



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

Школа: Инженерная школа информационных технологий и робототехники (ИШИТР)

Направление подготовки: 09.03.04 «Программная инженерия»

ООП/ОПОП: Разработка программно-информационных систем

Отделение школы: Отделение информационных технологий

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема работы
Программная реализация и анализ эффективности моделей сверточных нейронных сетей YOLO для решения задач классификации объектов на изображениях

УДК 004.032.26:004.932.72`1

Обучающийся

Группа	ФИО	Подпись	Дата
8К92	Беляев Семен Игоревич		

Руководитель ВКР

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

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОСГН ШБИП	Гасанов Магеррам Али оглы	д.э.н., профессор		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ООД	Мезенцева Ирина Леонидовна	-		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП/ОПОП, должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Чердынцев Евгений Сергеевич	к.т.н.		

**ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ООП/ОПОП
«Разработка программно-информационных систем»**

Код компетенции	Наименование компетенции
Универсальные компетенции	
УК(У)-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач
УК(У)-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений
УК(У)-3	Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде
УК(У)-4	Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(-ых) языке(-ах)
УК(У)-5	Способен воспринимать межкультурное разнообразие общества в социально-историческом, этическом и философском контекстах
УК(У)-6	Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни
УК(У)-7	Способен поддерживать должный уровень физической подготовленности для обеспечения полноценной социальной и профессиональной деятельности
УК(У)-8	Способен создавать и поддерживать безопасные условия жизнедеятельности, в том числе при возникновении чрезвычайных ситуаций
УК(У)-9	Способен проявлять предприимчивость в практической деятельности, в т.ч. в рамках разработки коммерчески перспективного продукта на основе научно-технической идеи
Общепрофессиональные компетенции	
ОПК(У)-1	Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности
ОПК(У)-2	Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности
ОПК(У)-3	Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности
ОПК(У)-4	Способен участвовать в разработке стандартов, норм и правил, а также технической документации, связанной с профессиональной деятельностью
ОПК(У)-5	Способен устанавливать программное и аппаратное обеспечение для информационных и автоматизированных систем
ОПК(У)-6	Способен разрабатывать алгоритмы и программы, пригодные для практического использования, применять основы информатики и программирования к проектированию, конструированию и тестированию программных продуктов

ОПК(У)-7	Способен применять в практической деятельности основные концепции, принципы, теории и факты, связанные с информатикой
ОПК(У)-8	Способен осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий
Профессиональные компетенции	
ПК(У)-1	Способен выполнять интеграцию программных модулей и компонент
ПК(У)-2	Владение навыками моделирования, анализа и использования формальных методов конструирования программного обеспечения
ПК(У)-3	Способен создавать техническую документацию на продукцию в сфере информационных технологий, управлять технической информацией
ПК(У)-4	Владение навыками использования операционных систем, сетевых технологий, средств разработки программного интерфейса, применения языков и методов формальных спецификаций, систем управления базами данных
ПК(У)-5	Способен проводить, оценивать и следить за выполнением концептуального, функционального и логического проектирования систем малого и среднего масштаба и сложности

Школа: Инженерная школа информационных технологий и робототехники (ИШИТР)
 Направление подготовки: 09.03.04 «Программная инженерия»
 Отделение школы: Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП/ОПОП
 _____ Чердынцев Е.С.
 (Подпись) (Дата) (ФИО)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

Обучающийся:

Группа	ФИО
8К92	Беляев Семен Игоревич

Тема работы:

Программная реализация и анализ эффективности моделей сверточных нейронных сетей YOLO для решения задач классификации объектов на изображениях	
Утверждена приказами директора ИШИТР	Приказ № 102-28/с от 12.04.2023

Срок сдачи обучающимся выполненной работы:	17.06.2023
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ

Исходные данные к работе	<p>Для детектирования (обнаружения, локализации и классификации) летающих объектов объектов трех классов на изображениях необходимо выбрать компактные модели сверточных нейронных сетей (СНС) семейства YOLO. Следует разработать датасет на основе набора изображений объектов классов: «Птица», «БПЛА вертолетного типа» и «БПЛА самолетного типа». Далее следует провести программную реализацию и обучение выбранных компактных моделей СНС на этом датасете. Исследование эффективности обученных моделей СНС необходимо провести по точности и скорости детектирования объектов каждого класса на изображениях с использованием тестовой выборки датасета.</p>
Перечень разделов пояснительной записки подлежащих исследованию,	1. Обзор проблемы детектирования объектов на изображениях с помощью

проектированию и разработке	мобильных СКЗ; 2. Выбор компактных моделей СНС семейства YOLO; 3. Подготовка датасета; 4. Программная реализация моделей СНС и их обучение с использованием датасета; 5. Исследование эффективности компактных моделей СНС семейства YOLO с использованием датасета; 6. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение; 7. Социальная ответственность.
Перечень графического материала	1. Графики, полученные по результатам обучения и исследования компактных моделей СНС 2. Матрица SWOT-анализа; 3. Диаграмма Ганта.

Консультанты по разделам выпускной квалификационной работы

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Гасанов Магеррам Али оглы, д.э.н., профессор
Социальная ответственность	Мезенцева Ирина Леонидовна, старший преподаватель

Названия разделов, которые должны быть написаны на русском и иностранном языках:

--	--

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	23.01.2023
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	д.т.н, профессор		23.01.2023

Задание принял к исполнению обучающийся:

Группа	ФИО	Подпись	Дата
8К92	Беляев Семен Игоревич		23.01.2023

Школа: Инженерная школа информационных технологий и робототехники (ИШИТР)
 Направление подготовки: 09.03.04 «Программная инженерия»
 Уровень образования: бакалавр
 Отделение школы: Отделение информационных технологий
 Период выполнения: весенний семестр 2022/2023 учебного года)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Обучающийся:

Группа	ФИО
8К92	Беляев Семен Игоревич

Тема работы:

Программная реализация и анализ эффективности моделей сверточных нейронных сетей YOLO для решения задач классификации объектов на изображениях

Срок сдачи обучающимся выполненной работы:	17.06.2023
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
14.04.2023	Разделы 1.,2. Задачи детектирования объектов на изображениях с помощью СНС. Задача исследований	20
25.04.2023	Раздел 3. Выбор компактных моделей СНС семейства YOLO	20
01.06.2023	Разделы 4, 5. Подготовка датасета, программная реализация компактных моделей СНС и их обучение	20
14.06.2023	Раздел 6. Результаты исследования компактных моделей СНС семейства YOLO	20
29.05.2023	Разделы 7, 8. Финансовый менеджмент, ресурсоэффективность; социальная ответственность	20

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Марков Николай Григорьевич	д.т.н, профессор		23.01.2023

СОГЛАСОВАНО:

Руководитель ООП/ОПОП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Чердынцев Евгений Сергеевич	к.т.н.		13.03.2023

Обучающийся

Группа	ФИО	Подпись	Дата
8К92	Беляев Семен Игоревич		13.03.2023

РЕФЕРАТ

Выпускная квалификационная работа выполнена на 99 страницах, содержит 24 рисунка, 23 таблицы, 46 источников литературы, 1 приложение.

Ключевые слова: **МОБИЛЬНАЯ СИСТЕМА КОМПЬЮТЕРНОГО ЗРЕНИЯ, ДЕТЕКТИРОВАНИЕ ЛЕТАЮЩИХ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ, СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ КЛАССА YOLO, ДАТАСЕТ, МАШИННОЕ ОБУЧЕНИЕ.**

Объектом исследования являются оптические изображения летающих объектов трех классов: «Птица», «БПЛА самолётного типа» и «БПЛА вертолётного типа».

Цель ВКР - программная реализация и исследование компактных моделей сверточных нейронных сетей (СНС) класса (семейства) YOLO для детектирования летающих объектов трех классов на изображениях и выбор в результате исследований наиболее эффективных моделей СНС для создания мобильных систем компьютерного зрения (СКЗ) реального времени. Под компактными моделями СНС понимаются модели класса YOLO, имеющие небольшое число сверточных слоев.

В результате исследования сформирован набор оптических изображений с летающими объектами трех классов, разработаны программы для создания датасета. С их помощью на основе этого набора сформирован датасет. В результате анализа моделей класса YOLO выбраны и обучены наиболее перспективные компактные модели СНС: YOLOv4-Tiny, YOLOv5n, YOLOv7-Tiny и YOLO Nano. Проведены исследования этих моделей СНС, получены новые результаты, позволившие оценить их эффективность. Сформированы рекомендации по использованию этих моделей СНС при создании мобильных СКЗ реального времени.

Область применения: мобильные СКЗ, позволяющие решать задачи детектирования объектов на изображениях в реальном времени.

Содержание

Введение.....	11
1 Задачи детектирования объектов на изображениях с помощью моделей СНС.....	13
1.1 Детектирование летающих объектов на изображениях.....	13
1.2 Однофазные и двухфазные детекторы объектов на изображениях.....	14
2 Постановка задачи исследований.....	19
2.1 Требования к моделям СНС для детектирования летающих объектов.....	19
2.2 Цель работы и решаемые задачи.....	20
3 Выбор компактных моделей семейства YOLO.....	22
3.1 Сравнительный анализ моделей семейства YOLO.....	22
3.2 Подходы к обнаружению мелких объектов на изображениях.....	26
3.3 Анализ компактных моделей СНС.....	28
3.3.1 Модель СНС YOLO Nano.....	29
3.3.2 Архитектура модели СНС YOLOv4-Tiny.....	32
3.3.3 Архитектура модели СНС YOLOv5.....	33
3.3.4 Архитектура модели СНС YOLOv7.....	36
4 Подготовка датасета.....	41
5 Программная реализация компактных моделей СНС и особенности их обучения.....	45
5.1 Общие положения.....	45
5.2 Особенности разработки и обучения моделей СНС семейства YOLO.....	46
6 Результаты обучения и исследования компактных моделей СНС семейства YOLO.....	48
6.1 Результаты обучения компактных моделей СНС.....	48

6.2 Результаты исследования компактных моделей СНС.....	52
7 Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения.....	57
7.1 Потенциальные потребители результатов исследования.....	57
7.2 Анализ конкурентных технических решений.....	57
7.3 Технология QuaD.....	59
7.4 SWOT-анализ.....	60
7.5 Структура работ в рамках научного исследования.....	61
7.6 Определение трудоемкости выполнения работ.....	62
7.7 Разработка графика проведения научного исследования.....	63
7.8 Бюджет научно-технического исследования.....	65
7.8.1 Расчет затрат на специальное оборудование для научных (экспериментальных) целей.....	65
7.8.2 Основная заработная плата исполнителей темы.....	66
7.8.3 Дополнительная заработная плата исполнителей темы.....	68
7.8.4 Отчисления во внебюджетные фонды (страховые отчисления).....	69
7.8.5 Накладные расходы.....	70
7.8.6 Формирование бюджета затрат научно-исследовательского проекта.....	71
7.8.7 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.....	71
7.9 Определение потенциального эффекта исследования.....	74
8 Социальная ответственность.....	77
8.1 Правовые и организационные вопросы обеспечения безопасности.....	78
8.1.1 Правовые нормы трудового законодательства.....	78
8.1.2 Организационные мероприятия при компоновке рабочей зоны.....	78

8.2 Производственная безопасность.....	79
8.2.1 Вредные производственные факторы.....	80
8.2.1.1 Нагрузка на зрительный аппарат.....	80
8.2.1.2 Отсутствие или недостаток необходимого искусственного освещения и пониженная световая и цветовая контрастность.....	81
8.2.1.3 Умственное перенапряжение, в том числе вызванное информационной нагрузкой.....	82
8.2.2 Опасные производственные факторы.....	83
8.3 Экологическая безопасность.....	84
8.4 Безопасность в чрезвычайных ситуациях.....	84
8.5 Вывод по разделу.....	86
Заключение.....	88
Список публикаций.....	90
Список источников.....	91
Приложение А.....	96

Введение

Обнаружение (детектирование) объектов на изображении представляет собой ключевую задачу в области компьютерного зрения и включает одновременное определение местоположения и классификацию объектов на изображении. Положение объекта определяется с помощью ограничивающего прямоугольника (англ. bounding box). Современные достижения в области глубокого обучения позволили сделать значительные успехи в решении многих задач компьютерного зрения, включая обнаружение объектов, благодаря применению глубоких сверточных нейронных сетей (СНС) [1, 2].

Большая часть современных моделей СНС для решения задач мультиклассификации объектов на изображениях сосредоточена на повышении точности детектирования объектов, не уделяя должного внимания вычислительной эффективности алгоритмов детектирования объектов. Такой подход приводит к моделям СНС с высокой вычислительной сложностью и низкой ресурсоэффективностью, которые поэтому сложно применять в случае развертывания таких моделей СНС на мобильных устройствах, например, на беспилотных летательных аппаратах (БПЛА) [3]. Таким образом, к моделям СНС, предназначенным для развертывания на мобильных устройствах предъявляются требования [4]:

1. Проведение мультиклассификации объектов на изображениях в режиме реального времени;
2. Высокая вычислительная эффективность СНС и эффективность по занимаемой памяти (малый объем весовых коэффициентов модели СНС);
3. Конкурентноспособность по сравнению с другими моделями СНС в части точности мультиклассификации объектов на изображениях.

В наибольшей степени предъявляемым требованиям соответствуют модели семейства (класса) YOLO [5]. В моделях этого класса мультиклассификация объектов рассматривается как задача регрессии

прямоугольных рамок и классов объектов. Модели СНС класса YOLO обрабатывают изображение за один проход. На сегодняшний день существует большое разнообразие моделей YOLO с разным количеством параметров, отличающихся друг от друга по точности и скорости обнаружения объектов [6]. Исследование соответствующих моделей СНС из этого класса, отвечающих вышеуказанным требованиям, с целью использования их в тех или иных мобильных системах компьютерного зрения (СКЗ) является актуальной и весьма сложной задачей.

Данная работа посвящена программной реализации и исследованию четырех компактных моделей СНС класса (семейства) YOLO для применения наиболее эффективных из них в мобильных СКЗ, детектирующих летающие объекты на изображениях в реальном времени.

В разделе 1 рассмотрены современные классы моделей СНС для детектирования летающих объектов на изображениях с целью использования их в мобильных СКЗ. В разделе 2 показаны требования к моделям СНС для детектирования летающих объектов и сформулированы цель и задачи исследования. Раздел 3 посвящен выбору компактных моделей СНС семейства YOLO для дальнейших исследований. В разделе 4 описаны особенности созданного датасета. В разделе 5 освещена программная реализация компактных моделей СНС и особенности их обучения. В разделе 6 приведены результаты обучения и исследования компактных моделей СНС. В разделе 7 изложены вопросы, связанные с финансовым менеджментом, ресурсоэффективностью и ресурсосбережением разработки. Раздел 8 содержит вопросы социальной ответственности: правовые и организационные вопросы, техники безопасности работников и экологической безопасности.

Результаты работы докладывались на трех международных научно–практических конференциях студентов, аспирантов и молодых учёных, получен диплом за лучший доклад. Основные результаты работы опубликованы в виде трех статей в сборниках этих конференций.

1 Задачи детектирования объектов на изображениях с помощью моделей СНС

1.1 Детектирование летающих объектов на изображениях

Сегодня мобильные СКЗ востребованы также при решении задач обнаружения и классификации летающих объектов в воздушном пространстве. На изображениях и видео, получаемых при мониторинге воздушного пространства с помощью, соответствующей фото или видеоаппаратуры в составе таких СКЗ, могут появляться один и более летающих объектов, причем в случае нескольких объектов они могут принадлежать к разным классам. Мобильные СКЗ на основе моделей СНС должны обнаруживать, локализовать, сопровождать (осуществлять трекинг при перемещении объектов) и классифицировать летающие объекты. По сути, каждая такая СКЗ с помощью той или иной модели СНС должна решать задачу распознаванию объекта (англ. Object Recognition) или задачу объектного детектирования на изображении. При создании мобильных СКЗ, в принципе, можно использовать модели СНС, позволяющие распознавать один летающий объект на изображении, то есть решать задачи по распознаванию объекта [1,2]. Суть задачи по распознаванию объекта заключается в следующем: модель СНС получает на вход изображение с ровно одним объектом, а на ее выходе появляются координаты содержащего объект прямоугольника и его класс; таким образом, решается задача локализации и классификации одного объекта на изображении. Однако СКЗ с такой моделью СНС может показать высокую точность распознавания только в случае одиночного объекта на анализируемом изображении. Поэтому в случае двух и более объектов на изображении мобильная СКЗ должна решать задачу объектного детектирования (англ. Object Detection), часто говорят детектирования объектов на изображении [5]. Это означает, что в случае нескольких объектов на изображении модель СНС должна обвести прямоугольником и предсказать класс каждого из объектов; эта задача

является одной из фундаментальных задач компьютерного зрения и заключается в одновременном обнаружении, локализации (определении местоположения) и классификации всех объектов на изображении. Местоположение объекта определяется координатами ограничивающего его прямоугольника (англ. bounding box).

Поскольку модели СНС, решающие задачи детектирования нескольких летающих объектов на изображении, могут использоваться и в случае одиночного летающего объекта, тогда имеет смысл применять только их при создании мобильных СКЗ, то есть решать с помощью соответствующих моделей СНС задачи объектного детектирования.

1.2 Однофазные и двухфазные детекторы объектов на изображениях

Рассмотрим современные модели СНС, позволяющие решать задачи объектного детектирования изображений. Одноэтапные (часто говорят, однофазные) и двухэтапные (говорят, двухфазные) детекторы объектов на изображениях являются двумя основными типами детекторов (классификаторов) в современных СКЗ. Детекторы объектов, основанные на глубоком обучении, извлекают признаки из входного изображения или видеокadra в целом [7]. Детектор объектов решает две крупные задачи:

1. Нахождение (обнаружение) произвольного количества объектов на изображении;
2. Классификация каждого объекта и помещение его в ограничительную рамку (англ. bounding box), то есть решается задача его классификации и локализации.

В общем случае можно разделить эти обязанности на две части, чтобы упростить процедуру.

В двухэтапных детекторах производится определение приблизительных областей нахождения объекта прежде чем они будут

использованы для классификации и регрессии ограничивающей рамки для объекта-кандидата. Примерами двухэтапных детекторов могут служить модели СНС R-CNN [8], SPP-NET [9] и Fast-RCNN [10]. В этих детекторах одна модель используется для извлечения областей объектов, а вторая модель используется для классификации и дальнейшего уточнения локализации объекта. Такие методы, как известно, являются относительно медленными, но очень мощными, но недавний прогресс, такой как совместное использование функций, улучшил двухступенчатые детекторы, чтобы иметь почти такую же вычислительную стоимость, что и одноступенчатые детекторы.

Исследователи, предложившие модель СНС YOLO [5], одними из первых применили другой подход к решению задачи детектирования: они рассматривали детектирование как задачу регрессии рамок и классов объектов. СНС класса YOLO, как и другие СНС с одноступенчатой архитектурой, обрабатывают изображение в исходном масштабе за один проход. В моделях СНС YOLO используются подсети для разных этапов обработки и детектирования объектов. На рисунке 1 схематично представлена работа однофазного детектора, использованного в СНС семейства YOLO. Кратко опишем эти подсети.

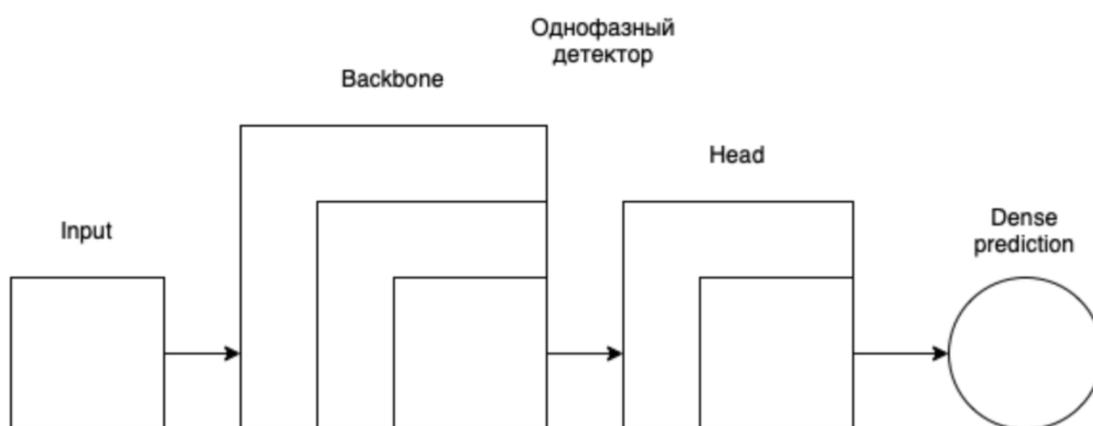


Рисунок 1 – Схема однофазного детектора в сетях семейства YOLO

1. Основная сверточная подсеть (Backbone): эта подсеть отвечает за

извлечение признаков из изображения. Обычно основная подсеть представляет собой СНС (англ. CNN) с несколькими сверточными и max-pooling слоями, которая выделяет высокоуровневые и низкоуровневые признаки изображения. Оригинальная версия YOLO использует Darknet-19 [11], а новые версии, такие как YOLOv3, используют Darknet-53 [11].

