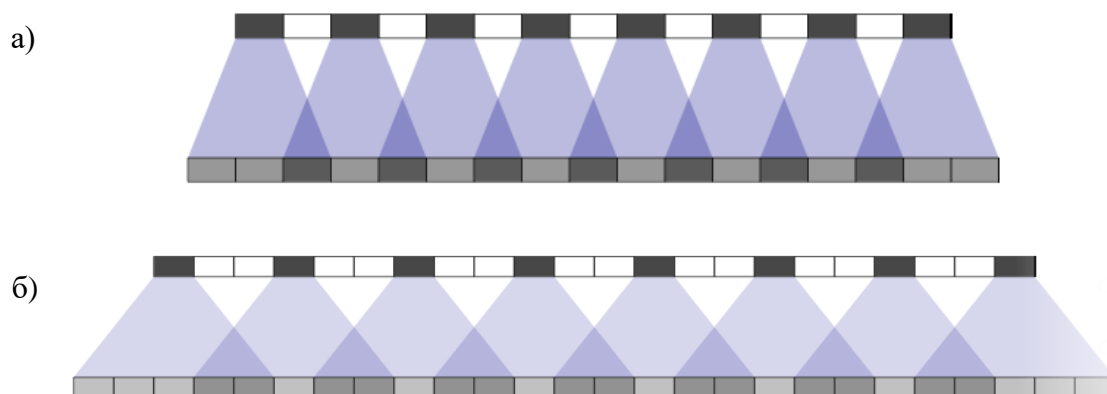


Рисунок 1.14 – Неравномерное покрытие: а) при размере окна 3, шаге 2, б) при размере окна 5, шаге 3.

Аналогично паттерн формируется при рассмотрении двумерной ситуации. Более того, наложения усиливаются, создавая характерную шахматную картину с перекрытиями разной плотности (рисунок 1.15).

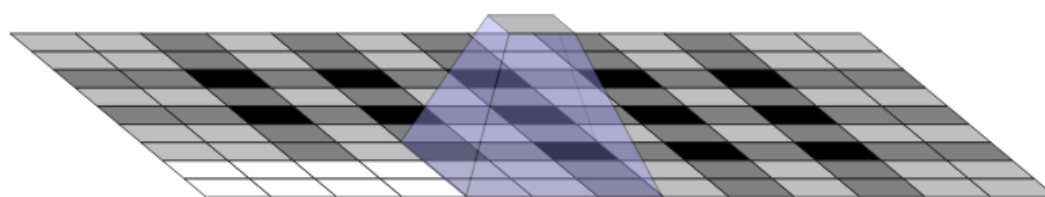


Рисунок 1.15 – Неравномерное покрытие на плоскости.

1.4.2 Методы устранения артефактов в сверточных нейронных сетях

В нейросетях несколько слоев деконволюции идут последовательно, итеративно создавая большее изображение на основе серии изображений с меньшим разрешением. С одной стороны, серия деконволюций смягчает

выраженность артефактов, с другой – этот эффект создается из-за сложения нескольких паттернов и возникновения артефактов с различными масштабом и периодичностью (рисунок 1.16).

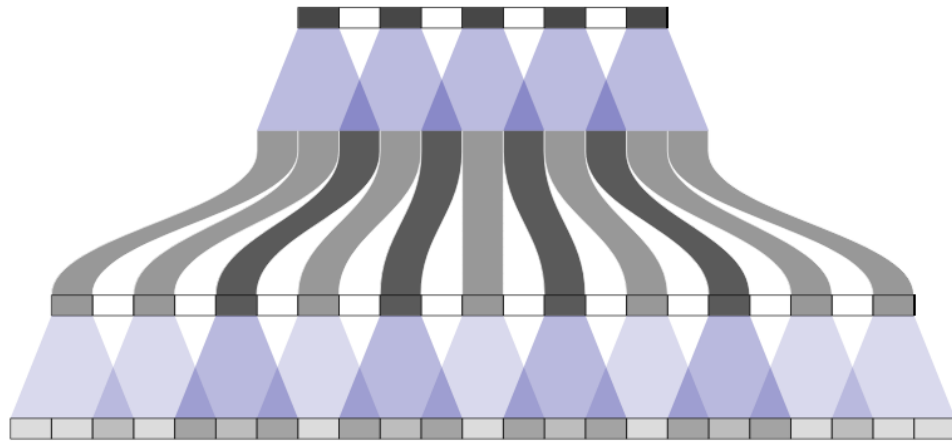


Рисунок 1.16 – Результат двух последовательных слоев деконволюции.

Можно предположить, что артефакты могут быть компенсированы во время обучения весов модели (рисунок 1.17).

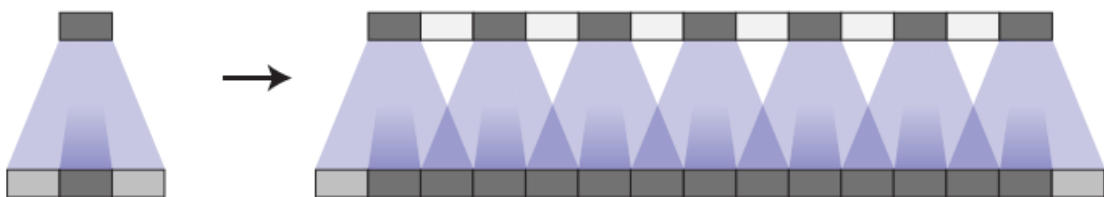


Рисунок 1.17 – Предположение о компенсации неравномерного покрытия с помощью настройки весов.

Однако это труднодостижимо, особенно при обработке нескольких каналов признаков. Избегание моделью паттерна артефактов в значительной степени ограничивает тренировку возможных фильтров, сказываясь на емкости модели. Поэтому на практике модель не обучается избегать создания артефактов. Наоборот, модели действуют таким образом, что даже при равномерном покрытии (рисунок 1.18) обучают фильтры свертки, которые приводят к возникновению артефактов (рисунок 1.19).

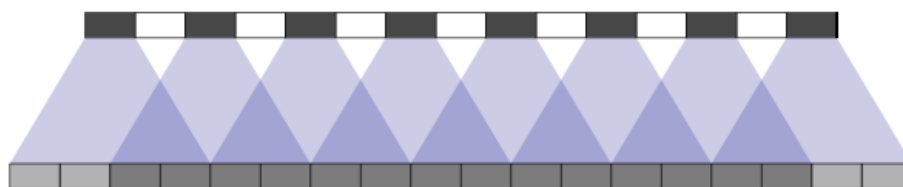


Рисунок 1.18 – Деконволюция с равномерным покрытием при размере окна 2, шаге 2.

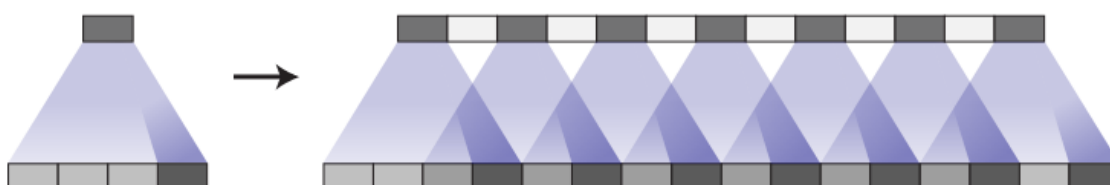


Рисунок 1.19 – Склонность модели обучать фильтры, вызывающие артефакты.

Таким образом, полное исключение артефактов является ограничением для обучения емких фильтров, и поведение с избеганием создания артефактов не является типичным для сверточных нейронных сетей. Поэтому аккуратный подбор параметров размера ядра свертки и размера шага является желательным, однако недостаточным средством устранения артефактов.

Для того чтобы избежать возникновения артефактов, следует выбрать альтернативу стандартной деконволюции. Альтернативный метод транспонированной свертки не должен по умолчанию обладать поведением, склонным к возникновению артефактов.

Таким методом является отделение операции повышающей дискредитации (повышение разрешения), от операции свертки. Деконволюция разделяется на два шага – изменение размера изображения и следующее за ним извлечение признаков в сверточном слое. Такой метод получил название *resize-convolution*. Изменение разрешения может проводиться с применением интерполяции. Самые распространенные виды – интерполяция по ближайшему соседу и билинейная интерполяция (рисунок 1.20).

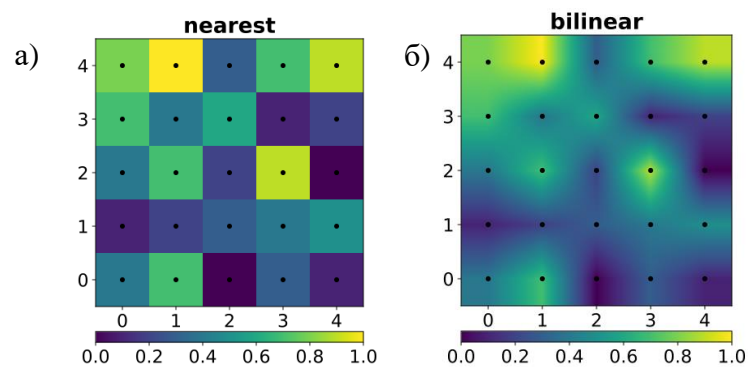


Рисунок 1.20 – Интерполяция по ближайшему соседу (а) и билинейная интерполяция (б).

Как деконволюция, так и различные подходы `resize-convolution` являются линейными операциями и могут быть интерпретированы как матрицы. В деконволюции каждому отдельному пикселю соответствует уникальная запись ядра свертки, тогда как `resize-convolution` неявно связывает веса между собой, что препятствует высокочастотным артефактам. Иллюстрация к сравнению деконволюции и подходов `resize-convolution` приведена на рисунке 1.21.

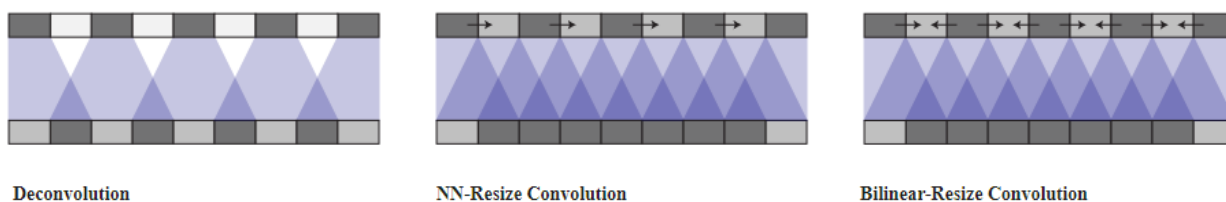


Рисунок 1.21 - Сравнение деконволюции и подхода `resize-convolution` с различными видами интерполяции в качестве метода повышающей дискретизации.

В данной работе при построении нейронной сети использовался подход `resize-convolution` для декодировщика. В нем для увеличения разрешения карты признаков применяется билинейная интерполяция. Сравнение масок, полученных при применении слоев деконволюции и слоев `resize-convolution` приведено в параграфе 3.3.

2 Проектирование метода решения задачи локализации объектов на изображении

2.1 Описание подхода к решению задачи

К рассматриваемой задаче локализации объектов на изображении ставится требование получения результата в виде двумерных относительных координат (в пикселях), соответствующих положению структурного центра каждого объекта. Таким образом, задачу можно сформулировать как *детекцию ключевых точек*.

Общим подходом к задаче детекции объектов на изображении является получение ограничивающих рамок (в английской литературе – *bounding boxes*) вокруг интересующих объектов (рисунок 2.1) [20].

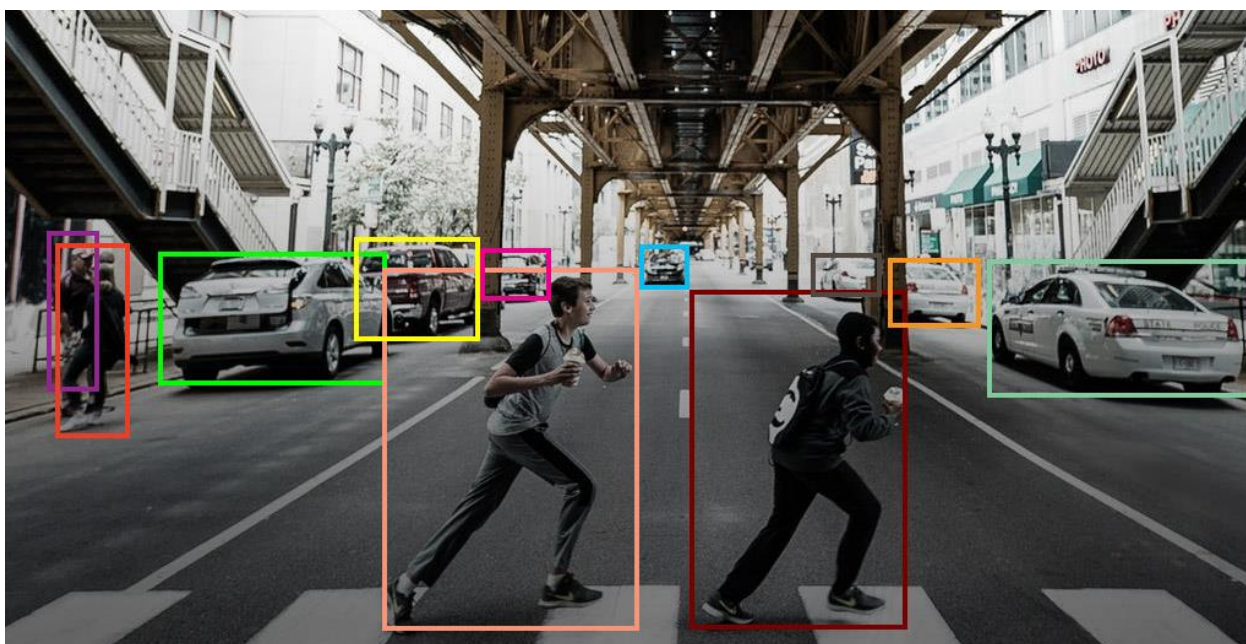


Рисунок 2.1 – Пример ограничивающих рамок в задаче детекции объектов на изображении.

К задаче, рассматриваемой в данной работе подобный подход не применим в силу нескольких причин:

- 1) целью является определение положения не целого объекта, а его структурного центра (не геометрического, который можно было бы определить из положения ограничивающих рамок);

2) разметка ограничивающих рамок для обучающей выборки обычно делается вручную, из-за чего эта процедура утомительна, дорога и занимает большое количество времени.

Данная работа посвящена методу оценки положения объектов на изображениях, не требующих аннотации ограничивающих рамок или использования методов, подобных скользящему окну и предложениям областей (в англоязычной литературе – *sliding window*, *proposal regions*), пример которых показан на рисунке 2.2.

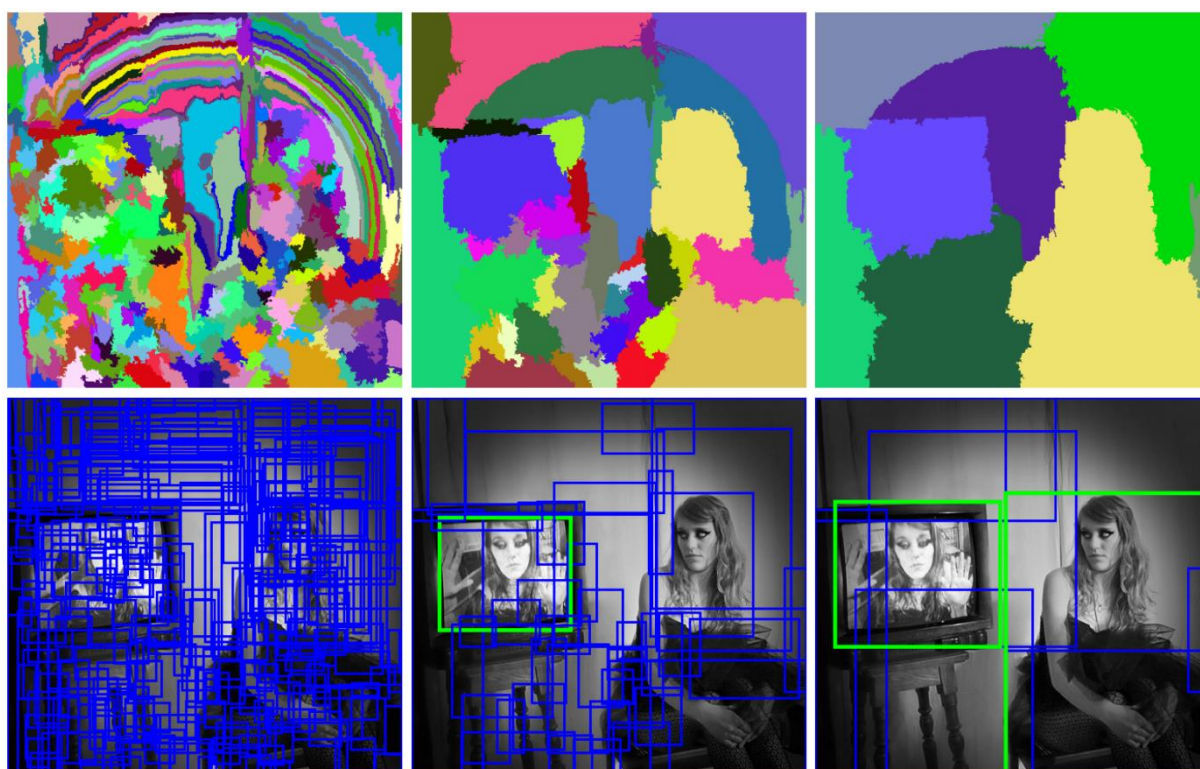


Рисунок 2.2 – Пример решения задачи детекции объектов с помощью метода предложений областей.

Говоря о разработке метода решения, задача детекции ключевых точек фактически является задачей сегментации изображений, так как обученная модель создает на выходе карту вероятностей (в англоязычной литературе – *probability map*), после обработки которой возможно получить результат в виде координат точек. Модель принимает на вход пару изображение – список

2D координат. При этом в данной задаче заранее неизвестно количество объектов на каждом изображении.

2.2 Определение структуры алгоритма

Основную часть алгоритма составляют два модуля, осуществляющие сегментацию изображений с картой вероятностей на выходе и вычисление относительных координат центров выделенных областей. Общая структура алгоритма представлена на рисунке 2.3.



Рисунок 2.3 – Общая структура алгоритма

Более подробная структура алгоритма представлена на рисунке 2.4. Модуль сегментации предполагает использование Unet-подобной нейронной сети – TernausNet, для обучения которой в качестве функции ошибки выбрано взвешенное расстояние Хаусдорфа. Подробнее проектирование модуля сегментации рассмотрено в параграфе 2.3.

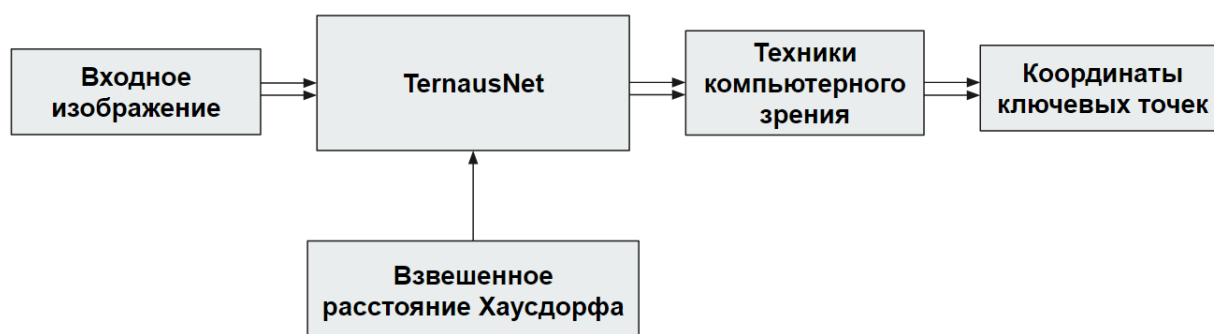


Рисунок 2.4 – Подробная структура алгоритма

На выходе нейронная сеть создает карту вероятностей, которая представляет собой матрицу равного размера с исходным изображением с элементами в диапазоне от 0 до 1. Эти значения, предсказанные моделью

попиксельно, являются вероятностями, что пиксель относится к интересующей области. В визуализации на рисунке 2.5 пиксели, выделенные желтым, имеют вероятность близкую к 1; области, имеющие более темный оттенок (зеленый, синий) обладают меньшей вероятностью, в результате чего при постобработке они «обрежутся»; фиолетовая область относится к фону.

