

Инженерная школа информационных технологий и робототехники
 Направление: «Информационные системы и технологии»
 Отделение: «Информационные технологии»

БАКАЛАВРСКАЯ РАБОТА

Тема работы
Исследование сверточных нейронных сетей класса YOLO при решении задачи детектирования объектов разных масштабов на изображениях

УДК 004.932.72'1:528.2:629.7

Студент

Группа	ФИО	Подпись	Дата
8И7Б	Игольников Никита Александрович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Марков Николай Григорьевич	Доктор технических наук		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН	Маланина Вероника Анатольевна	Кандидат экономических наук		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД	Аверкиев Алексей Анатольевич	-		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Цапко Ирина Валериевна	Кандидат технических наук		

РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (КОМПЕТЕНЦИИ ВЫПУСКНИКОВ)

по направлению 09.03.02 «Информационные системы и технологии»

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

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ol style="list-style-type: none"> 1. Анализ предметной области 2. Обучение СНС 3. Анализ полученных результатов 4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение 5. Социальная ответственность
<p>Перечень графического материала <i>(с точным указанием обязательных чертежей)</i></p>	<p>Презентация в формате *.pptx</p>

Консультанты по разделам выпускной квалификационной работы
(с указанием разделов)

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Маланина Вероника Анатольевна
Социальная ответственность	Аверкиев Алексей Анатольевич

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

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

Задание выдал руководитель / консультант (при наличии):

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Марков Николай Григорьевич	Доктор технических наук		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И7Б	Игольников Никита Александрович		

Инженерная школа информационных технологий и робототехники
 Направление: «Информационные системы и технологии»
 Отделение: «Информационные технологии»
 Период выполнения: (осенний / весенний семестр 2020 /2021 учебного года)

Форма представления работы:

бакалаврская работа

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	14.06.2021
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
	Основная часть	75
	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	15
	Социальная ответственность	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ	Марков Николай Григорьевич	Доктор технических наук		

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Цапко Ирина Валерьевна	Кандидат технических наук		

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

Студенту:

Группа	ФИО
8И7Б	Игольникову Никите Александровичу

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	ИСТ

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%.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды 30,2%

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	Анализ конкурентных технических решений
2. Планирование и формирование бюджета научных исследований	Формирование плана и графика разработки: -определение структуры и трудоемкости работ; -разработка графика Гантта. Формирование бюджета затрат: -затраты на специальное оборудование; -заработная плата (основная и дополнительная); -отчисления на социальные цели; -накладные расходы.
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	Определение потенциального эффекта исследования

Перечень графического материала (с точным указанием обязательных чертежей):

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

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

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОСГН	Маланина В.А.	к.э.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И7Б	Игольников Никита Александрович		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8И7Б	Игольникову Никите Александровичу

Школа	Инженерная школа информационных технологий и робототехники	Отделение (НОЦ)	Отделение информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	ИСТ

Тема ВКР:

Исследование сверточных нейронных сетей класса YOLO при решении задачи детектирования объектов разных масштабов на изображениях

Исходные данные к разделу «Социальная ответственность»:

<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Объект исследования – сверточные нейронные сети класса YOLO и задача детектирования объектов разных масштабов на изображениях. Рабочая зона – аудитория с естественным и искусственным освещением, оборудованная системой отопления и кондиционирования воздуха. Область применения – специалисты в области анализа данных мониторинга земной поверхности и, в частности, специалисты в области анализа данных мониторинга климатических и экологических систем.</p>
---	---

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Правовые и организационные вопросы обеспечения безопасности:</p> <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	<p>Требования к организации и оборудованию рабочих мест с ПЭВМ согласно требованиям СП 2.2.3670-20. Регулирования организации рабочего места при выполнении работы сидя проводятся согласно ГОСТ 12.2.032-78, а также расположение элементов рабочего места согласно ГОСТ 22269-76. Трудовые отношения регулируются согласно ТК РФ ФЗ-197 от 30.12.2001. Опасные и вредные производственные факторы устанавливаются согласно ГОСТ 12.0.003-2015. Нормы естественного, искусственного и совмещенного освещения зданий устанавливаются согласно СП 52.13330.2016"Естественное и искусственное освещение".</p>
--	---

	<p>Превышение уровня шума рассматривается согласно СП 2.2.3670-20.</p> <p>Опасность поражения электрическим током по ГОСТ 12.1.038–82 и ГОСТ 12.1.019-2017</p> <p>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.</p> <p>Обращение с отходами рассматривается согласно ГОСТ Р 53692–2009.</p>
<p>2. Производственная безопасность:</p> <p>2.1. Анализ выявленных вредных и опасных факторов</p> <p>2.2. Обоснование мероприятий по снижению воздействия</p>	<p>Вредные факторы:</p> <ol style="list-style-type: none"> 1. отклонение показателей микроклимата (температуры и влажности воздуха) 2. отсутствие или недостаток необходимого естественного и искусственного освещения 3. превышение уровня шума <p>Опасные факторы:</p> <ol style="list-style-type: none"> 1. опасность поражения электрическим током
<p>3. Экологическая безопасность:</p>	<p>Анализ негативного воздействия на окружающую природную среду: утилизация компьютеров и другой оргтехники</p> <p>Негативное воздействие на гидросферу и атмосферу совершается посредством производства различной оргтехники</p> <p>Воздействие на литосферу в результате образования отходов при поломке оргтехники</p>
<p>4. Безопасность в чрезвычайных ситуациях:</p>	<p>Возможные ЧС: землетрясение, наводнение, пожар</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
Ассистент	Аверкиев Алексей Анатольевич	-		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И7Б	Игольников Никита Александрович		

Реферат

Выпускная квалификационная работа содержит 72 страницы, 11 рисунков, 25 таблиц, 63 источника, 3 приложения.

Ключевые слова: детектирование объектов на изображении, сверточные нейронные сети, объекты разных масштабов, встраиваемые и мобильные системы компьютерного зрения, YOLO.

Объектом исследования являются сверточные нейронные сети для детектирования объектов разных масштабов на изображении.

Цель работы: исследование СНС класса YOLO на предмет возможности решения задачи детектирования объектов разных масштабов в рамках встраиваемых и мобильных СКЗ.

В процессе исследования проводились: обоснование актуальности работы, анализ предметной области, анализ и сравнение различных архитектур СНС для решения задачи детектирования объектов на изображении, обучение, тестирование и анализ эффективности СНС класса YOLO.

В результате проведено исследование эффективности СНС класса YOLO и исследование точности детектирования объектов разных масштабов такими СНС.

Область применения: решение задачи детектирования объектов разных масштабов в рамках встраиваемых и мобильных систем компьютерного зрения (СКЗ).

Список сокращений

МО – машинное обучение

КЗ – компьютерное зрение

СКЗ – система компьютерного зрения

СНС – сверточные нейронные сети

ГП – графический процессор

FLOP – float operation

БПЛА – беспилотный летательный аппарат

BoF – Bag-of-Freebies

BoS – Bag-of-Specials

Содержание

Введение	13
1 Задача детектирования объектов и ее решение с помощью СНС класса YOLO во встраиваемых и мобильных СКЗ	15
2 Предварительный анализ архитектур исследуемых СНС	25
2.1 Архитектура СНС YOLOv4	25
2.2 Архитектура СНС YOLO Nano	28
3 Постановка задачи	32
4 Используемое программное и аппаратное обеспечение	33
5 Результаты экспериментов	34
6 ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ	36
6.1 Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения	36
6.1.1 Потенциальные потребители результатов исследования	36
6.1.2 Анализ конкурентных технических решений	36
6.1.3 Технология QuaD	38
6.1.4 SWOT-анализ	39
6.2 Планирование научно-исследовательских работ	40
6.2.1 Структура работ в рамках научного исследования	40
6.2.2 Определение трудоемкости выполнения работ	40
6.2.3 Разработка графика проведения научного исследования	42
6.2.4 Бюджет научно-технического исследования	43
6.2.5 Расчет затрат на специальное оборудование для научных (экспериментальных) целей	43
6.2.6 Основная заработная плата исполнителей темы	44
6.2.7 Дополнительная заработная плата исполнителей темы	46
6.2.8 Отчисления во внебюджетные фонды (страховые отчисления)	47
6.2.9 Накладные расходы	48
6.2.10 Формирование бюджета затрат научно-исследовательского проекта	48
6.3 Определение потенциального эффекта исследования	49
7 СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ	50

Введение	50
7.1 Правовые и организационные вопросы обеспечения безопасности	50
7.1.1 Специальные правовые нормы трудового законодательства	50
7.1.2 Организационные мероприятия при компоновке рабочей зоны	51
7.2 Производственная безопасность	53
7.2.1 Вредные производственные факторы	54
7.2.1.1 Отклонение показателей микроклимата в помещении	54
7.2.1.2 Недостаточная освещенность рабочей зоны	55
7.2.1.3 Превышение уровня шума	56
7.2.2 Опасные производственные факторы	57
7.2.2.1 Опасность поражения электрическим током	57
7.3 Экологическая безопасность	58
7.4 Безопасность в чрезвычайных ситуациях	59
Вывод	60
Заключение	61
Список использованных источников	62
Приложение 1 – Эффективность в терминах МО	69
Приложение 2 – MS COCO	70
Приложение 3 – Точность и скорость современных СНС для детектирования	71

Введение

Задача детектирования объектов на изображении является одной из фундаментальных задач компьютерного зрения (КЗ) и заключается в одновременной локализации (определении местоположения) и классификации всех объектов на изображении. Местоположение объекта определяется ограничивающим его прямоугольником (англ. bounding box). На решении данной задачи построено множество приложений, включая робототехнику, беспилотные автомобили, анализ аэрокосмических снимков и др.

Современные достижения в глубоком обучении [1, 2, 3] позволили достичь значительных успехов в решении многих задач КЗ. В частности, применение глубоких сверточных нейронных сетей (СНС) привело к значительным успехам в решении задачи детектирования объектов [4].

Исследовательский интерес большинства ранних работ [5, 6], а также многих современных [7, 8], сосредоточен на достижении высокой точности детектирования, при этом эффективности алгоритмов детектирования не уделяется значительного внимания. Это приводит к архитектурам СНС с высокой вычислительной сложностью и низкой эффективностью. Применение таких СНС во многих реальных задачах затруднено. В робототехнике, беспилотных автомобилях, беспилотных летательных аппаратах (БПЛА) и других приложениях задачу детектирования объектов необходимо решать в режиме реального времени в условиях ограниченных ресурсов памяти и вычислительного устройства встраиваемых и мобильных систем компьютерного зрения (СКЗ). В этой связи растет интерес к эффективным и компактным архитектурам СНС. Примерами работ, нацеленных на разработку более эффективных архитектур СНС для детектирования, могут служить работы по исследованию одноступенчатых архитектур [9, 10, 11, 12, 13], безъякорных архитектур [14, 15, 16], сжатия существующих моделей [17, 18].

Для решения реальных задач КЗ с применением СНС в рамках встраиваемых и мобильных СКЗ необходимо соблюдение баланса между точностью и эффективностью¹ таких СНС.

Еще одной актуальной проблемой при решении задачи детектирования объектов на изображении является детектирование объектов разных масштабов, а в особенности детектирование объектов малого масштаба, которое представляет особую сложность для современных СНС. Не смотря на общие успехи в достижении высокой точности детектирования, все еще сохраняется колоссальный разрыв между точностью детектирования объектов малого и большого масштабов.

Данные проблемы делают актуальной задачу детектирования объектов разных масштабов в рамках встраиваемых и мобильных СКЗ.

В разделе 1 рассмотрены современные СНС для детектирования объектов в рамках встраиваемых и мобильных СКЗ, рассмотрена проблема детектирования объектов разных масштабов. В разделе 2 проведен предварительный анализ исследуемых СНС для детектирования. В разделе 3 поставлена цель и сформулированы задачи исследования. В разделе 4 описано используемое программное и аппаратное обеспечение. В разделе 5 приведены результаты проведенных экспериментов и их анализ.

¹ Определение алгоритмической эффективности в задачах машинного обучения (МО), а также некоторые тенденции данного показателя, описанные в [19], приведены в приложении 1.

1 Задача детектирования объектов и ее решение с помощью СНС класса YOLO во встраиваемых и мобильных СКЗ

Ранние алгоритмы детектирования объектов на изображении на основе глубокого обучения адаптировали классификаторы для решения данной задачи. СНС для классификации применялась последовательно к различным областям изображения по методу скользящего окна [20], либо, в более поздних работах, к определенным образом сгенерированным областям интереса на изображении [5]. Высокая оценка классификатора в некоторой области изображения расценивалась как детекция объекта в данной области. Такие двухступенчатые архитектуры СНС для детектирования объектов являются сложными и медленными — для обработки одного изображения требуется от нескольких сотен до нескольких тысяч проходов сети по различным областям изображения, в связи чем такие архитектуры непригодны к использованию во встраиваемых и мобильных СКЗ.