2. Подсеть для обработки признаков (Neck): в двухфазном классификаторе YOLO neck - часть отвечает за объединение и обработку признаков, извлеченных из основной сверточной подсети (backbone). В контексте YOLO, особенно начиная с модели YOLOv3, head - фаза представляет собой модуль Feature Pyramid Network (FPN) [12]. FPN состоит из последовательности сверточных слоев и предназначен для объединения признаков с разными уровнями масштаба и детализации. Результатом является набор многоуровневых признаков, которые могут быть использованы для обнаружения объектов различных размеров и масштабов. Используя признаки, полученные от backbone и FPN, "head" фаза передает их далее на подсети предсказания.

3. Генерация предсказаний (Dense Prediction): Dense prediction фаза в классификаторе YOLO отвечает за генерацию предсказаний ограничительных рамок и классификации объектов на основе признаков полученных после обработки через основную сверточную подсеть (backbone) и "шею" (neck stage, например, Feature Pyramid Network).

На этапе dense prediction СНС YOLO выполняет следующие задачи:

1. Предсказывает ограничительные рамки (Bounding Box): генерирует координаты (x, y) , высоту и ширину рамки - для предсказанных ограничивающих рамок объектов. Эти данные предоставляют информацию о положении и размере объектов на изображении.

2. Предсказывает уровень уверенности (Objectness score): Создает оценку уверенности для каждой ограничительной рамки, чтобы определить, насколько вероятно, что рамка содержит объект. Эта оценка учитывает вероятность присутствия объекта и точность предсказания рамки

3. Классификация объектов: на основе входных признаков, dense prediction фаза генерирует вероятности принадлежности каждой рамки к различным классам объектов. Это позволяет определить, к какому классу принадлежат объекты внутри ограничительных рамок.

Dense prediction - фаза в СНС YOLO имеет плотную структуру предсказаний, которая позволяет эффективно и быстро обрабатывать изображения для детектирования и классификации объектов. Эта особенность YOLO обеспечивает высокую скорость и точность предсказания по сравнению с другими методами обнаружения объектов, такими как R-CNN и Faster R-CNN [13].

В двухфазных детекторах добавляется еще один шаг «sparse prediction». На рисунке 2 представлено схематическое изображение такого двухфазного детектора.

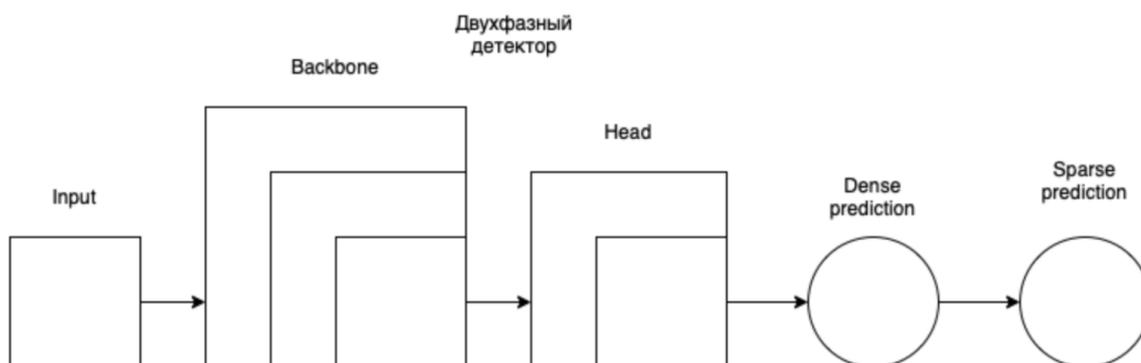


Рисунок 2 – Схематическое изображение двухфазного детектора

Sparse prediction в двухфазном классификаторе относится к подходу, при котором на первой стадии меньшее количество областей (anchors) на изображении анализируется на предмет наличия релевантных объектов. За счет этого достигается экономия вычислительных ресурсов и увеличение скорости обработки изображений. Этот подход является противоположностью dense prediction, которое анализирует большое количество областей, или пропускает промежуточные стадии, вроде region proposal в классификаторе YOLO.

Двухфазовые классификаторы, такие как СНС R-CNN и ее вариации, используют sparse prediction подход. На первой стадии (region proposal) такие алгоритмы генерируют предложения областей с использованием, например, методов поиска селективных связей или предсказания областей интереса (Region Proposal Network). Эти области являются предполагаемыми местами нахождения объектов на изображении.

На второй стадии (классификация и оценка положения) каждая предложенная область передается на обработку детектирующему и классифицирующему модулям, которые предсказывают и обновляют ограничительные рамки, вероятности объекта и классификацию.

Проведенный анализ различных классов моделей СНС позволяет вслед за авторами [5,12] считать, что наиболее приемлемыми для детектирования объектов на изображениях в мобильных СКЗ являются модели класса YOLO. Модели этого класса, как и другие однофазные модели СНС, обрабатывают изображение за один проход, что значительно повышает скорость анализа изображения. Ряд авторов исследований эффективности современных моделей СНС с учетом требований к точности и скорости детектирования объектов на изображениях также отдает предпочтение моделям СНС класса (семейства) YOLO [11,12].

2 Постановка задачи исследований

2.1 Требования к моделям СНС для детектирования летающих объектов

Вычислительные возможности мобильных СКЗ обычно весьма ограничены, что накладывает жёсткие требования на скорость вычисления моделей СНС. Это указывает на актуальность разработки и исследования новых моделей СНС и выбора из них тех, что отвечают требованиям по скорости вычислений и точности распознавания объектов на изображениях, выдвигаемым при проектировании конкретных мобильных СКЗ. Этим взаимоисключающим требованиям особенно непросто удовлетворить в случаях создания мобильных СКЗ, функционирующих в режиме реального времени. Таким взаимоисключающим требованиям должна также удовлетворять мобильная СКЗ, предназначенная для детектирования (обнаружения, локализации и классификации) летающих объектов в воздушном пространстве.

Пусть на основе модели СНС из класса YOLO проектируется мобильная СКЗ, которая должна позволять детектировать на изображениях летающие объекты следующих трёх классов:

- Птица;
- БПЛА самолётного типа;
- БПЛА вертолётного типа;

Обычно используются количественные характеристики вышеуказанных требований к мобильной СКЗ и, соответственно, к лежащей в ее основе той или иной модели СНС. Точность детектирования (классификации) объектов на изображении вычисляется по общепринятым метрикам Precision и (или) Recall. Наиболее часто в исследованиях используется метрика AP – площадь под кривой Precision - Recall (AUC PR), оцененная с использованием порогового значения IoU. Применяется также метрика mAP – среднее значение AP по всем классам объектов. Для $AP_{0,5}$ и $mAP_{0,5}$ значение IoU равно 0,5, в то время как для расчета $AP_{0,5;0,95}$ и

$mAP_{0,5;0,95}$ используется интервал значений IoU [0,5;0,95] с шагом 0,05. Пороговые значения метрик $AP_{0,5}$ и $mAP_{0,5}$ для оптических изображений должны быть 0,7. Требование по точности детектирования объектов на изображениях удовлетворяется, если превышено пороговое значение каждой из этих метрик.

Второй количественной характеристикой мобильной СКЗ является скорость детектирования объектов на изображении (скорость вычислений модели СНС), измеряемая как число анализируемых изображений в секунду в FPS (англ. Frames Per Second). Значение FPS, позволяющее считать, что мобильная СКЗ удовлетворяет требованию по скорости детектирования объектов с помощью модели СНС, должно быть не менее 25.

Третья количественная характеристика – алгоритмическая эффективность модели СНС, определяемая как количество ресурсов вычислительного устройства, необходимое для достижения определенных показателей модели СНС в решении поставленной задачи. Она рассчитывается на основе вычислительной сложности и размера модели СНС. Вычислительная сложность модели СНС, в FLOPs – это количество операций умножения, сложения и сравнения над 16-битными числами с плавающей точкой. Размер модели (компактность архитектуры) СНС, в МБ – это объем памяти вычислительного устройства СКЗ, необходимый для хранения весовых коэффициентов модели СНС. Пороговые значения вычислительной сложности и размера модели СНС, позволяющие считать, что мобильная СКЗ (и, соответственно, модель СНС в ее основе) удовлетворяет требованиям разработчиков, задаются в техническом задании на создание мобильной СКЗ.

2.2 Цель работы и решаемые задачи

Целью ВКР является программная реализация и исследование компактных моделей СНС класса YOLO для детектирования летающих

объектов трех классов на изображениях и выбор в результате исследований наиболее эффективных моделей СНС для создания мобильных СКЗ реального времени. Компактными моделями СНС здесь и далее будем считать модели класса YOLO, имеющие небольшое число сверточных слоев.

Для реализации поставленной цели необходимо решить следующие задачи:

1. Провести предварительный анализ существующих компактных моделей СНС семейства YOLO и выбрать из них наиболее подходящие для исследования;

2. Сформировать набор оптических изображений летающих объектов трех классов: «Птица», «БПЛА вертолетного типа» и «БПЛА самолетного типа».

3. Разметить и аннотировать объекты указанных классов на изображениях этого набора и сформировать в итоге датасет;

4. Провести программную реализацию выбранных компактных моделей YOLO и провести их обучение и тестирование с использованием обучающих и тестовых выборок из датасета;

5. Исследовать эффективность обученных моделей СНС с использованием тестовой выборки из датасета;

6. Провести анализ всех результатов исследований моделей СНС на предмет удовлетворения сформированным требованиям к мобильной СКЗ и, соответственно, к модели СНС; рекомендовать наиболее эффективную модель к использованию в мобильной СКЗ.

3 Выбор компактных моделей семейства YOLO

3.1 Сравнительный анализ моделей семейства YOLO

Предварительный проведенный нами анализ показывает, что часть моделей семейства YOLO имеет сложную архитектуру и поэтому превышает пороговые значения по вычислительной сложности и по размеру занимаемой моделью СНС памяти. При этом не подходят для мобильных СКЗ первые модели этого класса: YOLOv2, YOLOv3 и т.п. модели. Иными словами, для мобильных СКЗ необходимы модели СНС с более компактными архитектурами, в первую очередь, с небольшим числом сверточных слоев.

В связи с этим было предложено использовать модели СНС с компактными архитектурами семейства YOLO, известные как архитектуры СНС Tiny YOLO [14]. Такие модели, называемые далее компактными, обычно имеют небольшое число сверточных слоев и поэтому предъявляют значительно меньшие требования к памяти и производительности вычислительного устройства мобильной СКЗ. Однако точность детектирования (классификации) объектов на изображениях с помощью таких моделей ниже, чем у соответствующих им полноразмерных моделей СНС. Одной из первых компактных архитектур СНС для детектирования объектов является также архитектура СНС YOLO Nano [15]. На данный момент практически все модели YOLO имеют свою уменьшенную версию.

В моделях СНС семейства YOLO для агрегации карт признаков применяются такие СНС как FPN [16], PANet [17], Bi-FPN [18], SFAM [19], ASFF [20], NAS-FPN [21] и т.д. Также в подсети для агрегации признаков могут использоваться дополнительные модули, например SPP [22], ASPP [23], RFB [24], SAM [25] и др. Все эти механизмы пропусков признаков и подсети нацелены на повышение точности детектирования объектов на изображениях. На данный момент наивысшую точность детектирования на датасете MS COCO [26] показывают модели СНС подкласса YOLOv7. Однако, эти модели выставляют высокие требования к вычислительным

ресурсам при реализации их в мобильной СКЗ.

Кратко рассмотрим основные параметры современных моделей СНС семейства YOLO. В таблице 1 приведены такие параметры, полученные с использованием известных датасетов [27]. Как видно из нее, наименьшее число весовых коэффициентов и, соответственно, вес модели имеют СНС YOLOv4-Tiny, YOLOv5n, YOLOv7-Tiny, так как они имеют облегченную архитектуру (небольшое число сверточных слоев). Также одной из перспективных моделей для применения в мобильных СКЗ является YOLO Nano, так как согласно исследованиям она обходит СНС YOLOv4-Tiny [28] по объему требуемых ресурсов памяти и скорости вычислений. Однако для сравнения их по точности детектирования объектов необходима проверка на одном датасете.

Таблица 1 – Основные параметры современных моделей СНС класса YOLO

	Размер входного изображения	Число весовых коэффициентов модели (млн.)	GFLOPS	mAP _{0,5} (на валидационной выборке - val)
YOLOv4	640	64,4	142,8	68,2
YOLOv4-CSP	640	52,9	120,4	68,6
YOLOv4-Tiny	416	6,1	6,9	42,1
YOLOv5n	640	1,9	4,5	45,7
YOLOv5s	640	7,2	16,5	56,8
YOLOv5m	640	21,2	49,0	64,1
YOLOv5l	640	46,5	109,1	67,3
YOLOv5x	640	86,7	205,7	86,7
YOLOv7	640	36,9	104,7	69,7
YOLOv7-X	640	71,3	189,9	71,1
YOLOv7-Tiny	320	6,2	3,5	47,3
YOLOv7-E6	1280	97,2	515,2	73,5
YOLOv7-D6	1280	154,7	806,8	73,8
YOLOv7-E6E	1280	151,7	843,2	74,4

На рисунке 3 представлен график значений AP (англ. average precision), полученный на валидационной выборке датасета COCO [26] для различных моделей СНС четвертой, пятой и седьмой версии из семейства YOLO [27]. Как видно из него, наибольшего значения метрики AP достигают модели СНС версии четыре, однако при этом у них низкое значение характеристики FPS при мультиклассификации (модели не смогут обеспечить режим реального времени при функционировании мобильной СКЗ). Более того, они имеют большое число параметров, обусловленное их реализацией в фреймворке Darknet. Модели СНС YOLOv7 имеют самое большое значение FPS, при этом облегченная версия СНС YOLOv7-Tiny опережает YOLOv5n и по числу анализируемых кадров в секунду, и по значению APval.

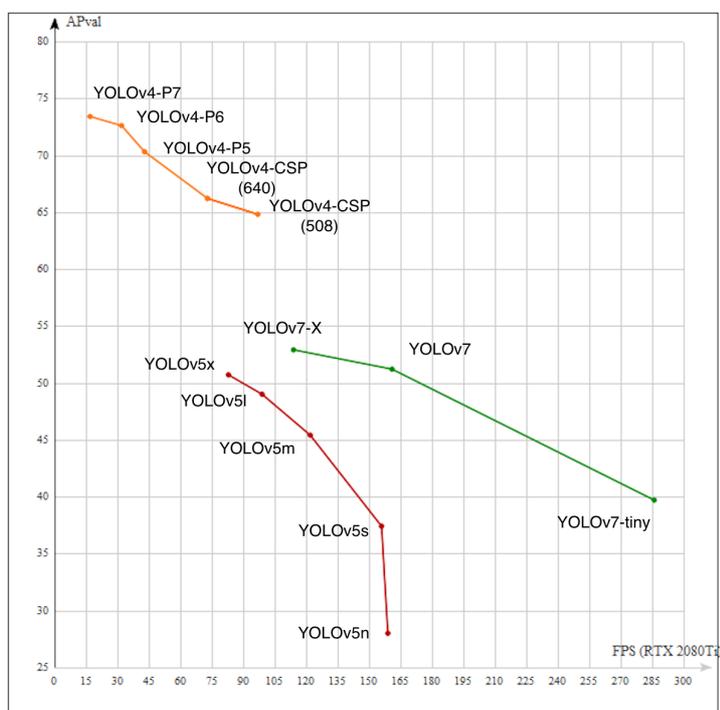


Рисунок 3 – Параметры APval в зависимости от FPS для моделей семейства YOLO

На рисунках 4, 5, 6 представлены графики зависимости точности классификации (детектирования) объектов, соответственно малого, среднего и большого размеров от FPS на тестовых (англ. test) выборках датасета COCO [26] для СНС YOLO версий (подклассов) четыре и семь [27]. На всех трех графиках сохраняется тенденция к превосходству моделей СНС YOLOv7.

Также стоит отметить обратно пропорциональную зависимость точности классификации (детектирования) от количества кадров в секунду FPS. При этом точность классификации мелких по размерам объектов на 30%-50% хуже точности классификации крупных (крупномасштабных) объектов [27].

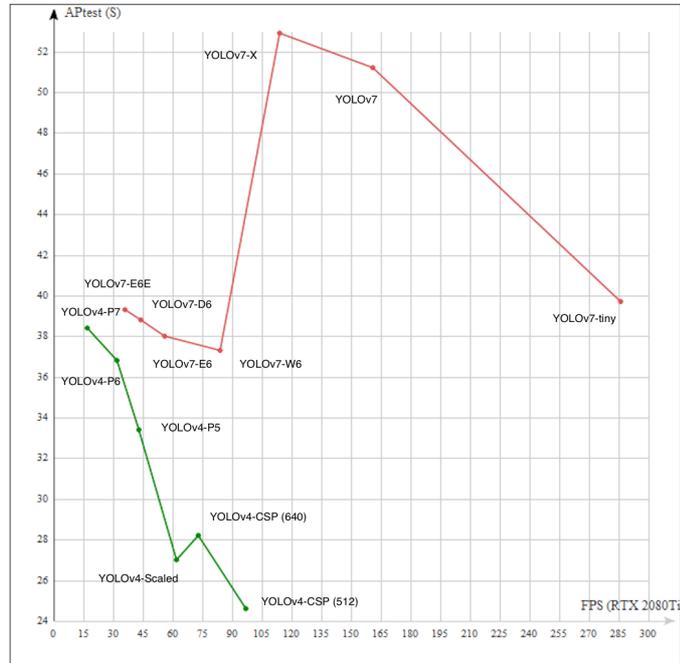


Рисунок 4 – Параметры APtest в зависимости от FPS для моделей YOLOv4 и YOLOv7 в случае мелких объектов

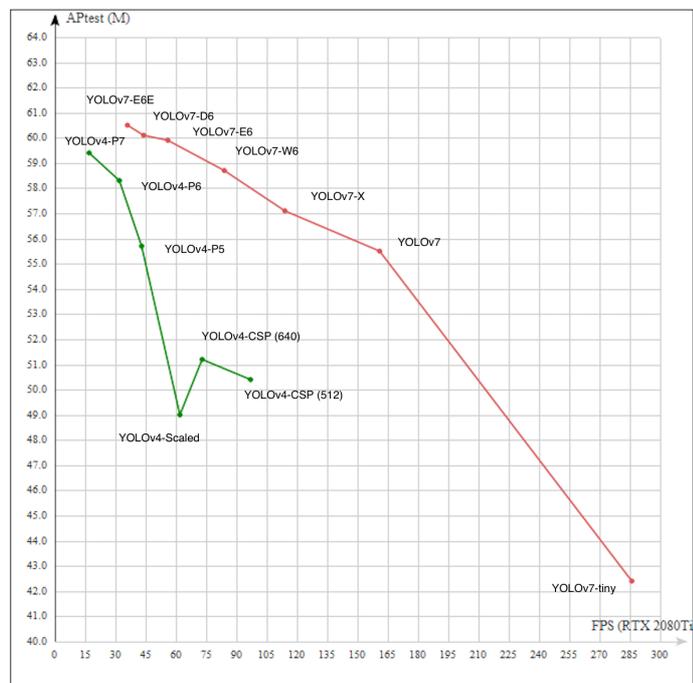


Рисунок 5 – Параметры APtest в зависимости от FPS для моделей YOLOv4 и YOLOv7 в случае средних объектов

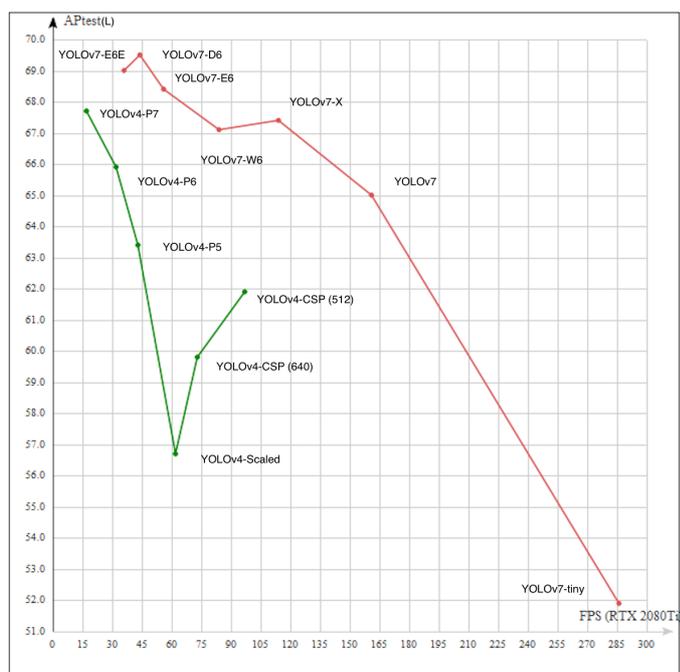


Рисунок 6 – Параметры APtest в зависимости от FPS для моделей YOLOv4 и YOLOv7 в случае крупных объектов

Учитывая все изложенное, можно считать, что модели YOLOv4-Tiny, YOLOv5n, YOLOv7-Tiny и YOLO Nano потенциально перспективны к использованию в мобильных СКЗ и подлежат дальнейшему исследованию в этой работе. Далее именно эти модели СНС из-за небольшого числа сверточных слоев именуется как компактные.