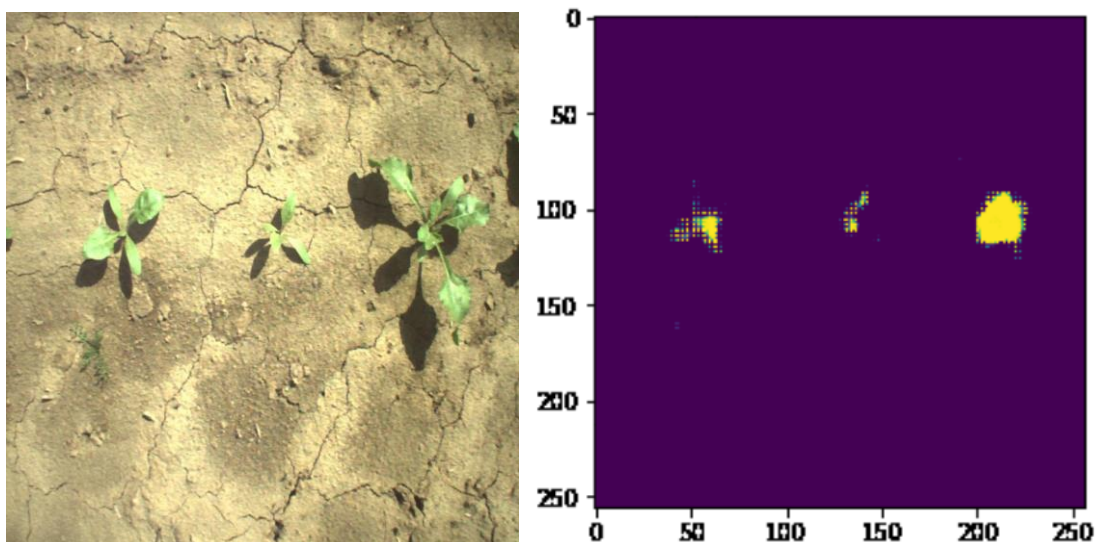


Рисунок 2.5 – Пример карты вероятности для изображения

На карте вероятностей оказываются выделены вероятные области, которые соответствуют положению ключевых точек объектов (в данной работе – структурных центров побегов). Далее полученная карта обрабатывается в модуле вычисления координат объекта. В данном модуле используются техники компьютерного зрения, подробно описанные в параграфе 2.4. На выходе модуль имеет список пар координат, выраженных в пикселях.

Выбор такой структуры алгоритма локализации объектов обосновывается несколькими аргументами:

- 1) Взвешенное расстояние Хаусдорфа в качестве функции ошибки позволяет тренировать модель в контексте задачи детекции ключевых точек, где истинные данные (в англоязычной литературе – *ground truth*) состоят из единственной точки [15].

2) Центры объектов на изображениях из датасета, на котором происходила реализация алгоритма, расположены друг от друга на расстояниях, значительно больших или сравнимых с размерами объектов, таким образом обеспечивается четкое разделение выделенных моделью центров. Вследствие этой причины, для вычисления координат было решено применить техники компьютерного зрения из библиотеки OpenCV.

3) При выборе нейронной сети для рассматриваемой задачи было учтено, что нейронная сеть должна быть способна выделять различное количество объектов на изображении. Выбор полностью сверточной сети (то есть такой, у которой все слои сверточные и нет полносвязного слоя в конце) обусловлен естественной интерпретацией ее выходов как весов (p_x) в функции взвешенного расстояния Хаусдорфа [21]. Последний слой такой сети – сверточный, выдающий на выходе одноканальный вектор p , в котором p_x – значение предсказания модели на пикселе x .

В результате, разработанный метод не ограничен максимальным числом объектов на каждом изображении и не требует разметки ограничивающих рамок.

2.3 Разработка модуля сегментации изображений

2.3.1 Функция ошибки для задачи детекции ключевых точек

В качестве функции ошибки (в англоязычной литературе – *loss function*) для обучения и оценки качества модели выбрано взвешенное расстояние Хаусдорфа (ВРХ). ВРХ может быть использовано в любой полностью сверточной сети [22]. Эта функция ошибки является модификацией метрики Хаусдорфа – среднего расстояния Хаусдорфа между двумя неупорядоченными наборами точек (рисунок 2.6) [23].

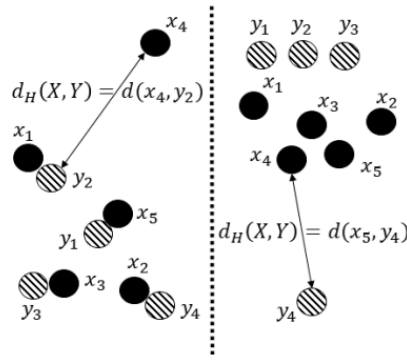


Рисунок 2.6 – Метрика Хаусдорфа для двух неупорядоченных множеств

Выбор взвешенного расстояния Хаусдорфа обусловлен следующими ограничениями архитектуры:

Сверточные нейронные сети с линейными слоями способны предсказывать непосредственно координаты ключевых точек, однако количество предсказываемых точек определяется размером последнего слоя – это является недостатком, так как фактическое количество точек зависит от содержания самого изображения.

Полностью сверточные нейронные сети способны определять присутствие центра объекта большим значением активации, но не способны возвращать координаты пикселей. Для того, чтобы проводить обучение модели с помощью обратного распространения ошибки, функция ошибки должна быть дифференцируемой по отношению к выходу сети.

ВРХ преодолевает эти ограничения и является дифференцируемой относительно выхода сети. Функция принимает карту вероятностей как предсказание модели и координаты ключевых точек как истинные данные. Формула ВРХ следующая:

$$d_{WH}(p, Y) = \frac{1}{S + \epsilon} \sum_{x \in \Omega} p_x \min_{y \in Y} d(x, y) + \frac{1}{|Y|} \sum_{y \in \Omega} M_{\alpha}^{x \in \Omega} [p_x d(x, y) + (1 - p_x) d_{max}],$$

где

$$S = \sum_{x \in \Omega} p_x,$$

$$M_{\alpha}^{a \in A} [f(a)] = \left(\frac{1}{|A|} \sum_{a \in A} f^{\alpha}(a) \right)^{\frac{1}{\alpha}}.$$

M_a – обобщенное среднее; ϵ принимается равным 10^{-6} . $p_x \in [0, 1]$ – значение предсказания на выходе сети для пикселя с координатой x . На последнем слое сети используется нелинейная функция активации (обычно сигмоида) для приведения значений от нуля до единицы.

ВРХ состоит из двух слагаемых:

- В первом слагаемом присутствует множитель p_x , который «штрафует» за высокие значения на карте активации в тех областях, близи которых нет истинных точек y . Другими словами, функция ошибки «штрафует» модель за предсказание точек в тех местах, где их быть не должно.
- Во втором слагаемом присутствует выражение $p_x d(x, y) + (1 - p_x) d_{max}$. Оно «штрафует» модель за низкие значения на карте активации в областях, где присутствуют истинные точки.

Если убрать первое слагаемое, тогда существует тривиальное решение $p_x = 1 \forall x \in \Omega$, то есть модель сможет выдавать низкую ошибку, активируя все пиксели изображения. Если же убрать второе слагаемое, то тривиальным решением будет $p_x = 0 \forall x \in \Omega$, тогда модель обучится не активировать ни одного пикселя.

Почему в задаче детекции ключевых точек необходимо использовать ВРХ, а не часто применяемые метрики наподобие отношения пересечения к объединению (в англоязычной литературе *IoU – Intersection over Union*)? Очевидной альтернативой также было бы использовать метрику попиксельного сопоставления предсказания истинными данными: если $x \in Y$, то $l_x = 1$, иначе $l_x = 0$, и далее вычислить ошибку как средне-квадратичное отклонение. Однако проблема попиксельных функций ошибок в том, что они не учитывают информацию, насколько близко находятся точки множеств $x \in \Omega$ и $y \in Y$, учитывая только случай $x = y$. Другими словами, они являются неинформативными для подавляющего большинства пикселей, что делает обучение невозможным.

2.3.2 Архитектура нейронной сети

Для сегментации изображений реализована TerausNet, архитектура которой основана на архитектуре U-Net с кодировщиком VGG, предобученным на базе изображений ImageNet для задачи сегментации [24]. Архитектура TerausNet изображена на рисунке 2.7. Вместо стандартных слоев транспонированной свертки в декодировщике используется подход *resize-convolution* с применением билинейной интерполяции и свертки для устранения артефактов на генерируемых сетью масках (подробнее в главе 1.4).

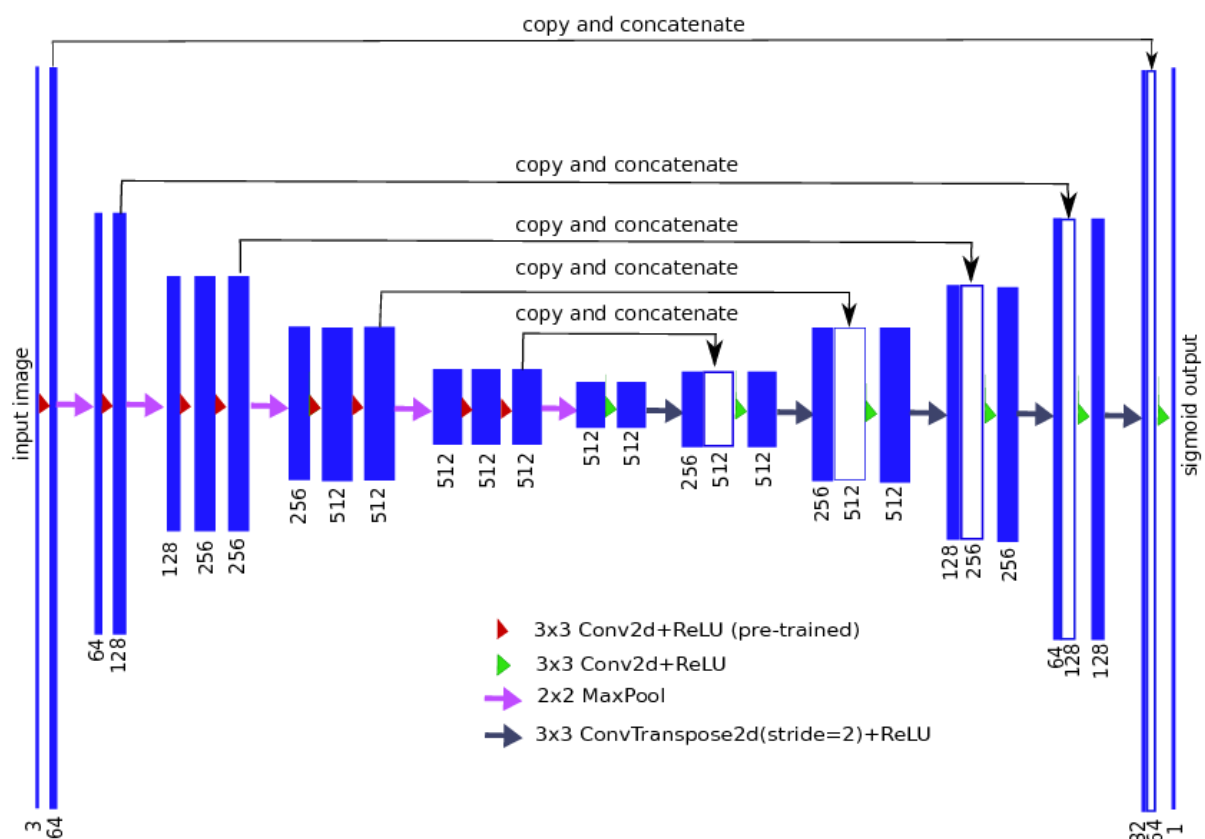


Рисунок 2.7 – Архитектура нейронной сети TerausNet

Архитектура U-Net стандартна для задач сегментации [18]. В общем виде основанные на ней архитектуры состоят из суживающегося пути для захвата контекста и симметрично расширяющегося пути для осуществления точной локализации.

Суживающийся путь отвечает типичной архитектуре сверточных сетей с поочередно идущими операциями свертки и субдискретизации (в

англоязычной литературе – *pooling*), благодаря которым постепенно снижает размерность карты признаков и увеличивает их количество на каждом слое (понижающая дискретизация с удваиванием каналов признаков). Каждый шаг расширяющегося пути состоит из повышения размерности карты признаков с последующей сверткой. Таким образом, расширяющаяся ветвь увеличивает разрешение выхода. Для того, чтобы локализовать признаки, расширяющийся путь объединяет их с признаками высокой размерности из сужающегося пути (в англоязычной литературе такая операция объединения называется *skip connections*). На выходе сеть создает маску, которая показывает класс каждого пикселя. Такая архитектура показывает хорошие результаты в задачах сегментации с ограниченным количеством данных [25].

Для рассматриваемой в работе задаче использована нейронная сеть по типу *encoder-decoder*, также известная как *U-Net*, с нейронной сетью *VGG* без полносвязных слоев в качестве кодировщика. На рисунке 2.7 каждый синий прямоугольный блок представляет собой многоканальные карты признаков, проходящие через серию преобразований. Высота блока отражает относительный размер карт (в пикселях), тогда как его ширина пропорциональна количеству каналов (количество явно указано рядом с соответствующим блоком). Количество каналов возрастает от шага к шагу в левой части и убывает в правой декодирующей части. Стрелки в верхней части схемы соответствуют передаче информации от каждого кодирующего слоя и объединению с соответствующим кодирующим слоем.

VGG-Encoder. Обычно нейронные сети при инициализации с весами из сети, предобученной на больших наборах данных наподобие *ImageNet*, показывают более высокие результаты, чем при инициализации со случайными весами и обучении с нуля на небольших датасетах. В некоторых практических приложениях, например, в медицине и дорожной безопасности, точность моделей критически важна.

В работе [24] продемонстрировано, как нейронная сеть с архитектурой, подобной *U-Net*, может быть улучшена с помощью использования

предобученного кодировщика. В качестве примера авторы работы представляют применение такого подхода в наборе данных Aerial Image Labeling Dataset [26], который содержит аэрокосмические снимки нескольких крупных городов в высоком разрешении. Каждый пиксель изображений размечен в соответствии с принадлежностью к классам «здание» и «не здание». Пример разметки и результаты предсказания модели с предобученными весами и модели, обученной с нуля, продемонстрированы на рисунке 2.8.

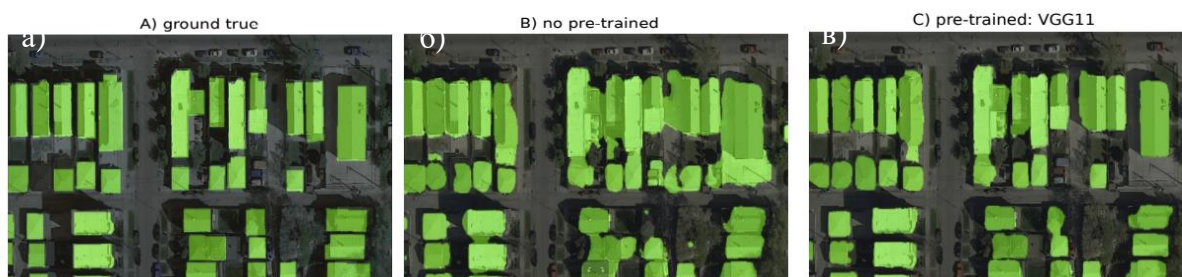


Рисунок 2.8 – Бинарные маски (зеленый цвет идентифицирует принадлежность пикселя к классу «здание»: а) оригинальное изображение с разметкой, б) результат предсказания модели, обученной с нуля, в) результат предсказания модели, инициализированной с весами предобученной сети VGG11.

На рисунке можно увидеть предсказания моделей, инициализированных по разной схеме и обученные за 100 эпох. Нейронная сеть на рисунке 2.8 б) инициализирована со случайными весами, модель на рисунке 2.8 в) имела инициализацию декодировщика со случайными весами и кодировщика с весами VGG11, предобученном на ImageNet. При сравнении заметно, что выделенные области на предсказании предобученной модели имеют формы, близкие к разметке, о более четко разграничены между собой. Метрики эффективности модели, оцениваемые в работе, выше для этой схемы инициализации сети, чем для модели без предобучения.

Предварительная тренировка на большом датасете, включающем изображения, размеченные по многим классам, позволяет модели научиться

выделять общие признаки объектов, что в последующем позволяет более качественно обучиться на выделении признаков, специфичных для решаемой задачи.

Использование весов VGG11 или VGG16 позволяет делать неслучайную инициализацию весов кодировщика. Помимо этого, декодировщик также можно настроить перед обучением на данных, подготовленных для целевой задачи. В таком случае его в начале инициализируют со случайными весами, затем обучают нейронную сеть на дополнительном датасете изображений из той же предметной области, но не полностью совпадающих по содержанию с данными из основной задачи. Подготовка завершается обновлением весов последнего слоя сети на случайные. Этот метод может использоваться, если необходимо добиться максимальной точности и при небольшом количестве данных.

Для решения рассматриваемой в работе задачи было решено сравнить эффективность модели при обучении только на основных данных, представляющих снимки ростков сахарной свеклы, и при предварительном обучении на дополнительном датасете, содержащем увеличенные спутниковые снимки фермерских угодий с другими произрастающими культурами.

2.3.3 Выбор метода оптимизации

Алгоритм оптимизации необходим для обучения модели посредством итеративного обновления весов сети на основе обучающих данных.

В данной работе при выборе техники оптимизации было протестировано четыре алгоритма:

- SGD – стохастический градиентный спуск (англ. Stochastic Gradient Descent). Алгоритм поддерживает постоянную скорость обучения на протяжении всех обновлений весов [27].

- Adam – расширение стохастического градиентного спуска, представляющее собой адаптивную оценку момента. Алгоритм адаптирует скорости обучения параметров на основе средних значений первого и второго момента градиентов [28].

- AdamMax – разновидность алгоритма Adam, основанный на норме бесконечности [28].

- Adadelata – алгоритм, название которого является сокращением «адаптивной дельты», где речь идет о дельте – разнице между текущим весом и недавно обновленным весом. Метод динамично адаптируется со временем, используя только информацию первого порядка и требует незначительно больше вычислительных затрат, чем стохастический градиентный спуск [29].

Модификации алгоритмов оптимизации призваны с одной стороны, способствовать более быстрому обучению модели, а с другой – увеличивать вероятность схождения функции ошибки к глобальному минимуму.

На рисунке 2.9 изображено сравнение процесса схождения функции ошибки к минимуму при использовании различных техник оптимизации. В сравнении с другими алгоритмами Adadelata обычно приводит к более быстрой сходимости процесса обучения.

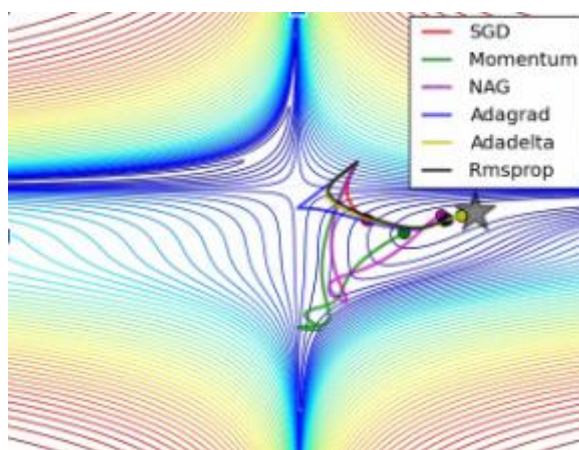


Рисунок 2.9 – Сравнительный обзор техник оптимизации

В данной работе Adadelata также показала наиболее подходящий результат. Процесс обучения сошелся при различных начальных коэффициентах скорости обучения при Adadelata, тогда как использование

других оптимизаторов потребовало более аккуратного выбора значений коэффициента скорости обучения и момента, в ином случае приводя к расхождению обучения.