Исследователи, предложившие архитектуру YOLO [10], 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. Преимущества таких одноступенчатых архитектур СНС в сравнении с двухступенчатыми:

- предсказания основаны на глобальном контексте, т. к. модель обрабатывает сразу всё изображение;

- предсказание рамок и классов всех объектов на изображении за один проход по сети значительно ускоряет процесс детектирования.

СНС для детектирования объектов с одноступенчатой архитектурой способны достигать эффективности, достаточной для встраиваемых СКЗ, работающих на ГП. Однако, большинство таких архитектур, хоть и значительно быстрее и эффективнее двухступенчатых, все же не подходят для мобильных

СКЗ, в связи с более жесткими ограничениями ресурсов памяти и вычислительного устройства таких СКЗ. СНС YOLOv3 [12], к примеру, занимает 236 МБ и требует около 140 GFLOP для своей работы. Для мобильных СКЗ необходимы более компактные архитектуры СНС.

В связи с этим были предложены уменьшенные версии архитектур СНС класса YOLO — архитектуры СНС Tiny YOLO [10], со значительно меньшими требованиями к памяти и производительности вычислительного устройства СКЗ, однако точность таких моделей ниже полноразмерных версий. Одной из таких компактных архитектур СНС для детектирования объектов является также архитектура СНС YOLO Nano [21].

Архитектуры современных СНС для детектирования объектов на изображениях обычно состоят из трех подсетей (рис. 1):

-СНС для извлечения признаков (англ. feature extraction network или backbone): базовая СНС, служащая для извлечения признаков входного изображения;

-СНС для агрегации признаков (англ. feature fusion network или neck) с разных уровней базовой сети — низкоуровневых карт признаков с локальной информацией и высокоуровневых с семантической;

-СНС для предсказания рамок и классов объектов (англ. box/class network или head) на основе агрегированных ранее карт признаков. В некоторых архитектурах СНС используются две отдельные подсети — одна для предсказания рамок, вторая - для предсказания классов.

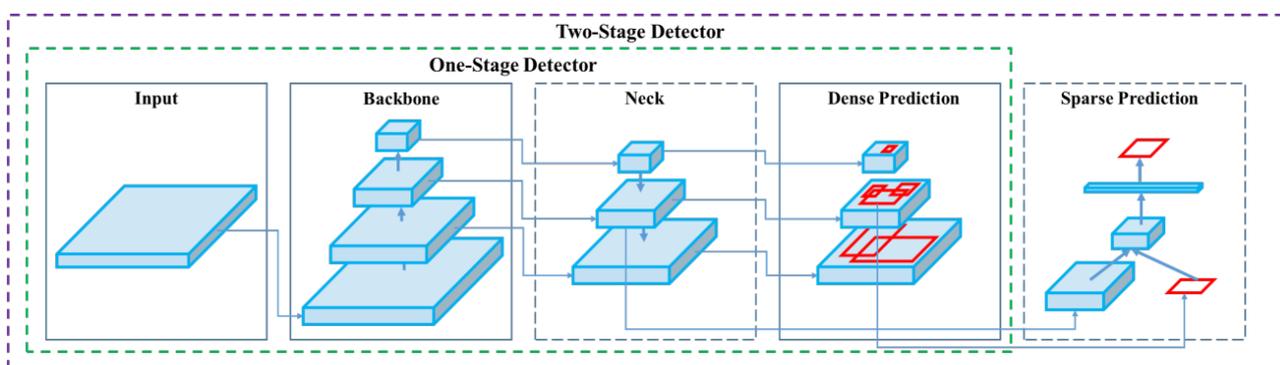


Рисунок 1 [13] — Архитектура современных СНС для детектирования в общем виде

Для извлечения признаков используются такие СНС, как VGG [22], ResNet [23], ResNeXt [24], Darknet [11, 12], DenseNet [25], SqueezeNet [26], MobileNet [27, 28, 29], MNASNet [30], ShuffleNet [31, 32], SpineNet [33], EfficientNet [3], CSPResNeXt [34], CSPDarknet [34] и др. СНС для извлечения признаков, как правило, предобучаются на датасетах для классификации изображений, чаще всего на датасете ImageNet [35].

Для агрегации карт признаков применяются такие СНС, как FPN [36], PANet [37], Bi-FPN [4], SFAM [38], ASFF [39], NAS-FPN [40] и др. Также в подсети для агрегации признаков могут использоваться дополнительные модули, например SPP [41], ASPP [42], RFB [43], SAM [44] и др.

СНС для предсказания рамок и классов делятся на одноступенчатые и двухступенчатые, якорные и безъякорные (табл.1).

Таблица 1 – СНС для предсказания рамок и классов

	Якорные	Безъякорные
Одноступенчатые	SSD [9], YOLO [10], RetinaNet [45]	CornerNet [46, 47], CenterNet [48], MatrixNet [49], FCOS [16]
Двухступенчатые	R-CNN [5], Fast R-CNN [50], Faster R-CNN [51], Mask R-CNN [6], Libra R-CNN [52], R-FCN [53]	RepPoints [54]

При обучении СНС для детектирования объектов в современных работах используются различные методы повышения точности детектирования (табл. 2). Они делятся на две группы: Bag-of-Freebies (BoF) и Bag-of-Specials (BoS).

Группа BoF включает методы, не оказывающие влияния на вычислительную сложность обученной модели и скорость ее работы. В группу BoF входят аугментация данных, различные функции потерь и методы регуляризации.

Группа BoS включает методы, которые незначительно уменьшают скорость работы обученной модели, при этом значительно улучшая точность детектирования. В группу BoS входят различные функции активации, методы нормализации, пробросы карт признаков, техники постобработки и

дополнительные модули (расширение области восприятия, введение механизма внимания и др.).

Таблица 2 [13] – Методы повышения точности детектирования

BoF	
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
Регуляризация	DropOut, Spatial DropOut, DropPath, DropBlock, DropConnect, Label Smoothing
BoS	
Функции активации	ReLU, LReLU, PReLU, SELU, ReLU6, Swish, hard-Swish, Mish
Нормализация	Batch Normalization (BN), Cross-GPU BN (CGBN или SyncBN), Filter Response Normalization (FRN), Cross-Iteration BN (CBN)
Пробросы признаков	Residual connections, Weighted residual connections, Multi-input weighted residual connections, Cross stage partial connections (CSP)

На данный момент наивысшую точность детектирования на датасете MS COCO² [55] показывают СНС класса EfficientDet [4]. При этом, несмотря свое название, СНС класса EfficientDet являются не самыми эффективными СНС для детектирования объектов.

В таблице 3 приведены показатели точности и вычислительной сложности СНС класса YOLO и СНС класса EfficientDet.

В приложении 3 приведена более полная таблица сравнения точности и эффективности СНС для детектирования объектов на изображении.

² MS COCO Object Detection Dataset [55] - наиболее широко используемый датасет для детектирования объектов. Здесь и далее использованы метрики AP, AP50, AP75, APS, APM, APL. Полный список метрик датасета COCO и их описание приведены в приложении 2.

Таблица 3 [13, 4] – Точность и эффективность СНС классов YOLO и EfficientDet

	Размер входного изображения	Размер модели, МБ ³	Количество операций, GFLOP	FPS ⁴	MS COCO test-dev AP, %
YOLOv3	320	236	38.97	45(M)	28.2
YOLOv3	416	236	65.86	35(M)	31
YOLOv3	608	236	140.69	20(M)	33
YOLOv3-SPP	608	240	141.45	20(M)	36.2
YOLOv4	320	245	35.5	- /123(V)	38
YOLOv4	416	245	60.1	38(M) / 96(V)	41.2
YOLOv4	512	245	91.1	31(M) / 83(V)	43.0
YOLOv4	608	245	128.5	23(M) / 62(V)	43.5
EfficientDet-D0	512	59.5	2.5	98(V)	34.6
EfficientDet-D1	640	100.7	6.1	74.1(V)	40.5
EfficientDet-D2	768	123.6	11	56.5(V)	43.9
EfficientDet-D3	896	183.1	25	34.5(V)	47.2
EfficientDet-D4	1024	320.4	55	23.4(V)	49.7
EfficientDet-D5	1280	518.8	135	13.8(V)	51.5
EfficientDet-D6	1280	793.5	226	10.8(V)	52.6
EfficientDet-D7	1536	793.5	325	8.2(V)	53.7
EfficientDet-D7x	1536	1174.9	410	6.5(V)	55.1

Как видно из таблицы 3, state-of-the-art СНС EfficientDet-D7x достигает AP 55.1% на датасете MS COCO test-dev. При этом данная СНС требует 1.17 ГБ памяти (в случае точности весовых коэффициентов в 16 бит) и 410 GFLOP вычислений и работает с кадровой частотой 6.5 FPS на ГП Tesla V100. Для использования во встраиваемых СКЗ необходимы СНС, способные работать в режиме реального времени. По данному критерию для таких систем могут подойти, например, СНС EfficientDet-D0~D3. Однако при сравнении СНС EfficientDet-D0~D3 и YOLOv4 последние показывают более высокую эффективность. Как видно из рисунка 2, СНС YOLOv4 расположена на парето-оптимальной кривой (границе Парето) и обходит СНС

³ 2[^]20 байт в случае точности весовых коэффициентов равной 16 бит (FP16).

⁴ В данном столбце использованы следующие обозначения: M - ГП архитектуры Maxwell, V - ГП архитектуры Volta.

EfficientDet-D0~D2 и по скорости, и по точности. В связи с чем для дальнейшего исследования выбор сделан в пользу СНС YOLOv4.

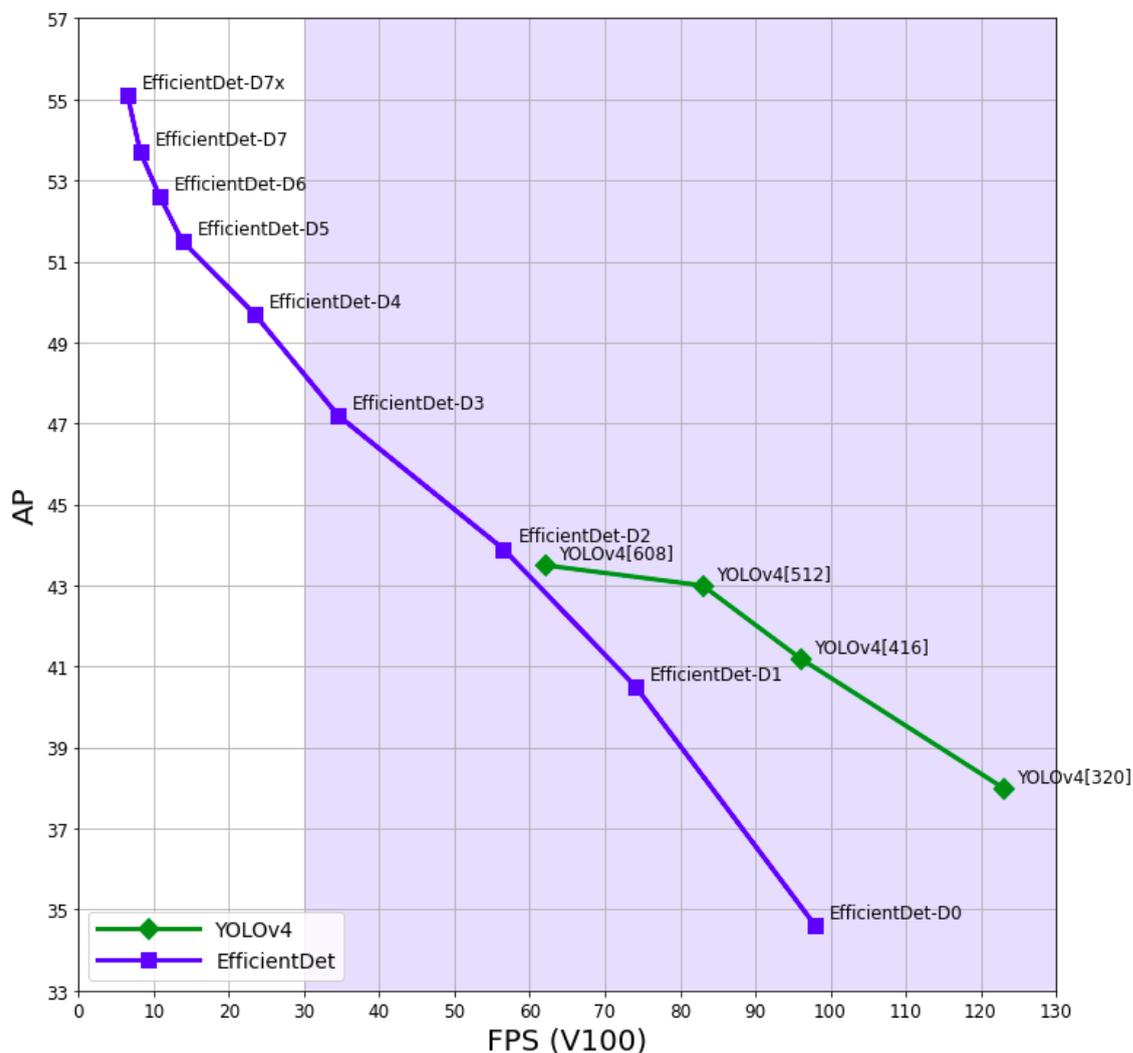


Рисунок 2 — Кадровая частота и точность СНС EfficientDet и YOLOv4

СНС класса YOLO [10, 11, 12, 13] являются одними из наиболее широко применяемых СНС для детектирования во встраиваемых СКЗ. Что касается мобильного детектирования, одними из наиболее популярных компактных СНС для использования в мобильных СКЗ являются СНС Tiny YOLO [10, 11, 12, 13].