3.2 Подходы к обнаружению мелких объектов на изображениях

Существует несколько известных подходов к увеличению точности детектирования мелких объектов на изображениях [29].

1. Увеличение разрешения изображений. Очень маленькие объекты могут содержать всего несколько пикселей в ограничительной рамке – это означает, что очень важно увеличить разрешение ваших изображений, чтобы увеличить богатство функций, которые ваш детектор может формировать из этой маленькой рамки. В связи с этим рекомендуется делать снимки с максимально высоким разрешением, если это возможно.

2. Синтезирование искусственных данных [30]. Аугментация

создает новые изображения из базового набора данных. Это может быть очень полезно, чтобы предотвратить перенасыщение модели СНС обучающим (тренировочным) набором. Некоторые особенно полезные дополнения для обнаружения мелких объектов включают случайную обрезку, случайное вращение и мозаичную аугментацию.

3. Разбиение изображений на несколько отдельных. Для увеличения числа объектов в исходной обучающей выборке можно разбивать изображения на части в качестве шага предварительной обработки. Мозаика эффективно приближает ваш детектор к небольшим объектам, но позволяет сохранить небольшое входное разрешение, необходимое для быстрого логического вывода.

4. Использование трансферного обучения [31]. Использование предварительно обученную модель СНС (например, обученную на датасете COCO) в качестве отправной точки для конкретной задачи обнаружения объектов может улучшить точность детектирования объектов. Путем точной настройки весов можно использовать имеющиеся знания для улучшения обнаружения мелких объектов.

5. Модификация блоков привязки [32]. Оптимизация размеров блоков привязки для более точного представления небольших объектов может улучшить результаты обнаружения. Можно использовать методологию кластеризации (например, кластеризацию k-средних), чтобы определить идеальные размеры ячеек привязки. В сети YOLOv5 и более новых вариациях такая настройка происходит автоматически.

6. Агрегация карт признаков и использование пирамид функций (FPN) [16]: использование сети пирамид функций (FPN) с YOLO для создания многомасштабных карт функций поможет идентифицировать небольшие объекты на более глубоких уровнях модели.

7. Модификация функции потерь и функций активации

В таблице 2 [29] представлены перечисленные выше подходы и конкретные технологии и метрики, необходимые для применения при

решении наших задач детектирования летающих объектов.

Таблица 2 – Методы повышения точности детектирования объектов

Функции потерь	MSE, IoU, GIoU, CIoU, DIoU
Аугментация данных	-случайные изменения яркости, контраста, оттенка, насыщенности, добавление шума (photometric distortions). Случайные повороты, перевороты, обрезка, масштабирование (geometric distortions) -CutOut, MixUp, CutMix, Mosaic. Модификация разрешения изображений. Модификация якорей привязки.
Агрегация карт признаков	FPN, PANet, Bi-FPN, SFAM, ASFF, NAS-FPN
Функции активации	ReLU, LReLU, PReLU, SELU, ReLU6, Swish, hard-Swish, Mish, LeakyReLU
Пробросы признаков	Residual connections, Weighted residual connections, Multi-input weighted residual connections, Cross stage partial connections (CSP)
Использование трансферного обучения	Весовые коэффициенты моделей СНС, обученных на датасетах COCO и ImageNet

Основные известные подходы к обнаружению мелких объектов на изображениях и ряд перечисленных технологий из этой таблицы будут использованы нами при создании датасета и при разработке и реализации выбранных четырех компактных моделей СНС. Правильное применение ряда описанных технологий должно повысить точность детектирования летающих объектов малых размеров на изображениях.

3.3 Анализ компактных моделей СНС

Ниже описаны выбранные для исследования четыре компактные модели СНС, в первую очередь, с точки зрения особенностей проектирования их архитектуры и обучения.

3.3.1 Модель СНС YOLO Nano

СНС YOLO Nano – высококомпактная модель для решения задачи детектирования объектов на изображениях в рамках мобильных СКЗ. СНС YOLO Nano наследует архитектурные принципы СНС класса YOLO.

Архитектура СНС YOLO Nano создана с использованием стратегии совместного проектирования человека и машины (англ. *human-machine collaborative design strategy*), включающей два основных этапа [15]:

- принципиальное прототипирование архитектуры СНС (англ. *principled network design prototyping*);
- машинный поиск архитектуры (англ. *machine-driven design exploration*).

На этапе принципиального прототипирования архитектуры сети создается исходный прототип архитектуры сети (обозначенный как ϕ), основанный на определенных человеком архитектурных принципах, которые далее используются на этапе машинного поиска архитектуры. Исходный прототип архитектуры сети базируется на архитектурных принципах СНС класса YOLO.

Конфигурация отдельных модулей и слоев, а также их количество, определяется на этапе машинного поиска архитектуры с использованием исходного прототипа архитектуры СНС ϕ , данных и определенных человеком требований и ограничений архитектуры, ориентированных на использование во встраиваемых и мобильных СКЗ с ограниченными ресурсами вычислений и памяти.

Машинный поиск архитектуры в работе достигается в форме генеративного синтеза, конечная цель которого заключается в обучении генеративных машин, способных генерировать архитектуры глубоких СНС, удовлетворяющих поставленным требованиям и ограничениям.

Архитектура СНС YOLO Nano, полученная путем машинного поиска на основе генеративного синтеза, характеризуется высокой гетерогенностью

и состоит из ряда модулей представления признаков (англ. feature representation modules) с уникальной микроархитектурой каждого модуля, а также из отдельных слоев свертки 3×3 и 1×1 . Также осуществляются пропуски карт признаков (англ. shortcut connections) (англ. skip connection) между различными уровнями СНС.

Архитектура СНС YOLO Nano включает следующие модули [15]:

1. модуль расширения-проекции (англ. expansion-projection, EP);
2. модули проекции-расширения-проекции (англ. projection-expansion-projection, PEP);
3. модуль полносвязного внимания (англ. fully-connected attention, FCA).

Модуль EP (рисунок 7(b)) состоит из:

1. расширяющего слоя – сверточного слоя 1×1 , проецирующего входные карты признаков в тензор большей размерности путем увеличения количества каналов;
2. слоя глубинной свертки (англ. depthwise convolution) – сверточного слоя 3×3 , производящего пространственную свертку с отдельным ядром для каждого канала входной карты признаков, количество каналов при этом остается неизменным;
3. проецирующего слоя – сверточного слоя 1×1 , проецирующего входные карты признаков в тензор меньшей размерности путем сокращения количества каналов.

Модуль PEP (рисунок 7(c)) состоит из: проецирующего слоя, расширяющего слоя, слоя глубинной свертки, проецирующего слоя.

Слои 2 и 3 модуля EP, как и слои 3 и 4 модуля PEP, в совокупности образуют слой глубинно разделимой свертки (англ. depthwise separable convolution).

Модуль FCA (рисунок 7(d)) состоит из двух последовательных полносвязных слоев. Во всех трех модулях используются остаточные связи – пропуски карт признаков со входа модуля на его выход, где производится

поэлементная операция над входными и выходными картами признаков модуля (поэлементное сложение в случае модулей EP и PEP, и поэлементное умножение в случае модуля FCA).

Общая схема архитектуры СНС YOLO Nano показана на рисунке 7(a) [15]. “x” в записи PEP(x) обозначает количество каналов в первом проекционном слое модуля PEP. “x” в записи FCA(x) обозначает коэффициент сокращения размерности равный x. Далее в скобках указаны размерности выходных карт признаков.

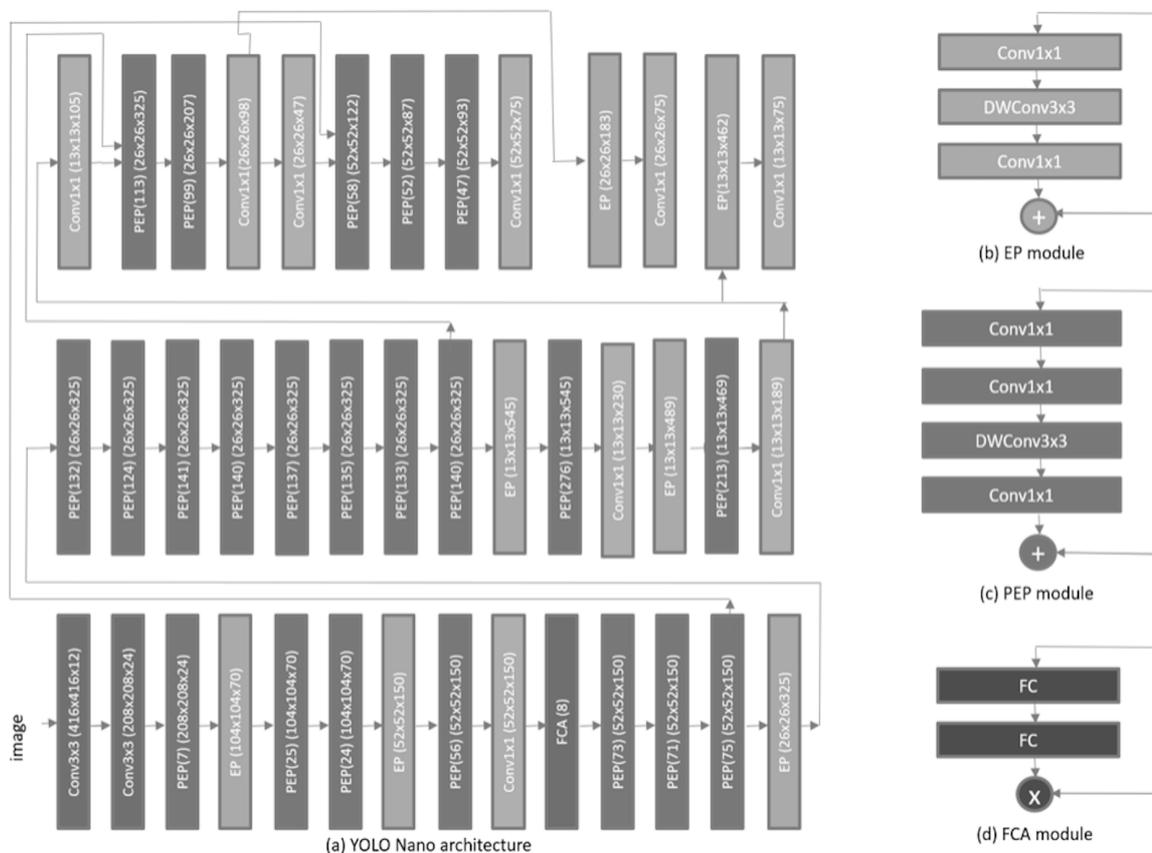


Рисунок 7 – Архитектура СНС YOLO Nano

За модулями представления признаков следуют несколько сверточных слоев, чьи выходы являются трехмерным тензорами, содержащими предсказания ограничивающих рамок, априорной вероятности наличия объекта и вероятностей классов для трех различных масштабов – 52x52, 26x26 и 13x13.

3.3.2 Архитектура модели СНС YOLOv4-Tiny

Изначально модель СНС YOLOv4 [34] - модель для решения задачи детектирования объектов, ориентированная на использование в рамках встраиваемых и мобильных СКЗ. Было предложено упростить структуру СНС YOLOv4 и уменьшить число параметров, чтобы ее можно было использовать в качестве основы для разработки мобильных и встроенных СКЗ. Так появилась упрощенная модель СНС YOLOv4-Tiny [28].

Можно использовать СНС YOLOv4-Tiny для более быстрого обучения и более быстрого детектирования объектов. У нее только две «головы» YOLO, а не три как в исходной модели YOLOv4. Она имеет 29 предварительно обученных сверточных слоев, в отличие от СНС YOLOv4, которая имеет 137 предварительно обученных сверточных слоев.

Модель YOLOv4-Tiny достигает точности детектирования 22,0% по метрике AP (42,0% по метрике AP_{0.5}) при скорости 443 FPS на графическом процессоре RTX 2080Ti, а при использовании TensorRT, размера партии, равного 4, и точности FP16 модель YOLOv4-Tiny достигает 1774 FPS.

Для обнаружения летающих объектов в реальном времени модель YOLOv4-Tiny является лучшим вариантом по сравнению с YOLOv4, поскольку более быстрое время вывода важнее, чем точность или правильность при работе со средой обнаружения объектов в реальном времени.

Архитектурно модель YOLOv4-Tiny использует несколько различных изменений по сравнению с исходной моделью YOLOv4, чтобы помочь ей достичь этих высоких скоростей. Прежде всего, количество сверточных слоев в backbone части CSP сжимается до 29 предварительно обученных сверточных слоев. Кроме того, количество слоев в этой модели было уменьшено до двух вместо трех, а также меньше стало полей привязки для прогнозирования. На рисунке 8 представлена архитектура YOLOv4-Tiny [28].

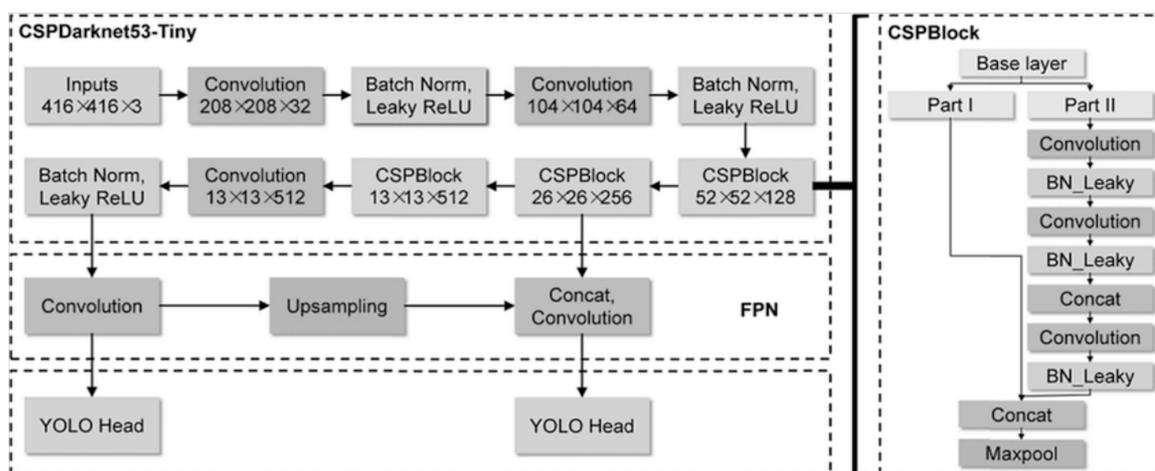


Рисунок 8 – Архитектура модели YOLOv4-Tiny

В качестве neck в архитектуре СНС YOLOv4-Tiny используется сочетание модулей SPP и PANet. Модуль SPP объединяет карты признаков разного масштаба выполняя несколько операций субдискретизации с различным размером окна (1×1 , 5×5 , 9×9 , и 13×13) и используется для расширения области восприятия и выделения наиболее значимых семантических признаков. При этом модуль SPP практически не влияет на скорость работы СНС. Вместо модуля FPN, использовавшегося в YOLOv3, в YOLOv4-Tiny для агрегации признаков применяется модуль PANet. В архитектуре СНС YOLOv4-Tiny также используются модули SAM [44].

В модифицированном модуле PANet сложение карт признаков при пробросе заменено конкатенацией. В модифицированном модуле SAM spatial-wise attention заменено point-wise attention.

3.3.3 Архитектура модели СНС YOLOv5

Модель СНС YOLOv5 [35] была выпущена с рядом отличий и улучшений относительно YOLOv4. Выпуск (подкласс) СНС YOLOv5 включает пять разных моделей: YOLOv5s (самая легковесная модель), YOLOv5m, YOLOv5l, YOLOv5x.

Это первый выпуск моделей в семействе YOLO, который изначально

написан с применением библиотеки для машинного обучения PyTorch [36], а не с использованием фреймворка Darknet. В связи со сменой среды реализации, использование модели YOLOv5 становится проще для исследовательского сообщества. Это также упрощает развертывание ее на мобильных устройствах.

Также, модель YOLOv5 является более легкой по сравнению с реализованной с использованием фреймворка Darknet моделью YOLOv4 [35]. Согласно исследованиям, YOLOv5 почти на 90 процентов меньше, чем YOLOv4. Это означает, что YOLOv5 можно гораздо проще развернуть на встроенных вычислительных устройствах.

Все модели подкласса YOLOv5 состоят из одних и тех же 3 компонентов: CSP-Darknet53 с модификацией BottleNeckCSP в качестве основы, CSP блоком с использованной стратегией CSPNet и SPPF модификацией в neck части и head частью как в модели YOLOv4. На рисунке 9 представлена архитектура подкласса моделей YOLOv5.

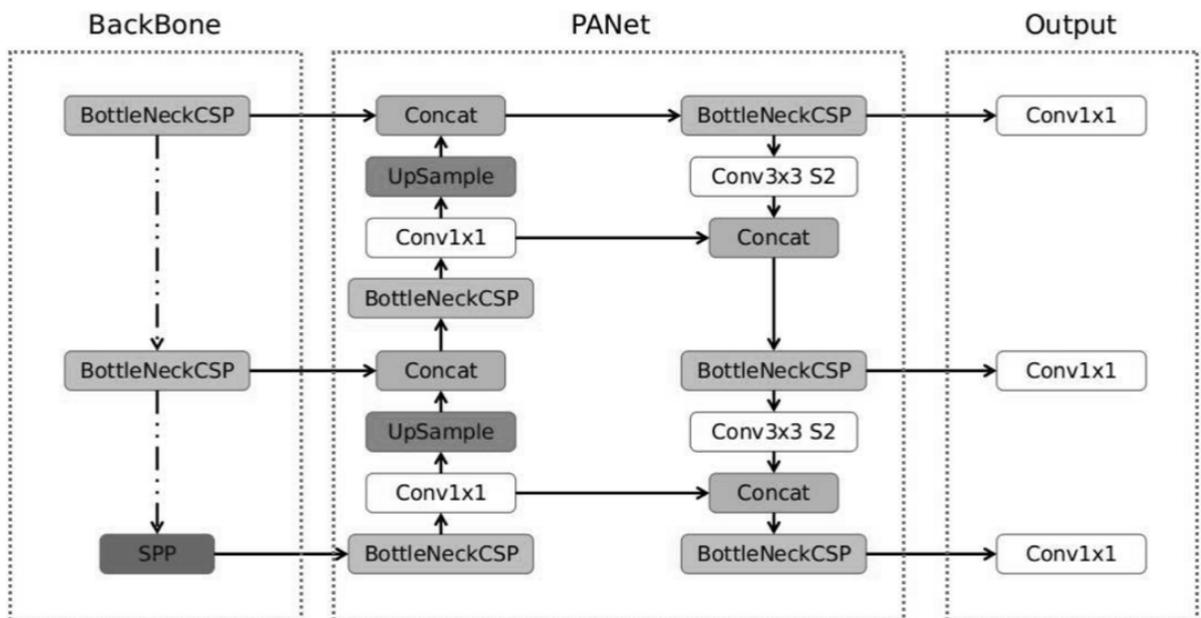


Рисунок 9 – Архитектура моделей подкласса YOLOv5

Поскольку модели подкласса YOLO v5 – это глубокие СНС, то в них используются остаточные и плотные блоки, чтобы обеспечить поток информации в самые глубокие слои и решить проблему исчезающего

градиента. Однако одним из преимуществ использования плотных и остаточных блоков является возможность решения проблемы избыточных градиентов. CSPNet помогает решить эту проблему, усекая градиентный поток. Каждая модель YOLOv5 использует стратегию CSPNet для разделения карты объектов базового уровня на две части, а затем объединяет их через иерархию между этапами. Схема работы BottleNeckCSP блока представлена на рисунке 10.

Применение этой стратегии дает большие преимущества для моделей YOLOv5, поскольку помогает уменьшить количество параметров и помогает сократить важный объем вычислений, что приводит к увеличению скорости вывода, которая является ключевым параметром в моделях обнаружения объектов в реальном времени [35].