2.4 Разработка модуля вычисления относительных координат объекта на изображении

В модуле сегментации изображений модель на выходе создает карту вероятностей, которая обрабатывается нелинейной функцией – сигмоидой в силу бинарности задачи. В результате получается маска, на которой области со значением 1 интерпретируются как регионы, окружающие ключевые точки. Пиксели со значениями 0 относятся к фону. Пример маски представлен на рисунке 2.10.

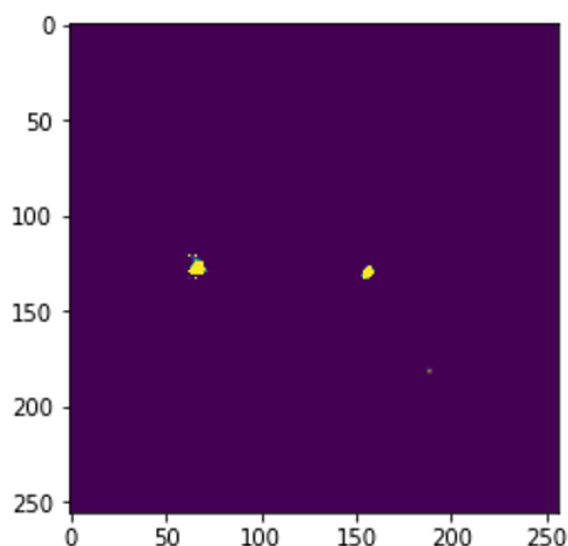


Рисунок 2.10 – Бинарная маска на выходе модуля сегментации

Обученная модель, как правило, активирует не отдельные пиксели, а регионы, соответствующие окружению ключевых точек. Такая особенность является приемлемой и объяснимой, так как интересующие признаки на физическом объекте имеют некоторую площадь, и любой пиксель, выделенный моделью в пределах этого окружения, можно считать верным предсказанием. Тогда положения ключевых точек можно рассчитать как координаты геометрического центра выделенных регионов.

Однако маска не содержит информации о координатах пикселей. Целью задачи является определение координат ключевых точек, поэтому бинарные маски требуют пост-обработки для нахождения замкнутых областей и вычисления координат центра каждого региона.

На рисунке 2.11 проиллюстрированы последовательные шаги обработки, которой подвергается бинарная маска, и результат вычисления координат точек, визуализированный на оригинальном изображении.

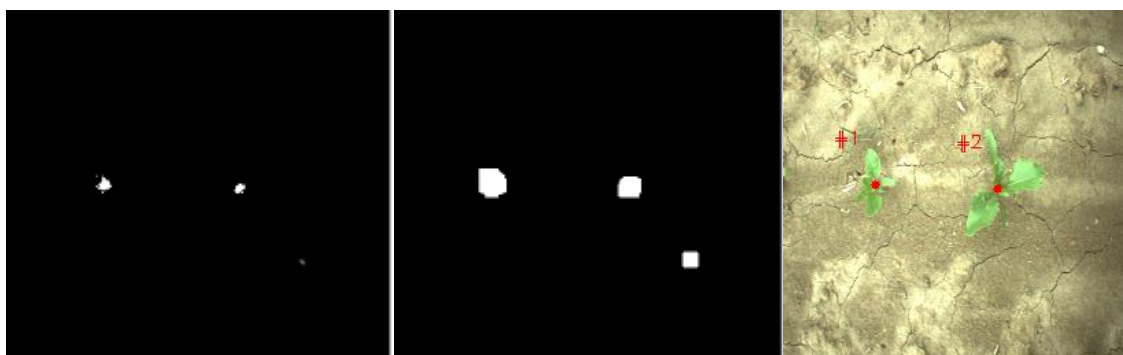


Рисунок 2.11 – Шаги обработки маски и визуализация результата вычисления координат ключевых точек на оригинальном изображении

Обработка маски состоит из следующих последовательных процедур:

1) Выравнивание изображения посредством размытия (blurring) с целью уменьшения высокочастотного шума.

2) Бинаризация по порогу (binary thresholding) для обнаружения самых ярких областей.

3) Растяжение областей (dilations) для избавления от крупного шума, в частности, для объединения небольших групп пикселей с крупными областями, вблизи которых они находятся.

4) Проведение анализа связных компонент для выделения замкнутых групп пикселей отдельных областей и фильтрации оставшегося шума (одиночных пикселей, находящихся далеко от регионов).

5) Детекция контуров на маске и их сортировка слева-направо. После того, как контуры отсортированы, можно итерироваться по ним по отдельности.

б) Определение ограничивающей рамки для каждого контура и вычисление его центра. Координаты центров найденных ограничивающих рамок являются искомыми.

7) Корректировка координат с учетом проведенного в начале изменения размера оригинального изображения (в нейронную сеть обычно подаются изображения определенного размера).

Обработка маски реализована на основе библиотеки компьютерного зрения *OpenCV* и библиотеки для обработки изображений *scikit-image*.

2.5 Выбор метрик оценки обученной модели

Для оценки эффективности модели использованы стандартные определения популярных метрик: Precision, Recall и F-score.

Метрика Precision интерпретируется как точность модели и определяет долю правильно предсказанных объектов положительного (искомого) класса среди всех объектов, отнесенных к положительному классу. Метрика Recall оценивает полноту модели и определяет долю правильно предсказанных объектов положительного класса среди всех правильно предсказанных объектов. F-score является средним гармоническим этих двух метрик.

В задаче детекции ключевых точек для вычисления метрик необходимо определить критерий, по которому предсказанные моделью точки можно отнести к одному из понятий, которыми фигурируют при расчете метрик: истинно-положительные (TP), ложно-положительные (FP), истинно-отрицательные (TN) и ложно-отрицательные объекты (FN). Для этого вводится параметр R, определяющий радиус окружности вокруг истинной ключевой точки (GT), при попадании внутрь которой предсказание считается относящимся к данной истинной точке [15].

Вычисление метрик производится по следующему алгоритму:

1) Для каждой *истинной точки* находится *наиближайшая точка предсказания* модели по методу k-ближайших-соседей (алгоритм kNN) с $k = 1$.

2) Если расстояние между *истинной точкой* и ее *ближайшим предсказанием* **меньше** R , то это предсказание считается истинно-положительным (TP). На рисунке 2.12 истинно-положительные предсказания выделены красным.

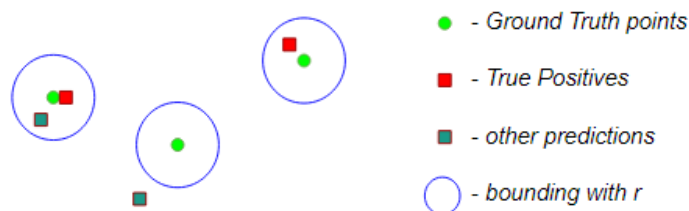


Рисунок 2.12 – Определение истинно-положительных предсказаний (красные квадраты). Синяя окружность – область, ограниченная параметром R .

Зеленые круги – истинные данные, серые квадраты – другие предсказания.

3) Количество ложно-отрицательных объектов (FN) определяется как разница между общим количеством *истинных точек* и истинно-положительными предсказаниями (TP): $FN = GT - TP$

4) Далее для каждого *предсказания* находится *наиближайшая истинная точка* по методу *k*-ближайших-соседей с $k = 1$.

5) Если расстояние между *предсказанием* и его *ближайшей истинной точкой* **больше** R , то предсказание считается ложно-положительным (FP). На рисунке 2.13 ложно-положительные предсказания выделены красным.

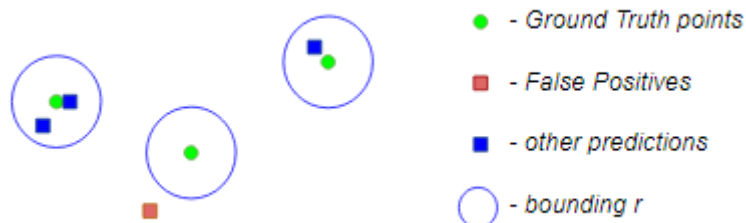


Рисунок 2.12 – Определение ложно-положительных предсказаний (красные квадраты). Синяя окружность – область, ограниченная параметром R .

Зеленые круги – истинные данные, синие квадраты – другие предсказания.

б) В конце метрики вычисляются с найденными значениями количества TP, FP, FN согласно стандартным формулам:

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}$$

Параметр R определяется индивидуально для условий каждой задачи. В описываемой работе он принят равным $R = 9$ пикселей, так как окружность такого радиуса охватывает интересующую часть растения (область проекции корня) и размер изображений во всем датасете одинаков.

Выводы

Задача локализации объектов на изображении определена в данной работе как задача детекции ключевых точек с нахождением координат. Для ее решения выбрана структура алгоритма, состоящая из двух модулей, предназначенных для сегментации изображений и вычисления координат ключевых точек.

Модуль сегментации изображений образован сверточной нейронной сетью TernausNet с предобученным кодировщиком VGG11. Для обучения модели в качестве функции ошибки выбрано взвешенное расстояние Хаусдорфа, позволяющее делать оценку качества модели в контексте задачи, где истинные данные представлены единичными пикселями. Для повышения эффективности нейронную сеть планируется предобучить на дополнительных данных из открытых источников с последующим обновлением последнего слоя. На выходе модуля создается бинарная маска.

Модуль вычисления координат ключевых точек предназначен для обработки полученной бинарной маски с целью вычисления относительных координат (в пикселях) выделенных моделью областей. Модуль реализован на основе техник обработки изображений и компьютерного зрения.

Оценка эффективности обученной модели планируется с помощью стандартных метрик Precision, Recall и F-score.

3 Реализация алгоритма локализации объектов на изображениях сахарной свеклы на разных стадиях вегетативного роста

3.1 Подготовка данных

3.1.1 Описание данных и аннотация изображений

Набор данных состоит из изображений побегов сахарной свеклы на разных стадиях вегетационного роста. В датасете присутствуют также снимки побегов, окруженные сорняками с различной плотностью. Снимки сделаны при различном освещении и разных погодных условиях. Примеры изображений представлены на рисунке 3.1.



Рисунок 3.1 – Примеры изображений из датасета

Снимки в датасете сделаны с помощью камеры автоматического культиватора. Общее количество изображений, отобранных для работы, составило 3283. Доли валидационной и тестовой выборок составили 18% и 10%, соответственно.

Для задачи детекции ключевых точек изображения датасета были размечены единичными пикселями. Ключевыми точками были отмечены положения розеток побегов, то есть область у основания листьев (Рисунок 3.2).



Рисунок 3.2 – Изображения с размеченными ключевыми точками

Данные аннотации изображений были сохранены в формате csv, в том числе координаты размеченных ключевых точек, число ключевых точек на каждом изображении и оригинальный размер изображения.

3.1.2 Аугментация данных в обучающей выборке

Для того, чтобы расширить датасет и увеличить количество изображений, используют аугментацию данных. Методика представляет собой изменение имеющихся в выборке изображений таким образом, чтобы симитировать присутствие в выборке новых данных с объектами, различающимися с имеющимися различными параметрами. Аугментация может включать операции с цветом (изменение яркости, контраста или оттенка), геометрические операции и операции с объектами (растяжения, сжатия, повороты, отражения, изменение масштаба).

При выборе операций аугментации, как правило, учитывают контекст и содержание изображений, так как неуместные операции могут исказить репрезентативность выборки. Например, в задаче распознавания чисел не следует делать отражение по горизонтальной оси, так как некоторые цифры, перевернутые вверх ногами, имеют схожий вид с написанием других символов. Кроме того, при аугментации данных не выполняют трансляции изображений, так как сверточные нейронные сети инвариантны к параллельному переносу объектов в силу своего принципа действия.

К изображениям в данной задаче были применены несколько видов операций: горизонтальное отражение, вертикальное отражение, изменение яркости. Для первых двух операций также требуется аналогичное изменение данных разметки. На рисунке 3.3 представлены оригинальное изображение и изображения, полученные посредством применения на нем операций аугментации.

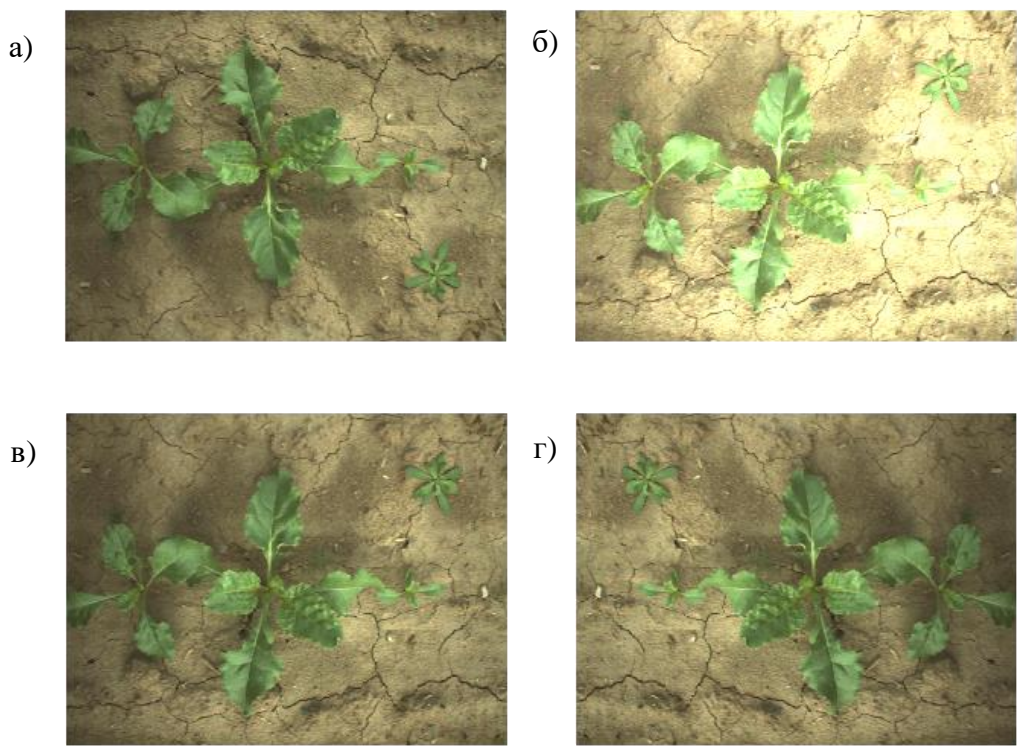


Рисунок 3.3 – Оригинальное изображение (а) и изображения, полученные путем применения к нему операций аугментации: б) изменение яркости, в) отражение вдоль относительно горизонтальной оси, г) отображение относительно вертикальной оси

К данным было решено не применять цветовые аугментации путем подкрутки цветовых каналов, так как они не являются естественными для подобного контекста и могут помешать обучению модели.

В результате аугментации изображений количество релевантных данных значительно увеличилось.

3.1.3 Описание дополнительной выборки изображений для предтренировки модели

В процессе обучения нейронной сети был использован дополнительный датасет “The Sorghum Plant Centers 2016 Dataset”, доступный из открытых источников [30]. Дополнительный набор изображений использован в качестве предобучающих данных с целью сглаживания

проблем, которые мог вызвать сравнительно небольшой размер основной обучающей выборки.

Стадия предобучения проводилась непосредственно перед стадией основного обучения и оканчивалась сбросом весов последнего сверточного слоя сети. Таким образом веса нейронной сети были обучены на узнавание объектов, схожих с теми, что содержатся в основных данных. Так как в рассматриваемой задаче стоит цель не только находить положение побегов сахарной свеклы, но и неявно отделять от них сорняки, то количество основных данных могло оказаться недостаточным для ее решения. В таком случае нейронная сеть, предварительно обученная на узнавание побегов растений на изображении, теоретически должна достигать лучшего качества сегментации (данное предположение подтверждается результатами оценки, представленными в параграфе 3.3.1).

“The Sorghum Plant Centers 2016 Dataset” состоит из 60 000 снимков растения сорго, сделанных беспилотным летательным аппаратом. Всего в датасете размечено 15 208 побегов растений. Разметка, аналогично рассматриваемой задаче, произведена в виде ключевых точек (рисунок 3.4).



Рисунок 3.4 – Изображения и пример разметки из набора данных “The Sorghum Plant Centers 2016 Dataset”

В данной работе для предобучения сети использовано 5 000 изображений из этого датасета.

3.2 Описание стека технологий

Supervisely – веб-платформа для машинного обучения, которая включает в себя инструменты для интеллектуальной обработки данных. Используется в том числе для аннотации изображений и включает проекты датасетов с открытым исходным кодом.

PyTorch – программное обеспечение, в частности, библиотека машинного обучения с открытым исходным кодом для языка программирования Python, основанная на библиотеке Torch. Используется в таких областях как глубокое обучение и обработка естественных языков. PyTorch предоставляет две высокоуровневые функции: тензорные вычисления (аналогично NumPy) с ускорением на графических процессорах (GPU), и построение глубоких нейронных сетей на базе системы autodiff.

Google Colab – бесплатная облачная среда на основе Jupyter Notebook, разработанный Google, который предоставляет свободный доступ к одному GPU (NVidia Tesla K80). Кроме того, он позволяет разрабатывать приложения для глубокого обучения с использованием популярных библиотек, таких как PyTorch, TensorFlow, Keras и OpenCV.

OpenCV (Open source computer vision) – библиотека компьютерного зрения с исходным кодом. Предназначена для алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения. OpenCV поддерживает фреймворки глубокого обучения: TensorFlow, Torch/PyTorch и Caffe.

Scikit-image – библиотека обработки изображений с открытым исходным кодом для языка программирования Python. Включает алгоритмы для сегментации, геометрических трансформаций, манипуляций с цветовыми пространствами, анализа, фильтрации, детекции признаков и многого другого. Предназначена для взаимодействия с Python библиотеками NumPy и SciPy, содержащих инструменты для числовых и научных вычислений.

3.3 Обсуждение результатов

3.3.1 Сравнение моделей

При выборе наиболее эффективного решения задачи локализации сравнивались несколько моделей. Оценка производилась для двух видов архитектуры TernausNet, основанных на разных модификациях U-Net: Unet11 и Unet16, отличающихся количеством слоев. Кроме того, производилась оценка эффективности моделей, предобученных на дополнительном датасете. Результаты вычисления метрик и значения средней ошибки, вычисленные в процессе обучения на тестовой выборке, представлены в таблице 3.1.

Таблица 3.1 – Значения метрик эффективности моделей и средней ошибки

Метрика	Unet11	Unet16	Unet16 + Sorghum Dataset
Precision	83,84	90,85	94,98
Recall	92,45	94,65	93,19
F-score	87,94	92,71	94,07
Average loss	38,18	26,65	22,6

Результаты демонстрируют более высокую эффективность модели Unet16, предобученной на дополнительном наборе данных “The Sorghum Plant Centers 2016 Dataset”. Результаты обучения и оценка эффективности на подвыборках для этой модели описаны в параграфе 3.3.2.

3.3.2 Результаты тренировки выбранной модели

На рисунке 3.5 представлен график обучения модели. Значения ошибки на обучающей и валидационной выборках уменьшаются на протяжении обучающих эпох. К концу обучения заметно, что валидационная ошибка

меньше, чем тренировочная ошибка, что является свидетельством хорошей настройки весов нейронной сети.