Как видно из таблицы 4, СНС YOLO Nano [21] превосходит СНС Tiny YOLOv2 [11] и Tiny YOLOv3 [12] по точности, требуя при этом значительно меньший объем ресурсов памяти и вычислений. СНС YOLO Nano также обходит СНС YOLOv4-tiny [13] по объему требуемых ресурсов памяти и вычислений, однако для сравнения их по точности необходима проверка на

одном датасете. В связи с чем, для дальнейшего исследования выбраны обе архитектуры СНС: YOLOv4-tiny и YOLO Nano.

Таблица 4 [21, 56] – Точность детектирования и вычислительная сложность компактных СНС класса YOLO (разрешение входного изображения - 416x416)

	Размер модели, МБ	Количество операций, GFLOP	PASCAL VOC 2007 [57] mAP, %	MS COCO test-dev [55] AP50, %
Tiny YOLOv2	60.5	6.97	57.1	-
Tiny YOLOv3	33.4	5.52	58.4	33.1
YOLOv4-tiny	23.1	6.90	-	40.2
YOLO Nano	4.0	4.57	69.1	-

Не смотря на высокую точность state-of-the-art СНС для детектирования, все еще наблюдается значительный разрыв между точностью детектирования больших и малых объектов.

Датасет MS COCO содержит большое количество объектов малого масштаба. Кроме того, распределение масштабов объектов неравномерное. Как показано в таблице 5, малых объектов в данном датасете больше, чем средних и больших. При этом в обучающей подвыборке всего 52% изображений, включающих объекты малого масштаба, в то время как распределение средних и больших объектов относительно более равномерное.

Таблица 5 [58] – Статистика объектов датасета MS COCO относительно совпавших якорей СНС Mask-RCNN

	Доля объектов	Доля изображений	Доля общей площади объектов	Доля совпавших якорей	Средние совпадающие якоря	Среднее максимальное IoU
Малый	41.43%	51.82%	1.23%	29.96%	1.00	0.29
Средний	34.32%	70.07%	10.18%	25.54%	1.03	0.57
Большой	24.24%	82.28%	88.59%	44.49%	2.54	0.66

Основную сложность при детектировании объектов на изображении представляют объекты малого масштаба. Сложности детектирования объектов малого масштаба связаны со следующими факторами:

- малые объекты занимают малую площадь изображения и несут меньше признаковой информации для СНС;

-информация об объекте малого масштаба может быть потеряна при прохождении слоев субдискретизации.

На графике (рис. 3) наглядно показана проблема разрыва между точностью детектирования больших и малых объектов. Помимо того, что кривая точности детектирования объектов малого масштаба расположена значительно ниже, чем кривые точности детектирования объектов среднего и большого масштабов, из графика на рисунке 3 также видно, что архитектуры YOLOv3, RetinaNet и YOLOv4 создают локальные максимумы на кривых точности малого и среднего масштабов, сокращая разрыв между точностью детектирования малых и средних объектов по отношению к точности детектирования больших объектов.

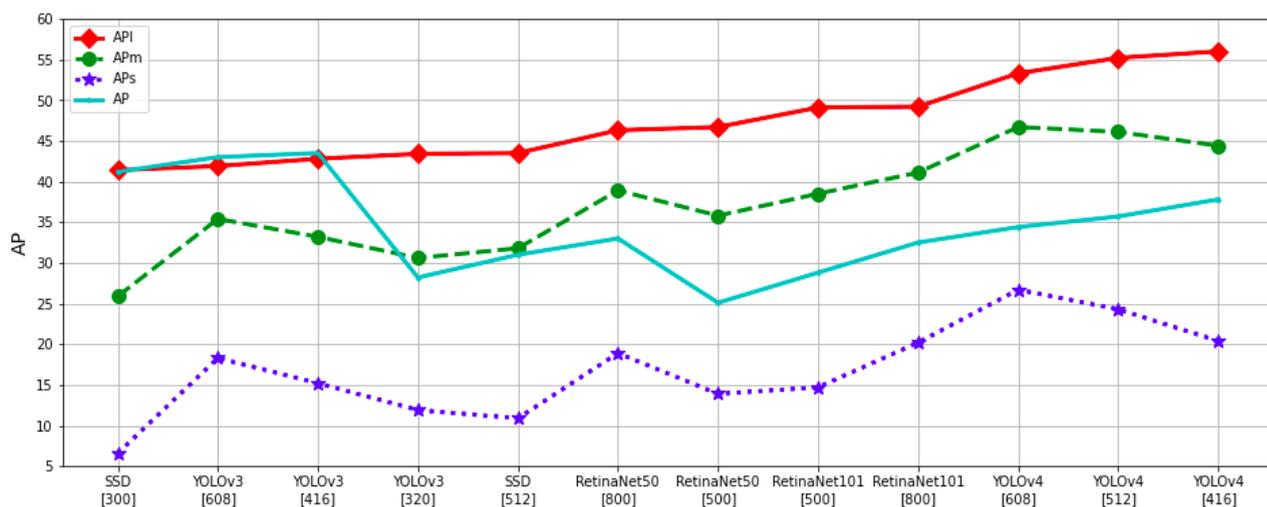


Рисунок 3 — Точность детектирования малых, средних, и больших объектов

На рисунке 4 приведен график точности детектирования объектов малого масштаба по отношению к частоте кадров при работе СНС на ГП архитектуры Maxwell. Из данного графика видно, что в режиме реального времени работают СНС SSD, YOLOv3 и YOLOv4. СНС YOLOv4 при этом выдает наивысшую точность при незначительных потерях в частоте кадров. На рисунке 5 приведен аналогичный график для всех масштабов объектов.

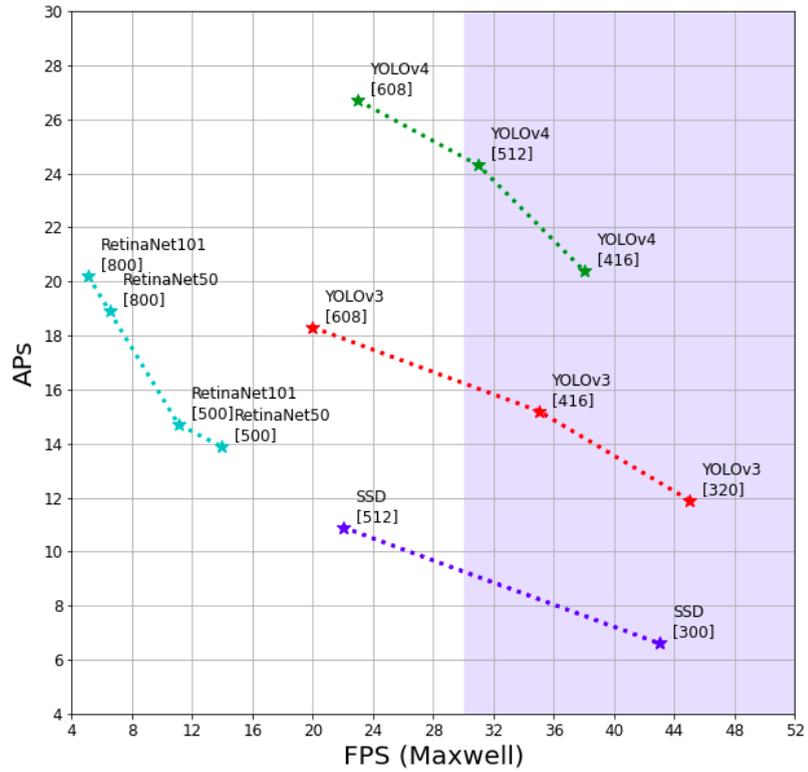


Рисунок 4 — Точность детектирования объектов малого масштаба

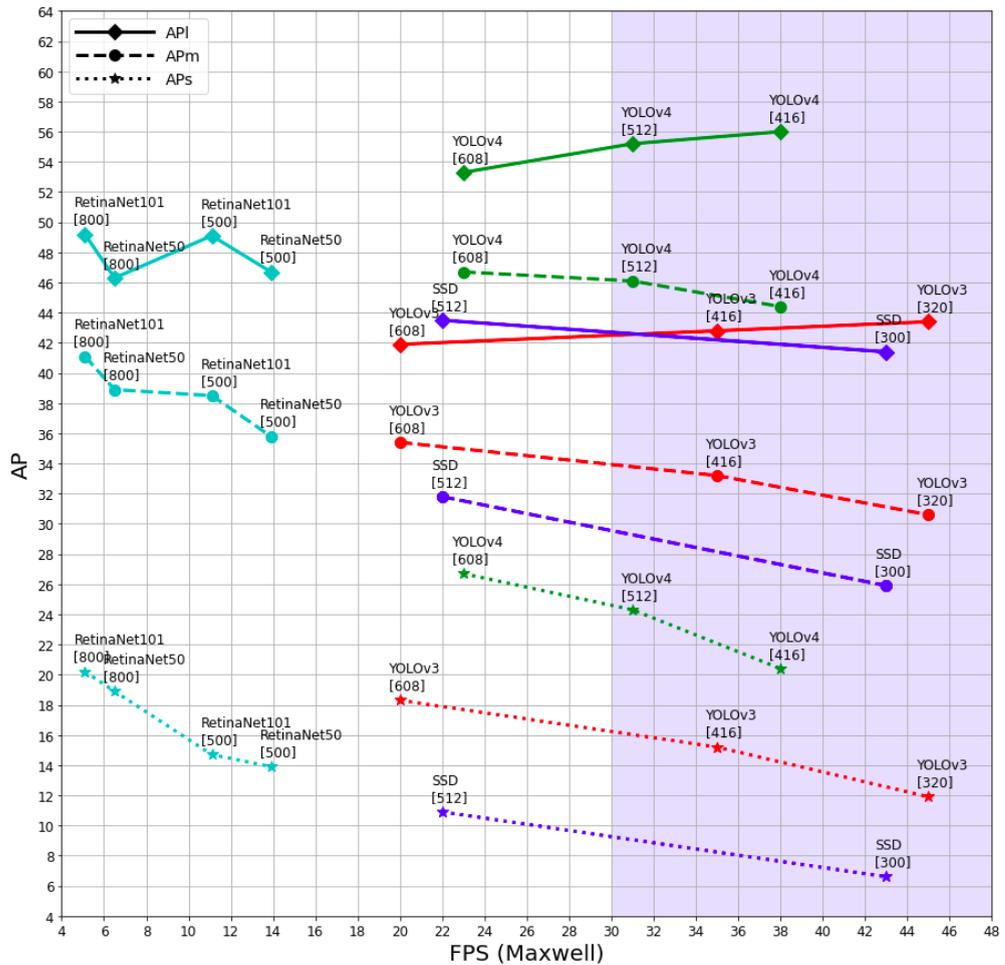


Рисунок 5 — Точность детектирования объектов разных масштабов

Точность детектирования объектов разных масштабов очень важна в частности при решении задачи детектирования объектов на снимках с БПЛА. Одним из наиболее известных датасетов для решения задачи детектирования объектов на снимках с БПЛА является Stanford Drones Dataset (SDD) [59]. Данный датасет содержит большое множество объектов разных масштабов. SDD представляет из себя ряд видеопоследовательностей, записанных с БПЛА — вид сверху на различные сцены, включающие пешеходов, велосипедистов, автомобили и др. На данных видеозаписях отличаются как относительные масштабы объектов — автомобиль, к примеру, занимает больший участок изображения, чем пешеход, так и абсолютные масштабы — высота съемки непостоянна, как и высота различных участков снимаемой сцены. Все эти факторы позволяют использовать данный датасет для оценки выбранных СНС на предмет возможности детектирования объектов разного масштаба. На рисунке 6 показаны примеры изображений (кадров видеопоследовательности) датасета SDD с выделенными рамками объектов.