При проектировании моделей подкласса YOLOv5 были внесены два основных изменения в neck часть модели. Сначала использовался вариант объединения пространственных пирамид (SPP), а сеть объединения путей PANet была изменена путем включения в ее архитектуру блоков BottleNeckCSP. Имеются два варианта включения:

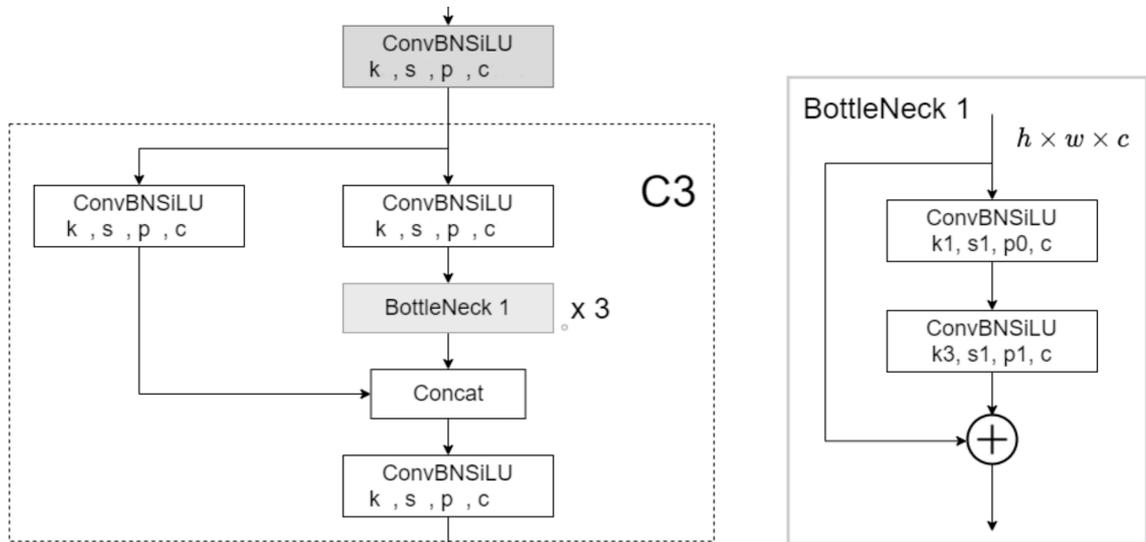


Рисунок 10 – Схема работы блока BottleNeckCSP

- CSPNet стратегия
- PANet – это сеть функциональных пирамид, она использовалась в

предыдущей версии YOLO (YOLOv4) для улучшения потока информации и помощи в правильной локализации пикселей в задаче предсказания маски. В YOLOv5 эта сеть была изменена путем применения к ней стратегии CSPNet (рисунок 10)

- SPPF модификация блока SPP

Блок SPP выполняет агрегирование информации, поступающей от входов, и возвращает на выходе фиксированную длину. Таким образом, он имеет то преимущество, что значительно увеличивает восприимчивое поле и выделяет наиболее релевантные контекстные функции без снижения скорости выполнения СНС. Этот блок использовался в предыдущих версиях СНС YOLO (YOLOv3 и YOLOv4) для отделения наиболее важных функций от backbone, однако в YOLOv5 (6.0/6.1) использовался SPPF, который является еще одним вариантом блока SPP, для улучшения скорость сети. На рисунке 11 представлена схема SPPF модуля [35].

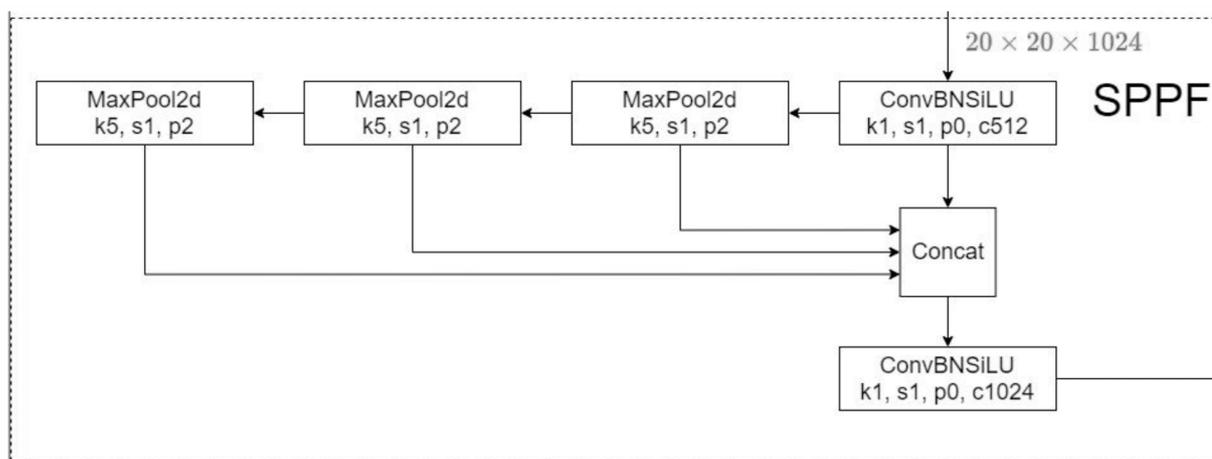


Рисунок 11 – Схема работы SPPF модуля

3.3.4 Архитектура модели СНС YOLOv7

Модель СНС YOLOv7 является одной из самых современных и эффективных в семействе моделей YOLO. Для достижения этих результатов авторы этой модели внесли ряд изменений в базовую СНС YOLO и

процедуры обучения. Среди изменений, сделанных в модели СНС YOLOv7, можно отметить следующие [27]:

1. Расширенная агрегация эффективных уровней

Эффективность сверточных слоев сети YOLO в backbone части имеет важное значение для эффективной скорости логического вывода. Авторы модели YOLOv7 начали модификацию сети с достижения максимальной эффективности уровней с помощью Cross Stage Partial Networks.

В модели YOLOv7 авторы основываются на исследованиях, проведенных по этой теме, с учетом объема памяти, необходимого для хранения слоев в памяти, а также расстояния, которое требуется градиенту для обратного распространения по слоям. Чем короче градиент, тем мощнее их сеть сможет обучаться. Последний уровень агрегации, который они выбирают, – это E-ELAN, расширенная версия вычислительного блока ELAN. На рисунке 12 представлена эволюция стратегий агрегации слоев в модели YOLOv7 [27].

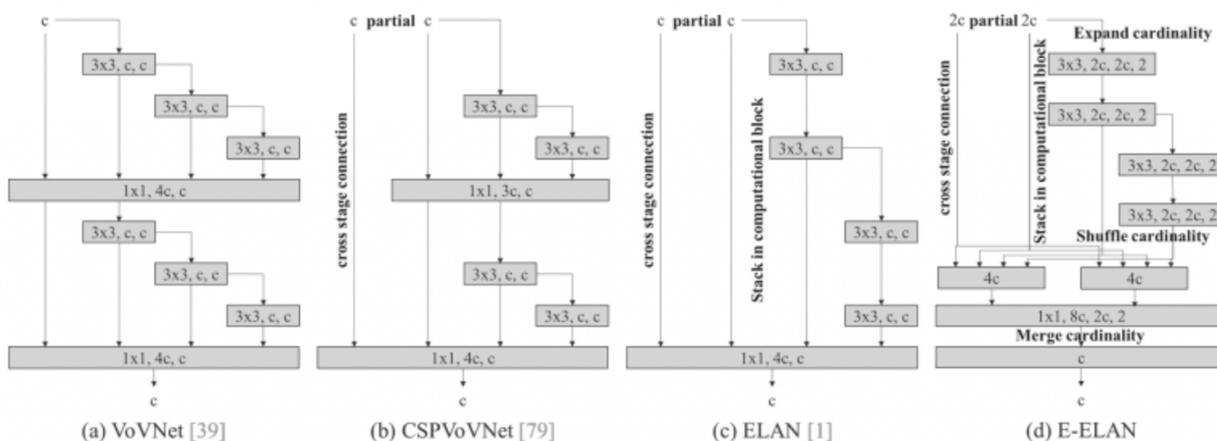


Рисунок 12 – Эволюция стратегий агрегации слоев в модели СНС YOLOv7

2. Методы масштабирования модели

Модели обнаружения объектов обычно выпускаются в виде серии моделей, увеличивающихся и уменьшающихся в размерах, поскольку для разных приложений требуются разные уровни точности и скорости вывода. Как правило, модели обнаружения объектов учитывают глубину сети,

ширину сети и разрешение, на котором обучается сеть. В YOLOv7 авторы согласованно масштабируют глубину и ширину сети, объединяя слои вместе. Исследования абляции показывают, что этот метод поддерживает оптимальную архитектуру модели при масштабировании для разных размеров. На рисунке 13 представлены методы масштабирования моделей в СНС YOLOv7 [27].

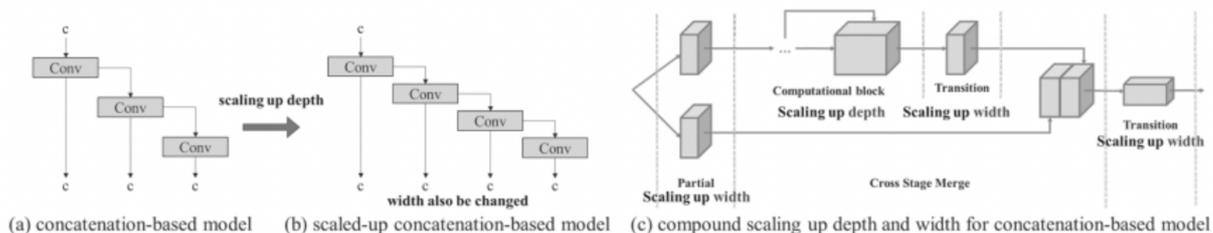


Рисунок 13 – Методы масштабирования моделей в СНС YOLOv7

3. Вспомогательный блок head

Head часть сети YOLO делает окончательные прогнозы для сети, но, поскольку она находится далеко в сети, может быть выгодно добавить вспомогательную head часть в сеть, которая находится где-то посередине. Вспомогательный head блок обучается не так эффективно, как последняя головка, потому что между ней и предсказанием меньше сети, поэтому авторы [27] модели YOLOv7 экспериментируют с различными уровнями наблюдения для этой головки, останавливаясь на определении от грубого к точному, где передается наблюдение назад от head блока с разной степенью детализации. На рисунке 14 представлено схематичное изображение блока head в YOLOv7.

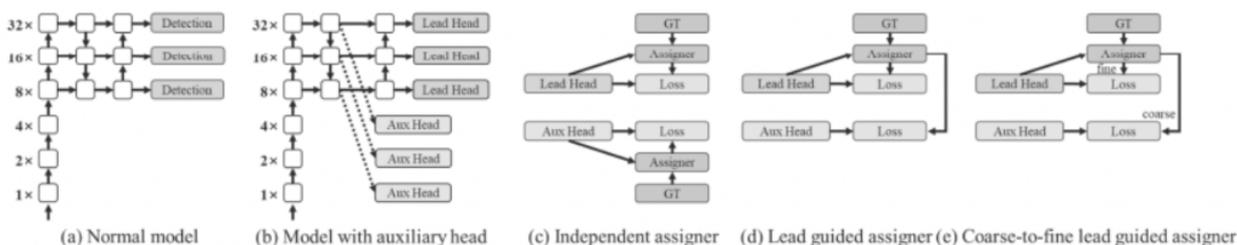


Рисунок 14 – Вспомогательный блок head в модели СНС YOLOv7

4. Планирование повторной параметризации

На рисунке 15 представлены блоки перепараметризации в YOLOv7.

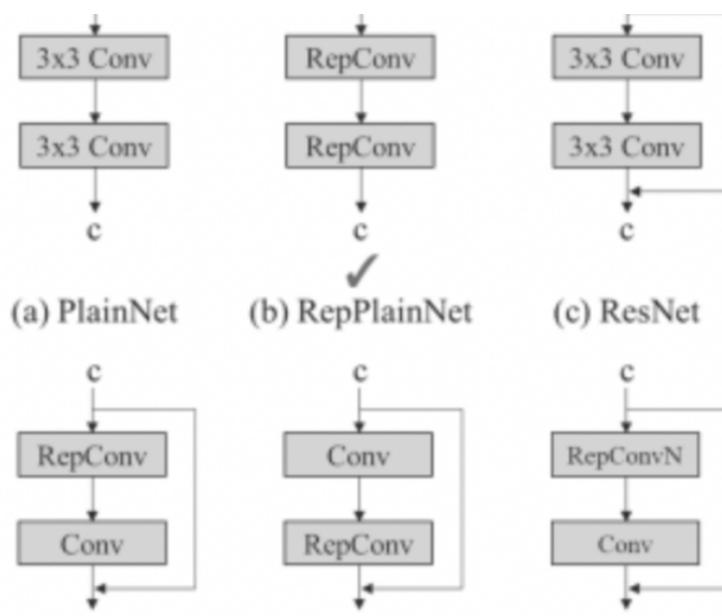


Рисунок 15 – Схема повторной параметризации

Методы повторной параметризации [27] включают усреднение набора весов модели для создания модели, более устойчивой к общим шаблонам, которые она пытается смоделировать. В исследованиях в последнее время основное внимание уделялось повторной параметризации (перепараметризации) на уровне модулей, когда часть СНС имеет свои собственные стратегии перепараметризации. Авторы модели YOLOv7 используют пути распространения градиентного потока, чтобы увидеть, какие модули в сети должны использовать стратегии повторной параметризации, а какие нет. На рисунке 15 представлены блоки перепараметризации в модели YOLOv7.

Модель СНС YOLOv7-Tiny является облегченной версией модели YOLOv7 и содержит всего около 6 миллионов параметров [27]. В отличие от других версий модели YOLOv7 в этой облегченной модели используется Leaky ReLU в качестве функции активации, в то время как в других версиях модели СНС YOLOv7 используется функция активации SiLU. Ввиду

меньшего числа сверточных слоев и параметров, СНС YOLOv7-Tiny является более пригодной для использования в мобильных СКЗ и поэтому выбрана нами для дальнейших исследований.

Отметим, что при реализации и настройке всех анализируемых выше компактных моделей СНС учитывались описанные выше особенности их архитектуры.

4 Подготовка датасета

Для обеспечения эффективной работы модели СНС необходимо обучить ее на достаточном количестве обучающих изображений. Поиск изображений для обучения является непростой задачей, так как получение качественных изображений многих типов объектов является затруднительным. В случае с БПЛА это объясняется сложностью получения изображений объектов, находящихся в движении и сложностью съемки таких объектов, находящихся на больших расстояниях. В ходе выполнения ВКР был осуществлен поиск готовых датасетов с размеченными изображениями трех классов. Ввиду малого числа найденных размеченных изображений с объектами классов «БПЛА вертолетного типа», «БПЛА самолетного типа» и класса «птица» осуществлялся дополнительный поиск исходных (неразмеченных) изображений, а также поиск видео с объектами этих классов и проводилась последующая разметка найденных изображений.

В связи с нижеперечисленными факторами, особенно важен контроль качества изображений, на которых будет обучаться модель СНС. Предварительная обработка и очистка данных должны проводиться до того, как набор изображений будет использоваться для обучения модели. Необработанные данные зачастую искажены и ненадежны, и в них могут быть пропущены значения. Использование таких данных при моделировании может приводить к неверным результатам. Эти задачи являются частью процесса обработки и анализа данных группы и обычно подразумевают первоначальное изучение набора данных, используемого для определения и планирования необходимой предварительной обработки.

Из этого следует, что необходимо придерживаться следующих правил [33]:

1. Эмпирические правила определения минимального количества изображений. Довольно сложно определить количество изображений, необходимых в наборе изображений для решения задачи классификации

изображений. Тем не менее, есть несколько хороших практических правил, которым мы здесь следовали. По одному из них, около 1000 примеров по каждому из классов объектов - это достаточное количество для начала.

2. Устранение дубликатов в исходных данных. Модель СНС не сможет выделить достаточно признаков из датасета, содержащего слишком похожие друг на друга обучающие данные. Одним из способов обнаружить дубликаты – представить изображения в виде векторов и рассчитать сходство между ними. Это более надежный способ вычисления сходства между двумя векторами, чем простое расстояние между векторами. Затем можно выбрать порог для применения и определить изображения, которые слишком похожи. Хорошее эмпирическое правило – отфильтровывать изображения с показателем сходства 0,9 и выше.

3. Отсутствие неразмеченных данных. В датасете обязательно все объекты каждого класса должны иметь метки. Проще говоря, в датасете не должно быть изображений с неразмеченными объектами.

В результате разметки данных с помощью утилиты Roboflow [37] был получен датасет изображений размером 416×416 пикселей с летающими объектами классов «БПЛА вертолетного типа (helicopteruav)», «БПЛА самолетного типа (winguav)» и «Птица (bird)». Для каждого изображения были получены координаты ограничивающей каждый объект рамки. Любое из размеченных изображений включает один и более объектов разных классов и разных размеров. Имеется довольно большое число изображений, включающих объекты малых размеров (площадью менее 32х32 пикселя). Последнее должно быть учтено при обучении моделей СНС. Всего в датасете содержится 3400 изображений, из них 2000 – в обучающей выборке, 800 – в тестовой выборке и 600 - изображений в валидационной.

На рисунке 16 представлен фрагмент процесса разметки изображений в интерактивной среде программного обеспечения Roboflow. Эта система имеет довольно дружелюбный пользовательский интерфейс, значительно облегчающий разметку больших объемов изображений.

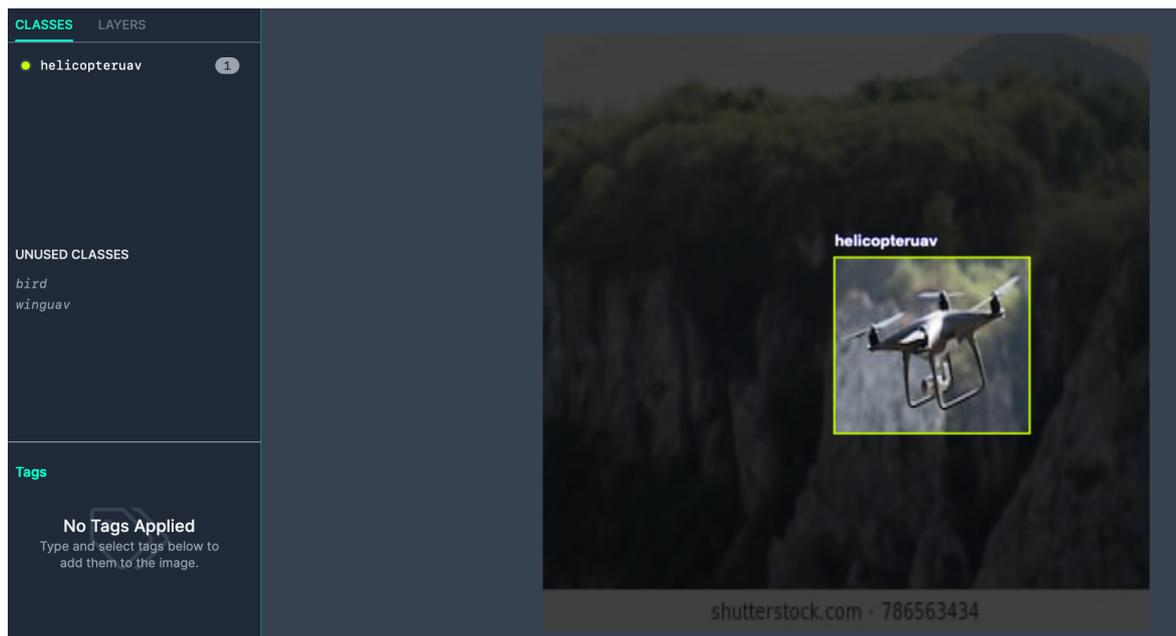


Рисунок 16 – Пример разметки изображений в интерактивной среде Roboflow

Ввиду того, что каждая выбранная модель СНС семейства YOLO требует датасет в своем формате, были написаны четыре скрипта для конвертации изображений в нужный формат: по одному скрипту для каждой из выбранных компактных моделей СНС. Ниже приведен пример скрипта для конвертации датасета из COCO формата в VOC формат, требуемый моделью YOLO Nano.

```
def convert_coco_to_voc(coco_path, img_folder, output_folder):
    coco = COCO(coco_path)
    categories = {cat['id']: cat for cat in coco.loadCats(coco.getCatIds())}

    for img_id in coco.getImgIds():
        ann_ids = coco.getAnnIds(imgIds=[img_id])
        img = coco.imgs[img_id]

        xml_data = create_xml(img, coco.loadAnns(ann_ids), img_folder, categories)
        xml_file = os.path.join(output_folder, img['file_name'].replace('.jpg', '.xml'))
```

```
shutil.copy2(os.path.join(img_folder, img['file_name']), output_folder)
with open(xml_file, 'wb') as f:
    f.write(tostring(xml_data))
```

```
coco_base = 'coco_dataset'
img_folder = os.path.join(coco_base, 'images', 'val2017')
coco_path = os.path.join(coco_base, 'annotations', 'instances_val2017.json')
output_folder = 'voc_output'
os.makedirs(output_folder, exist_ok=True)
convert_coco_to_voc(coco_path, img_folder, output_folder)
```

5 Программная реализация компактных моделей СНС и особенности их обучения

В этом разделе кратко рассмотрены основные вопросы реализации и проведения обучения компактных моделей СНС, выбранных в результате анализа основных современных моделей СНС семейства (класса) YOLO в части особенностей их архитектур (см. раздел 3). Все этапы работы проводились с использованием фреймворка PyTorch [36].

5.1 Общие положения

Ниже рассматриваются особенности программной реализации и обучения на обучающей выборке созданного датасета четырех выбранных потенциально перспективных для проектирования мобильных СКЗ компактных моделей СНС класса YOLO: YOLOv4 - Tiny, YOLOv5n, YOLOv7-Tiny и YOLO Nano. Для того, чтобы в исследованиях этих моделей получить сопоставимые результаты и правильно выбрать по этим результатам наиболее эффективную модель СНС необходимо при обучении всех моделей использовать схожие параметры. Размер каждого входного изображения из основного датасета для всех исследуемых моделей составлял 416×416 пикселей.