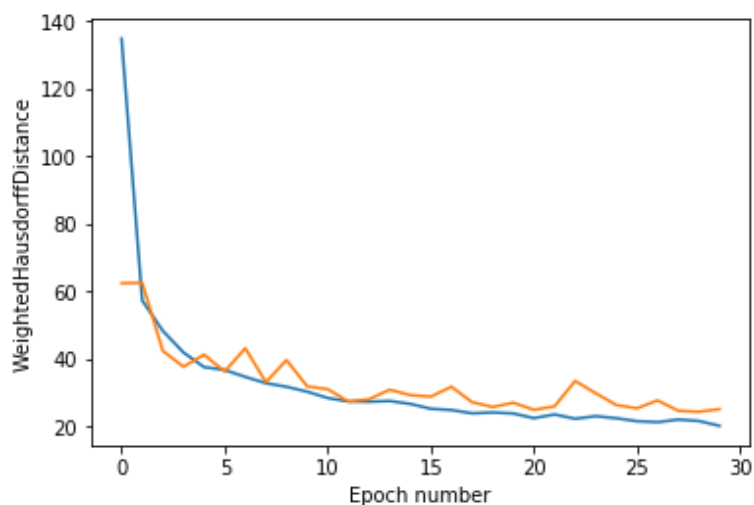


Рисунок 3.5 – График обучения модели: синяя кривая – ошибка на тренировочной выборке, желтая кривая – ошибка на валидационной выборке.

На рисунке 3.6 изображены результаты сегментации изображений моделью в разных эпохах обучения. Интересным наблюдением относительно представленных карт вероятностей является то, что модель сначала тренируется выделять растение целиком, назначая центральным пикселям большее значение активации, после чего становится способной находить структурный центр растения, отмеченный ключевыми точками.

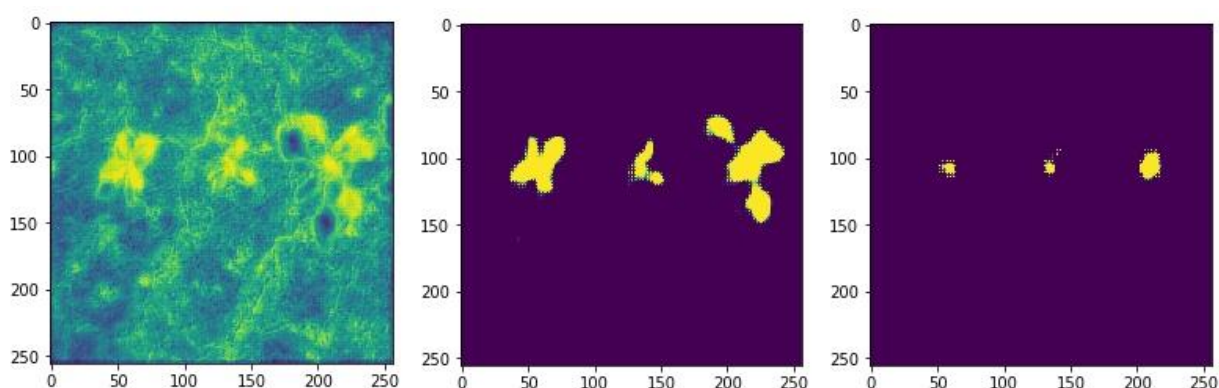


Рисунок 3.6 – Карта вероятностей, создаваемая сегментационной моделью после разных эпох обучения (первая, четвертая, двадцать восьмая)

На рисунке выше заметны артефакты в форме шахматной доски вблизи выделенных областей, возникающие при генерации изображения сверточной нейронной сетью. Для их устранения был применен подход *resize-convolution* с последовательными операциями билинейной интерполяции и свертки. Сравнение результатов при использовании данного подхода и классических слоев деконволюции приводится на рисунке 3.7.

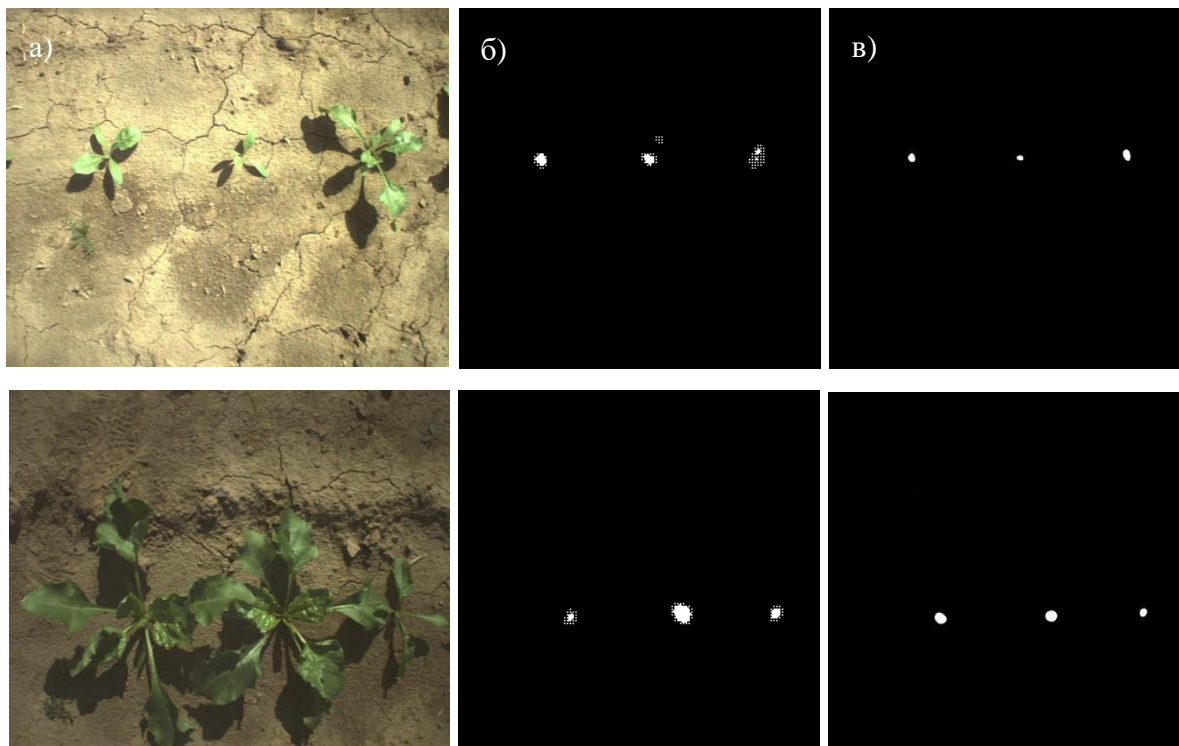


Рисунок 3.7 – Оригинальные изображения (а) и маски, полученные при использовании в декодировщике классических слоев деконволюции (б) и применении подхода *resize-convolution* (в).

Далее представлены результаты работы алгоритма (рисунок 3.8). В примерах приведены изображения, на которых представлены растения на разных стадиях вегетационного роста.

Алгоритм возвращает на выходе список координат, соответствующих положениям ключевых точек – структурных центров побегов сахарной свеклы.

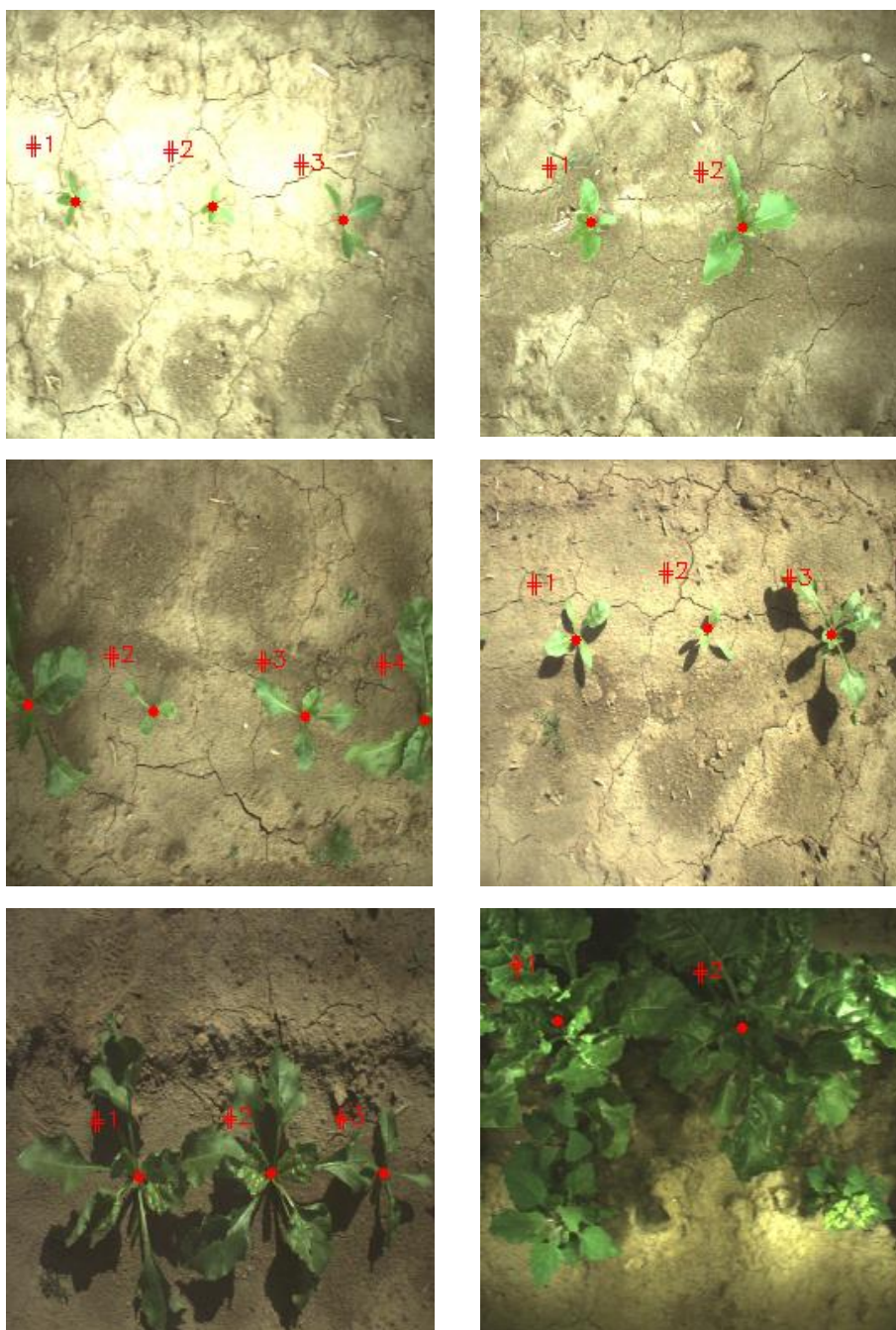


Рисунок 3.8 – Результаты работы алгоритма. Красные точки – визуализация вычисленных координат ключевых точек.

Эффективность алгоритма также проверена на трех подвыборках, содержащих изображения растений на начальной, средней либо поздней стадии роста. Результаты оценки приведены в таблице 3.2.

Таблица 3.2 – Значения метрик эффективности модели на подвыборках изображений

Название подвыборки	Precision	Recall	F-score
Ранняя стадия роста	96,3	90,9	93,5
Средняя стадия роста	97,8	95,7	96,8
Поздняя стадия роста	91,3	92,6	91,9

Как видно из результатов, значения метрик отличаются незначительно между подвыборками, соответствующими разным стадиям роста. F-score, определенный на полной тестовой выборке составил 94,1%, тогда как для подвыборок изображений значение метрики равно 93,5%, 96,8% и 91,9%, соответственно, для снимков на ранней, средней и поздней стадии роста. Наименьший результат наблюдается для взрослых побегов, что может быть объяснено особенностями вида побегов, например, пересечение листьев отдельных побегов, большее скрывание розетки листьями, если побег наклонен ветром, большее количество сорняков на некоторых снимках, что делает задачу более сложной в том числе для эксперта. В общем случае, результаты являются приемлемыми для всех трех подвыборок.

Из результатов следует, что на ранней и средней стадиях роста модель успешно справляется с выделением культурных растений, не охватывая при этом сорняки, о чем свидетельствует более высокое значение метрики Precision в сравнении с Recall. На поздней стадии роста модель характеризуется большей полнотой и лучше справляется с тем, чтобы выделять наибольшее количество центров побегов, что отражено в более высоком значении метрики Recall.

Выводы

Спроектированный алгоритм локализации объектов на изображениях реализован на данных, состоящих из снимков сахарной свеклы на разных стадиях вегетационного роста. Датасет был разделен на обучающую, валидационную и тестовую выборку, кроме того тестовая выборка, в свою очередь, разбита на три подвыборки в соответствии со стадиями роста.

Обучение модели нейронной сети было проведено с предварительной тренировкой на дополнительном наборе изображений проекта “The Sorghum Plant Centers 2016 Dataset”, доступном в открытых источниках. По окончании предобучения проведено обновление весов последнего слоя сети, после чего произведен основной цикл обучения модели, состоявший из тридцати эпох. В результате выбрана модель, полученная на двадцать восьмой эпохе обучения.

Эффективность модели, предобученной на дополнительном датасете, была сравнена с эффективностью модели, получившей обучение только на основном наборе изображений. Кроме того, при выборе модели проведено сравнение двух вариантов архитектуры TernausNet, которые отличаются друг от друга количеством сверточных слоев. В результате оценки эффективности среди моделей выбрана получившая наиболее высокие значения метрик, а именно модель нейронной сети с архитектурой U-Net16, предобученная на дополнительных данных.

Для выбранной модели получено значение метрики F-score равное 94,1%. Также оценена эффективность модели на подвыборках, соответствующих различным стадиям роста побегов. Значение метрики для ранней стадии роста составило 93,5%, для средней стадии – 96,8%, для поздней стадии – 91,9%.

Обученная модель показывает приемлемые результаты в задаче детекции ключевых точек, определяя координаты структурных центров растений. Помимо этого, результаты свидетельствуют об успешном отделении культурных растений от сорняков.

4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Целью дипломной работы является разработка алгоритма локализации сельскохозяйственных растений на изображении в течении вегетативного периода. В рамках настоящей работы предполагается создание алгоритмического обеспечения, позволяющего обнаруживать сельскохозяйственные культурные растения на снимках и исключать сорняки с помощью методов машинного обучения. Особенностью метода является возможность его использования на разных стадиях роста растений.

В разделе представлены предпроектный анализ и результаты следующих этапов управления научным проектом:

- 1) Инициация проекта.
- 2) Планирование проекта.
- 3) Исполнение проекта.

Целью данного раздела является определение перспективности разработки алгоритмы. Для достижения цели применены аналитические инструменты и выполнены следующие задачи:

- Анализ потенциальных потребителей.
- Анализ конкурентных технических решений.
- SWOT-анализ.
- Определение целей и ожидаемого результата проекта.
- Планирование организационной структуры проекта.
- Формулирование ограничений и допущений проекта.
- Планирование структуры работ проекта.
- Построение плана проекта.
- Формирование бюджета.
- Анализ рисков проекта.
- Расчет показателя финансовой эффективности.

4.1 Предпроектный анализ

4.1.1 Потенциальные потребители

Для определения потенциальных потребителей требуется выявить целевой рынок. Для того, чтобы определить организации, которым необходима данная разработка, было проведено сегментирование целевого рынка.

Рынок, на котором разработка потенциально может быть интересна – рынок производителей программного обеспечения для сельскохозяйственной техники.

Сегментация проводилась по следующим критериям: тип программного обеспечения (ПО), степень автономности программного продукта. Карта сегментирования представлена в таблице 4.1.

Таблица 4.1 – Сегментация рынка

	Алгоритм без обученной модели	Алгоритмический модуль с обученной моделью	Программный блок, готовый к интеграции
Производители ПО для массового оборудования			+
Разработчики универсального ПО для высокотехнологичного оборудования		+	+
Разработчики специализированного ПО для высокотехнологичного оборудования	+	+	

Разрабатываемый продукт предназначен для использования в автоматизированной или роботизированной сельскохозяйственной технике, предполагающей применение компьютерного зрения. Продукт предлагается использовать как часть аналитического модуля системы управления роботом или автоматического блока.

Потенциальными потребителями продукта могут быть компании, разрабатывающие программное обеспечение для высокотехнологичного

сельскохозяйственного оборудования и роботизированной техники обработки сельхозугодий общего назначения или специализированной под конкретный вид возделываемой культуры.

4.1.2 Анализ конкурентных технических решений

Анализ конкурентоспособности технического решения был проведен с помощью оценочной карты. В качестве конкурентных систем рассматривались разработки Naïo Technologies, Франция (K1) и FarmWise, США (K2). Результаты анализа представлены в таблице 4.2.

Позиция разработки и конкурентов оценивается по каждому показателю экспертным путем по пятибалльной шкале, где 1 – наиболее слабая позиция, а 5 – наиболее сильная. Веса показателей, определяемые экспертным путем, в сумме должны составлять 1.

Таблица 4.2 – Результаты анализа конкурентных систем анализа текста

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б _ф	Б _{к1}	Б _{к2}	К _ф	К _{к1}	К _{к2}
Технические критерии оценки ресурсоэффективности							
1. Требуемая квалификация для возможности конфигурирования решения	0,08	5	4	3	0,4	0,32	0,24
2. Возможность настройки для интеграции в оборудование различных производителей	0,15	5	3	3	0,75	0,45	0,45
2. Возможность настройки для различных видов возделываемых культур	0,13	5	5	5	0,65	0,65	0,65
Потенциальные возможности функционального расширения системы	0,08	4	3	3	0,32	0,24	0,24
4. Отказоустойчивость решения	0,12	3	5	5	0,36	0,6	0,6
5. Потребность в ресурсах памяти	0,05	3	5	4	0,15	0,25	0,2
6. Скорость обработки данных	0,15	4	5	5	0,6	0,75	0,75

Таблица 4.2 (Продолжение) – Результаты анализа конкурентных систем анализа текста

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б _ф	Б _{к1}	Б _{к2}	К _ф	К _{к1}	К _{к2}
Экономические критерии оценки ресурсоэффективности							
Доступность на отечественном рынке	0,12	5	3	1	0,6	0,36	0,12
8. Цена	0,12	4	3	2	0,48	0,36	0,24
Итого	1				4,31	3,98	3,39

Из таблицы сравнения собственной разработки с решениями Naïo Technologies, Франция (К1) и FarmWise, США (К2) можно сделать вывод, что уязвимость конкурентных решений связана с тем, что они разработаны зарубежными компаниями, в связи с чем выше стоимость их разработки и ниже доступность на российском рынке, а также с тем, что они разработаны под программное обеспечение определенной серии роботизированного оборудования

Таким образом, разработанное решение имеет преимущества на отечественном рынке производителей программного обеспечения за счет своей доступности, кроме того, решение обладает весомым достоинством в виде своей мобильности, возможности интеграции в программное обеспечение оборудования разных производителей и производимых для разнообразных целей и различных видов растений, что может быть интересно в том числе на зарубежном рынке экспериментальной разработки.

4.1.3 SWOT-анализ

SWOT – Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности) и Threats (угрозы) – это комплексный анализ научно-технического проекта. Такой анализ применяют для исследования внешней и внутренней среды проекта. В таблице 4.3 представлена матрица SWOT.