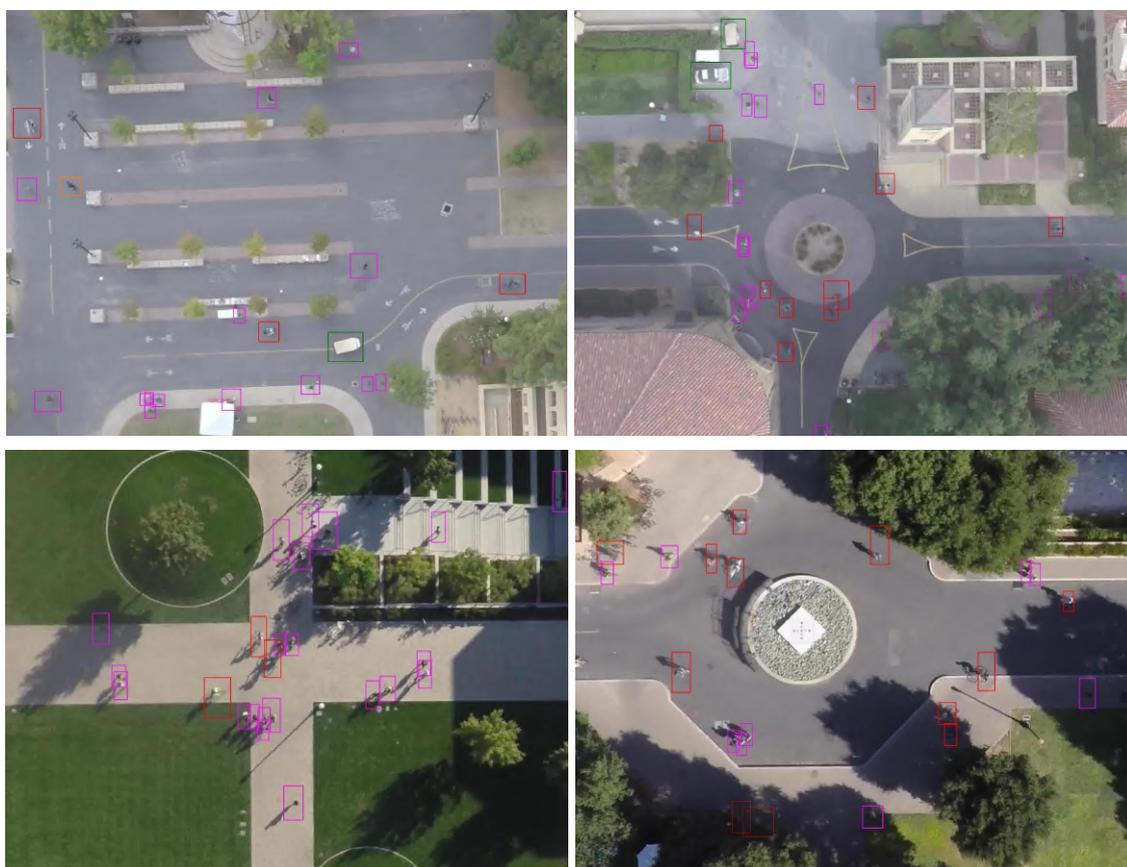


Рисунок 6 [59] — Примеры изображений SDD

2 Предварительный анализ архитектур исследуемых СНС

В данном разделе описаны архитектуры исследуемых СНС, а также особенности их проектирования и обучения.

2.1 Архитектура СНС YOLOv4

YOLOv4 - глубокая СНС для решения задачи детектирования объектов, ориентированная на использование в рамках встраиваемых СКЗ. Как отмечается в работе [13], основными задачами при проектировании архитектуры СНС YOLOv4 были высокая скорость работы СНС в промышленных системах и оптимизация для параллельных вычислений, а не низкий теоретический показатель объема вычислений (FLOP), а также пригодность для обучения на одном ГП. Архитектура СНС YOLOv4 ориентирована на обучение и работу на одном ГП среднего уровня с объемом графической памяти 8-16 ГБ (например, GTX 1080 Ti или RTX 2080 Ti).

Архитектура СНС YOLOv4 (рис. 7) состоит из следующих подсетей:

-backbone: CSPDarknet53 [34]

-neck: SPP [41], PANet [37]

-head: YOLOv3 [12]

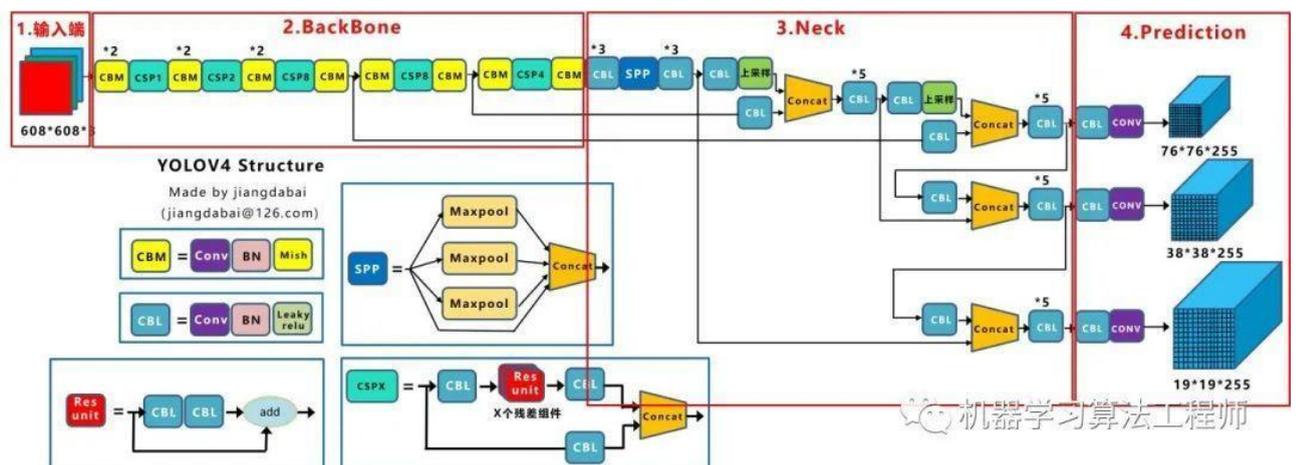


Рисунок 7 [60] — Схема архитектуры СНС YOLOv4

CSPDarknet53 — модифицированная с использованием стратегии CSPNet [34] версия архитектуры СНС Darknet53 [12].

В качестве neck в архитектуре CHC YOLOv4 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 CHC. Вместо модуля FPN, использовавшегося в YOLOv3, в YOLOv4 для агрегации признаков применяется модуль PANet. В архитектуре CHC YOLOv4 также используются модули SAM [44]. Модули PANet и SAM были модифицированы исследователями [13] (рис. 8 и рис. 9).

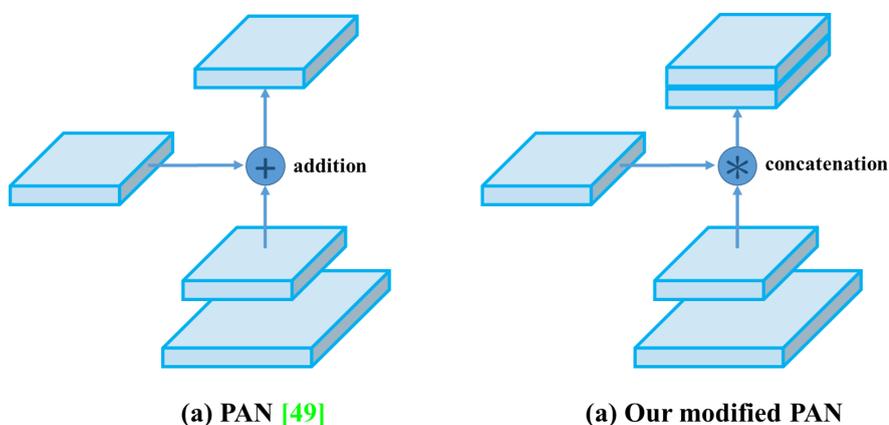


Рисунок 8 [13] — Исходный (а) и модифицированный (б) модули PANet

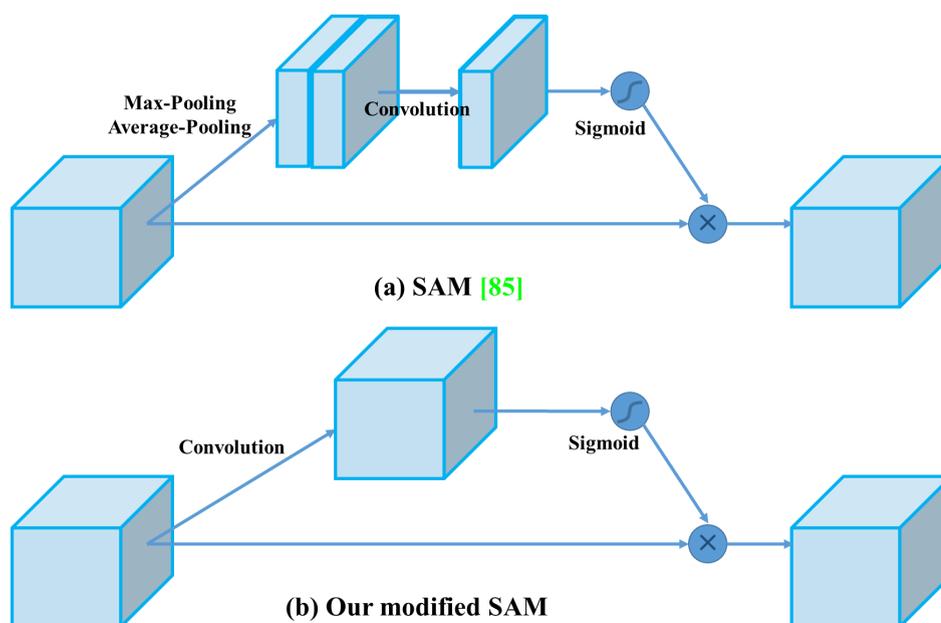


Рисунок 9 [13] — Исходный (а) и модифицированный (б) модули SAM

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

Также стоит отметить, что в работе [13] проведено исследование различных BoF и BoS, в результате которого был выбран ряд техник, использовавшийся при обучении YOLOv4 (табл. 6). Для подбора гиперпараметров применялись ГА.

Таблица 6 [13] – BoF и BoS, используемые при обучении СНС YOLOv4

	Bag of Freebies (BoF)	Bag of Specials (BoS)
backbone	<ul style="list-style-type: none"> -CutMix -Mosaic data augmentation -DropBlock regularization -Class label smoothing -dynamic mini-batch size 	<ul style="list-style-type: none"> -Mish activation -Cross-stage partial connections (CSP) -Multi-input weighted residual connections (MiWRC)
detector	<ul style="list-style-type: none"> -CIoU-loss -Cross mini-Batch Normalization (CmBN) -DropBlock regularization -Mosaic data augmentation -Self-Adversarial Training (SAT) -Eliminate grid sensitivity (solving the grid sensitivity problem) (grid sensitivity elimination) -Multiple anchors for a single ground truth -Cosine annealing scheduler -Optimal hyper-parameters -Random training shapes -IoU threshold -dynamic mini-batch size -Optimized Anchors -different IoU losses 	<ul style="list-style-type: none"> -Mish activation -SPP-block -SAM-block -PAN path-aggregation block -DIoU-NMS

2.2 Архитектура СНС YOLO Nano

YOLO Nano [21] — высоко компактная СНС для решения задачи детектирования объектов 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) [61], включающей два основных этапа:

1) принципиальное прототипирование архитектуры сети (англ. principled network design prototyping);

2) 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 [21] достигается в форме генеративного синтеза [62], конечная цель которого заключается в обучении генеративных машин, способных генерировать архитектуры глубоких СНС, удовлетворяющих поставленным требованиям и ограничениям. Генеративный синтез может быть сформулирован как задача условной оптимизации (оптимизации при наличии ограничений), заключающаяся в определении

генератора \mathcal{G} , способного генерировать сети $\{N_s | s \in S\}$ на основе набора значений S , решаемая путем максимизации универсальной функции эффективности \mathcal{U} (англ. universal performance function) с соблюдением требований и ограничений, определенных через индикаторную функцию (англ. indicator function) $l_r(\cdot)$ (формула 1).

$$\mathcal{G} = \max_{\mathcal{G}} \mathcal{U}(\mathcal{G}(s)) \text{ subject to } l_r(\mathcal{G}(s)) = 1, \forall s \in S. \quad (1)$$

Данная задача в работе [21] решается приближенно путем итеративной оптимизации.