В обучаемых моделях классы объектов представлены текстовыми идентификаторами:

- bird – Птица;
- helicopteruav – БПЛА вертолётного типа;
- winguav – БПЛА самолётного типа;

Для анализа эффективности этих моделей СНС нужно применять их программные реализации, созданные на одном фреймворке, поскольку скорость вычислений одной и той же модели СНС может зависеть от

используемого фреймворка. Также модели СНС могут различаться в использовании ресурсов памяти при их разработке и тестировании на разных фреймворках. Для настройки и обучения всех четырех моделей СНС нами использовался фреймворк PyTorch [36]. PyTorch – это фреймворк машинного обучения, который используется для создания нейронных сетей, в том числе свёрточных. Одной из основных особенностей PyTorch является простота и удобство использования. Этот фреймворк позволяет создавать и использовать модели нейронных сетей, не требуя больших усилий от разработчика.

Модель СНС YOLO Nano не имеет официальной программной реализации, поэтому автор реализовал ее программно в PyTorch на языке Python (Приложение А).

5.2 Особенности разработки и обучения моделей СНС семейства YOLO

Программная реализация, настройка и обучение моделей СНС проводились с использованием фреймворка PyTorch. В качестве основы при разработке и настройке использовались официально признанные международным сообществом архитектуры четырех компактных моделей СНС, описанные в разделе 3.

Обучение каждой из четырех компактных моделей СНС проводилось на обучающей выборке объемом 2000 размеченных изображений из датасета. При этом использовались следующие основные параметры:

- Начальная скорость обучения (lr0): 0,01;
- Конечная скорость обучения (lrf): 0,01;
- Импульс SGD (momentum): 0,937;
- Затухание веса оптимизатора (weight_decay): 0,0005;
- Количество эпох: 100 (YOLOv5n, YOLOv7-Tiny), 200 (YOLO Nano), 250 (YOLOv4-Tiny);

- Функция активации: Leaky ReLU;
- Функция оптимизации: Adam;
- Функция потерь: GIoU

Отметим, что при задании значений ряда параметров обучения учитывалось наличие у значительного числа изображений датасета летающих объектов малых размеров.

Обучение всех выбранных компактных моделей СНС велось на платформе для облачных вычислений Google Collab, на графических ускорителях NVIDIA A100-SXM4-40GB с процессором Intel(R) Xeon(R) CPU @ 2.30GHz. Хранение датасета организовывалось с помощью облачного хранилища. При обучении моделей использовалась библиотека CUDA версии 12.0. Для увеличения числа изображений в обучающей выборке использовался прием, называемый мозаичной аугментацией «на лету».

Каждая обученная модель СНС затем подлежала валидации на валидационной (проверочной) выборке из датасета. Объем этой выборки – 600 размеченных изображений из датасета.

6 Результаты обучения и исследования компактных моделей СНС семейства YOLO

6.1 Результаты обучения компактных моделей СНС

Для визуализации полученных результатов обучения моделей СНС в виде графиков использовалась платформа tensorboard, позволяющая отобразить графики значений метрик Precision, Recall, $AP_{0,5}$, $AP_{0,5:0,95}$, $mAP_{0,5}$ и $mAP_{0,5:0,95}$, полученные на обучающей выборке для каждой из четырех моделей СНС. На рисунке 17 приведены графики зависимости значений метрик Precision и Recall от числа эпох при обучении в случае модели СНС YOLOv5n. Видим, что метрики достигают значения 0,9 и 0,84 соответственно.

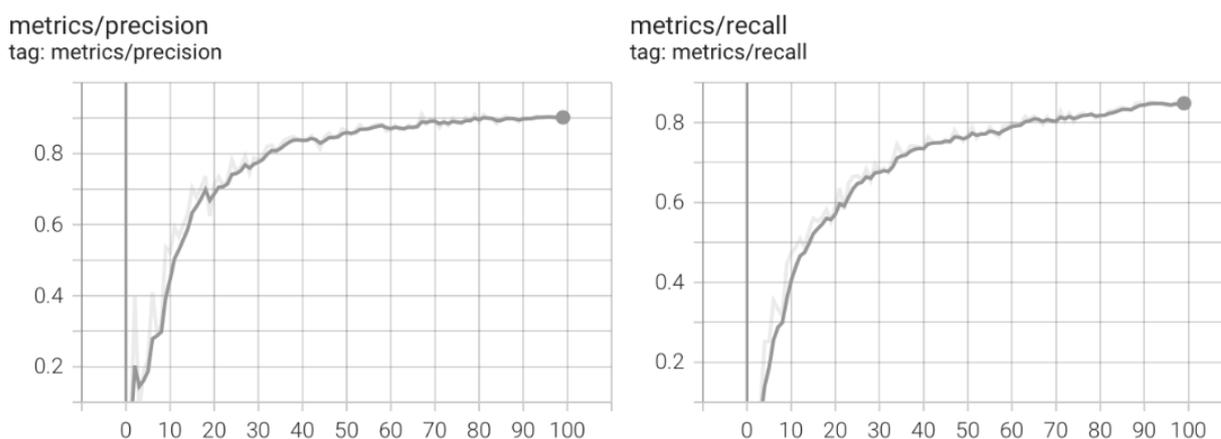


Рисунок 17 – Значения метрик Precision и Recall в случае модели СНС YOLOv5n в зависимости от числа эпох

На рисунке 18 приведены значения метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$, полученные при обучении этой же модели СНС. Эти метрики по результатам обучения имеют максимальные значения 0,89 и 0,60 соответственно.

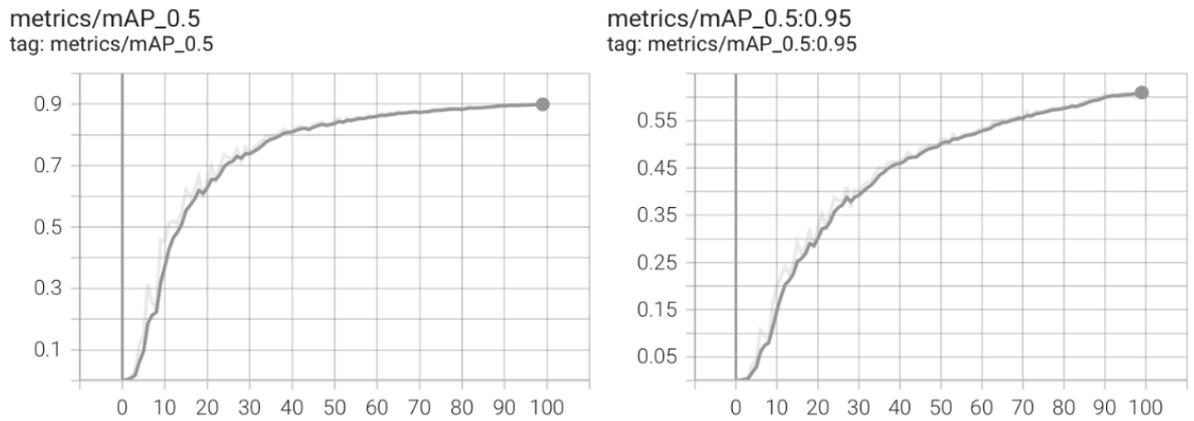


Рисунок 18 – Значения метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$ в случае модели СНС YOLOv5n в зависимости от числа эпох

На рисунке 19 представлен пример валидации обученного детектора-модели СНС YOLOv5n на изображении, содержащем БПЛА вертолетного типа и БПЛА самолетного типа при значении `threshold` (значение уверенности для детектора), равном 0,35.



Рисунок 19 – Пример валидации модели СНС YOLOV5n при детектировании на изображении БПЛА самолетного типа и БПЛА вертолетного типа

На рисунке 20 приведены графики зависимости метрик Precision и Recall от числа эпох при обучении модели СНС YOLOv7-Tiny. Видим, что

эти метрики принимают максимальные значения 0,91 и 0,80 соответственно.

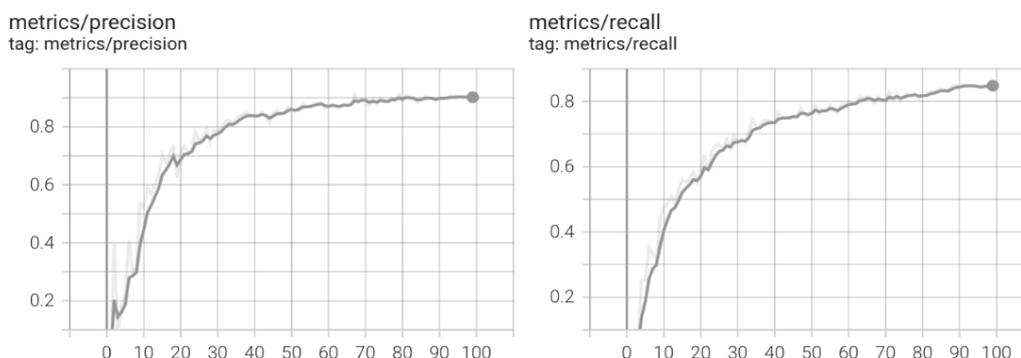


Рисунок 20 – Значения метрик Precision и Recall в случае модели СНС YOLOv7-Tiny в зависимости от числа эпох

На рисунке 21 приведены графики зависимости метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$ в зависимости от числа эпох для модели YOLOv7-Tiny. По результатам обучения максимальные значения метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$ составили 0,90 и 0,61 соответственно. Это означает, что для модели СНС YOLOv7-Tiny максимальные значения этих метрик незначительно превосходят максимальные значения соответствующих метрик для модели СНС YOLOv5n. Отметим, что по метрике $mAP_{0,5}$ обученные модели СНС YOLOv5n и СНС YOLOv7-Tiny значительно превышают пороговое значение 0,7, то есть удовлетворяют требованию по точности детектирования объектов на изображениях (см. подраздел 2.1).

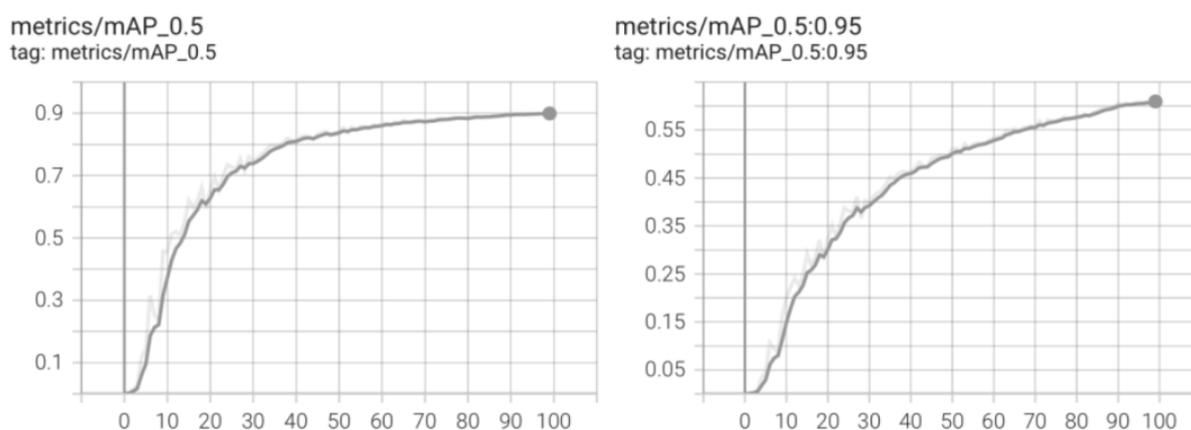


Рисунок 21 – Значения метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$ в случае модели СНС YOLOv7-Tiny в зависимости от числа эпох

На рисунке 22 представлен график зависимости метрики $mAP_{0.5}$ от числа эпох, полученный во время обучения модели YOLO Nano. Максимальное значение метрики $mAP_{0.5}$ равняется 0,66. График был построен с применением python-библиотеки matplotlib.

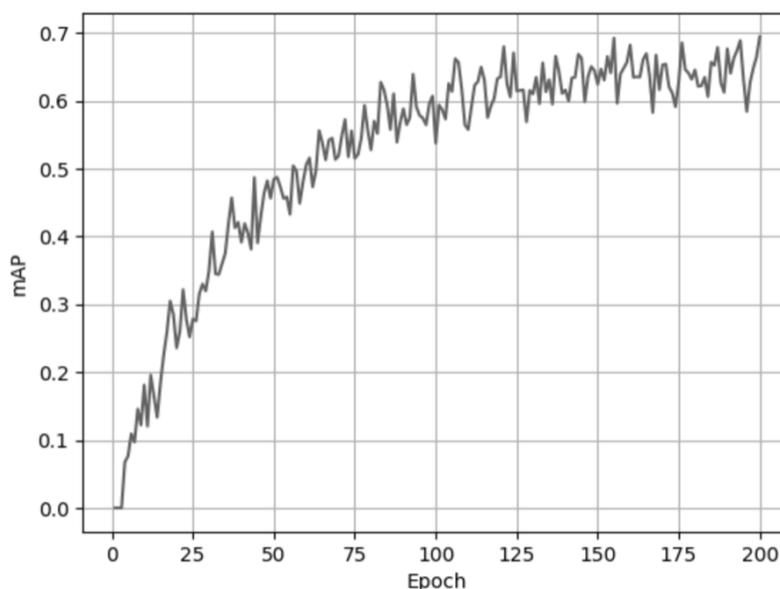


Рисунок 22 – Значения метрики $mAP_{0.5}$ в случае модели СНС YOLO Nano в зависимости от числа эпох

Аналогично, на рисунке 23 представлен график зависимости метрики $mAP_{0.5}$, полученный при обучении модели YOLOv4-Tiny, в зависимости от числа эпох обучения. Максимальное значение $mAP_{0.5}$ равняется 0,73.

Таким образом, для модели YOLOv4-Tiny на обучающей выборке удалось достигнуть большей усредненной точности детектирования объектов трех классов, чем в случае модели YOLO Nano. Однако обе эти модели значительно не достигают пороговое значение 0.8 по метрике $mAP_{0.5}$.

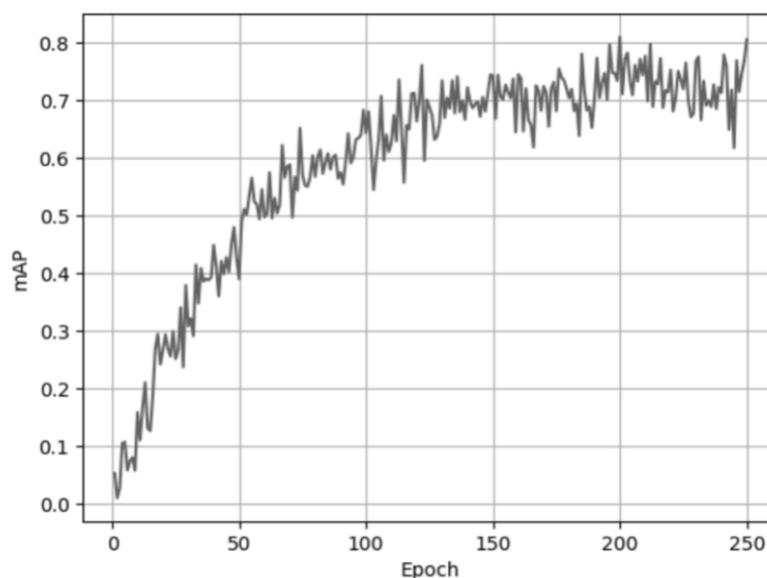


Рисунок 23 – Значения метрики $mAP_{0,5}$ в случае модели СНС YOLOv4-Tiny в зависимости от числа эпох

6.2 Результаты исследования компактных моделей СНС

Тестовая выборка составляет 800 размеченных изображений с тремя классами летающих объектов: «БПЛА самолетного типа», «БПЛА вертолетного типа» и «Птица». Сразу отметим, что результаты исследования моделей СНС по точности детектирования этих объектов на тестовой выборке следующие: максимальные значения метрик $mAP_{0,5}$, $mAP_{0,5:0,95}$ меньше, чем максимальные значения этих же метрик в случае обучающей и валидационной выборок, причем это имеет место для каждой из исследуемых компактных моделей СНС. В таблице 3 в качестве примера приведены значения метрик $AP_{0,5}$ для каждого из классов летающих объектов и значение метрики $mAP_{0,5}$, полученные при тестировании модели СНС YOLOv5n. Значение метрики $mAP_{0,5}$, равное 0,68, значительно меньше значения 0,89 для этой же модели в случае обучающей выборки. Заметим также, что это значение ниже порогового значения 0,7.

Как видно из таблицы 3, наименьшее значение точности детектирования по метрике $AP_{0,5}$ имеют объекты класса «Птица». Это возможно из-за разнообразия таких объектов (на изображениях в датасете

представлены как одна птица, так и стаи птиц, причем разных видов) и сложности их классификации, в том числе из – за малых размеров птиц.

Таблица 3 – Значения метрик $AP_{0,5}$ и $mAP_{0,5}$ при тестировании модели СНС YOLOv5n

Класс летающих объектов	$AP_{0,5}$, $mAP_{0,5}$
winguav (БПЛА самолетного типа)	0,67
helicopteruv (БПЛА вертолетного типа)	0,79
bird (Птица)	0,58
all (усредненное значение для всех классов, $mAP_{0,5}$)	0,68

Аналогичные исследования проведены с использованием тестовой выборки для модели YOLOv7-Tiny. В таблице 4 для этой модели приведены значения метрики $AP_{0,5}$ для каждого из классов объектов и метрики $mAP_{0,5}$, показывающей усредненную точность детектирования по всем трем классам.

Таблица 4 – Значения метрик $AP_{0,5}$ и $mAP_{0,5}$ при тестировании модели СНС YOLOv7-tiny

Название класса	$AP_{0,5}$, $mAP_{0,5}$
winguav (БПЛА самолетного типа)	0,76
helicopteruv (БПЛА вертолетного типа)	0,85
bird (Птица)	0,64
all (усредненное значение для всех классов, $mAP_{0,5}$)	0,75

Из таблицы 3 и таблицы 4 следует, что значения метрики $AP_{0,5}$ по классам объектов и метрики $mAP_{0,5}$ для всех классов в случае модели YOLOv7-tiny выше, чем в случае модели YOLOv5n.

Основные результаты исследования на тестовой выборке каждой из четырех компактных моделей по всем метрикам (по скорости вычислений модели СНС, компактности ее архитектуры, вычислительной сложности и

точности детектирования объектов на изображениях) приведены в таблице 5. Они позволяют оценить, соответствует ли та или иная модель СНС требованиям к моделям СНС, изложенным в подразделе 2.1.

Таблица 5 – Результаты исследований компактных моделей СНС YOLO

Модель СНС	Размер входного изображения, пикселей	Размер модели, МБ	Количество операций, GFLOP	Скорость вычислений, FPS	mAP _{0,5}	mAP _{0,5:0,95}
YOLO Nano	416	4	4,90	311,3	0,53	0,32
YOLOv4-Tiny	416	24	6,9	280,8	0,59	0,41
YOLOv5n	416	14,7	6,2	197,3	0,68	0,47
YOLOv7-Tiny	416	12,3	5,5	210,4	0,75	0,53

Скорость вычислений каждой модели СНС, описываемая метрикой FPS, оценивалась исходя из времени инференса одного изображения с помощью этой модели СНС на графических ускорителях NVIDIA A100-SXM4-40GB с процессором Intel(R) Xeon(R) CPU @ 2.30GHz (см. подраздел 5.2). Отметим, что все модели СНС по скорости детектирования объектов на изображении (по скорости вычисления модели) значительно превышают пороговое значение 25.

Размер модели определялся как итоговое число параметров обученной модели СНС (количество весовых коэффициентов в слоях модели). Из таблицы 5 видим, каждая из моделей СНС имеет весьма невысокие размер модели и вычислительную сложность (количество операций в GFLOP). Поэтому любая из этих моделей удовлетворяет требованию по алгоритмической эффективности модели, поскольку может выполняться на вычислительных устройствах даже с весьма скромными ресурсами. Однако наименьшие размер модели и вычислительную сложность имеет СНС YOLO Nano, при этом уступая другим исследуемым моделям СНС по точности детектирования объектов на изображениях.

Из таблицы 5 следует, что наибольшее значение метрики $mAP_{0,5}$ по точности детектирования объектов, превышающее пороговое значение 0,7, дает модель СНС YOLOv7-tiny. Отметим, что требование по точности детектирования объектов на изображениях удовлетворяется только для этой модели СНС. Однако для объектов класса «Птица» даже для этой модели точность детектирования таких объектов по метрике $AP_{0,5}$ не достигла порогового значения 0,7.

Анализируя все полученные результаты исследования компактных моделей СНС класса YOLO при решении задачи детектирования летающих объектов трех классов можно рекомендовать модель СНС YOLOv7-tiny к реализации в составе мобильной СКЗ реального времени. При этом предварительно следует увеличить число размеченных изображений с объектами класса «Птица» в датасете и дообучить эту модель, достигнув порогового значения 0,7 по метрике $AP_{0,5}$. В случае крайне малых вычислительных ресурсов мобильной СКЗ и при наличии невысоких требований с ее стороны к точности детектирования летающих объектов можно использовать модель СНС YOLO Nano.