Таблица 4.3 – Матрица SWOT разработки

	<p>Сильные стороны: S1. Использование современных методов сегментации изображения S2. Результатом работы алгоритма являются координаты точки крепления листьев, а не бинарная маска всего растения, что позволяет установить положение корнеплода S3. Возможность использовать алгоритм в с/х технике различного назначения S4. Возможность настройки для работы с разными культурами и отсеивание сорняков</p>	<p>Слабые стороны: W1. Необходимость датасета для настройки под различные культуры W2. Работа алгоритма проверена только на входных данных, представляющих снимки с видом растений сверху W3. Данный алгоритм не является полноценным блоком для решения задач, а лишь служит частью аналитического модуля</p>
<p>Возможности: O1. Зарождение рынка роботов для сельского хозяйства и рост экспериментальной разработки в этой области O2. Большое количество видов техники для растениеводства</p>	<p>Результаты анализа интерактивной матрицы проекта полей «Сильные стороны и возможности» – Универсальность алгоритма для различных задач позволяет охватить значительную часть рынка – Акцент на возможности применения алгоритма на различных видах растений и отсеивании сорняков позволит привлечь клиентов, разрабатывающих ПО для специализированной техники – Привлечение новых клиентов благодаря конфигурированию продукта под их предметную область и задачу</p>	<p>Результаты анализа интерактивной матрицы проекта полей «Слабые стороны и возможности» – Проверка и доработка алгоритма для работы со снимками вида растений сбоку – Сбор и разметка датасетов различных видов растений – Доработка алгоритма для повышения удобства его настройки и конфигурации пользователем, а также для наиболее простой интеграции в систему пользователя – Разработка обучающих материалов: подробная документации с примерами, проведение семинаров</p>

Таблица 4.3 (Продолжение) – Матрица SWOT разработки

<p>Угрозы:</p> <p>T1. Появление на рынке других разработчиков с аналогичными продуктами</p> <p>T2. Появление новых, принципиально отличающихся методов решения</p>	<p>Результаты анализа интерактивной матрицы проекта полей «Сильные стороны и угрозы»</p> <p>– Алгоритм, локализуемый корнеплод вместо области, охватывающей листья, представляет собой более гибкий инструмент для потребителя, что делает его более конкурентоспособным</p> <p>– Получение результата в виде координат делает продукт более интересным и удобным для клиента</p> <p>– Привлечение специалистов, внешних консультантов, а также профессиональное сообщество для масштабирования с точки зрения функционала и технических требований</p>	<p>Результаты анализа интерактивной матрицы проекта полей «Слабые стороны и угрозы»</p> <p>– Оказание услуг по конфигурации под конкретную предметную область и задачу</p> <p>– Поддержание стабильной ценовой политики</p>
---	--	--

Необходимо отметить, что особенностью данного решения является его универсальность и возможность конфигурирования под каждую конкретную предметную область. Однако данный алгоритм не является конечным решением, поэтому для решения каждой задачи по отдельности требуется дополнительное конфигурирование, либо предоставление обучающих материалов для специалистов, разрабатывающих конечное решение. Расширение функционала со временем позволит сделать процесс конфигурирования наиболее удобным и эффективным.

Еще одной особенностью разработки является вид представления результата: алгоритм позволяет на изображении с камер робота определить точное положение корнеплода, тогда как большая часть методов находят лишь область, охватывающую видимую часть растения (в основном, листья), на выходе алгоритм формирует координаты. Эта особенность способна

сформировать интерес потенциальных пользователей технологии, что повышает конкурентоспособность.

Поддержание стабильной ценовой политики и оказание услуг по настройке под конкретную предметную область или задачу будет напрямую способствовать конкурентоспособности разрабатываемого решения.

4.2 Инициация проекта

В процессе инициации проекта определяются начальные цели и содержание, а также фиксируются финансовые ресурсы. Определяются внутренние и внешние заинтересованные стороны проекта, которые будут взаимодействовать и влиять на общий результат научного проекта.

4.2.1 Цели и результат проекта

В данном разделе приводится информация о заинтересованных сторонах проекта, иерархии целей проекта и критериях достижения целей.

Под заинтересованными сторонами проекта понимаются лица или организации, которые активно участвуют в проекте или интересы которых могут быть затронуты как положительно, так и отрицательно в ходе исполнения или в результате завершения проекта (таблица 4.4).

Таблица 4.4 – Заинтересованные стороны проекта

Заинтересованные стороны	Ожидание сторон
Организация-заказчик (производитель ПО)	Решение по локализации объектов на изображении
Производитель сельхозтехники	ПО для роботизированной техники, оснащенное методами компьютерного зрения
Фермерские компании	Автономно работающая техника

В таблице 4.5 представлена цель проекта, а также критерии ее достижения.

Таблица 4.5 – Цели и результат проекта

Цели проекта:	Разработать алгоритм локализации объектов на изображении
Ожидаемые результаты проекта:	Возможность определять координаты центра крепления листьев растения, соответствующих нахождению корнеплода. Разделение культурных растений от сорняков.
Критерии приемки результата проекта:	Прохождение оценки эффективности модели на тестовой выборке.
Требования к результату проекта:	Выполняются оба условия из ожидаемых результатов. Значения метрик оценки модели выше установленных пороговых величин.

4.2.2 Организационная структура проекта

В таблице 4.6 представлена информация о рабочей группе проекта.

Таблица 4.6 – Рабочая группа проекта

№	ФИО, место работы, должность	Роль в проекте	Функции
1	Сергеева Наталья Дмитриевна, магистрант ТПУ, инженер по анализу данных и машинному обучению ООО Эко-Томск	Инженер	1. Проектирование 2. Реализация 3. Внедрение
2	Аксенов Сергей Владимирович, ТПУ, кандидат технических наук	Научный руководитель	1. Составление научных целей и задач 2. Проверка документации

4.2.3 Ограничения и допущения проекта

Ограничения проекта представлены в таблице 4.7.

Таблица 4.7 - Ограничения проекта

Фактор	Ограничения
Бюджет проекта	300 000 рублей
Источник финансирования	ТПУ
Сроки проекта	21.01.2020-16.05.2020
Дата утверждения плана управления проектом	21.01.2020
Дата завершения проекта	16.05.2020

Максимальный бюджет настоящего проекта установлен в сумме 300 000 рублей, а сроки составляют с 21 января по 16 мая 2020.

4.3 Планирование управления проектом

Планирование проекта предполагает определение условий выполнения всех этапов и задач для установления порядка и последовательности.

Основные этапы планирования:

- определение структуры работ в рамках проекта;
- определение участников каждой работы;
- установление продолжительности работ;
- построение графика выполнения проекта.

4.3.1 Структура работ в рамках проекта

Для составления структуры работ определяются ключевые события проекта, затем детальный перечень этапов и работ. На каждый вид работ определяется исполнитель. Распределение исполнителей по данным видам работ приведено в таблице 4.8.

Таблица 4.8 – Распределение исполнителей по работам

Основные этапы	№ этапа (код работ)	Содержание работ	Исполнители
Разработка задания	1	Постановка задачи	Сергеева Н.Д. Аксенов С.В.
Выбор направления исследования	2	Обзор научно-технической базы	Сергеева Н.Д.
	3	Разработка и утверждение ТЗ	Сергеева Н.Д. Аксенов С.В.
	4	Составление календаря проекта	Сергеева Н.Д.
	5	Разработка вариантов исполнения проекта	Сергеева Н.Д. Аксенов С.В.

Таблица 4.8 (Продолжение) – Распределение исполнителей по работам

Основные этапы	№ этапа (код работ)	Содержание работ	Исполнители
Разработка продукта	6	Разработка модуля предобработки графических данных	Сергеева Н.Д.
	7	Подбор модели сегментации изображений	Сергеева Н.Д.
	8	Разработка функции оценки	Сергеева Н.Д.
	9	Разработка модуля постобработки	Сергеева Н.Д.
	10	Проверка на тестовых данных	Сергеева Н.Д. Аксенов С.В.
Оформление отчетной документации	11	Составление пояснительной записки	Сергеева Н.Д.

4.3.2 Определение трудоемкости выполнения работ

Для определения ожидаемых сроков выполнения проекта необходимо оценить его трудоемкость. Воспользуемся формулой:

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

где $t_{ож\ i}$ – ожидаемая трудоемкость выполнения i -ой работы чел.-дн.;

$t_{min\ i}$ – минимально возможная трудоемкость выполнения заданной i -ой работы (оптимистическая оценка: в предположении наиболее благоприятного стечения обстоятельств), чел.-дн.;

$t_{max\ i}$ – максимально возможная трудоемкость выполнения заданной i -ой работы (пессимистическая оценка: в предположении наиболее неблагоприятного стечения обстоятельств), чел.-дн.

Исходя из ожидаемой трудоемкости работ, определяется продолжительность каждой работы в рабочих днях T_{pi} , учитывающая параллельность выполнения работ несколькими исполнителями:

$$t_{pi} = \frac{t_{ож\ i}}{Ч_i}$$

где t_{pi} – продолжительность одной работы раб.-дн.; $t_{ожi}$ – ожидаемая трудоемкость выполнения одной работы, чел.-дн; $Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел. Для удобства составления календарного плана и графика работ необходимо перевести длительность каждого из этапов из рабочих дней в календарные дни. Для этого воспользуемся следующей формулой:

$$T_{ki} = T_{pi} \cdot k_{кал}$$

где t_{ki} – продолжительность выполнения i -й работы в календарных днях; t_{pi} – продолжительность выполнения i -й работы в рабочих днях; $k_{кал}$ – коэффициент календарности. Коэффициент календарности определяется по следующей формуле:

$$k_{кал} = \frac{T_{кал}}{T_{кал} - T_{вых} - T_{пр}} = \frac{366}{366 - 66} = 1,22$$

где $T_{кал}$ – количество календарных дней в году; $T_{вых}$ – количество выходных дней в году; $T_{пр}$ – количество праздничных дней в году. В соответствии с производственным календарем (для 6-дневной рабочей недели) в 2020 году 366 календарных дней, 300 рабочих дней, 66 выходных/праздничных дней. В таблице 4.9 представлены подробные временные расчеты этапов отдельных видов работ.

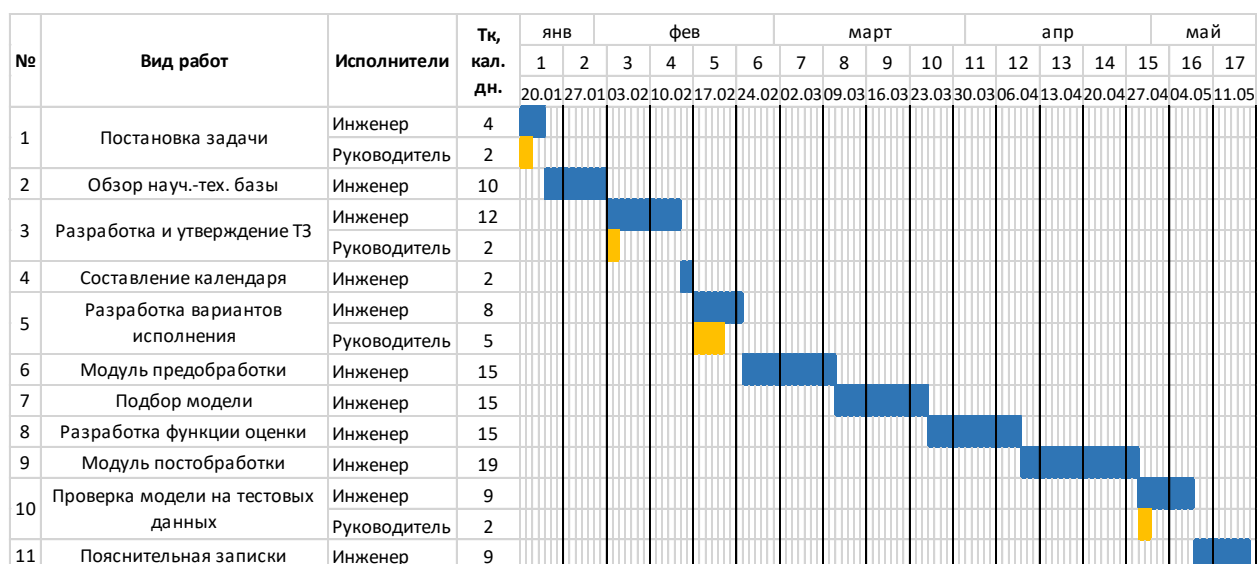
Таблица 4.9 – Временные показатели проведения проекта

Наименование работы	Исполнители работы	Трудоемкость работ, чел-дни			Длительность работ, дни	
		t_{min}	t_{max}	$t_{ож}$	T_p	T_k
Постановка задачи	Инженер	2	4	3,2	3	4
	Научный руководитель	1	3	1,8	2	2
Обзор научно-технической базы	Инженер	7	9	8,2	8	10
Разработка и утверждение ТЗ	Инженер	9	11	9,8	10	12
	Научный руководитель	1	3	1,8	2	2
Составление календаря проекта	Инженер	1	3	1,8	2	2
Разработка вариантов исполнения проекта	Инженер	5	7	6,2	6	8
	Научный руководитель	3	5	4,2	4	5

Таблица 4.9 (Продолжение) – Временные показатели проведения научно-технического проекта

Наименование работы	Исполнители работы	Трудоемкость работ, чел-дни			Длительность работ, дни	
		t_{min}	t_{max}	$t_{ож}$	T_p	T_k
Разработка модуля предобработки графических данных	Инженер	11	13	12,2	12	15
Подбор модели сегментации изображений	Инженер	11	13	12,2	12	15
Разработка функции оценки	Инженер	11	13	12,2	12	15
Разработка модуля постобработки	Инженер	14	16	15,2	15	19
Проверка модели на тестовых данных	Инженер	6	8	7,2	7	9
	Научный руководитель	1	3	1,8	2	2
Составление пояснительной записки	Инженер	6	8	7,2	7	9

Таблица 4.10 – Календарный план-график проекта



4.3.3 Бюджет проекта

Для полноты и достоверности учета всех расходов сгруппируем все затраты по следующим статьям:

- затраты на материалы;

- затраты на амортизацию;
- основная заработная плата исполнителей;
- дополнительная заработная плата исполнителей темы;
- отчисления во внебюджетные фонды (страховые отчисления);
- накладные расходы.

Материальные затраты

В расчет взяты только затраты на канцелярские товары в размере 1000 рублей.

Амортизационные отчисления

Для работы над проектом использовался ПК. Амортизацию рассчитаем линейным способом.

Первоначальная стоимость ПК 60000 рублей; срок полезного использования для машин офисных код 330.28.23.23 составляет 2-3 года, берем 3 года; планируется использовать ПК для написания работы в течение 4 месяцев. Тогда:

- месячная норма амортизации:

$$A_n = \frac{1}{n} \cdot 100 = \frac{1}{12 \cdot 3} \cdot 100 = 2,8$$

где n - количество месяцев полезного срока эксплуатации ОС.

Ежемесячные амортизационные отчисления:

$$A_r = 60000 \cdot 0,028 = 1680 \text{руб}$$

Итоговая сумма амортизации основных средств:

$$A_r = 1680 \cdot 4 = 6720 \text{руб}$$

Таким образом, в материальные затраты необходимо включить сумму амортизации основных средств в сумме 6720 руб.

Зарботная плата исполнителей проекта

Зарботная плата рассчитывается из суммы зарботной платы исполнителя и научного руководителя исходя из трудоемкости каждого этапа и занятости каждого из них на данном этапе. Расходы по статье зарботной платы:

$$C_{зп} = Z_{осн} + Z_{доп}$$

где $Z_{осн}$ – основная заработная плата; $Z_{доп}$ – дополнительная заработная плата.

Основная заработная плата рассчитывается по формуле:

$$Z_{осн} = Z_{дн} \cdot T_p$$

где $Z_{дн}$ – среднедневная заработная плата, руб.;

T_p – продолжительность работ, выполняемых работником, раб.дн.

Среднедневная заработная плата рассчитывается по формуле:

$$Z_{дн} = \frac{Z_m \cdot M}{F_d}$$

где Z_m – месячный должностной оклад работника, руб.;

M – количество месяцев работы без отпуска в течение года:

при отпуске в 24 раб. дня $M = 11,2$ месяца, 5-дневная неделя;

при отпуске в 48 раб. дней $M = 10,4$ месяца, 6-дневная неделя;

F_d – действительный годовой фонд рабочего времени научно-технического персонала, раб. дн.

Таблица 4.11 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	366	366
Количество нерабочих дней	66	66
- выходные дни		
- праздничные дни		
Потери рабочего времени	48	48
- отпуск		
- невыходы по болезни		
Действительный годовой фонд рабочего времени	252	252

Месячный должностной оклад работника:

$$Z_m = Z_б \cdot (1 + k_{пр} + k_d) \cdot k_p$$

где $Z_б$ – базовый оклад, руб.;

$k_{пр}$ – премиальный коэффициент, равный 0,3;

k_d – коэффициент доплат и надбавок составляет примерно 0,2 – 0,5 (в НИИ и на промышленных предприятиях – за расширение сфер обслуживания, за профессиональное мастерство, за вредные условия: 15-20 % от Z_6);

k_p – районный коэффициент, равный 1,3 (для Томска).

Для расчета основной заработной платы инженера берем оклад, равный окладу 21760 руб. Для расчета основной заработной платы руководителя в расчет возьмем оклад, равный 35120 руб.

Таблица 4.12 – Расчет основной заработной платы

Исполнители	Z_6 , руб.	$k_{пр}$	k_d	k_p	Z_m , руб	$Z_{дн}$, руб.	T_p , раб. дн.	$Z_{осн}$, руб.
Руководитель	35 120	0,3	0	1,3	59 353	2 449	10	24 990
Инженер	21 760	0	0	1,3	28 288	1 167	94	109 698

В дополнительную заработную плату входят суммы выплат, предусмотренные трудовым кодексом, например, оплата ежегодных и дополнительных отпусков, оплата времени, связанного с выполнением государственных и общественных обязанностей и т.д. (в среднем – 12 % от суммы основной заработной платы):

$$Z_{доп} = k_{доп} \cdot Z_{осн},$$

где $k_{доп}$ – коэффициент дополнительной заработной платы (на стадии проектирования принимается равным 0,12 – 0,15);

$Z_{осн}$ – основная заработная плата, руб.

В таблице 4.13 представлен расчет затрат на заработную плату исполнителей.

Таблица 4.13 – Заработная плата исполнителей

Заработная плата	Руководитель	Инженер
Основная зарплата	24 990	109 698
Дополнительная зарплата	2 999	23 164
Зарплата исполнителя	27 989	132 862
Итого по статье $C_{зп}$	160 851	

Таким образом, зарплата научного руководителя за период исполнения проекта составляет 27 989 рублей, инженера – 132 862 рублей. Всего расходов по статье заработной платы: 160 851 рублей.

Отчисления во внебюджетные фонды

В данной статье расходов отражаются обязательные отчисления по установленным законодательством Российской Федерации нормам органам государственного социального страхования (ФСС), пенсионного фонда (ПФ) и медицинского страхования (ФФОМС) от затрат на оплату труда работников.

Величина отчислений во внебюджетные фонды определяется исходя из следующей формулы:

$$З_{\text{внеб}} = k_{\text{внеб}} \cdot (З_{\text{осн}} + З_{\text{доп}})$$

где $k_{\text{внеб}}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

Общие тарифы страховых взносов в 2020 году в ИФНС: 22% — на пенсионное страхование; 2,9% — страхование по временной нетрудоспособности; 5,1% — медицинское страхование.

Отчисления во внебюджетные фонды составляют:

$$З_{\text{внеб}} = 0,3 \cdot 160\,851 = 48255 \text{ рублей}$$

Накладные расходы

При выполнении проекта могут возникнуть косвенные издержки – накладные расходы, возникающие дополнительно к основным затратам, например, на консультационные услуги, оплату коммунальных услуг, расходы на услуги связи (телефон, интернет) и так далее.

Их величина определяется по следующей формуле:

$$С_{\text{накл}} = k_{\text{накл}} \cdot (З_{\text{осн}} + З_{\text{доп}})$$

где $k_{\text{накл}}$ – коэффициент, учитывающий накладные расходы.

Величину коэффициента накладных расходов можно взять в размере 16%.

$$С_{\text{накл}} = 0,16 \cdot 160\,851 = 25736 \text{ рублей}$$

Прямые затраты

Данная статья включает затраты на электроэнергию, потребляемую оборудованием. Потребляемая мощность компьютера с видеокартой 0,4 кВт/ч, время работы 94 дня по 8 ч/день. Стоимость 1 кВт/час – составляет 2,45 руб.

$$C_{пр} = 0,4 \cdot 94 \cdot 8 \cdot 2,45 = 737 \text{ рублей}$$

Формирование бюджета

После выполнения всех расчетов по статьям можно определить плановую общую себестоимость проекта. В таблице 4.14 представлены статьи расходов проекта.