Индикаторная функция $l_r(\cdot)$ определяется следующими условиями:

- 1) средняя точность на VOC 2007 mAP $\geq 65\%$;
- 2) вычислительная сложность $\leq 5 \times 10^9$ операций GFLOPs;
- 3) точность весов 8 бит.

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:

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

Модуль EP (рис. 10(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 PER (рис. 10(c)) состоит из:

- 1) проецирующего слоя;
- 2) расширяющего слоя;
- 3) слоя глубинной свертки;
- 4) проецирующего слоя.

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

Модуль FCA (рис. 10(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 PER, и поэлементное умножение в случае модуля FCA).

Общая схема архитектуры СНС YOLO Nano показана на рисунке 10(a). “ x ” в записи PER(x) обозначает количество каналов в первом проекционном слое модуля PER. “ x ” в записи FCA(x) обозначает коэффициент сокращения размерности равный x . Далее в скобках указаны размерности выходных карт признаков.

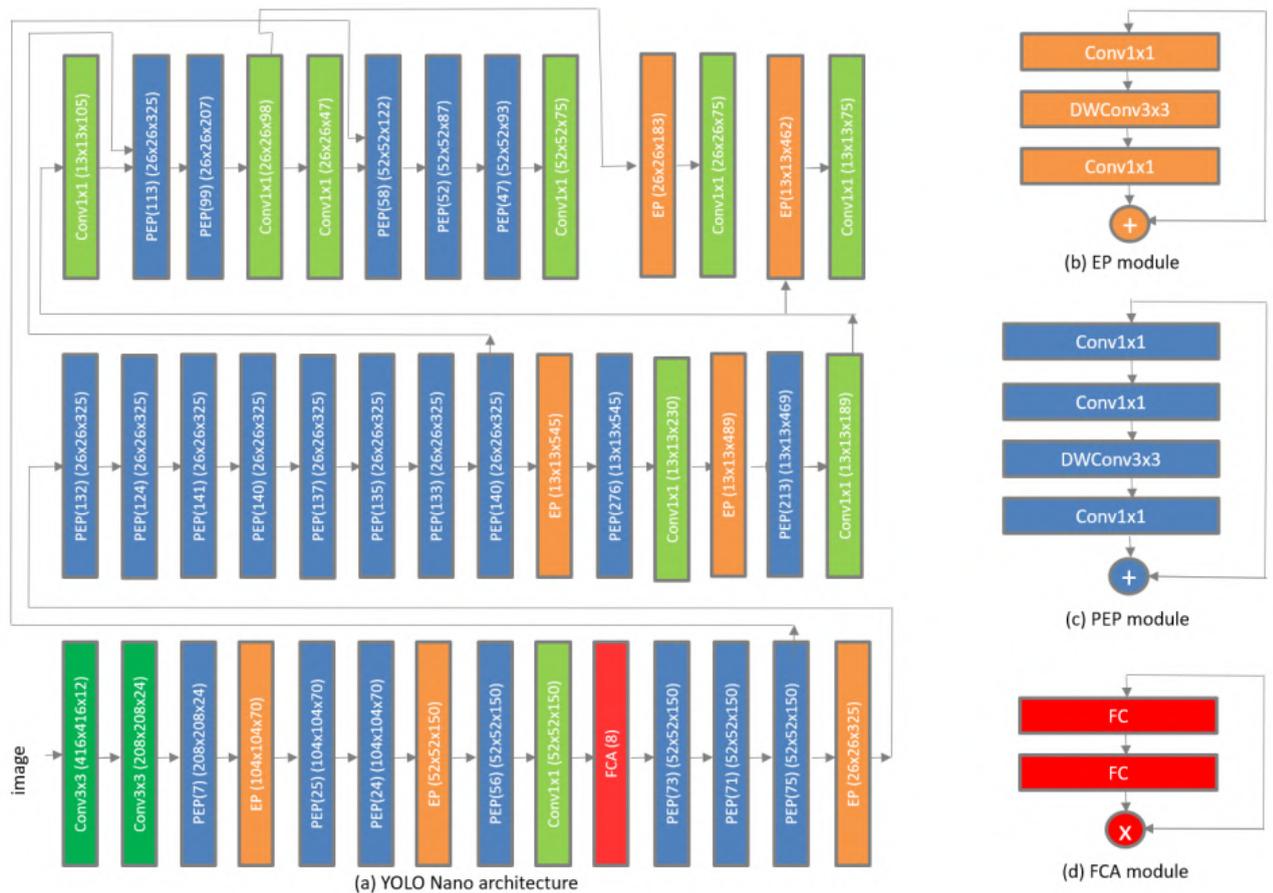


Рисунок 10 [21] — Схема архитектуры CHC YOLO Nano

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 Постановка задачи

Целью данной работы является исследование СНС класса YOLO на предмет возможности решения задачи детектирования объектов разных масштабов в рамках встраиваемых и мобильных СКЗ.

Две основные задачи — исследование эффективности СНС класса YOLO и исследование точности детектирования объектов разных масштабов такими СНС. Для решения первой задачи необходимо провести анализ и сравнение СНС YOLOv4 и YOLO Nano по следующим критериям:

-скорость работы (количество кадров в секунду);

-компактность архитектуры (объем памяти, необходимый для хранения весовых коэффициентов сверточных фильтров и промежуточных буферов при работе СНС);

-вычислительная сложность (количество операций умножения, сложения и сравнения с плавающей запятой).

Для решения второй задачи необходимо провести сравнение точности СНС YOLOv4 и YOLO Nano на датасете SDD по следующим метрикам: AP, APS, APM, APL. В таблице 7 приведены определения объектов разных масштабов.

Таблица 7 [58] – Определения объектов разных масштабов датасета MS COCO

	Минимальный размер	Максимальный размер
Малый объект	0 x 0	32 x 32
Средний объект	32 x 32	96 x 96
Большой объект	96 x 96	∞ x ∞

4 Используемое программное и аппаратное обеспечение

Для исследования эффективности СНС необходимо использовать реализации данных СНС на одном фреймворке, т. к. скорость работы одной и той же СНС в различных фреймворках отличается, также могут отличаться и ресурсы памяти.

Идеальным случаем исследования является сравнение моделей в их официальных реализациях на одном фреймворке. Однако авторы работ не всегда предлагают официальную реализацию. Даже при наличии официальных реализаций исследуемых СНС зачастую отличаются используемые авторами фреймворки.

СНС YOLOv4 и YOLOv4-tiny имеют официальные реализации на оупен-сорс фреймворке глубокого обучения Darknet, написанном на C и CUDA [56]. Официальная реализация СНС YOLO Nano отсутствует. В связи с чем принято решение об использовании фреймворка Darknet. Для СНС YOLO Nano используется собственная реализация на фреймворке Darknet.

Для обучения и тестирования исследуемых СНС используется ГП GTX 1080 Ti архитектуры Pascal.

5 Результаты экспериментов

Результаты исследования СНС YOLOv4 и СНС YOLO Nano по скорости работы, компактности архитектуры, вычислительной сложности и точности детектирования объектов приведены в таблице 8.

Таблица 8 – Результаты исследования СНС YOLOv4 и YOLO Nano

	Размер входного изображения, пикселей	Размер модели, МБ	Количество операций, GFLOP	Скорость работы, FPS	SDD AP, %
YOLOv4	416	245.0	60.10	63.1	10.7
YOLO Nano	416	4.0	4.57	340.2	5.3
YOLOv4	512	245.0	91.07	37.34	22.5
YOLO Nano	512	4.0	6.93	201.3	13.2
YOLOv4	1024	245.0	364.28	9.35	34.8
YOLO Nano	1024	4.0	27.72	50.33	20.1

Количество операций с плавающей точкой и скорость обработки изображений моделями исследуемых СНС измерены при размере батча, равном единице. Значения усреднены по всему набору данных.

Размер модели рассчитан для точности значений чисел с плавающей точкой, равной 16 бит (FP16). В таблице 8 в столбце с данным критерием показан объем памяти, необходимый для хранения весовых коэффициентов сверточных фильтров соответствующей модели.

По результатам, приведенным в таблице 8, можно сделать вывод, что при размере входного изображения, равном 416 x 416 пикселей, обе СНС показывают довольно низкие значения точности.

При максимальном размере входного изображения, равном 1024 x 1024 пикселей, СНС YOLO Nano сохраняет скорость работы в режиме реального времени (более 50 кадров в секунду). СНС YOLOv4 при таком размере входного изображения работает со скоростью ~9 кадров в секунду, что уже не является приемлемым для использования данной архитектуры СНС в мобильных СКЗ, работающих в режиме реального времени. Кроме того, СНС YOLOv4 требует значительно больше ресурсов памяти и вычислительного устройства (в 61.25

раз больше ресурсов памяти и, в среднем, в 13.14 раз больше ресурсов вычислительного устройства), что также важно при выборе архитектуры СНС для использования в мобильных СКЗ.

В таблице 9 приведены результаты исследования СНС YOLOv4 и СНС YOLO Nano по точности детектирования объектов разных масштабов. Все значения точности в таблице 9 получены при размере входного изображения, равном 1024 x 1024 пикселей.

Таблица 9 – Результаты исследования СНС YOLOv4 и YOLO Nano

	APS	APM	APL	AP
YOLOv4	25.3	35.2	45.2	34.8
YOLO Nano	13.47	21.0	30.1	20.1

Как видно из таблицы 9, по всем трем параметрам точности детектирования объектов разных масштабов лучшие значения показывает СНС YOLOv4. Более того, если посчитать относительную разницу между показателями AP и APS для двух СНС, то получится, что в случае СНС YOLOv4 APS меньше AP на 27.3%, а в случае СНС YOLO Nano – на 33.0%, что уже более очевидно говорит о том, что СНС YOLOv4 справляется с задачей детектирования объектов разных масштабов лучше, чем СНС YOLO Nano.

Однако, совмещая этот вывод с предыдущим, стоит отметить, что YOLO Nano является более подходящей архитектурой СНС для использования в мобильных СКЗ, т.к. в таких системах ресурсы памяти и вычислительного устройства сильно ограничены.

Анализируя все полученные результаты исследования СНС класса YOLO при решении задачи детектирования объектов разных масштабов, предпочтение следует отдать архитектуре СНС YOLO Nano. Данная СНС незначительно уступает СНС YOLOv4 по точности детектирования объектов разных масштабов, но, при этом, значительно менее требовательна к ресурсам памяти и вычислительного устройства СКЗ.

СНС YOLO Nano также является более подходящей архитектурой для использования в мобильных СКЗ, работающих в режиме реального времени.

РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

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

6.1.1 Потенциальные потребители результатов исследования

В рамках данной работы выполняется исследование, включающее в себя анализ и сравнение сверточных нейронных сетей YOLOv4 и YOLONano 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.

6.1.2 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 (K1)
- ручная обработка снимков специалистами (K2)

Анализ конкурентных технических решений был проведен с помощью оценочной карты, представленной в таблице 10.

Таблица 10. Оценочная карта критериев эффективности конкурентных технических решений (разработок)

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Бф	Бк1	Бк2	Кф	Кк1	Кк2
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 2:

$$K = \sum V_i \times B_i \quad K = \sum V_i \times B_i \quad (2)$$

где K – конкурентоспособность научной разработки или конкурента;

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

6.1.3 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 11.

Таблица 11. Оценка критериев в соответствии с технологией QuaD

Критерии оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение	Средневзвешенное значение
Скорость обработки данных	0,1	80	100	0,8	8
Простота использования системы	0,15	90	100	0,9	13,5
Удобный интерфейс пользователя	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.

6.1.4 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, а также его возможности и угрозы. Все параметры SWOT анализа данного проекта представлены в таблице 12.

Таблица 12. SWOT-анализ проекта

		Внутренние факторы	
		Сильные стороны проекта: 1. Надежность 2. Мобильность 3. Персонализация 4. Безопасность 5. Простота эксплуатации	Слабые стороны проекта: 1. Отсутствие прототипа научной разработки 2. Отсутствие финансирования 3. Повышенные требования к «железу» 4. Отсутствие репутации на рынке
Внешние факторы	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.	Появление дополнительного спроса на рынке может способствовать финансированию проекта. 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.

6.2 Планирование научно-исследовательских работ

6.2.1 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 13 представлен перечень работ, а также распределение исполнителей по ним.