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСООБЪЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
8K92	Беляеву Семену Игоревичу

Школа	ИШИТР	Отделение	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.04 Программная инженерия

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

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад инженера – 21760 руб. Оклад руководителя – 33664 руб.
2. Нормы и нормативы расходования ресурсов	Премияльный коэффициент 30%; Коэффициент доплат и надбавок 20%; Районный коэффициент 30%; Коэффициент дополнительной заработной платы 12%; Накладные расходы 16%. Районный коэффициент Томска равный 1.3
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды равный 30 %

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	Определение конечного потребителя. Проведение SWOT-анализа реализации проекта.
2. Планирование и составление бюджета проекта	Определение целей и ожидаемых результатов, требований проекта. Расчет трудоемкости выполнения работ. Подсчет бюджета проекта.
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной и экономической эффективности	Расчет показателей финансовой эффективности, ресурсоэффективности и эффективности исполнения.

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT
3. График проведения и бюджет НИ
4. Расчёт денежного потока
5. Оценка ресурсной, финансовой, бюджетной и экономической эффективности НИ

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОСГН	Гасанов Магеррам Али оглы	Доктор экономических наук		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8K92	Беляев Семен Игоревич		

7 Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения

7.1 Потенциальные потребители результатов исследования

В рамках данной работы выполняется исследование, включающее в себя анализ и сравнение сверточных нейронных сетей YOLOv4-Tiny, YOLOv5n, YOLOv7-Tiny, и YOLO Nano при решении задачи детектирования объектов разных масштабов на изображениях. Сравнение проводится по ряду параметров, которые можно объединить в две группы – точность детектирования и вычислительная сложность. Результаты исследования возможностей таких нейросетей при детектировании объектов разных масштабов на изображениях будут востребованы при анализе данных мониторинга земной поверхности. Мониторинг осуществляется с помощью спутниковых систем или беспилотных летательных аппаратов. Особый интерес представляет для специалистов в области анализа данных мониторинга земной поверхности и, в частности, специалистов в области анализа данных мониторинга климатических и экологических систем.

Исполнителями проекта являются научный руководитель (НР) – Марков Н.Г., профессор, д.т.н., а также бакалавр (Б) – Беляев С.И, студент группы 8К92.

7.2 Анализ конкурентных технических решений

При рассмотрении различных подходов к решению задачи детектирования объектов разных масштабов, в качестве конкурентных решений можно выделить:

- ПО, автоматизирующее детектирование объектов с помощью четких алгоритмов (К1)
- ручная обработка снимков специалистами (К2)

Анализ конкурентных технических решений был проведен с помощью оценочной карты, представленной в таблице 6.

Таблица 6 – Оценочная карта критериев эффективности конкурентных технических решений (разработок)

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
Технические критерии оценки ресурсоэффективности							
Скорость обработки данных	0,1	4	5	3	0,4	0,5	0,3
Простота использования системы	0,15	4	2	5	0,6	0,3	0,75
Удобный интерфейс пользователя	0,1	5	4	5	0,5	0,4	0,5
Функциональные возможности	0,05	4	5	3	0,2	0,25	0,15
Потребность в ресурсах памяти	0,15	4	3	5	0,6	0,45	0,75
Экономические критерии оценки эффективности							
Послепродажное обслуживание	0,15	4	5	3	0,6	0,75	0,45
Предполагаемый срок эксплуатации	0,1	4	3	4	0,4	0,3	0,4
Цена	0,2	4	5	2	0,8	1	0,4
Итого	1	33	35	30	4,1	3,95	3,7

Анализ конкурентных решений определяется по формуле 1:

$$K = \sum (B_i \times \text{В}_i), \quad (1)$$

где K – конкурентоспособность научной разработки или конкурента; В_i – вес показателя (в долях единицы); B_i – балл i -го показателя.

Основываясь на рассчитанных значениях, можно предположить, что простота использования исследуемого решения и его низкая потребность в ресурсах памяти в сочетании с выгодной ценой формирует конкурентное преимущество. Существующим альтернативам не хватает более низкой конечной стоимости в одном случае (ручная обработка снимков специалистами)

- низкой потребности в ресурсах памяти в другом случае (ПО, автоматизирующее детектирование объектов с помощью четких алгоритмов).

7.3 Технология QuaD

Технология QuaD (QUality ADvisor) представляет собой гибкий инструмент измерения характеристик, описывающих качество новой разработки и ее перспективность на рынке и позволяющие принимать решение целесообразности вложения денежных средств в научно-исследовательский проект.

Анализ имеющихся конкурентных продуктов необходимо проводить с достаточной регулярностью, так как рынок ИТ находится в постоянном и активном движении на сегодняшний день. Данный анализ позволяет производить изменения текущего внедрения, чтобы сделать его наиболее перспективным и конкурентоспособным в дальнейшем. Результат QuaD-анализа представлен в таблице 7.

Таблица 7. Оценка критериев в соответствии с технологией QuaD

Критерии оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение	Средневзвешенное значение
Скорость обработки изображений	0,1	80	100	0,8	8
Простота использования системы	0,15	90	100	0,9	13,5
Удобный интерфейс	0,1	90	100	0,9	9
Функциональные возможности	0,05	70	100	0,7	3,5
Потребность в ресурсах памяти	0,15	90	100	0,9	13,5
Послепродажное обслуживание	0,15	60	100	0,6	9
Предполагаемый срок эксплуатации	0,1	90	100	0,9	9
Цена	0,2	90	100	0,9	18
Итого:					83,5

Средневзвешенное значение показателя качества и перспективности научной разработки равен 83,5 (попадает в промежуток 80–100), то есть такая разработка считается перспективной.

7.4 SWOT-анализ

SWOT-анализ – метод стратегического планирования, заключающийся в выявлении факторов внутренней и внешней среды организации и разделении их на четыре категории: Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности), Threats (угрозы).

Сильные (S) и слабые (W) стороны являются факторами внутренней среды объекта анализа, (то есть тем, на что сам объект способен повлиять); возможности (O) и угрозы (T) являются факторами внешней среды (то есть тем, что может повлиять на объект извне и при этом не контролируется объектом).

В рамках данного анализа выявлены сильные и слабые стороны проекта, также его возможности и угрозы. Все параметры SWOT анализа данного проекта представлены в таблице 8.

Таблица 8 – SWOT-анализ проекта

	Внутренние факторы	
	Сильные стороны проекта:	Слабые стороны проекта:
	1. Надежность 2. Мобильность 3. Персонализация 4. Безопасность 5. Простота эксплуатации	1. Отсутствие прототипа научной разработки 2. Отсутствие финансирования 3. Повышенные требования к «железу» 4. Отсутствие репутации на рынке
Возможности: 1. Использование инновационной инфраструктуры ТПУ 2. Повышение качества детектирования 3. Появление дополнительного спроса на новый продукт	При использовании инфраструктуры ТПУ можно снизить стоимость проекта. Использование новых подходов (глубокое обучение) может повысить качество детектирования объектов разных масштабов на изображениях.	Проводя исследования на базе ТПУ, можно разработать прототип и получить известность на рынке

Угрозы: 1. Высокие требования к «железу» разработанного продукта 2. Неудовлетворительное качество работы разработки	Высокие требования к «железу» обусловлены спецификой исследуемых алгоритмов. Качество работы можно повысить за счет сбора качественной выборки.	Необходимо разработать прототип алгоритма и пройти сертификацию для того, чтобы выйти на рынок и заполучить репутацию.
--	---	--

7.5 Структура работ в рамках научного исследования

Важным этапом проведения научно-исследовательских работ является необходимость планирования работ, которое включает в себя определение полного перечня работ, а также их распределение между всеми исполнителями проекта. Исполнителями проекта являются студент и научный руководитель. Научный руководитель определяет цели и задачи для студента, направляет и контролирует работу его работу, оценивает результаты проделанной работы и дает рекомендации студенту. Студент полностью отвечает за выполняемую работу. В таблице 9 представлен перечень работ, а также распределение исполнителей по ним.

Таблица 9 – Перечень работ и распределение исполнителей

№ работы	Наименование работы	Исполнители работы
1	Выбор научного руководителя бакалаврской работы	НР, Б
2	Составление и утверждение темы бакалаврской работы	НР, Б
3	Подбор и изучение литературы по теме бакалаврской работы	Б
4	Составление календарного плана-графика выполнения бакалаврской работы	НР, Б
5	Анализ предметной области	Б
6	Выбор архитектур СНС	Б
7	Обучение СНС	Б
8	Сравнительный анализ результатов обучения СНС	Б
9	Согласование выполненной работы с научным руководителем	НР, Б.
10	Выполнение других частей работы (финансовый менеджмент, социальная ответственность)	Б
11	Подведение итогов, оформление работы	Б

7.6 Определение трудоемкости выполнения работ

Как правило, трудовые затраты образуют основную часть стоимости исследования, поэтому важным этапом планирования научно-исследовательской деятельности является определение трудоемкости работ.

Определение трудоёмкости выполнения работ осуществляется на основе экспертной оценки ожидаемой трудоемкости выполнения каждой работы путём расчёта длительности работ в рабочих и календарных днях каждого этапа работ.

Трудоемкость оценивается по следующей формуле 2:

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

где $t_{ож\ i}$ – это ожидаемая трудоемкость i -ой работы (чел.-дни);

$t_{min\ i}$ – это минимально возможная трудоемкость выполнения заданной i -ой работы (оптимистическая оценка) (чел.-дни);

$t_{max\ i}$ – это максимально возможная трудоемкость выполнения заданной i -ой работы (пессимистическая оценка) (чел.-дни).

После оценки ожидаемой трудоемкости работ, производится определение продолжительности каждой работы в рабочих днях по формуле:

$$T_{pi} = \frac{t_{ож\ i}}{Ч_i}, \quad (3)$$

где T_{pi} – это продолжительность одной работы (раб. дни),

$t_{ож\ i}$ – это ожидаемая трудоемкость выполнения одной работы (чел.-дни),

$Ч_i$ – это численность исполнителей, выполняющих одновременно одну и ту же работу на этом этапе (чел.).

Для того чтобы в дальнейшем построить график работ с помощью диаграммы Ганта, необходимо также произвести перевод длительности работ из рабочих дней в календарные по формуле 4:

$$T_{ki} = T_{pi} \cdot k_{\text{кал}}, \quad (4)$$

где T_{ki} – продолжительность выполнения i -й работы в календарных днях;
 T_{pi} – это продолжительность выполнения i -й работы в рабочих днях;
 $k_{\text{кал}}$ – это коэффициент календарности, равный 1,48.

Коэффициент календарности $k_{\text{кал}}$ рассчитывается по формуле 5:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - (T_{\text{вых}} + T_{\text{пр}})} \quad (5)$$

где $k_{\text{кал}}$ – коэффициент календарности;
 $T_{\text{кал}}$ – количество календарных дней в году;
 $T_{\text{пр}}$ – количество праздничных дней в году.

С учётом того, что календарных дней в 2023 году 365, а сумма выходных и праздничных дней составляет 118 дней, коэффициент календарности равен $k_{\text{кал}} = 1,48$.

7.7 Разработка графика проведения научного исследования

В таблице 10 приведены временные показатели научного исследования.

Таблица 10 – Временные показатели проведения научного исследования

Наименование работы	Исполнители работы	Трудоемкость работ, чел-дни		Длительность работ, дни		
		t_{min}	t_{max}	$t_{\text{ож}}$	T_p	T_k
Выбор научного руководителя бакалаврской работы	НР, Б	1	2	1,4	1	1
Составление и утверждение темы бакалаврской работы	НР, Б	1	3	1,8	1	1
Подбор и изучение литературы по теме бакалаврской работы	Б	5	10	7,5	4	6

Составление календарного плана-графика	НР, Б	2	8	6	3	4
Анализ предметной области	Б	1	3	1,5	1	1
Выбор архитектур СНС	Б	2	8	6	3	4
Обучение СНС	Б	10	20	15	14	19
Сравнительный анализ результатов обучения СНС	Б	1	2	1,1	1	1
Согласование выполненной работы с научным руководителем	НР, Б	2	4	3,5	3	3
Выполнение других частей работы (финансовый менеджмент, социальная ответственность)	Б	2	8	6	3	4
Подведение итогов, оформление работы	Б	2	4	3,5	3	3
Итого:	Б	28	72	53,3	37	51
	НР	9	19	14,2	9	11

На рисунке 24 представлена диаграмма Ганта, составленная на основе графика проведения научного исследования.

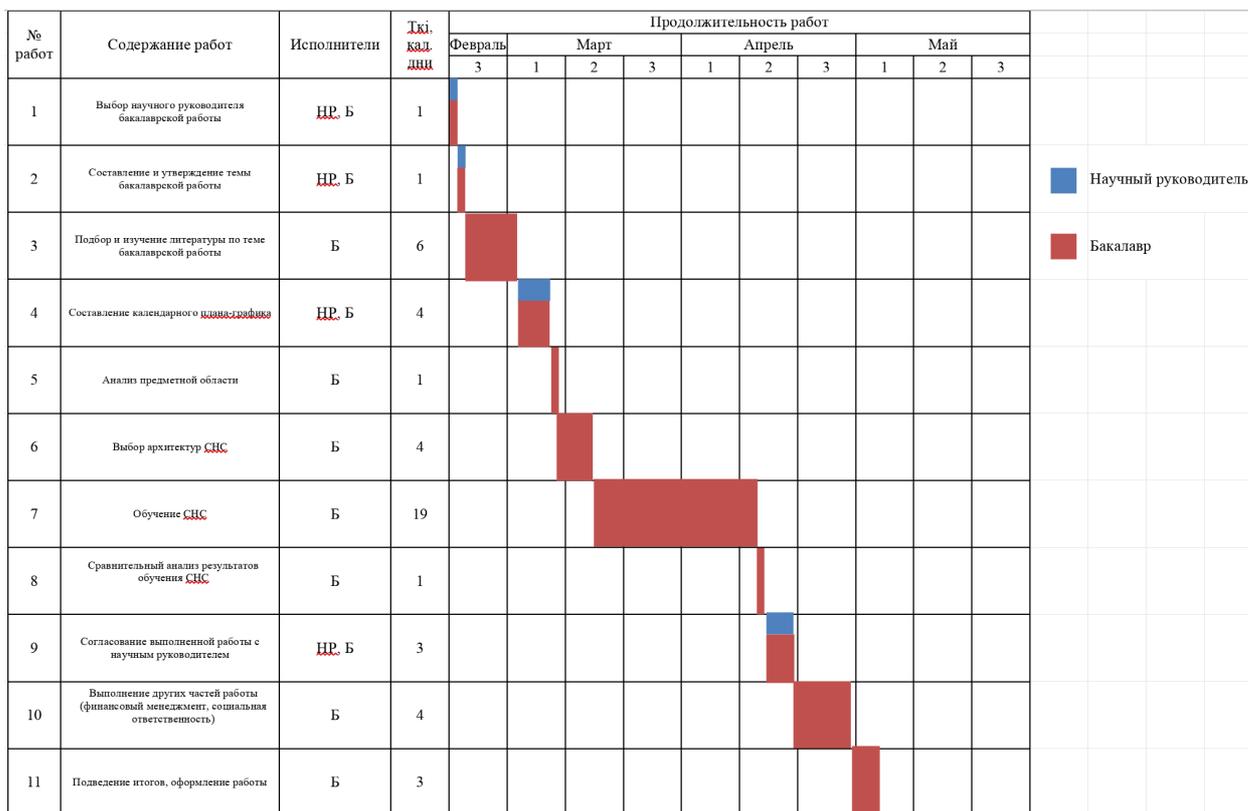


Рисунок 24 – Диаграмма Ганта для исследовательской работы

7.8 Бюджет научно-технического исследования

В состав бюджета входит стоимость всех расходов, необходимых для выполнения работ по проекту. При формировании бюджета используется группировка затрат по следующим статьям:

- затраты на специальное оборудование;
- основная заработная плата исполнителей;
- дополнительная заработная плата исполнителей;
- отчисления во внебюджетные фонды;
- накладные расходы.

7.8.1 Расчет затрат на специальное оборудование для научных (экспериментальных) целей

Данная статья затрат включает в себя затраты на приобретение

специального оборудования. Также в эту статью включаются затраты по доставке и монтажу оборудования, равные 15% от его стоимости.

В ходе работы над проектом использовалось оборудование, имеющееся у исполнителей, соответственно необходим расчет его амортизации.

При исследовании СНС были использованы два персональных компьютера, стоимость 20000 и 80000 рублей, суммарная стоимость 100000 рублей.

Расчет амортизации ПК: первоначальная стоимость ПК 100000 рублей; срок полезного использования для машин офисных код 330.28.23.23 составляет 36 месяцев. Планируемое время использования ПК для написания ВКР - 6 месяцев.

Норма амортизация основных средств линейным способом рассчитывается по формуле 6:

$$A_n = \frac{1}{n} \times 100, \quad (6)$$

где n – установленный срок в месяцах;

A_m – норма амортизации.

Тогда расчет амортизации ПК:

Норма амортизации:

$$A_n = \frac{1}{n} \times 100 = \frac{1}{36} \times 100 = 2,68$$

Ежемесячные амортизационные отчисления (в рублях):

$$A_m = 10000 \times 0,0278 = 2780$$

Итоговая сумма амортизации основных средств (в рублях):

$$A = 2780 \times 6 = 16680$$

Таким образом, сумма затрат на специальное оборудование составляет 16680 рублей, в виде амортизационных отчислений.

7.8.2 Основная заработная плата исполнителей темы

Данная статья затрат включает основную заработную плату, премии и доплаты всех исполнителей проекта. В качестве исполнителей проекта

выступают студенты и научный руководитель.

Зарботная плата рассчитывается по формуле 7:

$$Z_{зп} = Z_{осн} + Z_{доп}, \quad (7)$$

где $Z_{зп}$ – зарботная плата исполнителя;

$Z_{осн}$ – основная зарботная плата исполнителя;

$Z_{доп}$ – дополнительная зарботная плата исполнителя (12%-15% от размера основной зарботной платы).

Основную зарботную плату можно получить по формуле 8:

$$Z_{осн} = Z_{дн} \times T_p \times (1 + K_{пр} + K_d), \quad (8)$$

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

$K_{пр}$ – премиальный коэффициент (0,3);

K_d – коэффициент доплат и надбавок (0,2-0,5);

K_p – районный коэффициент (для Томска 1,3);

T_p – продолжительность работ, выполняемых работником, раб. дни;

T_p – продолжительность работ, выполняемых исполнителем.

Среднедневную зарботную плату можно получить по формуле 9:

$$Z_{дн} = \frac{Z_m \cdot M}{F_d}, \quad (9)$$

где Z_m – месячный должностной оклад исполнителя, рубли; M – количество месяцев работы равно:

При отпуске в 24 рабочих дня $M = 11,2$ месяца, 5 – дневная неделя; При отпуске в 48 рабочих дней $M = 10,4$ месяца, 6 дневная неделя;

F_d – действительный годовой фонд рабочего времени персонала по разработке.

Должностные оклады исполнителей проекта согласно приказу ТПУ представлены в таблице 11.

Таблица 11 – Месячные должностные оклады исполнителей

Исполнитель	Районный коэффициент (для Томска)	Размер месячного должностного оклада без учета коэффициента, рубли
Научный руководитель (должность – профессор, степень – доктор технических наук)	1,3	33664
Студент (ассистент, без степени)	1,3	21760

Баланс рабочего времени для 6-дневной рабочей недели представлен в таблице 12.

Таблица 12 – Баланс рабочего времени (для 6-дневной недели)

Показатели рабочего времени	Дни
Календарные дни	365
Нерабочие дни (праздники/выходные)	118
Действительный годовой фонд рабочего времени	247

На основе формулы 9 и таблиц 15–16 была рассчитана среднедневная заработная плата:

$$З_{\text{дн}}(\text{Б}) = \frac{21760 \times 10,4}{247} = 927,48$$

$$З_{\text{дн}}(\text{НР}) = \frac{33664 \times 10,4}{247} = 1434,86$$

Расчет затрат на основную заработную плату приведен в таблице 13.

Таблица 13 – Затраты на основную заработную плату

Исполнители	Здн, руб.	Кпр	Кд	Кр	Тр	Зосн, руб.
Б	927,48	0,3	0,2	1,3	59	106706,57
НР	1434,86	0,3	0,2	1,3	7	19585,84
Итого:						126292,41

Итоговая сумма затрат на основную заработную плату составила 126292,41 руб.

7.8.3 Дополнительная заработная плата исполнителей темы

Данная статья расходов учитывает величину предусмотренных

Трудовым кодексом РФ доплат за отклонение от нормальных условий труда и выплат, связанных с обеспечением гарантий и компенсаций.

Расчёт дополнительной заработной платы осуществляется по формуле 10:

$$Z_{\text{доп}} = k_{\text{доп}} \cdot Z_{\text{осн}}, \quad (10)$$

где $Z_{\text{доп}}$ – дополнительная заработная плата, рубли;

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

$Z_{\text{осн}}$ – основная заработная плата, рубли.

Расчет затрат на дополнительную зарплату приведен в таблице 14.