Таблица 4.14 – Бюджет затрат проекта

Наименование	Сумма, руб.	Удельный вес, %
Затраты на материалы	1 000	0,4
Затраты на амортизацию	6 720	2,7
Затраты на основную заработную плату	134 688	55,4
Затраты на дополнительную заработную плату	26 163	10,8
Страховые взносы	48 255	19,8
Накладные расходы	25 736	10,6
Прямые затраты	737	0,3
Общий бюджет	243 299	100

Исходя из расчета бюджета затрат следует, что наибольшая его часть приходится на основную и дополнительную заработную плату исполнителей (66,2 %). Также необходимо отметить, что расходы на страховые взносы (19,8 %) составляют значительную часть расходов. Затраты на амортизацию, материалы и накладные расходы составляют небольшую долю (суммарно 14 %). Это связано с отсутствием необходимости использования дорогостоящего оборудования и материалов.

4.3.4 Реестр рисков проекта

Риски в реализации проекта включают в себя возможные неопределенные события, которые могут возникнуть в проекте и вызвать

последствия, которые повлекут за собой нежелательные эффекты. Оценка рисков проекта представлена в таблице 4.15. Для каждого из них даны рекомендации по смягчению их воздействия.

Таблица 4.15 – Реестр рисков

№	Риск	Потенциальное воздействие	Вероятность наступления	Влияние риска (1-5)	Уровень риска	Способы смягчения риска	Условия наступления
1	Несоответствие разработанной и требуемой функциональности	Недостаточная функциональность может привести к неконкурентоспособности устройства	2	3	средний	Прототипирование, разработка сценариев использования, участие потенциальных пользователей	Ошибки при постановке задачи, недостаточный анализ качества разработки и ее перспективности на рынке
2	Постоянный поток изменений требований	Задержки выполнения работ	2	2	низкий	Установка ограничений для внесения изменений, итеративность разработки (внесения изменений в следующих итерациях)	Ошибки при постановке задачи
3	Технологическое отставание	Неконкурентоспособность устройства	2	2	низкий	Технический анализ, анализ стоимости, прототипирование	Недостаточная оценка существующих аналогов
4	Недостаточная производительность	Неконкурентоспособность устройства	1	3	средний	Проведение сравнительного тестирования, прототипирование	Ошибки при постановке задачи, недостаточный анализ качества разработки и ее перспективности на рынке

В результате данного этапа были рассмотрены возможные риски при реализации настоящей работы. Основная часть рисков может привести к неконкурентоспособности разработанного решения. Однако их воздействие можно минимизировать благодаря проведению прототипирования, итеративности разработки, проведению технического анализа стоимости и проведению сравнительного тестирования.

4.4 Определение экономической эффективности

Общая трудоемкость разработки решения составила 129 человеко-дней. Общий бюджет проекта составил 243 299 рублей. Исходя из ограничений, накладываемых на проект, максимальный бюджет не должен превышать 300 000 рублей.

Определим эффективность проекта с помощью интегрального финансового показателя эффективности:

$$I_{\text{фин}} = \frac{\Phi_p}{\Phi_{\text{max}}}$$

Где $I_{\text{фин}}$ – интегральный финансовый показатель разработки, Φ_p – стоимость исполнения работ, Φ_{max} – максимально допустимая стоимость исполнения проекта.

Таким образом значения финансового показателя составляет:

$$I_{\text{фин}} = \frac{243299\text{руб.}}{300000\text{руб.}} = 0,81$$

Значения финансового показателя составляет 0,81, что свидетельствует об эффективном использовании финансовых ресурсов.

Выводы

Результаты оценки востребованности разработки можно считать положительными, поскольку, во-первых, были выявлены потенциальные потребители разрабатываемого решения. Во-вторых, в результате анализа конкурентоспособности выяснилось, что разработанное решение является более предпочтительным для широкого класса задач и обладает достаточными конкурентными преимуществами благодаря новизне метода и особенностям получаемых результатов. В-третьих, проведенный SWOT анализ показал перспективность разработки. Расширение функционала, оказание услуг по настройке или консультированию под каждую конкретную предметную область совместно с поддержанием стабильной ценовой политики позволит сохранять свою конкурентоспособность.

В данном разделе разработан план и сформирован бюджет технического решения. Продолжительность проекта составила 118 календарных дней, а общий бюджет затрат составил 243 299 рубль. Таким образом, план-график и бюджет проекта успешно укладываются в ограничения. Разработанный реестр рисков отражает потенциальные пути преодоления внешних и внутренних рисков и способствует успешной реализации проекта, а также его дальнейшее существование, а рассчитанный интегральный финансовый показатель эффективности ($I_{фин} = 0,81$) свидетельствует о возможности реализации проекта.

5 Социальная ответственность

Целью дипломной работы является разработка алгоритма локализации сельскохозяйственных растений на изображении в течении вегетативного периода. В рамках настоящей работы предполагается создание алгоритмического и программного обеспечения, позволяющего обнаруживать сельскохозяйственные культурные растения на снимках и исключать сорняки с помощью методов машинного обучения. Особенностью метода является возможность его использования на разных стадиях роста растений.

В разделе будут рассмотрены опасные и вредные факторы, оказывающие влияние на производственную деятельность инженера-программиста. Исследовано рабочее место программиста и помещение, в котором он находится. Разработка осуществлялась в компьютерном классе Кибернетического центра ТПУ. Основные средства работы – персональный компьютер и локальная вычислительная сеть с выходом в Интернет. Рассмотрены воздействия объекта исследования на окружающую среду, правовые и организационные вопросы, а также мероприятия в чрезвычайных ситуациях.

5.1 Правовые и организационные вопросы обеспечения безопасности

Нормативное регулирование охраны труда при осуществлении трудовой деятельности за компьютерами осуществляется посредством следующих документов:

Трудовой кодекс РФ.

Приказ Минздравсоцразвития России "Об утверждении перечней вредных и (или) опасных производственных факторов и работ, при выполнении которых проводятся обязательные предварительные и периодические медицинские осмотры (обследования), и Порядка проведения обязательных предварительных и периодических медицинских осмотров

(обследований) работников, занятых на тяжелых работах и на работах с вредными и (или) опасными условиями труда" от 12.04.2011 N 302н.

Федеральный закон "О специальной оценке условий труда" от 28 декабря 2013 г. N 426.

СанПиН 2.2.2/2.4.1340-03 "Гигиенические требования к персональным электронно-вычислительным машинам и организации работы" (с изменениями на 21 июня 2016 года).

Типовая инструкция по охране труда при работе на персональном компьютере (ПК, ПЭВМ) ТОИ Р-45-084-01 и др.

Согласно "Трудовому кодексу Российской Федерации" от 30.12.2001 N 197-ФЗ работодатель обязан обеспечить нормальные условия для выполнения работниками норм выработки. К таким условиям, в частности, относятся:

исправное состояние помещений, сооружений, машин, технологической оснастки и оборудования;

своевременное обеспечение технической и иной необходимой для работы документацией;

надлежащее качество материалов, инструментов, иных средств и предметов, необходимых для выполнения работы, их своевременное предоставление работнику;

условия труда, соответствующие требованиям охраны труда и безопасности производства;

создание и функционирование системы управления охраной труда;

соответствующие требованиям охраны труда условия труда на каждом рабочем месте;

режим труда и отдыха работников в соответствии с трудовым законодательством и иными нормативными правовыми актами, содержащими нормы трудового права;

обучение безопасным методам и приемам выполнения работ и оказанию первой помощи пострадавшим на производстве, проведение

инструктажа по охране труда, стажировки на рабочем месте и проверки знания требований охраны труда;

недопущение к работе лиц, не прошедших в установленном порядке обучение и инструктаж по охране труда, стажировку и проверку знаний требований охраны труда;

организацию контроля за состоянием условий труда на рабочих местах, а также за правильностью применения работниками средств индивидуальной и коллективной защиты;

информирование работников об условиях и охране труда на рабочих местах, о риске повреждения здоровья, предоставляемых им гарантиях, полагающихся им компенсациях и средствах индивидуальной защиты;

принятие мер по предотвращению аварийных ситуаций, сохранению жизни и здоровья работников при возникновении таких ситуаций, в том числе по оказанию пострадавшим первой помощи;

санитарно-бытовое обслуживание и медицинское обеспечение работников в соответствии с требованиями охраны труда, а также доставку работников, заболевших на рабочем месте, в медицинскую организацию в случае необходимости оказания им неотложной медицинской помощи.

Федеральный закон "О специальной оценке условий труда" от 28.12.2013 N 426-ФЗ регламентирует проведение спецоценки, если деятельность работников предприятия предусматривает непрерывную работу за компьютеризированными системами. Результаты проведенной спецоценки влияют на установление гарантий и компенсаций работникам согласно Трудовому кодексу РФ. Так, сотрудники, условия труда на рабочих местах, которых признаны вредными, в зависимости от степени вредности имеют право на сокращенную рабочую неделю не более 36 часов, дополнительный отпуск не менее семи календарных дней и/или компенсацию в размере 4% от оклада.

Нормативные положения СанПиНа 2.2.2/2.4.1340-03 предъявляют определенные требования к оснащению рабочего места, предусматривающего длительную работу за ПК:

Таблица 5.1 – Нормы оборудования рабочих мест

Высота перегородок, разделяющих рабочие места	Не менее 1,5 метров
Ширина рабочего стола	От 80 до 140 см
Глубина рабочего стола	От 80 до 100 см
Высота рабочего стола	7,25 см
Расстояние от глаз до монитора	От 60 до 70 см
Расстояние клавиатуры от края стола	От 10 до 30 см
Сидение	Должно позволять регулировку по высоте, повороту и углу наклона спинки (регулировки должны быть независимыми друг от друга)
Подставка для ног	Ширина – от 30 см, глубина – от 40 см, с углом наклона до 20 градусов

5.2 Производственная безопасность

5.2.1 Анализ выявленных вредных и опасных факторов

Идентифицируем вредные и опасные факторы, возникающие при разработке проекта, в соответствии с ГОСТом 12.0.003-2015 «Опасные и вредные производственные факторы. Классификация».

Таблица 5.2 – Негативные факторы при работе с компьютером

Факторы (ГОСТ 12.0.003-2015)	Этапы работ		Нормативные документы
	Разработка	Эксплуатация	
Отклонение показателей микроклимата	+	+	СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату производственных помещений
Превышение уровня шума	+	+	СанПиН 2.2.4.3359-16 "Санитарно-эпидемиологические требования к физическим факторам на рабочих местах"
Отсутствие или недостаток естественного света	+	+	СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95

Таблица 5.2 (Продолжение) – Негативные факторы при работе с компьютером

Факторы (ГОСТ 12.0.003-2015)	Этапы работ		Нормативные документы
	Разработка	Эксплуатация	
Повышенная напряженность магнитного поля	+	+	СаНПиН 2.2.2/2.4.1340-03 "Гигиенические требования к персональным электронно-вычислительным машинам и организации работы" (с изменениями на 21 июня 2016 года)
Психофизиологический фактор	+	+	СаНПиН 2.2.2/2.4.1340-03 "Гигиенические требования к персональным электронно-вычислительным машинам и организации работы" (с изменениями на 21 июня 2016 года)

5.2.2 Обоснование мероприятий по снижению воздействия

1) Отсутствие или недостаток естественного света

В помещении при работе с ПК должно быть естественное и искусственное освещение. Естественное освещение обеспечивается через оконные проемы с коэффициентом естественного освещения КЕО не ниже 1,2% в зонах с устойчивым снежным покровом и не ниже 1,5% на остальной территории. Световой поток из оконного проема должен падать на рабочее место оператора с левой стороны.

Искусственное освещение в помещениях эксплуатации компьютеров должно осуществляться системой общего равномерного освещения.

Освещенность на поверхности стола в зоне размещения документа должна быть 300-500 лк. Допускается установка светильников местного освещения для подсветки документов. Местное освещение не должно создавать бликов на поверхности экрана и увеличивать освещенность экрана более 300 лк. Прямую блескость от источников освещения следует

ограничить. Яркость светящихся поверхностей (окна, светильники), находящихся в поле зрения, должна быть не более 200 кд/м².

Отраженная блескость на рабочих поверхностях ограничивается за счет правильного выбора светильника и расположения рабочих мест по отношению к естественному источнику света. Яркость бликов на экране монитора не должна превышать 40 кд/м². Показатель ослепленности для источников общего искусственного освещения в помещениях должен быть не более 20, показатель дискомфорта в административно-общественных помещениях не более 40. Соотношение яркости между рабочими поверхностями не должно превышать 3:1 – 5:1, а между рабочими поверхностями и поверхностями стен и оборудования 10:1.

Для искусственного освещения помещений с персональными компьютерами следует применять светильники типа ЛПО36 с зеркализированными решетками, укомплектованные высокочастотными пускорегулирующими аппаратами. Допускается применять светильники прямого света, преимущественно отраженного света типа ЛПО13, ЛПО5, ЛСО4, ЛПО34, ЛПО31 с люминисцентными лампами типа ЛБ. Допускается применение светильников местного освещения с лампами накаливания. Светильники должны располагаться в виде сплошных или прерывистых линий сбоку от рабочих мест параллельно линии зрения пользователя при разном расположении компьютеров. При периметральном расположении — линии светильников должны располагаться локализованно над рабочим столом ближе к его переднему краю, обращенному к оператору. Защитный угол светильников должен быть не менее 40 градусов. Светильники местного освещения должны иметь непросвечивающийся отражатель с защитным углом не менее 40 градусов.

Для обеспечения нормативных значений освещенности в помещениях следует проводить чистку стекол оконных проемов и светильников не реже двух раз в год и проводить своевременную замену перегоревших ламп.

Расчет уровня освещенности рабочего места в аудитории

Аудитория оснащена светильниками ЛВО 4×18 CSVT, с люминесцентными лампами типа L 18W/640. Количество светильников составляет 6 штук. Размер аудитории равен 6,5м × 5,5м. Высота подвеса светильников 3м (расстояние от рабочей поверхности до светильника).

Освещенность помещения рассчитывается по следующей формуле [1]:

$$E_{\Phi} = \frac{N \cdot \eta \cdot \Phi}{S \cdot k_3 \cdot z},$$

где N – число светильников; η – коэффициент использования светового потока; Φ – световой поток светильника, лм; S – площадь помещения, м²; k_3 – коэффициент запаса, учитывающий загрязнение светильника; z – коэффициент неравномерности освещения.

Для приведенного типа светильников (с мощностью 4×18 Вт) световой поток Φ составляет 4600 лм. Количество светильников $N = 6$. Площадь помещения $S = 35,75$ м². Значение коэффициента неравномерности освещения для люминесцентных ламп при расчётах берётся равным $z = 1,1$. Коэффициент запаса k_3 определяется из значений, приведенных в таблице 5.3.

Таблица 5.3. Коэффициент запаса светильников с люминесцентными лампами

Характеристика объекта	Коэффициент запаса
Помещения с большим выделением пыли	2,0
Помещения со средним выделением пыли	1,8
Помещения с малым выделением пыли	1,5

Для компьютерной аудитории примем значение коэффициента как для помещения с малым выделением пыли, то есть $k_3 = 1,5$.

Коэффициент использования светового потока η показывает, какая часть светового потока ламп попадает на рабочую поверхность. Он зависит от индекса помещения i , типа светильника, высоты светильников над рабочей поверхностью h и коэффициентов отражения стен ρ_c и потолка ρ_n .

Индекс помещения i определяется по формуле:

$$i = \frac{S}{h(A + B)},$$

где S – площадь помещения, м²; A – длина комнаты, м; B – ширина комнаты, м; h – высота подвеса светильников, м.

Для рассматриваемого помещения индекс равен:

$$i = \frac{6,5 \cdot 5,5}{3 \cdot (6,5 + 5,5)} = 0,993.$$

Коэффициенты отражения стен ρ_c и потолка ρ_n определяются по таблице 5.4.

Таблица 5.4. Значение коэффициентов отражения потолка и стен

Состояние потолка	$\rho_n, \%$	Состояние стен	$\rho_{ст}, \%$
Свежепобеленный	70	Свежепобеленные с окнами, закрытыми шторами	70
Побеленный, в сырых помещениях	50	Свежепобеленные с окнами без штор	50
Чистый бетонный	50	Бетонные с окнами	30
Светлый деревянный (окрашенный)	50	Оклеенные светлыми обоями	30
Бетонный грязный	30	Грязные	10
Деревянный неокрашенный	30	Кирпичные неоштукатуренные	10
Грязный (кузницы, склады)	10	С темными обоями	10

Потолок в аудитории покрыт белой потолочной плиткой из пенопласта, коэффициент отражения которой близок к коэффициенту отражения свежепобеленного потолка. Стены выкрашены белой краской, в аудитории присутствуют окна, примем коэффициент отражения как для свежепобеленных стен с окнами без штор. Таким образом, $\rho_n = 0,7$ и $\rho_c = 0,5$.

При таких значениях показателей коэффициент использования светового потока $\eta = 0,48$ [31].

Тогда освещенность помещения равна:

$$E_{\Phi} = \frac{9 \cdot 0,48 \cdot 4600}{35,75 \cdot 1,5 \cdot 1,1} = 336,9 \text{ лм.}$$

Данное значение освещенности соответствует допустимому уровню освещенности согласно СНиП 23-05-95.

2) Микроклимат

Длительное воздействие на человека неблагоприятных метеорологических условий резко уменьшает его самочувствие, снижает производительность труда и часто приводит к заболеваниям.

Влажность воздуха оказывает значительное влияние на терморегуляцию организма человека. Высокая относительная влажность воздуха при его высокой температуре способствует перегреванию организма. Низкая влажность вызывает пересыхание слизистых оболочек дыхательных путей. Подвижность воздуха весьма эффективно способствует теплоотдаче, что является положительным явлением при высокой температуре окружающей среды и отрицательным – при низкой.

Оптимальные и допустимые показатели температуры, относительной влажности и скорости движения воздуха в рабочей зоне производственных помещений должны соответствовать значениям, согласно СанПиН 2.2.4.548 – 96 для категории тяжести работ 1а (к категории 1а относятся работы с интенсивностью энерготрат до 120 ккал/ч (до 139 Вт), производимые сидя и сопровождающиеся незначительным физическим напряжением).

Таблица 5.5 – Оптимальные нормы микроклимата.

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

Таблица 5.6 – Допустимые величины показателей микроклимата на рабочих местах производственных помещений

Период года	Температура воздуха, °С		Относительная влажность воздуха, %	Скорость движения воздуха, не более, м/с	
	Диапазон ниже оптимальных величин	Диапазон выше оптимальных величин		при температуре воздуха ниже оптимальной	при температуре воздуха выше оптимальной
Холодный	20,0 – 21,9	24,1 – 25,0	15 – 75	0,1	0,1
Теплый	21,0 – 22,9	25,1 – 28,0	15 - 75	0,1	0,2

При обеспечении допустимых величин микроклимата на рабочих местах перепад температуры воздуха по высоте должен быть не более 3°C, перепад температуры воздуха по горизонтали, а также ее изменения в течение смены не должны превышать – 4°C. При этом абсолютные значения температуры воздуха не должны выходить за пределы оптимальных величин.

При температуре воздуха на рабочих местах 25 °С и выше максимально допустимые величины относительной влажности воздуха не должны выходить за пределы:

- 70% – при температуре воздуха 25 °С,
- 65% – при температуре воздуха 26 °С,
- 60% – при температуре воздуха 27 °С,
- 55% – при температуре воздуха 28 °С.

3) Уровень шума

Шум на рабочем месте оказывает раздражающее влияние на работника, повышает его утомляемость, а при выполнении задач, требующих внимания и сосредоточенности, способен привести к росту ошибок и увеличению продолжительности выполнения задания. Согласно СанПиН 2.2.4.3359-16 нормативным эквивалентным уровнем звука на рабочих местах является 50 дБА.