Таблица 13. Перечень работ и распределение исполнителей

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

6.2.2 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 трудоёмкости выполнения каждой работы путём расчёта длительности работ в рабочих и календарных днях каждого этапа работ.

Трудоемкость оценивается по следующей формуле 3:

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

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 4:

$$T_{p_i} = \frac{t_{ож\ i}}{Ч_i} \quad T_{p_i} = \frac{t_{ож\ i}}{Ч_i}, \quad (4)$$

где T_{p_i} – 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 Гантта, необходимо также произвести перевод длительности работ из рабочих дней в календарные по формуле 5:

$$T_{к_i} = T_{p_i} \times k_{кал} \quad T_{к_i} = T_{p_i} \times k_{кал}, \quad (5)$$

где $T_{к_i}$ – продолжительность выполнения i -й работы в календарных днях;

T_{p_i} – это продолжительность выполнения i -й работы в рабочих днях;

$k_{кал}$ – это коэффициент календарности, равный 1,48.

Коэффициент календарности $k_{кал}$ рассчитывается по формуле 6:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}} \quad k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}}, \quad (6)$$

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 2021 году 365, а сумма выходных и праздничных дней составляет 118 дней, коэффициент календарности равен $k_{\text{кал}} = 1,48$.

6.2.3 Разработка графика проведения научного исследования

В таблице 14 приведены временные показатели научного исследования.

Таблица 14. Временные показатели проведения научного исследования

Наименование работы	Исполнители работы	Трудоёмкость работ, чел-дни		Длительность работ, дни		
		tmin	tmax	tож	Тр	Тк
Выбор научного руководителя бакалаврской работы	Игольников Н.А. Марков Н.Г.	1	2	1,4	1	1
Составление и утверждение темы бакалаврской работы	Игольников Н.А. Марков Н.Г.	1	3	1,8	1	1
Составление календарного плана-графика выполнения бакалаврской работы	Игольников Н.А.	1	2	1,4	1	1
Подбор и изучение литературы по теме бакалаврской работы	Игольников Н.А. Марков Н.Г.	5	10	7	4	6
Анализ предметной области	Игольников Н.А.	5	10	7	7	10
Выбор архитектур СНС	Игольников Н.А.	8	14	10,4	10	15
Обучение СНС	Игольников Н.А.	5	12	7,8	8	12
Сравнительный анализ результатов обучения СНС	Игольников Н.А.	6	12	8,4	8	12
Согласование выполненной работы с научным руководителем	Игольников Н.А. Марков Н.Г.	2	2	2	1	1
Выполнение других частей работы (финансовый менеджмент, социальная ответственность)	Игольников Н.А.	5	12	7,8	8	12
Подведение итогов, оформление работы	Игольников Н.А.	8	14	10,4	10	15
Итого	Игольников Н.А.	47	93	65,4	59	86
	Марков Н.Г.	9	17	12,2	7	9

На основе данных таблицы 14 был построен календарный план-график представленный на рисунке 11.

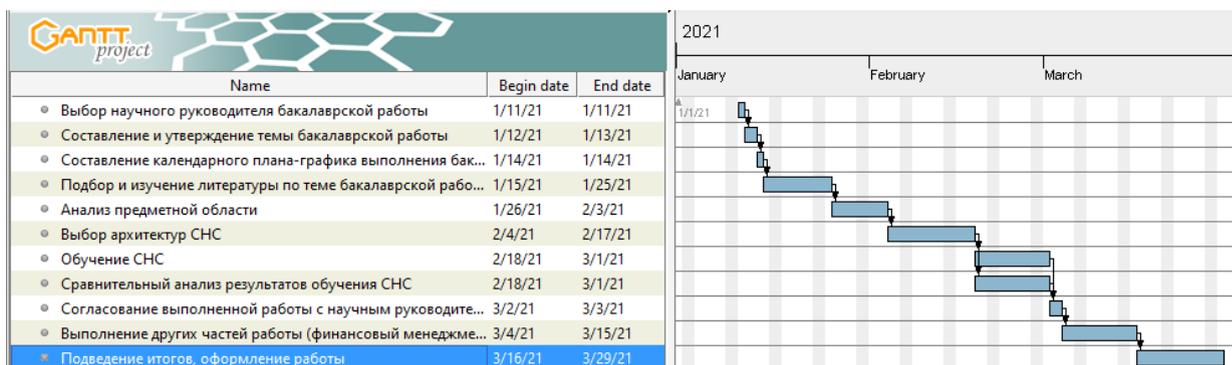


Рисунок 11 – Календарный план-график проведения работ

6.2.4 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.

6.2.5 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 7:

$$A_n = \frac{1}{n} \times 100\% \quad A_n = \frac{1}{n} \times 100\% \quad , \quad (7)$$

n – установленный срок в месяцах;

A_n – норма амортизации.

Тогда расчет амортизации ПК:

Норма амортизации:

$$A_n = \frac{1}{n} \times 100\% = \frac{1}{36} \times 100\% = 2,78\%$$

Ежемесячные амортизационные отчисления:

$$A_m = 100000 \times 0,0278 = 2780 \text{ рублей}$$

Итоговая сумма амортизации основных средств:

$$A = 2780 \times 6 = 16680 \text{ рублей}$$

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

6.2.6 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 8:

$$Z_{зп} = Z_{осн} + Z_{доп} \quad Z_{зп} = Z_{осн} + Z_{доп} \quad , \quad (8)$$

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 заработную плату можно получить по формуле 9:

$$Z_{\text{осн}} = Z_{\text{дн}} \times T_p \times (1 + K_{\text{пр}} + K_{\text{д}}) \times K_r \quad Z_{\text{осн}} = Z_{\text{дн}} \times T_p \times (1 + K_{\text{пр}} + K_{\text{д}}) \times K_r \quad (9)$$

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

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

$K_{\text{д}}$ – коэффициент доплат и надбавок (0,2-0,5);

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

T_p – продолжительность работ, выполняемых работником, раб. дни

где $Z_{\text{дн}}$ – среднедневная заработная плата исполнителя;

T_p – продолжительность работ, выполняемых исполнителем.

Среднедневную заработную плату можно получить по формуле 10:

$$Z_{\text{дн}} = \frac{Z_{\text{м}} \times M}{F_{\text{д}}} \quad Z_{\text{дн}} = \frac{Z_{\text{м}} \times M}{F_{\text{д}}}, \quad (10)$$

где $Z_{\text{м}}$ – месячный должностной оклад исполнителя, рубль;

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;

$F_{\text{д}}$ – действительный годовой фонд рабочего времени персонала по разработке.

Должностные оклады исполнителей проекта согласно приказу ТПУ представлены в таблице 15.

Таблица 15. Месячные должностные оклады исполнителей

Исполнитель	Районный коэффициент (для Томска)	Размер месячного должностного оклада без учета коэффициента, рубль
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 16.

Таблица 16. 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:

$$Z_{\text{дн}}(\text{студент}) = \frac{21760 \times 10,4}{247} = 927,48 \text{ рубля}$$

$$Z_{\text{дн}}(\text{научный руководитель}) = \frac{33664 \times 10,4}{247} = 1434,86$$

Расчет затрат на основную заработную плату приведен в таблице 17.

Таблица 17. Затраты на основную заработную плату

Исполнители	Здн, руб.	Кпр	Кд	Кр	Тр	Зосн, руб.
Студент	927,48	0,3	0,2	1,3	59	106706,57
Научный руководитель	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.

6.2.7 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 11:

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

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 18.

Таблица 18. Затраты на дополнительную заработную плату

Исполнители	Зосн, руб.	Кдоп	Здоп, руб.
Студент	106706,57	0,12	12804,79
Научный руководитель	19585,84	0,12	2350,30
Итого:			15155,09

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

6.2.8 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 12.

$$Z_{внеб} = k_{внеб} \times (Z_{осн} + Z_{доп}) \quad Z_{внеб} = k_{внеб} \times (Z_{осн} + Z_{доп}), \quad (12)$$

где $k_{внеб}$ – 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 30,2%.

Расчет затрат на отчисления во внебюджетные фонды приведен в таблице 19.

Таблица 19. Отчисления во внебюджетные фонды

Исполнители	Зосн, руб.	Здоп, руб.	Квнеб	Звнеб, руб.
Студент	106706,57	12804,79	0,302	36092,43
Научный руководитель	19585,84	2350,30	0,302	6624,71
Итого:				42717,14

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

6.2.9 Накладные расходы

Накладные расходы – расходы на организацию, управление и обслуживание процесса производства товара, оказания услуги; носят комплексный характер. Накладные расходы вычисляются по формуле 13:

$$Z_{\text{нкл}} = \text{сумма статей}(1 - 5) \times k_{\text{нр}} \quad Z_{\text{нкл}} = \text{сумма статей}(1 - 5) \times k_{\text{нр}}, \quad (13)$$

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

Расчет накладных расходов приведен в таблице 20.

Таблица 20. 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.

6.2.10 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 21.

Таблица 21. Расчет бюджета затрат НИИ

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 НИИ составляет 233217,41 рубля.

6.3 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.

7 СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Введение

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 и YOLO Nano при решении задачи детектирования объектов разных масштабов на изображениях.

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.

7.1 Правовые и организационные вопросы обеспечения безопасности

7.1.1 Специальные правовые нормы трудового законодательства

В трудовом кодексе РФ содержатся основные положения отношений между организацией и сотрудниками, включая оплату и нормирование труда, выходных, отпуска и так далее. Работа в офисе относится ко второй категории тяжести труда – работы выполняются при оптимальных условиях внешней производственной среды и при оптимальной величине физической, умственной и нервно-эмоциональной нагрузки.

Нормальная продолжительность рабочего времени не может превышать 40 часов в неделю. При работе в ночное время продолжительность рабочей смены сокращается на один час.

Сокращенная продолжительность рабочего времени устанавливается:

- для работников в возрасте до шестнадцати лет - не более 24 часов в неделю;

- для работников в возрасте от шестнадцати до восемнадцати лет - не более 35 часов в неделю;
- для работников, являющихся инвалидами I или II группы не более 35 часов в неделю;
- для работников, условия труда на рабочих местах которых по результатам специальной оценки условий труда отнесены к вредным условиям труда 3 или 4 степени или опасным условиям труда, - не более 36 часов в неделю.

Согласно трудовому кодексу РФ продолжительность рабочего времени конкретного работника устанавливается трудовым договором на основании отраслевого (межотраслевого) соглашения и коллективного договора с учетом результатов специальной оценки условий труда.

Работа в ночное время считается с 22 часов до 6 часов. Продолжительность работы (смены) в ночное время сокращается на один час без последующей отработки. Не сокращается продолжительность работы (смены) в ночное время для работников, которым установлена сокращенная продолжительность рабочего времени, а также для работников, принятых специально для работы в ночное время, если иное не предусмотрено коллективным договором.

7.1.2 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. Конструкция рабочего места и взаимное расположение всех его элементов (сиденье, органы управления, средства отображения информации и т. д.) должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы.

Конструкцией производственного оборудования и рабочего места должно быть обеспечено оптимальное положение работающего, которое достигается регулированием:

- высоты рабочей поверхности, сиденья и пространства для ног;
- высоты сиденья и подставки для ног (при нерегулируемой высоте рабочей поверхности).

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 22, согласно СП 2.2.3670-20 и ГОСТ 12.2.032-78.

Таблица 22 – 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 (700мм)
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.

Лицевые поверхности мониторов должны располагаться перпендикулярно взору работника. Допускаемое отклонение от этой плоскости – не более 45°; допускаемый угол отклонения линии взора от нормальной - не более 25° для стрелочных индикаторов и 30° для индикаторов с плоским изображением (ГОСТ 22269–76).

Клавиатуру следует располагать на поверхности стола на расстоянии 100—300 мм от края, обращенного к пользователю, или на специальной регулируемой по высоте рабочей поверхности, отделенной от основной столешницы (СП 2.2.3670-20).

7.2 Производственная безопасность

Первопричиной всех травм и заболеваний, связанных с процессом труда, является неблагоприятное воздействие на организм занятого трудом человека тех или иных факторов производственной среды и трудового процесса. Это воздействие, приводящее в различных обстоятельствах к различным результирующим последствиям, зависит от наличия в условиях труда того или иного фактора, его потенциально неблагоприятных для организма человека свойств, возможности его прямого или опосредованного действия на организм, характера реагирования организма в зависимости от интенсивности и длительности воздействия (экспозиции) данного фактора. Производственные факторы являются частным случаем факторов окружающей человека среды обитания и человеческой деятельности, связанных и (или) порождаемых производственной и трудовой деятельностью.