Таблица 14 – Затраты на дополнительную заработную плату

Исполнители	$Z_{\text{осн}}$, руб.	$k_{\text{доп}}$	$Z_{\text{доп}}$, руб.
Б	106706,57	0,12	12804,79
НР	19585,84	0,12	2350,30
	Итого:		15155,09

Итоговая сумма затрат на дополнительную заработную плату составила 15155,09 руб.

7.8.4 Отчисления во внебюджетные фонды (страховые отчисления)

К отчислениям во внебюджетные фонды относятся отчисления:

- отчисления органам государственного социального страхования (ФСС);
- отчисления в пенсионный фонд (ПФ);
- отчисления медицинского страхования (ФФОМС).

Сумма отчислений во внебюджетные фонды рассчитывается на основе затрат на оплату труда исполнителей и может быть вычислена по формуле 11.

$$Z_{\text{внеб}} = k_{\text{внеб}} \times (Z_{\text{осн}} + Z_{\text{доп}}), \quad (11)$$

где $k_{\text{внеб}}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и др.); Размер коэффициента определяется законодательно и в настоящее время согласно Федеральному закону от 24.07.2009 №212-ФЗ установлен в размере 30%.

Расчет затрат на отчисления во внебюджетные фонды приведен в таблице 15.

Таблица 15 – Отчисления во внебюджетные фонды

Исполнители	Зосн, руб.	Здоп, руб.	Квнеб	Звнеб, руб.
Студент	106706,57	12804,79	0,30	36092,43
Научный руководитель	19585,84	2350,30	0,30	6624,71
Итого:				42717,14

Итоговая сумма отчислений во внебюджетные фонды составила 42434,25 руб.

7.8.5 Накладные расходы

Накладные расходы – расходы на организацию, управление и обслуживание процесса производства товара, оказания услуги; носят комплексный характер. Накладные расходы вычисляются по формуле 12:

$$Z_{\text{накл}} = \left(\sum \text{статей} \right) \cdot k_{\text{нр}}, \quad (12)$$

где $k_{\text{нр}}$ – коэффициент накладных расходов (16% от суммы затрат, подсчитанных выше).

Расчет накладных расходов приведен в таблице 16.

Таблица 16 – Расчет накладных расходов

Статьи затрат	Сумма, руб.
Затраты на специальное оборудование	16680
Затраты на основную заработную плату	126292,41
Затраты на дополнительную заработную плату исполнителям проекта	15155,09
Затраты на отчисления во внебюджетные фонды	42717,14
Коэффициент накладных расходов	0,16
Накладные расходы	32372,77

Итоговая сумма накладных расходов составила 32372,77 руб.

7.8.6 Формирование бюджета затрат научно-исследовательского проекта

После того, как была подсчитана каждая из статей расходов, можно приступить к формированию общего бюджета затрат проекта. Итоговый бюджет затрат представлен в таблице 17.

Таблица 17 – Расчет бюджета затрат НТИ

Наименование статьи	Сумма, руб.	Удельный вес, %
Затраты на специальное оборудование	16680	7,2
Затраты на основную заработную плату	126292,41	54,3
Затраты на дополнительную заработную плату	15155,09	6,5
Страховые взносы	42717,14	18,2
Накладные расходы	32372,77	13,8
Общий бюджет	233217,41	100

Таким образом, общий бюджет НТИ составляет 233217,41 рубля.

7.8.7 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Определение эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Его нахождение связано с определением двух средневзвешенных величин:

финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования определяется как:

$$I_{\text{фин.р}}^{\text{исп.}i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}, \quad (13)$$

где $I_{\text{фин.р}}^{\text{исп.}i}$ – интегральный финансовый показатель разработки;

Φ_{pi} – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения научно-исследовательского проекта.

Полученная величина интегрального финансового показателя разработки отражает соответствующее численное увеличение бюджета затрат разработки в размах (значение больше единицы), либо соответствующее численное удешевление стоимости разработки в размах (значение меньше единицы, но больше нуля).

Тогда, рассчитывая значение интегрального финансового показателя, учтем, что исп. 1 – обучение модели сверточной нейронной сети для системы компьютерного зрения, исп. 2 – разметка объектов на изображениях вручную.

$$I_{\text{фин.р}}^{\text{исп1}} = \frac{233217,41}{250000} = 0,93$$

$$I_{\text{фин.р}}^{\text{исп2}} = \frac{250000}{250000} =$$

Интегральный показатель ресурсоэффективности вариантов исполнения объекта исследования можно определить следующим образом:

$$I_{pi} = \sum_{i=1}^n a_i \times b_i, \quad (14)$$

где I_{pi} – интегральный показатель ресурсоэффективности для i -го варианта исполнения разработки;

a_i – весовой коэффициент i -го варианта исполнения разработки;

b_i^a, b_i^k – балльная оценка i -го варианта исполнения разработки, устанавливается экспертным путем по выбранной шкале оценивания;

n – число параметров сравнения.

В таблице 18 приведена сравнительная оценка характеристик вариантов исполнения проекта

Таблица 18 – Сравнительная оценка характеристик вариантов исполнения проекта

Критерии \ Объект исследования	Весовой коэффициент параметра	Исп.1	Исп.2
1. Удобство конечным пользователям	0,3	5	4
2. Удобство администраторам	0,3	3	4
3. Масштабируемость	0,1	5	2
4. Гибкость	0,1	4	3
5. Отказоустойчивость	0,2	5	4
Итого	1	4,8	4,1

$$I_{p-исп.1} = 5 \cdot 0,3 + 3 \cdot 0,3 + 5 \cdot 0,1 + 4 \cdot 0,1 + 5 \cdot 0,2 =$$

$$I_{p-исп.2} = 4 \cdot 0,3 + 4 \cdot 0,3 + 2 \cdot 0,1 + 3 \cdot 0,1 + 4 \cdot 0,2 =$$

Интегральный показатель эффективности вариантов исполнения разработки ($I_{испi}$) определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп1} = \frac{I_{p-исп1}}{I_{фин п}} = \frac{4,3}{0,932} = 4,6$$

$$I_{исп2} = \frac{I_{p-исп2}}{I_{фин п}} = \frac{3,7}{1} = 3,7$$

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволит определить сравнительную эффективность проекта и выбрать наиболее целесообразный вариант из предложенных.

Сравнительная эффективность проекта ($\mathcal{E}_{ср}$):

$$\mathcal{E}_{ср} = \frac{I_{исп}}{I_{исп,}}, \quad (15)$$

В таблице 19 приведена сравнительная эффективность разработки

Таблица 19 – Сравнительная эффективность разработки

№	Показатели	Исп.1	Исп.2
1	Интегральный финансовый показатель разработки	0,932	1
2	Интегральный показатель ресурсоэффективности разработки	4,3	3,7
3	Интегральный показатель эффективности	4,61	3,7
4	Сравнительная эффективность вариантов исполнения	1,24	

Сравнив значения интегральных показателей эффективности, можно

сделать вывод, что использование модели СНС является более эффективным вариантом решения задачи, поставленной в данной работе с позиции финансовой и ресурсной эффективности.

7.9 Определение потенциального эффекта исследования

В рамках данного раздела был проведен комплексный SWOT-анализ проекта, который позволил выявить его сильные и слабые стороны, а также определить соответствие его характеристик внешним факторам. Более того, были выявлены возможные пути дальнейшего развития и совершенствования проекта для повышения его конкурентоспособности. Также была произведена оценка качества и перспективности данного проекта.

Ещё одним ключевым моментом, рассмотренным в данном разделе, является планирование работ, выполняемых в рамках проекта. Составленный список необходимых работ с распределением исполнителей, а также вычисленная трудоемкость, позволили построить наглядный план-график работ в виде диаграммы Ганта. Это позволило рационально распорядиться временными ресурсами, отведенными на выполнение проекта.

Кроме того, был определен бюджет проекта. Определены затраты на заработные платы исполнителям с учетом страховых отчислений, рассчитаны накладные расходы, а также амортизация оборудования, задействованного в процессе выполнения проекта.

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа		ФИО	
8K92		Беляев Семен Игоревич	
Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.04 Программная инженерия

Тема ВКР:

Программная реализация и анализ эффективности моделей сверточных нейронных сетей YOLO для решения задач классификации объектов на изображениях	
Исходные данные к разделу «Социальная ответственность»:	
<p>Введение</p> <ul style="list-style-type: none"> Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения. Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации 	<p><i>Объект исследования:</i> Программная реализация моделей сверточных нейронных сетей семейства YOLO для решения задач классификации объектов на изображениях</p> <p><i>Область применения:</i> промышленные и частные системы компьютерного зрения для распознавания объектов в видимом диапазоне</p> <p><i>Рабочая зона:</i> офис</p> <p><i>Размеры помещения:</i> 5*10 м.</p> <p><i>Количество и наименование оборудования рабочей зоны:</i> компьютер с графическим процессором</p> <p><i>Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне:</i> Проектирование системы, обучение модели нейронной сети и ее тестирование</p>
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p>1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения:</p> <ul style="list-style-type: none"> специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; организационные мероприятия при компоновке рабочей зоны. 	<p>СанПиН 1.2.3685-21. Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания; ТК РФ Статья 91. Понятие рабочего времени. Нормальная продолжительность рабочего времени; ГОСТ 12.2.032-78. Рабочее место при выполнении работ сидя;</p>
<p>2. Производственная безопасность при разработке проектного решения:</p> <ul style="list-style-type: none"> Анализ выявленных вредных и опасных производственных факторов 	<p>Вредные производственные факторы:</p> <ul style="list-style-type: none"> Умственное перенапряжение, в том числе вызванное информационной нагрузкой; Нагрузка на зрительный аппарат; Отсутствие или недостаток необходимого искусственного освещения и пониженная световая и цветовая контрастность; <p>Опасные факторы: Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий.</p> <p>Требуемые средства индивидуальной и коллективной защиты от выявленных факторов: осветительные приборы, заземляющие устройства для электронного оборудования, предохранительные устройства в случае короткого замыкания.</p>
<p>3. Экологическая безопасность при разработке проектного решения</p>	<p>Воздействие на селитебную зону отсутствует Воздействие на литосферу возможно в следствии</p>

	<p>утилизации бытовых отходов в процессе разработки</p> <p>Воздействие на гидросферу возможно в следствие утилизации бытовых отходов в процессе разработки</p> <p>Воздействие на атмосферу возможно в следствие утилизации бытовых отходов в процессе разработки</p>
4. Безопасность в чрезвычайных ситуациях при разработке проектного решения	<p>Возможные ЧС:</p> <p>Природные катастрофы (наводнения и т.д.);</p> <p>Геологические воздействия (землетрясения и т.д.);</p> <p>Техногенные аварии (пожары);</p> <p>Наиболее типичная ЧС: пожар</p>
Дата выдачи задания для раздела по линейному графику	

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Мезенцева Ирина Леонидовна			

Задание приняли к исполнению студенты:

Группа	ФИО	Подпись	Дата
8К92	Беляев Семен Игоревич		

8 Социальная ответственность

Объектом исследования являются сверточные нейронные сети класса YOLO и задача детектирования объектов разных масштабов на изображениях.

Выпускная квалификационная работа представляет собой исследование сверточных нейронных сетей YOLOv4-Tiny, YOLOv5n , YOLO Nano и YOLOv7-Tiny при решении задачи детектирования объектов разных масштабов на изображениях.

В настоящее время количество актуальных приложений компьютерного зрения продолжает расти. В частности, для решения становятся доступными задачи, связанные с анализом видеоданных в режиме реального времени, основанные на применении моделей сверточных нейронных сетей. Таким образом, областью применения работы являются промышленные и частные системы компьютерного зрения для распознавания объектов в видимом диапазоне.

Все работы, выполняемые пользователем, производятся за столом в офисном помещении и заключаются в проектировании, обучении и тестировании моделей сверточных нейронных сетей.

Офисное помещение длиной 10 метров и шириной 5 метров оборудовано компьютером с графическим процессором и источником бесперебойного питания.

В данном разделе выпускной квалификационной работы рассмотрены опасные и вредные факторы, оказывающие влияние на производственную деятельность программиста. Также рассмотрены вопросы техники пожарной безопасности, правила поведения во время чрезвычайных ситуаций и способы предотвращения их возникновения.

8.1 Правовые и организационные вопросы обеспечения безопасности

8.1.1 Правовые нормы трудового законодательства

В статье 108 Трудового Кодекса РФ «Перерывы для отдыха и питания» сказано, что в течение рабочего дня работнику должен быть предоставлен перерыв продолжительностью не более двух часов и не менее 30 минут, который в рабочее время не включается.

Согласно статье 163 ТК РФ «Обеспечение нормальных условий работы для выполнения норм выработки», работодатель обязан обеспечить:

- исправное состояние помещений и оборудования;
- условия труда, соответствующие требованиям охраны труда и безопасности производства.

Согласно статье 212 ТК РФ «Обязанности работодателя по обеспечению безопасных условий и охраны труда», работодатель обязан обеспечить:

- безопасность работников при эксплуатации зданий, оборудования, осуществлении технологических процессов, применяемых материалов;
- создание и функционирование системы управления охраной труда;

8.1.2 Организационные мероприятия при компоновке рабочей зоны

Выполнение работ производится в основном в сидячем положении, поэтому следует выполнять все рекомендуемые требования ГОСТ 12.2.032–78 о рабочем месте при выполнении работ сидя. В соответствии с ГОСТ 12.2.032-78. «Система стандартов безопасности труда». Рабочее место при выполнении работ сидя» рабочий стол может быть любой конструкции,

отвечающей современным требованиям эргономики и позволяющей удобно разместить на рабочей поверхности оборудование с учетом его количества, размеров и характер выполняемой работы

Выполнение требований на данном рабочем месте отражено ниже в таблице 20, согласно ГОСТ 12.2.032-78.

Таблица 20 – Требования к организации рабочего места при работе с ПЭВМ

Требование	Требуемое значение	Значение параметров в помещении
Высота рабочей поверхности стола	Регулируемая высота (680-800мм) Нерегулируемая высота (725мм)	Нерегулируемая высота (690мм)
Рабочий стул	Подъемно-поворотный, регулируемый по высоте и углу наклона спинки	Соответствует
Расположение монитора от глаз пользователя	600-700мм	Соответствует

Средства отображения информации необходимо группировать и располагать группы относительно друг друга в соответствии с последовательностью их использования. При этом средства отображения информации необходимо размещать в пределах групп так, чтобы последовательность их использования осуществлялась слева направо или сверху вниз.

8.2 Производственная безопасность

Согласно производственным факторам ГОСТ 12.0.003-2015 ССБТ, на оператора ПЭВМ в течение рабочего дня воздействует множество различных производственных факторов, каждый из которых влияет на производительность, работоспособность и физическое состояние.

В таблице 21 представлены возможные вредные и опасные факторы,

возникающие при работе за ПК.

Таблица 21 – Возможные опасные и вредные производственные факторы на рабочем месте инженера-программиста по ГОСТ 12.0.003-2015

Факторы ГОСТ 12.0.003-2015)	Нормативные документы
Нагрузка на зрительный аппарат	СП 52.13330.2016 «Естественное и искусственное освещение». Актуализированная редакция СНиП 23-05-95
Отсутствие или недостаток необходимого искусственного освещения и пониженная световая и цветовая контрастность	СП 52.13330.2016 «Естественное и искусственное освещение». Актуализированная редакция СНиП 23-05-95
Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий	ГОСТ 12.1.019-2017 «Электробезопасность. Общие требования и номенклатура видов защиты»
Умственное перенапряжение, в том числе вызванное информационной нагрузкой	Р 2.2.2006-05 «Гигиена труда. Руководство по гигиенической оценке факторов рабочей среды и трудового процесса. Критерии и классификация условий труда»

8.2.1 Вредные производственные факторы

8.2.1.1 Нагрузка на зрительный аппарат

Работа на ПК сопровождается постоянным и значительным напряжением функций зрительного анализатора. Как и недостаточная освещенность, чрезмерная работы за ПК может вызвать ухудшение зрения работника.

Спектр излучения компьютера включает в себя рентгеновскую и ультрафиолетовую области спектра, а также широкий диапазон электромагнитных волн других частот. Опасность рентгеновских лучей считается сейчас специалистами пренебрежимо малой, поскольку этот вид

лучей поглощается веществом экрана.

Допустимые уровни ультрафиолетового излучения для мониторов регулируются в соответствии с СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» и указаны в таблице 22.

Таблица 22 – Допустимые уровни ультрафиолетового излучения

Вид изделий	Спектральный диапазон длин волн, нм	Допустимая интенсивность облучения, Вт/м ²
Экраны телевизоров, видеомониторов, осциллографов измерительных и других приборов, средств отображения информации с визуальным контролем	Свыше 315 до 400	Не более 0,1
	Свыше 280 до 315	Не более 0,0001
	От 200 до 280	Не допускается

Чтобы снизить зрительное напряжение, необходимо выбрать такой монитор, параметры которого удовлетворяли бы параметрам таблицы 4. Соблюдение этих норм позволит снизить зрительное напряжение.

8.2.1.2 Отсутствие или недостаток необходимого искусственного освещения и пониженная световая и цветовая контрастность

Недостаточная освещенность приводит к понижению работоспособности, а также может вызвать проблемы со здоровьем, а именно может негативно повлиять на качество зрения работников.

Согласно СП 52.13330.2016 зрительную работу пользователя программного обеспечения можно характеризовать как работу разряда Б – высокой точности (наименьший эквивалентный размер объекта различения составляет 0,3-0,5 мм), под разряда 1 (относительная продолжительность зрительной работы при направлении зрения на рабочую поверхность не менее 70%).

Рабочие столы следует размещать таким образом, чтобы мониторы были расположены боковой стороной к источникам естественного света. Следует применять системы комбинированного освещения.

Контраст объекта различения с фоном K . Является отношением абсолютной величины разности между яркостью объекта и фона к яркости фона. Контраст объекта различения с фоном считается большим при K более 0,5; средним - при K от 0,2 до 0,5; пониженным - при K менее 0,2 согласно СанПиН 1.2.3685-21. Для работы за ПК ввиду класса работы высокой точности контрастность нужно устанавливать среднюю.

В таблице 23 приведены требования к искусственному освещению рабочего помещения для вышеуказанного разряда

Таблица 23 – Требования к искусственному освещению рабочего помещения

Освещенность на рабочей поверхности от системы общего освещения, лк	Цилиндрическая освещенность, лк	Объединенный показатель UGR, не более	Коэффициент пульсации освещенности, Кп, %, не более
400	100	19	15

8.2.1.3 Умственное перенапряжение, в том числе вызванное информационной нагрузкой

Большую часть своего рабочего времени разработчики занимаются умственным трудом, заключающемся в решении интеллектуальных задач. Избыток подобных задач в течении дня приводит к появлению нервно-психических перегрузок у рабочего, что может вести к появлению самых разных симптомов, таких как приступы головной боли, повышенная раздражительность и заторможенность, а также крайне негативно сказывается на производительности разработчиков.

Согласно СанПиН 1.2.3685-21 для более эффективного выполнения работы разработчику следует по возможности делать регулярные перерывы

в работе за ПК.

8.2.2 Опасные производственные факторы

Персональный компьютер предоставляет для человека большую потенциальную опасность, так как в процессе эксплуатации или проведения профилактических работ человек может коснуться комплектующих компьютера, находящихся под напряжением.

Прохождение электрического тока через тело человека сопровождается нарушением функций внутренних органов или ожогом тканей. Поражение током приводит к остановке сердца, судорогам, потере сознания.

Электробезопасность подразумевает под собой систему мероприятий, технических и организационных, направленных на защиту людей от опасного воздействия электрического тока, статического электричества и электромагнитного поля. Значения вышеперечисленных факторов регулируются ГОСТ Р 58698-2019.

Меры предосторожности для основной защиты от повреждения электрическим током:

- использование защитных ограждений или оболочек и заземляющего устройства;
- размещение опасных для жизни и здоровья человека участков электропроводов и приборов вне зоны досягаемости рукой;
- автоматическое отключение питания (защитное устройство, которое будет отключать систему, питающую электрическое оборудование или установку в случае замыкания).

Меры защиты:

- посредством системы безопасного сверхнизкого напряжения (БСНН) и защитного сверхнизкого напряжения (ЗСНН).

8.3 Экологическая безопасность

Правильная утилизация компьютерной техники, электроники и аккумуляторов является важным фактором для сохранения окружающей среды и здоровья людей. Компьютерная техника и электроника включают в себя тяжелые металлы такие как ртуть, которые могут быть опасными для окружающей среды, если они не утилизируются правильно.

Предельная допустимая концентрации (ПДК) ртути в почве регулируется СанПиН 1.2.3685-21 и равняется 2.1 мг/кг. В жидкокристаллических мониторах с диагональю 26 дюймов содержится от 2 до 4 U-образных ламп массой по 13 г, которые содержат ртуть в количестве 62,14 мг/кг, что превышает норму допустимого содержания в почве в 30 раз. ПДК ртути в воде равна 500 нг/л, для атмосферы ПДК равна 0,0003 мг/м³. Неутилизированные компьютеры и оргтехника вызывают ртутное загрязнение литосферы и гидросферы, а при сжигании – и атмосферы.