Источником шумовых помех на рабочем месте могут выступать вентиляционные установки, кондиционеры, ЭВМ и его периферийные устройства, а также серверные комнаты.

4) Повышенная напряженность магнитного поля

Электромагнитные поля, характеризующиеся напряженностями электрических и магнитных полей, наиболее вредны для организма человека. Основным источником этих проблем, связанных с охраной здоровья людей, использующих в своей работе автоматизированные информационные системы на основе персональных компьютеров, являются дисплеи (мониторы), они

представляют собой источники наиболее вредных излучений, неблагоприятно влияющих на здоровье человека.

Предельно допустимые значения излучений от ЭВМ в соответствии с СанПиН 2.2.2/2.4.1340-03 приведены в таблице 5.7.

Таблица 5.7. Допустимые уровни ЭМП, создаваемых

Наименование параметров		ВДУ ЭМП
Напряженность электрического поля	В диапазоне частот 5 Гц – 2 кГц	25 В/м
	В диапазоне частот 2 кГц – 400 кГц	2,5 В/м
Плотность магнитного потока	В диапазоне частот 5 Гц – 2 кГц	250 нТл
	В диапазоне частот 2 кГц – 400 кГц	25 нТл
Электростатический потенциал экрана видеомонитора		500 В

5) Психофизиологический фактор

Настоящей работе сопутствует ряд вредных психофизиологических факторов: напряжение зрения и внимания; эмоциональные, интеллектуальные и длительные статические нагрузки; большая монотонность труда; большой объем обрабатываемой информации в единицу времени; нерациональная организация рабочего места.

В результате воздействия данных факторов к концу рабочего дня рабочий испытывает неприятные ощущения: переутомление глаз, головная боль, тянущие боли в мышцах шеи, рук и спины, снижение концентрации внимания.

В целом продолжительность непрерывной работы за компьютером не должна превышать 2-х часов. Основная работа за компьютером предусматривает не менее 50 % времени в течение рабочей смены или рабочего дня нахождения за ним. Время перерыва зависит от вида и сложности осуществляемой работы путем деления на группы. Выделяют 3 группы: А (работа по считыванию информации с экрана компьютера с предварительным запросом), Б (работа по вводу информации), В (творческая работа в режиме диалога с компьютером). Во время перерывов следует выполнять

специальную гимнастику для снятия напряжения с глаз. Рекомендуемый комплекс упражнения представлен в Приложении 8 к СанПиН 2.2.2/2.4.1340-03. Выполнять какую-либо работу, не связанную с компьютером, во время перерыва нельзя. Потому как перерыв приравнивается к времени отдыха. А в соответствии со ст. 106 ТК РФ время отдыха – это свободное от исполнения трудовых обязанностей время, которое работник может использовать по своему усмотрению.

5.3 Экологическая безопасность

Деятельность по разработке ПО не связана с производством, поэтому влияние на окружающую среду минимально. При работе над проектом применяются рекомендации по минимизации влияния на окружающую среду.

С учетом того, что наиболее значительную часть твердых бытовых отходов (до 40 % в развитых странах) составляет бумага и картон — бумага для печати, упаковка и упаковочные материалы, актуальным является осуществление их утилизации для повторного использования в промышленном производстве, что является наиболее экономически эффективным способом обращения с отходами согласно ГОСТ Р 55090-2012 «Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги» [32].

При завершении срока службы ПК их можно отнести к отходам электронной промышленности. Переработка таких отходов осуществляется разделением на однородные компоненты, химическим выделением пригодных для дальнейшего использования компонентов и направлением их для дальнейшего использования согласно ГОСТ Р 55102-2012 «Ресурсосбережение. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования, за исключением ртульсодержащих устройств и приборов» [33].

Перечень элементов и содержащее их отработанное электротехническое и электронное оборудование, которые должны быть раздельно собраны при выводе отработавшего электротехнического и электронного оборудования из эксплуатации:

- конденсаторы, содержащие ПХБ;
- печатные платы и других устройств с площадью поверхности больше 10 см²;
- картриджи;
- пластик;
- электронно-лучевые трубки;
- элементы отработавшего электротехнического и электронного оборудования;
- газоразрядные лампы;
- жидкокристаллические экраны (если необходимо, вместе с корпусом) с поверхностью более 100 см² и все экраны с подсветкой газоразрядными лампами;
- внешние электрические кабели; - элементы, содержащие огнеупорные керамические слои;
- конденсаторы, содержащие электролит (размер хотя бы одной из сторон конденсатора должен быть 25 мм или более).

Люминесцентные лампы относят к ртутьсодержащим отходам, и для их утилизации действует Постановление Правительства РФ от 03.09.2010 № 681 (ред. от 01.10.2013) «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [34]. Согласно постановлению, устанавливается порядок обращения с отходами производства и потребления в части осветительных устройств,

электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде.

Не допускается самостоятельное обезвреживание, использование, транспортирование и размещение отработанных ртутьсодержащих ламп потребителями отработанных ртутьсодержащих ламп, а также их накопление в местах, являющихся общим имуществом собственников помещений многоквартирного дома, за исключением размещения в местах первичного сбора и размещения и транспортирования до них.

Сбор отработанных ртутьсодержащих ламп у потребителей осуществляют специализированные организации.

Отходы, не подлежащие переработке и вторичному использованию подлежат захоронению на полигонах.

5.4 Безопасность в чрезвычайных ситуациях

5.4.1 Перечень возможных ЧС на объекте

Компьютерная аудитория кафедры, согласно НПБ 105-03 «Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности» относится к категории В3 по пожароопасности, содержит вещества и материалы, способные при взаимодействии с водой, кислородом воздуха или друг с другом только гореть.

Для минимизации возможности возникновения фактора пожара необходимо проводить пожарную профилактику. Пожарная профилактика представляет собой комплекс организационных и технических мероприятий, направленных на обеспечение безопасности людей, на предотвращении пожара, ограничение его распространения, а также создание условий для успешного тушения пожара. Для профилактики пожара чрезвычайно важна

правильная оценка пожароопасности, определение опасных факторов и обоснование способов и средств пожар предупреждения и защиты.

Одно из условий обеспечения пожаробезопасности - ликвидация возможных источников воспламенения.

Обогревание помещения открытыми электронагревательными приборами могут привести к пожару, т.к. в помещении находятся бумажные документы и справочная литература. Следовательно, использование открытого нагревательного прибора неприемлемо.

5.4.2 Меры по предотвращению и ликвидации ЧС и их последствий

В целях предотвращения пожара предлагается:

проводить с сотрудниками противопожарный инструктаж;

проводить плановый осмотр и своевременно устранять все неисправности в электроприборах;

предотвращать небезопасное хранение легковоспламеняющихся жидкостей;

оснащение помещения автоматической системой обнаружения пожара;

оснащение помещения автоматической системой оповещения о пожаре.

Согласно СП 5.13130.2009 «Системы противопожарной защиты. Установки пожарной сигнализации и пожаротушения автоматические.

Нормы и правила проектирования (с Изменением N 1)» [5] и СП 3.13130.2009 «Системы противопожарной защиты. Система оповещения и управления эвакуацией людей при пожаре. Требования пожарной безопасности» [6] помещения с категорией ВЗ должны оснащаться системой автоматической пожарной сигнализации и система оповещения и управления эвакуацией людей при пожарах.

Компьютерная аудитория кафедры оснащена одним пожарным, а также двумя дымовыми извещателями, для оповещения о пожаре установлена звуковая сирена.

Выводы

В результате проведенного анализа были выявлены вредные и опасные производственные факторы для компьютерной аудитории кафедры и работы, связанной с проектом.

Для вредных и опасных факторов, из нормативных документов, были определены значения нормативных показателей. Приведены рекомендации по улучшению условий труда и минимизации влияния на работника вредных и опасных факторов. Проведен анализ воздействия на окружающую среду и обозначены проблемы утилизации отходов. Определена категория помещения по пожароопасности, а также действия по минимизации риска возникновения пожара.

Проведен расчет уровня освещенности рабочего места аудитории, в которой выполнялась разработка. Полученное значение освещенности соответствует допустимым показателям, определенным соответствующим нормативным документом.

Также были рассмотрены правовые нормы трудового законодательства, применимые к условиям настоящего проекта, определены основные требования к организации рабочего места.

После проведенного анализа можно сделать вывод, что рабочее место соответствует нормативным требованиям производственной безопасности и охраны труда.

Заключение

В рамках настоящей работы разработан алгоритм локализации объектов на изображении и произведена его реализация на датасете снимков сахарной свеклы на разных стадиях вегетационного роста. Выполнены следующие задачи:

- 1) Проведен обзор литературы по теме работы и рассмотрены методы решения задачи распознавания объектов на изображении.
- 2) Составлены подход к решению поставленной задачи и структура алгоритма.
- 3) Спроектирован модуль сегментации изображений.
- 4) Разработан модуль вычисления относительных координат объекта на изображении.
- 5) Выбрана методика оценки качества обученной модели.
- 6) Алгоритм локализации реализован на имеющихся данных.
- 7) Проведено обсуждение результатов.

Рассмотренная задача локализации структурных центров растений была сформулирована как задача детекции ключевых точек объектов на изображении. Разработанный алгоритм состоит из двух модулей, выполняющих сегментацию изображений и вычисление координат ключевых точек на сгенерированной маске. Сегментация выполняется с помощью сверточной нейронной сети TernausNet с предобученным кодировщиком сети VGG16, взвешенным расстоянием Хаусдорфа в качестве функции ошибки и заменой слоев деконволюции в декодировщике на последовательные операции билинейной интерполяции и свертки. Обучение модели проведено в два этапа: обучение всей сети на дополнительном датасете изображений из близкой предметной области с последующим обновлением весов последнего слоя и обучение на основных данных – снимках сахарной свеклы в период вегетационного роста.

Эффективность модели подтверждена значениями метрик качества: F-score обученной модели составил 94,1%. Данное значение превышает значения метрик, полученные другими архитектурами, которые участвовали в сравнении при выборе модели. В том числе полученная величина F-score оказалась выше, чем для модели, обученной в один этап исключительно на основном наборе данных, что подтвердило эффективность двухэтапной тренировки с дополнительным набором данных.

Эффективность модели также была доказана для обработки снимков растений на различных стадиях роста. Оценка качества модели проведена на подвыборках изображений с побегами на ранней, средней и поздней стадиях вегетативного периода – значения F-score для них составили, соответственно, 93,5%, 96,8%, и 91,9%.

В работе приведена оценка востребованности и экономической эффективности разработки, разработан бюджет проектного решения и сформирован реестр рисков. Проведен анализ вредных и опасных производственных факторов, произведена оценка рабочего места на соответствие нормативным требованиям производственной безопасности и охраны труда.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Thies J. et al. Face2face: Real-time face capture and reenactment of rgb videos //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 2387-2395.
2. Using Markerless Augmented Reality for Digital Avatars [Электронный ресурс] // Deepmotion: [сайт] – 2019. – Режим доступа: <https://blog.deepmotion.com/2019/05/21/markerless-augmented-reality-for-ar-avatars/>, свободный (дата обращения: 15.04.2020).
3. Belghit H. et al. Vision-based Pose Estimation for Augmented Reality: A Comparison Study //arXiv preprint arXiv:1806.09316. – 2018.
4. Zhang S. H. et al. Pose2seg: detection free human instance segmentation //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2019. – С. 889-898.
5. Zhang Y. et al. Aircraft type recognition in remote sensing images based on feature learning with conditional generative adversarial networks //Remote Sensing. – 2018. – Т. 10. – №. 7. – С. 1123.
6. Girshick R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2014. – С. 580-587.
7. Ren S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks //Advances in neural information processing systems. – 2015. – С. 91-99.
8. Girshick R. Fast r-cnn //Proceedings of the IEEE international conference on computer vision. – 2015. – С. 1440-1448.
9. He K. et al. Mask r-cnn //Proceedings of the IEEE international conference on computer vision. – 2017. – С. 2961-2969.
10. Redmon J. et al. You only look once: Unified, real-time object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 779-788.

11. Moon G., Chang J. Y., Lee K. M. Posefix: Model-agnostic general human pose refinement network //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2019. – C. 7773-7781.
12. Sun K. et al. Deep high-resolution representation learning for human pose estimation //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2019. – C. 5693-5703.
13. Li W. et al. Rethinking on multi-stage networks for human pose estimation //arXiv preprint arXiv:1901.00148. – 2019.
14. Pavlakos G. et al. 6-dof object pose from semantic keypoints //2017 IEEE international conference on robotics and automation (ICRA). – IEEE, 2017. – C. 2011-2018.
15. Ribera, J., Guera, D., Chen, Y., & Delp, E. J. (2019). Locating Objects Without Bounding Boxes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6479-6489).
16. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks //arXiv preprint arXiv:1511.06434. – 2015.
17. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks //arXiv preprint arXiv:1511.06434. – 2015.
18. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation //International Conference on Medical image computing. – Springer, Cham, 2015. – C. 234-241.
19. Odena A., Dumoulin V., Olah C. Deconvolution and checkerboard artifacts //Distill. – 2016. – T. 1. – №. 10. – C. e3.
20. Dasiopoulou, Stamatia, et al. "Knowledge-assisted semantic video object detection." IEEE Transactions on Circuits and Systems for Video Technology 15.10 (2005): 1210-1224.
21. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine

Intelligence, vol. 39, no. 4, pp. 640-651, 1 April 2017, doi: 10.1109/TPAMI.2016.2572683.

22. Ribera, Javier, et al. "Weighted hausdorff distance: A loss function for object localization." arXiv preprint arXiv:1806.07564 (2018).

23. Скворцов, В. А. (2002). Примеры метрических пространств. М.: МЦНМО, 24.

24. Iglovikov, V., & Shvets, A. (2018). Terausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746.

25. Shvets, Alexey & Iglovikov, Vladimir & Rakhlin, Alexander & Kalinin, Alexandr. (2018). Angiodysplasia Detection and Localization Using Deep Convolutional Neural Networks.

26. Maggiori, Emmanuel, et al. "Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark." 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2017.

27. Ning Qian. On the momentum term in gradient descent learning algorithms. Neural networks: the official journal of the International Neural Network Society, 12(1):145–151, 1999

28. Kingma D. P., Ba J. Adam: A method for stochastic optimization //arXiv preprint arXiv:1412.6980. – 2014.

29. Zeiler M. D. Adadelta: an adaptive learning rate method //arXiv preprint arXiv:1212.5701. – 2012.

30. Sorghum Plant Centers 2016 Dataset. The PhenoSorg Project // VIPER Video and Image Processing Laboratory [Электронный ресурс] – Режим доступа: <https://engineering.purdue.edu/~sorghum/dataset-plant-centers-2016/>, свободный (дата обращения: 16.12.2019).

31. Назаренко О. Б. Расчёт искусственного освещения. Методические указания к выполнению индивидуальных заданий для студентов дневного и заочного обучения всех направлений и специальностей ТПУ. – Томск: Изд. ТПУ, 2008.

32. ГОСТ 55090-2012. Ресурсосбережение. Обращение с отходами. Рекомендации по утилизации отходов бумаги // Электронный фонд правовой и нормативно-технической документации. [Электронный ресурс]. URL: <http://docs.cntd.ru/document/1200103182> (дата обращения: 06.05.2020).

33. ГОСТ 55102-2012. Ресурсосбережение. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования, за исключением ртульсодержащих устройств и приборов // Электронный фонд правовой и нормативно-технической документации. [Электронный ресурс]. URL: <http://docs.cntd.ru/document/1200104723> (дата обращения: 06.05.2020).

34. Постановление Правительства РФ от 03.09.2010 N 681 (ред. от 01.10.2013) "Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде // Государственная система правовой информации [Электронный ресурс]. URL: <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102141053> (дата обращения: 06.05.2020).

35. СП 5.13130.2009 Системы противопожарной защиты. Установки пожарной сигнализации и пожаротушения автоматические. Нормы и правила проектирования (с Изменением N 1) // Электронный фонд правовой и нормативно-технической документации. [Электронный ресурс]. URL: <http://docs.cntd.ru/document/1200071148> (дата обращения: 06.05.2020).

36. СП 3.13130.2009 Системы противопожарной защиты. Система оповещения и управления эвакуацией людей при пожаре. Требования пожарной безопасности // Электронный фонд правовой и нормативно-технической документации. [Электронный ресурс]. URL: <http://docs.cntd.ru/document/1200071145> (дата обращения: 06.05.2020).

Приложение I

(справочное)

Designing method for the problem of objects localization in an image

Студент

Группа	ФИО	Подпись	Дата
8ПМ8И	Сергеева Н.Д.		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Аксенов С.В.	к.т.н.		

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Пичугова И.Л.			

1 Approach used

We defined the task as obtaining a single 2D coordinate corresponding to the location of the structural center for each object. The task could be formulated as *Keypoint Detection*.

A common approach in object detection was to obtain bounding boxes around the objects of interest. In this work, we were not interested in obtaining bounding boxes due to some reasons:

- 1) The goal of the task was to obtain position of not the whole object, but its structural center (different from the geometrical center that could be obtained from known bounding box).

- 2) Bounding-box annotation is tedious, time-consuming and expensive.

The proposed method has no notion of bounding boxes, region proposals, or sliding windows.

Regarding to the solution development, Keypoint Detection is segmentation problem, as the model produces saliency map as output. Processing the saliency map allow obtaining the result in the form of point coordinate. The model takes the input in form of pair image – set of single 2D coordinates. Furthermore, we do not know in advance the number of keypoints in the image.

2 Pipeline description

The algorithm included two modules providing image segmentation and coordinate calculation. The general pipeline is presented in Figure 1.

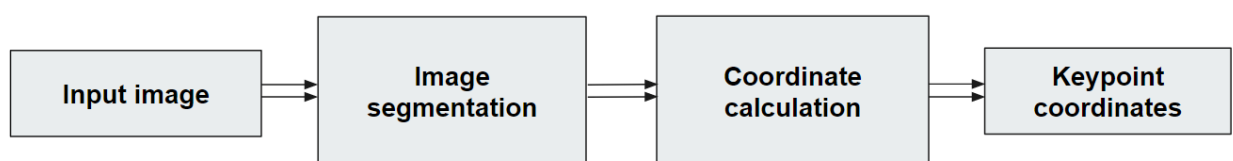


Figure 1. The general pipeline

The detailed pipeline is shown in Figure 2. The module of image segmentation supposed usage of Unet-like neural network – TernausNet with

Weighted Hausdorff Distance as the loss function. The design of the segmentation module is discussed in more detail in paragraph 3.

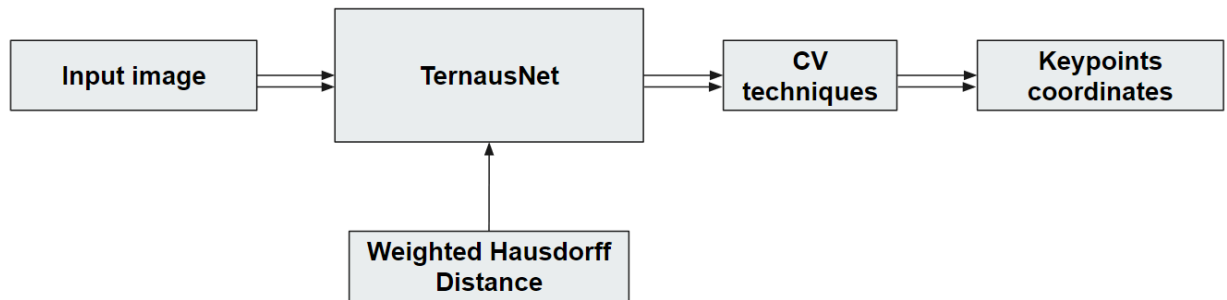


Figure 2. The detailed pipeline

The output of the neural network is the probability map which is the matrix of the size equal to the input image with elements from 0 to 1. These values predicted pixel-wise by the model are the probabilities of pixels to be in the region of interest. Figure 3 presents the visualization where pixels in yellow areas have probabilities close to 1, minor dark areas (green and blue) have less values and the violet area corresponds to background. The regions with values considerably less than 1 would be “cut” while post-processing.