Согласно ГОСТ 12.0.003-2015 неблагоприятные производственные факторы по результирующему воздействию на организм работающего человека подразделяют на:

- вредные производственные факторы, то есть факторы, приводящие к заболеванию, в том числе усугубляющие уже имеющиеся заболевания;

- опасные производственные факторы, то есть факторы, приводящие к травме, в том числе смертельной.

В таблице 23 представлены возможные вредные и опасные факторы, возникающие при работе за ПК.

Таблица 23 – Вредные и опасные факторы по ГОСТ 12.0.003-2015, возникающие при работе за персональным компьютером

Наименование видов работ	Факторы (ГОСТ 12.0.003-2015)	Нормативные документы
Вредные факторы		
Работа за персональным компьютером	Отклонение показателей микроклимата (температуры и влажности воздуха)	СП 2.2.3670-20
	Отсутствие или недостаток необходимого естественного и искусственного освещения	СП 52.13330.2016
	Превышение уровня шума	СП 2.2.3670-20
Опасные факторы		
Работа за персональным компьютером	Опасность поражения электрическим током	ГОСТ 12.1.038–82 ГОСТ 12.1.019-2017

7.2.1 Вредные производственные факторы

7.2.1.1 Отклонение показателей микроклимата в помещении

Показателями, характеризующими микроклимат, являются: температура воздуха, температура поверхностей (ограждающих конструкций, устройств, технологического оборудования), влажность воздуха, скорость движения воздуха, тепловое облучение (при наличии источников лучистого тепла).

Оптимальные микроклиматические условия установлены по критериям оптимального теплового и функционального состояния человека. Они обеспечивают общее и локальное ощущение теплового комфорта в течение 8-часовой рабочей смены при минимальном напряжении механизмов терморегуляции, не вызывают отклонений в состоянии здоровья, создают

предпосылки для высокого уровня работоспособности и являются предпочтительными на рабочих местах.

Работа программиста относится к категории Ia, которые производятся сидя и сопровождаются незначительным физическим напряжением. Оптимальные параметры микроклимата на рабочих местах должны соответствовать величинам, приведённым в таблице 24, применительно к выполнению работ в холодный и тёплый периоды года.

Таблица 24 – Оптимальные параметры микроклимата на рабочих местах

Период года	Категория работ по уровню энергозатрат, Вт	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	Ia (до 139)	22-24	21-25	60-40	0,1
Тёплый	Ia (до 139)	23-25	22-26	60-40	0,1

В кабинете поддерживается температура равная 22-24 С°, при относительной влажности в 50-60%, что является нормой по требованиям СП 2.2.3670-20. Для поддержки такого состояния необходимо проводить в помещении ежедневную влажную уборку и систематическое проветривание.

7.2.1.2 Недостаточная освещенность рабочей зоны

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

Освещенность на поверхности стола в зоне размещения рабочего документа должна быть 300–500 лк. Освещение не должно создавать бликов на поверхности экрана. Освещенность поверхности экрана не должна быть более 300 лк.

Следует ограничивать прямую блескость от источников освещения, при этом яркость светящихся поверхностей (окна, светильники и др.), находящихся в поле зрения, должна быть не более 200 кд/м². Светильники местного освещения должны иметь непросвечивающий отражатель с защитным углом не менее 40 градусов.

В качестве источников света при искусственном освещении следует применять преимущественно люминесцентные лампы типа ЛБ и компактные люминесцентные лампы (КЛЛ).

7.2.1.3 Превышение уровня шума

Звуковое давление – переменная составляющая давления воздуха или газа, возникающая в результате звуковых колебаний, Па.

Допустимый уровень шума – это уровень, который не вызывает у человека значительного беспокойства и существенных изменений показателей функционального состояния систем и анализаторов, чувствительных к шуму.

Характеристикой постоянного шума на рабочих местах являются уровни звукового давления в дБ в октавных полосах со среднегеометрическими частотами 31,5; 63; 125; 250; 500; 1000; 2000; 4000; 8000 Гц, определяемые по формуле:

$$L = \frac{20 \lg P}{P_0} \quad L = \frac{20 \lg P}{P_0} \quad (14)$$

где P – среднеквадратичная величина звукового давления, Па;

P_0 – исходное значение звукового давления в воздухе, равное 2×10^{-5} Па.

Предельно допустимые уровни звукового давления в октавных полосах частот, уровни звука и эквивалентные уровни звука для трудовой деятельности программиста, разработанные с учетом категории тяжести и напряженности труда, представлены в таблице 25.

Таблица 25 – Предельно допустимые уровни звукового давления для программиста

Уровни звукового давления, дБ, в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука и эквивалентные уровни звука (дБА)
31,5	63	125	250	500	1000	2000	4000	8000	
86	71	61	54	49	45	42	40	38	50

Для исследуемого объекта (персональный компьютер) основными источниками шумов являются составляющие компьютера:

- вентилятор блока питания;
- вентилятор центрального процессора;
- вентилятор на высокопроизводительной видеокарте.

Для снижения уровня шума в помещении обеспечена вентиляция системного блока. Со стороны вентиляционных отверстий системного блока обеспечено не менее 20-30 см свободного пространства.

7.2.2 Опасные производственные факторы

7.2.2.1 Опасность поражения электрическим током

При работе за компьютером необходимо учитывать требования электробезопасности, так как это может привести к негативным последствиям, таким как, поражение электрическим током и возникновение пожара.

Электробезопасность должна обеспечиваться:

- конструкцией электроустановок и архитектурно-планировочными решениями;
- организацией технологических процессов;
- техническими способами и средствами защиты;
- организационными и техническими мероприятиями при производстве работ;
- электрозщитными средствами, средствами защиты от электрических и магнитных полей и другими средствами индивидуальной защиты, применяемыми при эксплуатации электроустановок;
- организацией технического обслуживания электроустановок.

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

В помещении используются для питания приборов напряжение 220 В переменного тока с частотой 50 Гц. Обязательны следующие предосторожности:

- перед началом работы убедиться, что выключатели, розетки закреплены и не имеют оголенных токоведущих частей;
- не включать в сеть компьютеры и другую оргтехнику со снятыми крышками;
- при обнаружении неисправности компьютера необходимо выключить его и отключить от сети;
- запрещается загромождать рабочее место лишними предметами;
- при несчастном случае необходимо немедленно отключить питание электроустановки, вызвать скорую помощь и оказать пострадавшему первую помощь до прибытия врача, согласно правилам;
- дальнейшее продолжение работы возможно только после устранения причины поражения электрическим током.

7.3 Экологическая безопасность

Большое количество процессов, операций и материалов, используемых при изготовлении электронных средств, являются источниками огромного количества веществ, имеющих неблагоприятное воздействие на человека и биосферу. При изготовлении элементной базы, электронных изделий, при обработке, выращивании полупроводниковых кристаллов, при изготовлении интегральных схем, в процессе гальванического производства утилизация исходных материалов часто происходит с низким коэффициентом использования, огромное количество их идет в отходы, попадая в атмосферу, гидросферу, загрязняя почву. Таким образом, наряду с истощением природных запасов дефицитных материалов происходит загрязнение окружающей среды, что ведет к губительным последствиям для отдельных экосистем и биосферы в целом.

На данном рабочем месте выявлен предполагаемый источник загрязнения окружающей среды, а именно воздействие на литосферу в результате образования отходов при поломке предметов вычислительной техники и оргтехники.

Согласно ГОСТ Р 53692—2009, вышедшее из строя ПЭВМ и сопутствующая оргтехника относится к IV классу опасности и подлежит специальной утилизации.

Первым этапом является утилизация обезвреженных (инертных) отходов. Во время утилизации может быть произведена переработка бракованных или вышедших из употребления видов продукции, изделий, их составных частей и отходов от них путем разборки (разукрупнения), переплавки, использования других технологий с обеспечением рециркуляции (восстановления) органической и неорганической составляющих.

Вторым этапом является безопасное размещение отходов I—IV классов опасности на соответствующих полигонах или уничтожение.

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

Причинами возгорания при работе с вычислительной техникой могут быть:

- токи короткого замыкания;
- перегрев и дальнейшее воспламенение;
- неисправность работы техники;
- повреждённая электросеть.

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.

Вывод

В данном разделе были рассмотрены правовые нормы трудового законодательства, организационные мероприятия при компоновке рабочей зоны, различные вредные и опасные факторы в производственной среде, а также рассмотрена типичная чрезвычайная ситуация, учитывая специфику выпускной квалификационной работы при работе с вычислительной техникой. Каждый подраздел составлен с учётом государственных стандартов, сводов правил, трудового кодекса РФ и санитарных норм.

Заключение

В рамках данной работы выявлены наиболее подходящие архитектуры СНС для решения задачи детектирования объектов разных масштабов на изображениях в рамках встраиваемых и мобильных СКЗ среди современных state-of-the-art СНС на основе официальных результатов точности на датасете MS COCO, а также официальных показателей требуемых вычислительных ресурсов для данных СНС.

Поставлена актуальная задача исследования СНС класса YOLO на предмет возможности решения задачи детектирования объектов разных масштабах в рамках встраиваемых и мобильных СКЗ.

В качестве исследуемых СНС выбраны СНС YOLOv4 и СНС YOLO Nano. В качестве датасета для решения задачи детектирования объектов разных масштабов выбран датасет SDD. Проведен сравнительный анализ точности детектирования объектов разных масштабов на изображениях моделями исследуемых СНС на датасете SDD.

В результате исследований выявлено, что архитектура СНС YOLO Nano наиболее пригодна для решения задачи детектирования объектов разных масштабов на изображениях в рамках встраиваемых и мобильных СКЗ. СНС YOLO Nano также является наиболее подходящей архитектурой для использования в СКЗ, работающих в режиме реального времени. Данная СНС незначительно уступает СНС YOLOv4 по точности детектирования объектов разных масштабов, но, при этом, значительно менее требовательна к ресурсам памяти и вычислительного устройства СКЗ.

Однако разработчикам встраиваемых и мобильных СКЗ необходимо учитывать, что СНС YOLOv4 справляется с задачей детектирования объектов разных масштабов лучше, чем СНС YOLO Nano. Если при разработке СКЗ точность детектирования объектов разных масштабов является более приоритетной, чем минимизация ресурсов памяти и вычислительного устройства СКЗ, то предпочтение стоит отдать архитектуре СНС YOLOv4.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Y. LeCun, Y. Bengio, G. Hinton. Deep learning. *Nature*, 2015.
2. A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, pages 1097–1105, 2012.
3. M. Tan, Q. V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of International Conference on Machine Learning (ICML)*, 2019.
4. 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.
5. R. Girshick, J. Donahue, T. Darrell, J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, pages 580–587, 2014.
6. K. He, G. Gkioxari, P. Dollár, R. Girshick. Mask R-CNN. *ICCV*, pages 2980–2988, 2017.
7. B. Zoph, E. D. Cubuk, G. Ghiasi, T.-Y. Lin, J. Shlens, Q. V. Le. Learning data augmentation strategies for object detection. *arXiv preprint arXiv:1804.02767*, 2019.
8. Y. Li, Y. Chen, N. Wang, Z. Zhang. Scale-aware Trident Networks for Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 6054–6063, 2019.
9. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg. SSD: Single shot multibox detector. In *European Conference on Computer Vision*, pages 21–37, 2016.
10. J. Redmon, S. Divvala, R. Girshick, A. Farhadi. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

11. J. Redmon, A. Farhadi. YOLO9000: Better, Faster, Stronger. arXiv preprint arXiv:1612.08242, 2016.
12. J. Redmon, A. Farhadi. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767, 2018.
13. A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934, 2020.
14. H. Law, J. Deng. CornerNet: Detecting Objects as Paired Keypoints. ECCV, 2018.
15. X. Zhou, D. Wang, P. Krhenbhl. Objects as points. arXiv preprint arXiv:1904.07850, 2019.
16. Z. Tian, C. Shen, H. Chen, T. He. FCOS: Fully convolutional one-stage object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 9627–9636, 2019.
17. Z. Liu, M. Sun, T. Zhou, G. Huang, T. Darrell. Rethinking the value of network pruning. ICLR, 2019.
18. J. Pedoem, R. Huang. YOLO-LITE: a real-time object detection algorithm optimized for non-gpu computers. arXiv preprint arXiv:1811.05588, 2018.
19. D. Hernandez, T. B. Brown. Measuring the Algorithmic Efficiency of Neural Networks. arXiv preprint arXiv:2005.04305.
20. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, D. Ramanan. Object detection with discriminatively trained part based models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, 2010.
21. A. Wong, M. Famuori, M. J. Shafiee, F. Li, B. Chwyl, J. Chung. YOLO Nano: a Highly Compact You Only Look Once Convolutional Neural Network for Object Detection. arXiv preprint arXiv:1910.01271, 2019.
22. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

23. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 2016.
24. Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1492–1500, 2017.
25. Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4700–4708, 2017.
26. Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint arXiv:1602.07360, 2016.
27. Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017.
28. Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4510–4520, 2018.
29. Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, et al. Searching for MobileNetV3. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.
30. Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, and Quoc V Le. MNASnet: Platform-aware neural

architecture search for mobile. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2820–2828, 2019.

31. Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 6848–6856, 2018.

32. Ningning Ma, Xiangyu Zhang, Hai-Tao Zheng, and Jian Sun. ShuffleNetV2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), pages 116–131, 2018.

33. Xianzhi Du, Tsung-Yi Lin, Pengchong Jin, Golnaz Ghiasi, Mingxing Tan, Yin Cui, Quoc V Le, and Xiaodan Song. SpineNet: Learning scale-permuted backbone for recognition and localization. arXiv preprint arXiv:1912.05027, 2019.

34. Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPR Workshop), 2020.

35. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. CVPR, 2009.

36. 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.

37. 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.

38. Qijie Zhao, Tao Sheng, Yongtao Wang, Zhi Tang, Ying Chen, Ling Cai, and Haibin Ling. M2Det: A Single-Shot Object Detector based on Multi-Level

Feature Pyramid Network. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), volume 33, pages 9259–9266, 2019.

39. Songtao Liu, Di Huang, and Yunhong Wang. Learning spatial fusion for single-shot object detection. arXiv preprint arXiv:1911.09516, 2019.

40. Golnaz Ghiasi, Tsung-Yi Lin, and Quoc V Le. NAS-FPN: Learning scalable feature pyramid architecture for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 7036–7045, 2019.

41. 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.

42. 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.

43. 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.

44. 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.

45. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal Loss for Dense Object Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 2980–2988, 2017.

46. Hei Law and Jia Deng. CornerNet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV), pages 734–750, 2018.

47. Hei Law, Yun Teng, Olga Russakovsky, and Jia Deng. CornerNet-Lite: Efficient keypoint based object detection. arXiv preprint arXiv:1904.08900, 2019.
48. Kaiwen Duan, Song Bai, Lingxi Xie, Honggang Qi, Qing-ming Huang, and Qi Tian. CenterNet: Keypoint triplets for object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 6569–6578, 2019.
49. Abdullah Rashwan, Agastya Kalra, and Pascal Poupart. Matrix Nets: A new deep architecture for object detection. In Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCV Workshop), 2019.
50. Ross Girshick. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 1440–1448, 2015.
51. S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Advances in neural information processing systems (NIPS), pages 91–99, 2015.
52. Jiangmiao Pang, Kai Chen, Jianping Shi, Huajun Feng, Wanli Ouyang, and Dahua Lin. Libra R-CNN: Towards Balanced Learning for Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 821–830, 2019.
53. Jifeng Dai, Yi Li, Kaiming He, and Jian Sun. R-FCN: Object Detection via Region-based Fully Convolutional Networks. In Advances in Neural Information Processing Systems (NIPS), pages 379–387, 2016.
54. Ze Yang, Shaohui Liu, Han Hu, Liwei Wang, and Stephen Lin. RepPoints: Point Set Representation for Object Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 9657–9666, 2019.
55. 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.
56. Darknet: Open Source Neural Networks in C [Электронный ресурс] / URL: <https://github.com/AlexeyAB/darknet>.

57. The PASCAL Visual Object Classes Challenge 2007 [Электронный ресурс] / URL: <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/>.
58. M. Kisantal, Z. Wojna, J. Murawski, J. Naruniec, K. Cho. Augmentation for small object detection. arXiv preprint arXiv:1902.07296, 2019.
59. A. Robicquet, A. Sadeghian, A. Alahi, S. Savarese, Learning Social Etiquette: Human Trajectory Prediction In Crowded Scenes in European Conference on Computer Vision (ECCV), 2016.
60. A simple explanation of YOLOv3 and YOLOv4 [Электронный ресурс] / URL: <https://www.programmingsought.com/article/13544183937/>.
61. A. Wong, Z. Q. Lin, and B. Chwyl. Attonets: Compact and efficient deep neural networks for the edge via human-machine collaborative design. arXiv preprint arXiv:1903.07209, 2019.
62. A. Wong, M. J. Shafiee, B. Chwyl, and F. Li. Ferminets: Learning generative machines to generate efficient neural networks via generative synthesis. Advances in neural information processing systems Workshops (NIPS Workshops), 2018.
63. G. E. Moore. Cramming more components onto integrated circuits. Electronics 38(8), 1965.

Приложение 1 – Эффективность в терминах МО

Эффективность — основной путь измерения алгоритмического прогресса в классических задачах информатики [19]. Алгоритмическая эффективность может быть определена как количество ресурсов вычислительного устройства, необходимое для достижения определенных показателей в решении поставленной задачи.

Исследование тенденций эффективности решения задачи классификации ImageNet [35] в работе [19] показывает, что с 2012 года количество вычислительных ресурсов, необходимое для достижения уровня точности AlexNet [2] снижается вдвое каждые 16 месяцев. На основе этих результатов выдвигается предположение о том, что в высоко инвестируемых доменах машинного обучения (МО) изменение алгоритмической эффективности в настоящее время имеет эффект, по масштабу похожий на эффект закона Мура [63] на аппаратную эффективность.

Так же предполагается, что алгоритмический прогресс вносит больший вклад, чем производительность аппаратного обеспечения. При этом исследователи отмечают, что аппаратная и алгоритмическая эффективности неразрывны друг от друга и нельзя пренебрегать ни тем, ни другим фактором.

Приложение 2 – MS COCO

Для определения качества детектирования на датасете COCO используются 12 метрик, приведенных в таблице ниже [55].

Average Precision (AP)	
AP	% AP at IoU=.50:.05:.95 (primary challenge metric)
AP _{IoU=.50}	% AP at IoU=.50 (PASCAL VOC metric)
AP _{IoU=.75}	% AP at IoU=.75 (strict metric)
AP Across Scales	
AP _{small}	% AP for small objects: area < 32 ²
AP _{medium}	% AP for medium objects: 32 ² < area < 96 ²
AP _{large}	% AP for large objects: area > 96 ²
Average Recall (AR)	
AR _{max=1}	% AR given 1 detection per image
AR _{max=10}	% AR given 10 detections per image
AR _{max=100}	% AR given 100 detections per image
AR Across Scales	
AR _{small}	% AR for small objects: area < 32 ²
AR _{medium}	% AR for medium objects: 32 ² < area < 96 ²
AR _{large}	% AR for large objects: area > 96 ²

Приложение 3 – Точность и скорость современных СНС для детектирования

В таблице ниже приведены результаты точности детектирования на датасете MS COCO (test-dev 2017) [55] и скорости работы современных СНС для детектирования объектов. Для СНС класса EfficientDet [4] использованы результаты из [4], для всех остальных — из [13], в связи с тем, что результаты работы [4] были обновлены после выхода работы [13].

Method	Backbone	Size	FPS ⁵	AP, %	AP50, %	AP75, %	APS, %	APM, %	APL, %
MAXWELL									
YOLOv4	CSPDarknet-53	416	38	41.2	62.8	44.3	20.4	44.4	56
YOLOv4	CSPDarknet-53	512	31	43	64.9	46.5	24.3	46.1	55.2
YOLOv4	CSPDarknet-53	608	23	43.5	65.7	47.3	26.7	46.7	53.3
YOLOv3	Darknet-53	320	45	28.2	51.5	29.7	11.9	30.6	43.4
YOLOv3	Darknet-53	416	35	31	55.3	32.3	15.2	33.2	42.8
YOLOv3	Darknet-53	608	20	33	57.9	34.4	18.3	35.4	41.9
YOLOv3-SPP	Darknet-53	608	20	36.2	60.6	38.2	20.6	37.4	46.1
SSD	VGG-16	300	43	25.1	43.1	25.8	6.6	25.9	41.4
SSD	VGG-16	512	22	28.8	48.5	30.3	10.9	31.8	43.5
RetinaNet	ResNet-50	500	13.9	32.5	50.9	34.8	13.9	35.8	46.7
RetinaNet	ResNet-101	500	11.1	34.4	53.1	36.8	14.7	38.5	49.1
RetinaNet	ResNet-50	800	6.5	35.7	55	38.5	18.9	38.9	46.3
RetinaNet	ResNet-101	800	5.1	37.8	57.5	40.8	20.2	41.1	49.2

⁵ Результаты получены при размере батча равном 1 без использования tensorRT. Используются следующие обозначения:

-M: архитектура Maxwell (ГП GTX Titan X (Maxwell) или Tesla M40)

-P: архитектура Pascal (ГП Titan X (Pascal), Titan Xp, GTX 1080 Ti, или Tesla P100)

-V: архитектура Volta (ГП Titan Volta или Tesla V100)

Method	Backbone	Size	FPS	AP, %	AP50, %	AP75, %	APS, %	APM, %	APL, %
PASCAL									
YOLOv4	CSPDarknet-53	416	54	41.2	62.8	44.3	20.4	44.4	56
YOLOv4	CSPDarknet-53	512	43	43	64.9	46.5	24.3	46.1	55.2
YOLO v4	CSPDarknet-53	608	33	43.5	65.7	47.3	26.7	46.7	53.3
RetinaNet	ResNet-50	-	10.8	37.1	56.9	40	20.1	40.1	48
Faster R-CNN	ResNet-50	-	9.4	39.8	59.2	43.5	21.8	42.6	50.7
RetinaMask	ResNet-50-FPN	800x	8.1	39.4	58.6	42.3	21.9	42	51
RetinaMask	ResNet-101-FPN	800x	6.9	41.4	60.8	44.6	23	44.5	53.5
RetinaMask	ResNet-101-FPN-GN	800x	6.5	41.7	61.7	45	23.5	44.7	52.8
Retina Mask	ResNeXt-101-FPN-GN	800x	4.3	42.6	62.5	46	24.8	45.6	53.8
Cascade R-CNN	ResNet-101	-	8	42.8	62.1	46.3	23.7	45.5	55.2
TridentNet	ResNet-101	-	2.7	42.7	63.6	46.5	23.9	46.6	56.6
TridentNet	ResNet-101-DCN	-	1.3	46.8	67.6	51.5	28	51.2	60.5
VOLTA									
YOLOv4	CSPDarknet-53	416	96	41.2	62.8	44.3	20.4	44.4	56
YOLOv4	CSPDarknet-53	512	83	43	64.9	46.5	24.3	46.1	55.2
YOLOv4	CSPDarknet-53	608	62	43.5	65.7	47.3	26.7	46.7	53.3
EfficientDet-D0	EfficientNet-B0	512	98	34.6	53	37.1	-	-	-
EfficientDet-D1	EfficientNet-B1	640	74.1	40.5	59.1	43.7	-	-	-
EfficientDet-D2	EfficientNet-B2	768	56.5	43.9	62.7	47.6	-	-	-
EfficientDet-D3	EfficientNet-B3	896	34.5	47.2	65.9	51.2	-	-	-
EfficientDet-D4	EfficientNet-B4	1024	23.4	49.7	68.4	53.9	-	-	-
EfficientDet-D5	EfficientNet-B5	1280	13.8	51.5	70.5	56.1	-	-	-
EfficientDet-D6	EfficientNet-B6	1280	10.8	52.6	71.5	57.2	-	-	-
EfficientDet-D7	EfficientNet-B6	1536	8.2	53.7	72.4	58.4	-	-	-
EfficientDet-D7x	EfficientNet-B7	1536	6.5	55.1	74.3	59.9	-	-	-
YOLOv3+ASFF*	Darknet-53	320	60	38.1	57.4	42.1	16.1	41.6	53.6
YOLOv3 + ASFF*	Darknet-53	416	54	40.6	60.6	45.1	20.3	44.2	54.1
YOLOv3 + ASFF*	Darknet-53	608x	45.5	42.4	63	47.4	25.5	45.7	52.3
YOLOv3 + ASFF*	Darknet-53	800x	29.4	43.9	64.1	49.2	27	46.6	53.4
RetinaNet	ResNet-50	640	37	37	-	-	-	-	-
RetinaNet	ResNet-101	640	29.4	37.9	-	-	-	-	-
RetinaNet	ResNet-50	1024	19.6	40.1	-	-	-	-	-
RetinaNet	ResNet-101	1024	15.4	41.1	-	-	-	-	-
NAS-FPN	ResNet-50	640	24.4	39.9	-	-	-	-	-
NAS-FPN	ResNet-50	1024	12.7	44.2	-	-	-	-	-