Вышедшие из строя ПЭВМ и сопутствующая оргтехника относится к IV классу опасности и подлежит специальной утилизации. Для оказания наименьшего влияния на окружающую среду, необходимо проводить специальную процедуру утилизации ПЭВМ и оргтехники, при которой более 90% отправится на вторичную переработку и менее 10% будут отправлены на свалки. При этом она должна соответствовать процедуре утилизации ГОСТ Р 53692-2009 «Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов».

На селитебную зону влияние не оказывается.

8.4 Безопасность в чрезвычайных ситуациях

Возможными чрезвычайными ситуациями на рабочем месте могут быть землетрясение, наводнение, пожар и другие. Наиболее типичной чрезвычайной ситуацией является пожар, так как специфика работы

заклучается во взаимодействии с вычислительной техникой, что может являться причиной возгорания и угрожать жизни людей.

На основании Федерального закона от 22.07.2008 N 123-ФЗ (ред. от 30.04.2021) "Технический регламент о требованиях пожарной безопасности" был определен класс возможного пожара по виду горючего материала – Класс Е (пожары горючих веществ и материалов электроустановок, находящихся под напряжением). Для тушения пожара такого класса, причиной которого стало возгорание ПЭВМ, достаточно воспользоваться переносным и передвижным огнетушителем, находящимся в здании офиса.

Причинами возгорания при работе с вычислительной техникой могут быть:

- токи короткого замыкания;
- перегрев и дальнейшее воспламенение;
- неисправность работы техники или электросети;

Причинами возникновения КЗ могут быть ошибки при проектировании, старение изоляции, увлажнение изоляции и механические перегрузки. Пожарная опасность при перегрузках – чрезмерное нагревание отдельных элементов, которое может происходить при ошибках проектирования в случае длительного прохождения тока, превышающего номинальное значение.

Рабочее место программиста должно соответствовать требованиям ФЗ Технический регламент по ПБ и норм пожарной безопасности (НПБ 105-03) и удовлетворять требованиям по предотвращению и тушению пожара по ГОСТ 12.1.004-91 и СНиП 21-01-97.

В случае возникновения пожара в здании автоматически срабатывают датчики дыма, и звуковая система оповещает всех сотрудников о немедленной эвакуации из здания, и далее сотрудники направляются на выход в соответствии с планом эвакуации при пожарах и других ЧС.

Для предотвращения возникновения пожара необходимо:

- регулярно проводить инструктажи сотрудников предприятия по пожарной безопасности;
- разместить в помещении план эвакуации и плакаты с краткой информацией с действиями при возникновении пожара;
- соблюдать правила и нормы при монтаже электронных приборов и проведении электрической проводки;
- оборудовать помещение пожарной сигнализацией и красными кнопками, а также средствами тушения пожара.

Если все же не удалось предотвратить пожар, то каждый сотрудник должен:

- незамедлительно сообщить об этом в пожарную охрану;
- принять меры по эвакуации людей, каких-либо материальных ценностей согласно плану эвакуации;
- отключить электроэнергию, приступить к тушению пожара первичными средствами пожаротушения.

8.5 Вывод по разделу

В ходе работы по разделу были выявлены основные законодательные акты, которые обеспечивают безопасность жизнедеятельности на рабочем месте. Также были рассмотрены наиболее значимые опасные и вредные факторы, связанные с разработкой программного решения, а также описано влияние этого процесса на окружающую среду и необходимые меры для уменьшения негативного воздействия. Кроме того, были проанализированы возможные чрезвычайные ситуации и принятые меры для их предотвращения.

Рабочее место выполнения выпускной квалификационной работы соответствует выявленным требованиям. Для обеспечения необходимого уровня освещения используются несколько энергосберегающих ламп и дополнительное настольное освещение. Для защиты от воздействия

электрического тока проводится проверка состояния ПЭВМ и соблюдаются правила безопасности при работе с ней.

Согласно правилам установок электроустановок, рабочее помещение инженера-программиста относится к помещениям без повышенной опасности поражения электрическим током.

Рабочее помещение оборудовано в соответствии с требованиями электро- и пожарной безопасности. Работа программиста относится к категории тяжести труда Ia и требует 1 группы по электробезопасности.

Рабочее помещение является пожароопасным и относится к категории В, а возможный пожар может относиться к классу А, так и классу Е.

Эксплуатация ПЭВМ не оказывает значительного вреда экологии, однако неправильная утилизация комплектующих ПЭВМ и макулатуры может оказывать негативное воздействие на литосферу, атмосферу и литосферу.

Заключение

При выполнении ВКР получены следующие наиболее важные результаты и сделаны выводы.

1. Проанализированы модели СНС для детектирования летающих объектов на изображениях с целью создания мобильных СКЗ. Выявлено, что наиболее перспективными для реализации этой цели являются модели СНС класса YOLO. Для последующего исследования выбраны четыре компактные модели из этого класса СНС как наиболее отвечающие сформированным требованиям.

2. Разработаны программы для подготовки датасета. Программно реализована модель СНС YOLO Nano.

3. Сформирован набор оптических изображений с летающими объектами трех классов: «Птица», «БПЛА вертолетного типа» и «БПЛА самолетного типа». На изображениях находятся один или более объектов разных размеров и разных классов. На основе этого набора создан датасет в виде размеченных изображений. В него включено 3400 изображений и файлов аннотаций.

4. Проведен детальный анализ архитектуры компактных СНС YOLOv4 - Tiny, YOLOv5n, YOLOv7-Tiny и YOLO Nano, позволивший задать оптимальные значения параметров этих моделей при обучении.

5. Каждая из четырех компактных моделей СНС настроена, обучена и валидирована с использованием обучающей и валидационной выборок датасета.

6. Проведены исследования эффективности каждой из четырех моделей СНС с использованием тестовой выборки датасета. Показано, что точность детектирования летающих объектов на изображениях по метрикам $AP_{0,5}$ и $mAP_{0,5}$ наибольшая при использовании модели СНС YOLOv7-tiny. При этом требование по точности детектирования объектов на изображениях удовлетворяется только для этой модели. Выявлено, что для всех моделей

СНС скорость детектирования объектов значительно превышает пороговое значение FPS, равное 25, а алгоритмическая эффективность этих моделей позволяет им выполняться на вычислительных устройствах мобильных СКЗ даже с весьма скромными ресурсами.

7. По результатам всех исследований модель СНС YOLOv7-tiny рекомендована к использованию в мобильных СКЗ реального времени.

8. Результаты работы докладывались на трех международных научно–практических конференциях студентов, аспирантов и молодых учёных, получен диплом за лучший доклад. Основные результаты работы опубликованы в виде трех статей в сборниках этих конференций.

Список публикаций

Список публикаций по теме ВКР

1. Ротарь В. Г., Беляев С. И. Обнаружение автомобильных аварий в реальном времени с применением нейросети архитектуры YOLOv3 //Молодежь и современные информационные технологии: сборник трудов XVIII Международной научно-практической конференции студентов, аспирантов и молодых учёных, 22-26 марта 2021 г., г. Томск. – Томск: изд-во Томского политехнического университета, 2021. – С. 49-50.

2. Беляев С. И., Мыцко Е. А. Детектирование трудноразличимых объектов с применением сверточной нейронной сети YOLO Nano //Молодежь и современные информационные технологии: сборник трудов XIX Международной научно-практической конференции студентов, аспирантов и молодых учёных, 21-25 марта 2022 г., г. Томск. – Томск: изд-во Томского политехнического университета, 2022. – С. 14-15.

3. Беляев С. И. Анализ эффективности компактных моделей сверточных нейронных сетей класса YOLO при решении задач классификации объектов на изображениях//Молодежь и современные информационные технологии: сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых учёных, 20-22 марта 2023 г., г. Томск. – Томск: изд - во Томского политехнического университета, 2023. – С. 60-63.

Список иных публикаций

4. Беляев С. И. Распознавание автомобильных номеров методом Виолы-Джонса //Перспективы развития фундаментальных наук: сборник научных трудов XVII Международной конференции студентов, аспирантов и молодых ученых, г. Томск, 21-24 апреля 2020 г. , секция IT-технологии и электроника.– Томск: изд-во ТУСУР, 2020. – Т. 7. – С. 29-31.

Список источников

1. Voulodimos A. et al. Deep learning for computer vision: A brief review //Computational intelligence and neuroscience. – 2018. – Т. 2018.
2. Alzubaidi L. et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions //Journal of big Data. – 2021. – Т. 8. – С. 1-74.
3. Hassan S. A., Rahim T., Shin S. Y. Real-time uav detection based on deep learning network //2019 International Conference on Information and Communication Technology Convergence (ICTC). – IEEE, 2019. – С. 630-632.
4. Tang W., Hua G., Wang L. How to train a compact binary neural network with high accuracy? //Proceedings of the AAAI conference on artificial intelligence. – 2017. – Т. 31. – №. 1.
5. 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.
6. Wang S. Research towards Yolo-series algorithms: comparison and analysis of object detection models for real-time UAV applications //Journal of Physics: Conference Series. – IOP Publishing, 2021. – Т. 1948. – №. 1. – С. 012021.
7. Lu X. et al. MimicDet: bridging the gap between one-stage and two-stage object detection //Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16. – Springer International Publishing, 2020. – С. 541-557.
8. 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.
9. Purkait P., Zhao C., Zach C. SPP-Net: Deep absolute pose regression with synthetic views //arXiv preprint arXiv:1712.03452. – 2017.
10. Girshick R. Fast r-cnn //Proceedings of the IEEE international

conference on computer vision. – 2015. – C. 1440-1448.

11. Redmon J., Farhadi A. Yolov3: An incremental improvement //arXiv preprint arXiv:1804.02767. – 2018.

12. Lin T. Y. et al. Feature pyramid networks for object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – C. 2117-2125.

13. Girshick R. Fast r-cnn //Proceedings of the IEEE international conference on computer vision. – 2015. – pp. 1440-1448.

14. Ma J., Chen L., Gao Z. Hardware implementation and optimization of tiny-YOLO network //Digital TV and Wireless Multimedia Communication: 14th International Forum, IFTC 2017, Shanghai, China, November 8-9, 2017, Revised Selected Papers 14. – Springer Singapore, 2018. – pp. 224-234.

15. Wong A. et al. YOLO nano: A highly compact you only look once convolutional neural network for object detection //2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMC2-NIPS). – IEEE, 2019. – pp. 22-25.

16. T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2117–2125, 2017.

17. Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 8759–8768, 2018.

18. M. Tan, R. Pang, Q. V. Le. EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

19. Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE

Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 40(4):834–848, 2017.

20. Ali J. B. et al. Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations //Engineering Applications of Artificial Intelligence. – 2015. – T. 42. – C. 67-81.

21. Qiu M., Huang L., Tang B. H. ASFF-YOLOv5: multielement detection method for road traffic in UAV images based on multiscale feature fusion //Remote Sensing. – 2022. – T. 14. – №. 14. – C. 3498.

22. Ghiasi G., Lin T. Y., Le Q. V. Nas-fpn: Learning scalable feature pyramid architecture for object detection //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – C. 7036-7045.

23. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 37(9):1904–1916, 2015.

24. Songtao Liu, Di Huang, et al. Receptive Field Block Net for Accurate and Fast Object Detection. In Proceedings of the European Conference on Computer Vision (ECCV), pages 385–400, 2018.

25. Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. CBAM: Convolutional Block Attention Module. In Proceedings of the European Conference on Computer Vision (ECCV), pages 3–19, 2018.

26. T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, P. Dollár. Microsoft COCO: Common Objects in Context. arXiv preprint arXiv:1405.0312, 2014.

27. Wang C. Y., Bochkovskiy A., Liao H. Y. M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2023. – C. 7464-7475.

28. A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao. YOLOv4: Optimal

Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934, 2020.

29. Liu Y. et al. A survey and performance evaluation of deep learning methods for small object detection //Expert Systems with Applications. – 2021. – Т. 172. – С. 114602.

30. Shorten C., Khoshgoftaar T. M. A survey on image data augmentation for deep learning //Journal of big data. – 2019. – Т. 6. – №. 1. – С. 1-48.

31. Weiss K., Khoshgoftaar T. M., Wang D. D. A survey of transfer learning //Journal of Big data. – 2016. – Т. 3. – №. 1. – С. 1-40.

32. Duan S. et al. An anchor box setting technique based on differences between categories for object detection //International Journal of Intelligent Robotics and Applications. – 2022. – Т. 6. – №. 1. – С. 38-51.

33. Подготовка данных для расширенного машинного обучения [Электронный ресурс] // Microsoft: [сайт] – Режим доступа: <https://docs.microsoft.com/ru-ru/azure/architecture/data-science-process/prepare-data> свободный. – Загл. с экрана (дата обращения 15.02.2023)

34. Bochkovskiy A., Wang C. Y., Liao H. Y. M. Yolov4: Optimal speed and accuracy of object detection //arXiv preprint arXiv:2004.10934. – 2020.

35. Jocher G. et al. ultralytics/yolov5: v5. 0-YOLOv5-P6 1280 models, AWS, Supervise. ly and YouTube integrations //Zenodo. – 2021.

36. PyTorch [Электронный ресурс] – Режим доступа: <https://pytorch.org> свободный. – Загл. с экрана (дата обращения 15.05.2023)

37. Roboflow [Электронный ресурс] – Режим доступа: <https://roboflow.com> свободный. – Загл. с экрана (дата обращения 15.05.2023)

38. Кузьмина Е.А, Кузьмин А.М. Методы поиска новых идей и решений «Методы менеджмента качества» №1 2003 г.

39. Кузьмина Е.А, Кузьмин А.М. Функционально-стоимостный анализ. Экскурс в историю. «Методы менеджмента качества» №7 2002 г.

40. Основы функционально-стоимостного анализа: Учебное пособие / Под ред. М.Г. Карпунина и Б.И. Майданчика. - М.: Энергия, 1980. - 175 с.

41. Скворцов Ю.В. Организационно-экономические вопросы в

дипломном проектировании: Учебное пособие. – М.: Высшая школа, 2006. – 399 с.

42. СП 52.13330.2016 (ред. от 28.12.2021) [Электронный ресурс]. – URL: <https://ac-mos.ru/about/price-expert/chatbot/market-price/download/rent/44-ФЗ/acts/12>. (дата обращения: 07.04.2023)

43. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. От 19.12.2022, с изм. от 11.04.2023) [Электронный ресурс]. – URL: https://www.consultant.ru/document/cons_doc_LAW_34683/ (дата обращения: 23.04.2023)

44. ГОСТ 12.1.019-2017 [Электронный ресурс]. – URL: <https://docs.cntd.ru/document/1200161238> (дата обращения: 15.05.2023)

45. СП 52.13330.2016 [Электронный ресурс]. – URL: https://energy.midural.ru/images/Upload/2017/101/SPEIO_07.11.2016_777.pdf (дата обращения: 12.04.2023)

46. СНиП 23-05-95 [Электронный ресурс]. – URL: <https://kola.rosavtodor.gov.ru/storage/app/media/kola/uploaded-files/08-16-sp-52133302016-estestvennoe-i-iskusstvennoe-osveshchenie-aktualizirovannaya-redaktsiya-snip-23-05-95-1.pdf> (дата обращения: 11.04.2023)

Приложение А

(обязательное)

Программная реализация модели СНС YOLO Nano на языке Python

```
import torch

import torch.nn as nn
import torch.nn.functional as F

from .basic_layers import conv1x1, conv3x3, EP, PEP, FCA, YOLOLayer
from utils.stats import build_targets, to_cpu, non_max_suppression

class YOLONano(nn.Module):
    def __init__(self, num_classes, image_size):
        super(YOLONano, self).__init__()
        self.num_classes = num_classes
        self.image_size = image_size
        self.num_anchors = 3
        self.yolo_channels = (self.num_classes + 5) * self.num_anchors

        anchors52 = [[10,13], [16,30], [33,23]]
        anchors26 = [[30,61], [62,45], [59,119]]
        anchors13 = [[116,90], [156,198], [373,326]]

        self.conv1 = conv3x3(3, 12, stride=1)
        self.conv2 = conv3x3(12, 24, stride=2)
        self.pep1 = PEP(24, 24, 7, stride=1)
        self.ep1 = EP(24, 70, stride=2)
        self.pep2 = PEP(70, 70, 25, stride=1)
        self.pep3 = PEP(70, 70, 24, stride=1)
        self.ep2 = EP(70, 150, stride=2)
        self.pep4 = PEP(150, 150, 56, stride=1)
        self.conv3 = conv1x1(150, 150, stride=1)
        self.fca1 = FCA(150, 8)
        self.pep5 = PEP(150, 150, 73, stride=1)
        self.pep6 = PEP(150, 150, 71, stride=1)

        self.pep7 = PEP(150, 150, 75, stride=1)
        self.ep3 = EP(150, 325, stride=2)
        self.pep8 = PEP(325, 325, 132, stride=1)
        self.pep9 = PEP(325, 325, 124, stride=1)
        self.pep10 = PEP(325, 325, 141, stride=1)
        self.pep11 = PEP(325, 325, 140, stride=1)
```

```

self.pep12 = PEP(325, 325, 137, stride=1)
self.pep13 = PEP(325, 325, 135, stride=1)
self.pep14 = PEP(325, 325, 133, stride=1)

self.pep15 = PEP(325, 325, 140, stride=1)
self.ep4 = EP(325, 545, stride=2) # output: 13x13x545
self.pep16 = PEP(545, 545, 276, stride=1)
self.conv4 = conv1x1(545, 230, stride=1)
self.ep5 = EP(230, 489, stride=1)
self.pep17 = PEP(489, 469, 213, stride=1)

self.conv5 = conv1x1(469, 189, stride=1)
self.conv6 = conv1x1(189, 105, stride=1)
self.pep18 = PEP(430, 325, 113, stride=1)
self.pep19 = PEP(325, 207, 99, stride=1)

self.conv7 = conv1x1(207, 98, stride=1)
self.conv8 = conv1x1(98, 47, stride=1)
self.pep20 = PEP(197, 122, 58, stride=1)
self.pep21 = PEP(122, 87, 52, stride=1)
self.pep22 = PEP(87, 93, 47,
stride=1)
self.conv9 = conv1x1(93, self.yolo_channels, stride=1, bn=False)
self.yolo_layer52 = YOLOLayer(anchors52, num_classes,
img_dim=image_size)

self.ep6 = EP(98, 183, stride=1)
self.conv10 = conv1x1(183,
self.yolo_channels, stride=1, bn=False) # output: 26x26x yolo_channels
self.yolo_layer26 = YOLOLayer(anchors26, num_classes,
img_dim=image_size)

self.ep7 = EP(189, 462, stride=1)
self.conv11 = conv1x1(462, self.yolo_channels, stride=1, bn=False) # output:
13x13x yolo_channels
self.yolo_layer13 = YOLOLayer(anchors13, num_classes,
img_dim=image_size)

for m in self.modules():
    if isinstance(m, nn.Conv2d):
        m.weight = nn.init.xavier_normal_(m.weight, gain=0.02)
    elif isinstance(m, nn.BatchNorm2d):
        nn.init.normal_(m.weight.data, 1.0, 0.02)
        m.bias.data.zero_()

```

```

def forward(self, x, targets=None):
    loss = 0
    yolo_outputs = []
    image_size = x.size(2)

    out = self.conv1(x)
    out = self.conv2(out)
    out = self.pep1(out)
    out = self.ep1(out)
    out = self.pep2(out)
    out = self.pep3(out)
    out = self.ep2(out)
    out = self.pep4(out)
    out = self.conv3(out)
    out = self.fca1(out)
    out = self.pep5(out)
    out = self.pep6(out)

    out_pep7 = self.pep7(out)
    out = self.ep3(out_pep7)
    out = self.pep8(out)
    out = self.pep9(out)
    out = self.pep10(out)
    out = self.pep11(out)
    out = self.pep12(out)
    out = self.pep13(out)
    out = self.pep14(out)

    out_pep15 = self.pep15(out)
    out = self.ep4(out_pep15)
    out = self.pep16(out)
    out = self.conv4(out)
    out = self.ep5(out)
    out = self.pep17(out)

    out_conv5 = self.conv5(out)
    out = F.interpolate(self.conv6(out_conv5), scale_factor=2)
    out = torch.cat([out, out_pep15], dim=1)
    out = self.pep18(out)
    out = self.pep19(out)

```

```
out_conv7 = self.conv7(out)
out = F.interpolate(self.conv8(out_conv7), scale_factor=2)
out = torch.cat([out, out_pep7], dim=1)
out = self.pep20(out)
out = self.pep21(out)
out = self.pep22(out)
out_conv9 = self.conv9(out)
temp, layer_loss = self.yolo_layer52(out_conv9, targets, image_size)
loss += layer_loss
yolo_outputs.append(temp)
```

```
out = self.ep6(out_conv7)
out_conv10 = self.conv10(out)
temp, layer_loss = self.yolo_layer26(out_conv10, targets, image_size)
loss += layer_loss
yolo_outputs.append(temp)
```

```
out = self.ep7(out_conv5)
out_conv11 = self.conv11(out)
temp, layer_loss = self.yolo_layer13(out_conv11, targets, image_size)
loss += layer_loss
yolo_outputs.append(temp)
```

```
yolo_outputs = to_cpu(torch.cat(yolo_outputs, 1))
```

```
return yolo_outputs if targets is None else (loss, yolo_outputs)
```