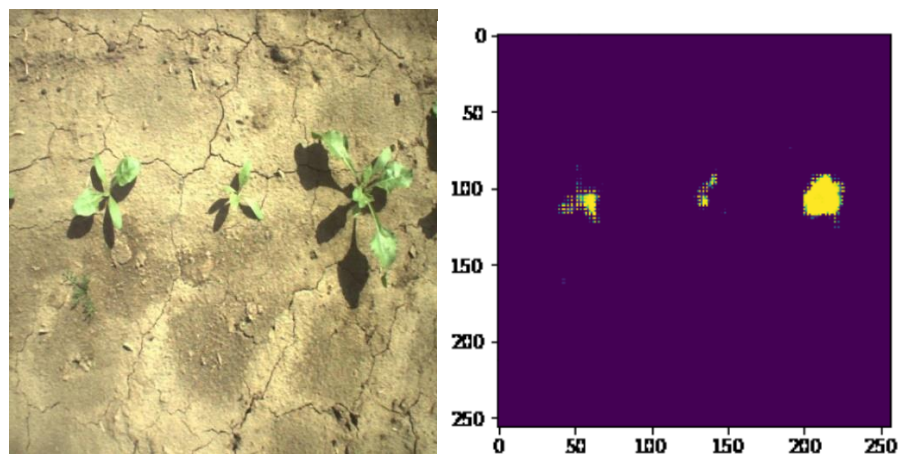


Figure 3. Example of the image probability map

The probability map has bright regions that correspond to the keypoints positions (in the current work – structural centers of the sprouts). Then the module of coordinate calculation processes the predicted probability map. The module applies computer vision techniques (which is discussed in detail in paragraph 4) and outputs the list of 2D coordinates.

Such a pipeline for object localization problem arose from several arguments:

1) Weighted Hausdorff distance as the loss function allowed training model with single points coordinates as the ground truth.

2) Objects in the images from our dataset had distances between their cores larger than or comparable with their size, therefore, plant keypoints were sharply separated by the model. This lead to the implementation of the coordinate calculation based on CV techniques.

3) The network for this task must be able to estimate a variable number of object locations. The choice of Fully Convolutional Network arises from the natural interpretation of its outputs as the weights (p_x) in the function of Weighted Hausdorff Distance (WHD). The last layer is a convolutional layer that maps to the single-channel output of the network p ; p_x – single-valued output of network at pixel coordinate x .

Consequently, the proposed method was not restricted to a maximum number of objects in the image and did not require bounding boxes.

3 Image segmentation module development

Loss function

As a loss function we applied Weighted Hausdorff Distance (WHD). The WHD can be used in any fully convolutional network (FCN) to estimate object locations [1]. This loss function is a modification of the average Hausdorff distance between two unordered sets of points (Figure 4).

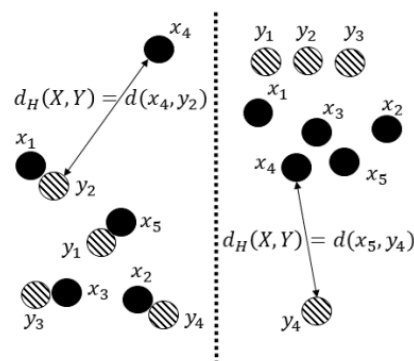


Figure 4. Hausdorff distance between two unordered sets of points

The choice of WHD is determined by the following limitations of the architecture:

4) Convolutional Neural Networks with linear layers implicitly determine the estimated number of points as the size of the last layer – this is a drawback because the actual number of points depends on the content of the image itself;

5) Fully Convolutional Networks indicate the presence of an object center with higher activation, but they do not return the pixel coordinates. In order to learn with backpropagation, the loss function must be differentiable with respect to the network output.

The WHD overcomes these limitations and is differentiable with respect to the network output. It takes probability map as prediction to evaluate and coordinates as ground truth. The equation of WHD is:

$$d_{WH}(p, Y) = \frac{1}{S + \epsilon} \sum_{x \in \Omega} p_x \min_{y \in Y} d(x, y) + \frac{1}{|Y|} \sum_{y \in \Omega} M_{\alpha}^{x \in \Omega} [p_x d(x, y) + (1 - p_x) d_{max}],$$

where

$$S = \sum_{x \in \Omega} p_x,$$

$$M_{\alpha}^{a \in A} [f(a)] = \left(\frac{1}{|A|} \sum_{a \in A} f^{\alpha}(a) \right)^{\frac{1}{\alpha}}.$$

M_{α} is the generalized mean, and ϵ is set to 10^{-6} . $p_x \in [0, 1]$ is the single-valued output of the network at pixel coordinate x . The last activation of the network can be bounded between zero and one by using a sigmoid nonlinearity.

WHD consists of two terms each of that is crucial for training model:

- In the first term, there is multiplier p_x to penalize high activations in areas of the image where there is no ground truth point y nearby. In other words, the loss function penalizes estimated points that should not be there.

- In the second term there is an equation $p_x d(x, y) + (1 - p_x) d_{max}$. It penalizes low activations around ground truth points.

If the first term is removed, then the trivial solution is $p_x = 1 \forall x \in \Omega$. If the second term is removed, then the trivial solution is $p_x = 0 \forall x \in \Omega$.

A naive alternative is to use a one-hot map as label, defined as $l_x = 1$ for $x \in Y$ and $l_x = 0$ otherwise, and then use a pixelwise loss such as the Mean Squared Error (MSE) or the L^2 norm. The issue with pixelwise losses is that they are not informative of how close two points $x \in \Omega$ and $y \in Y$ are unless $x = y$. In other words, it is flat for the vast majority of the pixels, making training unfeasible.

Neural Network Architecture

In our project we used TerausNet, which is U-Net-based architecture with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation [2]. Its scheme is shown in the Figure 5.

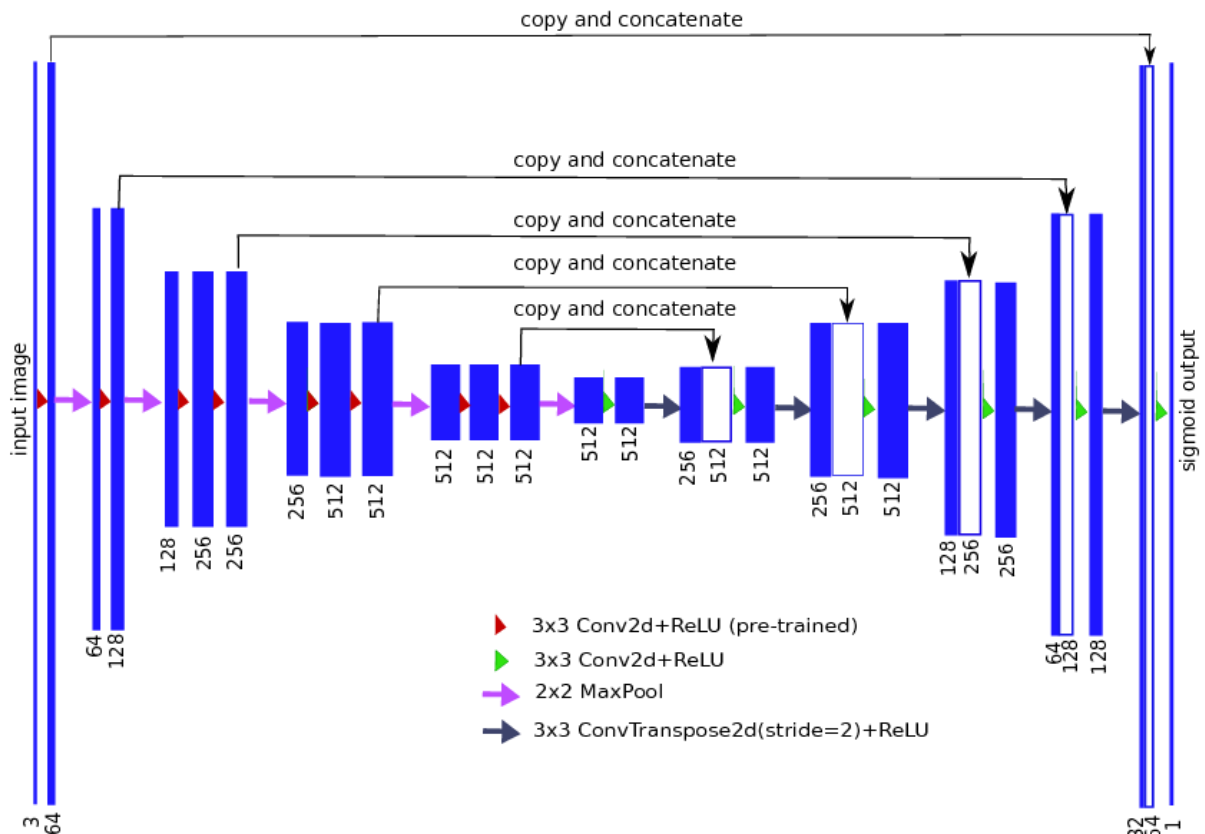


Figure 5. TerausNet network architecture

In general, a U-Net architecture consists of a contracting path to capture context and of a symmetrically expanding path that enables precise localization [3].

The contracting path follows the typical architecture of a convolutional network with alternating convolution and pooling operations and progressively downsamples feature maps, increasing the number of feature maps per layer at the same time. Every step in the expansive path consists of an upsampling of the feature map followed by a convolution. Hence, the expansive branch increases the resolution of the output. In order to localize, upsampled features, the expansive path combines them with high-resolution features from the contracting path via skip-connections. The output of the model is a pixel-by-pixel mask that shows the class of each pixel (as we have binary problem and use WHD we apply sigmoid function after the final convolutional layer). This architecture proved itself very useful for segmentation problems with limited amounts of data [4].

TerausNet is an encoder-decoder architecture, where VGG11 neural network without fully connected layers was used as an encoder. In Figure 5 each blue rectangular block represents a multi-channel features map passing through a series of transformations. The height of the rod shows a relative map size (in pixels), while their widths are proportional to the number of channels (the number is explicitly subscribed to the corresponding rod). The number of channels increases stage by stage on the left part while decrease stage by stage on the right decoding part. The arrows on top show transfer of information from each encoding layer and concatenating it to a corresponding decoding layer.

VGG-Encoder. Typically, neural network initialized with weights from a network pre-trained on a large data set like ImageNet shows better performance than those trained from scratch on a small dataset. In some practical applications, particularly in medicine and traffic safety, the accuracy of the models is of utmost importance. In the paper [2], it is demonstrated how the U-Net type architecture can be improved by the use of the pretrained encoder.

As an example, they show the application of such approach to Aerial Image Labeling Dataset, that contains aerospace images of several cities with high resolution. Each pixel of the images is labeled as belonging to either "building" or "not-building" classes (Figure 6).

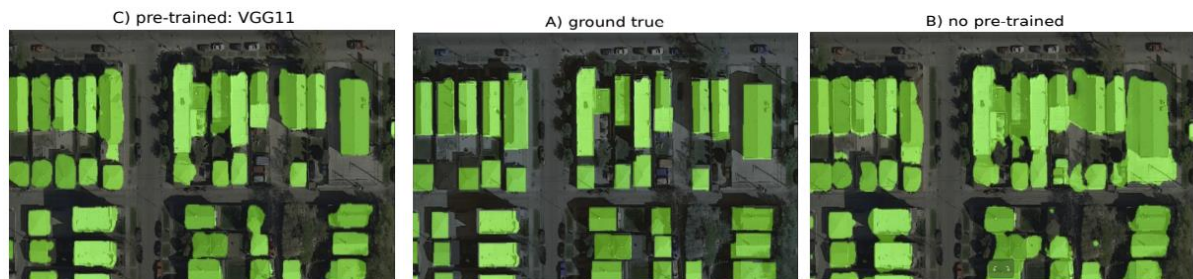


Figure 6. Binary masks with green pixels indicate class membership (buildings)

Binary masks with green pixels indicate class membership (buildings). Image A) shows an original image with the superimposed ground true mask; Images B) to D) show predictions, initialized with different schemas and trained for 100 epochs. Network in image B) had randomly initialized weights. The model in image C) used randomly initialized decoder weights and encoder weights initialized with VGG11, pre-trained on ImageNet.

Optimization

We tested four different optimization techniques:

- Adam – an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments [5].
 - AdamMax – a variant of Adam based on the infinity norm [5].
 - SGD – an algorithm, that implements stochastic gradient descent (optionally with momentum) [6].
 - Adadelata – the method dynamically adapts over time using only first order information and has minimal computational overhead beyond vanilla stochastic gradient descent [7].

In Figure 7 it is shown that usually Adadelata lead to faster convergence of the training process in comparison with other techniques.

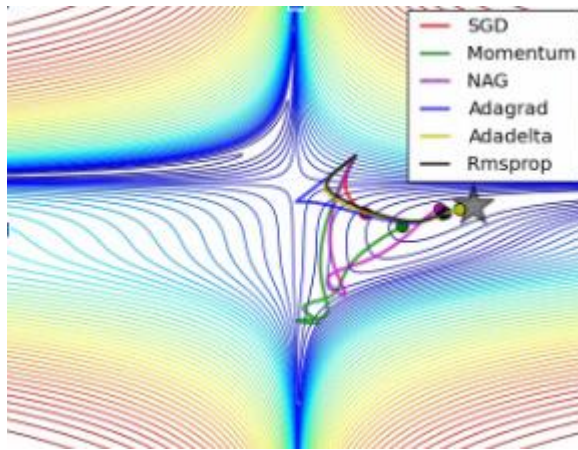


Figure 7. Comparison of optimizers

In the project, Adadelta showed the most reliable results. Training converged with different values of learning rate when using of Adadelta, whereas other optimizers required precise selection of learning rate and momentum values as the training diverged in more than a half of experiments.

4 Development of a module for calculating the relative coordinates of an object in an image

The model output looks like probability map represented in Figure 8.

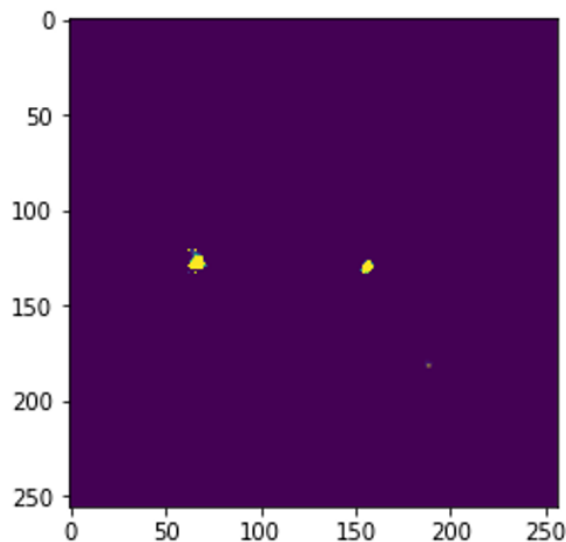


Figure 8. Probability map

Figure 9 illustrates the stages of post processing and the result of coordinate calculation that are drawn on the original image.

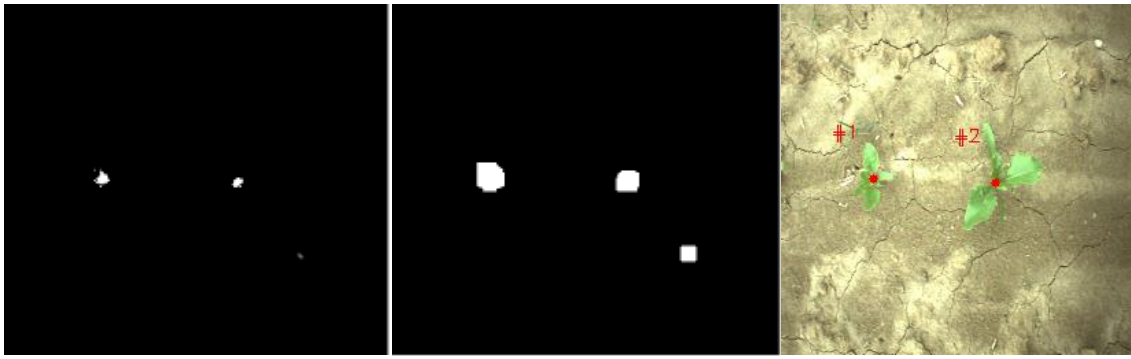


Figure 9. Data post processing

- 1) To start detecting activated regions in an image, we first smooth (i.e., blurring) it to reduce high frequency noise.
- 2) To reveal the brightest regions in the blurred image we apply binary thresholding.
- 3) However, there is a noise in this image (i.e., small blobs), so we perform dilations in order to join regions.
- 4) In order to filter remaining noise we perform a connected-component analysis (region extraction). The code performs the actual connected-component analysis using the scikit-image library.
- 5) Then we detect the contours in the mask image and then sort them from left-to-right. Once the contours have been sorted and then we loop over them individually.
- 6) For each of these contours we compute the bounding box and calculate its center. The center coordinates of the bounding boxes we define as target coordinates.
- 7) The last step is correction of coordinates, as previous processing was conducted with resized image.

5 Evaluation metrics

In evaluation of models, we used standard definitions of popular metrics – precision, recall and F-score.

The calculation of the metrics was the following:

1) For every *ground truth* point we find the *closest prediction* using k-Nearest-Neighbors algorithm (with k=1).

2) If the distance between *ground truth* point and its *closest prediction* is less than pre-defined parameter *r*, number of True Positives (TP) increases by one (Figure 10).

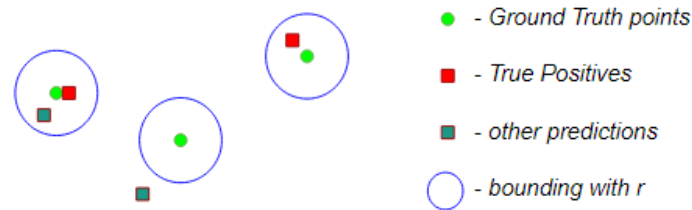


Figure 10. True Positives (TP) identifying

3) Number of False Negatives (FN) equals total number of ground truth points minus TP.

4) For every *prediction* we find the *closest ground truth* point using k-Nearest-Neighbors algorithm (with k=1).

5) If the distance between *prediction* point and its closest *ground truth* is less than pre-defined parameter *r*, number of False Positives (FP) decreases by one (initialized with total number of predictions). The result is illustrated in Figure 11.



Figure 11. False Positives (FP) identifying

6) Then metrics are calculated with found TP, FP, FN by the formulas:

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}$$

Conclusion

During the research the algorithm of object localization in the image has been developed. It includes two structural modules – image segmentation module and module for coordinate calculation. Image segmentation is provided by neural network TernaNet which has Unet-like architecture, applying Weighted Hausdorff Distance as the loss function due to its suitability to the keypoint detection problem. The coordinate calculation module uses computer vision techniques in order to handle the probability map produced by the segmentation model and obtain the coordinates of objects.

REFERENCES

1. Ribera, J., Guera, D., Chen, Y., & Delp, E. J. (2019). Locating Objects Without Bounding Boxes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6479-6489).
2. Iglovikov, V., & Shvets, A. (2018). Ternaunet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746.
3. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation //International Conference on Medical image computing. – Springer, Cham, 2015. – C. 234-241.
4. Shvets, Alexey & Iglovikov, Vladimir & Rakhlin, Alexander & Kalinin, Alexandr. (2018). Angiodysplasia Detection and Localization Using Deep Convolutional Neural Networks.
5. Kingma D. P., Ba J. Adam: A method for stochastic optimization //arXiv preprint arXiv:1412.6980. – 2014.
6. Ning Qian. On the momentum term in gradient descent learning algorithms. Neural networks: the official journal of the International Neural Network Society, 12(1):145–151, 1999
7. Zeiler M. D. Adadelata: an adaptive learning rate method //arXiv preprint arXiv:1212.5701. – 2012.