

**УТВЕРЖДЕНА**

Решением Ученого совета  
АНО ВО «Центральный университет»  
«07» марта 2024 г.  
Протокол №1

**Рабочая программа дисциплины (модуля)  
«Computer vision (Компьютерное зрение)»**

**Направление подготовки:** 02.04.01 Математика и компьютерные науки

**Направленность (профиль) подготовки:** Backend-разработка

**Квалификация (степень) выпускника:** магистр

**Форма обучения:** очная

**Срок освоения программы:** 2 года

**Год набора:** 2024

**Москва  
2024**

## Содержание

1. Краткая характеристика дисциплины (модуля) .....	3
2. Перечень планируемых результатов обучения.....	4
3. Тематический план.....	6
4. Содержание дисциплины (модуля).....	6
5. Учебно-методическое обеспечение .....	7
6. Материально-техническое обеспечение .....	7
7. Методические и оценочные материалы .....	9

## 1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Computer vision (Компьютерное зрение)» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по специальности 02.04.01 Математика и компьютерные науки, профиль Backend-разработка, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Computer vision (Компьютерное зрение)» позволяет понять, как машины могут интерпретировать и анализировать визуальную информацию, что открывает новые возможности в таких областях, как автономные транспортные средства, медицинская диагностика и безопасность. Эта дисциплина (модуль) сочетает в себе элементы машинного обучения, обработки изображений и искусственного интеллекта, что делает её ключевой для развития современных технологий и инновационных решений.

### Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Backend-разработка и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений.

Дисциплина (модуль) изучается на 2 курсе в 3 семестре, доступна для прохождения при условии успешного завершения дисциплины (модуля) «Deep Learning (Глубокое обучение)».

**Цель изучения дисциплины (модуля):** освоение методов и алгоритмов для автоматического анализа, обработки и понимания визуальной информации с помощью компьютерных систем.

### Задачи изучения дисциплины (модуля):

— формирование знаний по темам: принципы получения и хранения изображения в компьютере, классические необучаемые методы обработки изображений, архитектуры бэкбонов нейронных сетей для задач классификации, сегментации, детекции, адаптация трансформерных архитектур для области CV, устройство и принципы работы архитектур нейросетей для задач сегментации, pose estimation и распознавания действий, актуальные подходы в задаче детектирования объектов: two-stage/anchor-based, anchor-free;

— приобретение знаний и развитие понимания по темам: устройство и принципы работы фундаментальных моделей в CV, таких как CLIP, SAM, Grounding DINO, подходы для обучения моделей эффективного извлечения признаков, в частности для задач FaceRec/ReID/Retrieval, основные методы для трекинга одного/множества объектов, пайплайн оптического распознавания символов, современные методы обучения сетей без учителя, подходы к генерации изображений, в том числе генеративные и диффузионные модели;

— освоение умений применять классические методы обработки изображений, обучать нейронные сети для задач классификации/ детекции/ сегментации/ нахождения ключевых точек/распознавания действий, решать задачи эффективно за счет построения лучшего пайплайна обучения и использования современных библиотек, применять фундаментальные модели для решения задач, решать задачи не только на статичных кадрах, но и на видеопотоке;

— формирование навыков проведения и оценки экспериментов по обучению сетей разнообразных поставленной задаче, декомпозирования сложносоставных задач на более мелкие и создания пайплайна, решающего целевую задачу.

## 2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-6.	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1.	Знает основные методы самооценки и анализа своей деятельности, а также принципы управления временем и целеполагания
		УК-6.2	Умеет ставить реалистичные и достижимые цели, определять приоритеты в своей деятельности, а также разрабатывать и внедрять планы по совершенствованию своих навыков и компетенций на основе полученной самооценки
		УК-6.3	Имеет практический опыт применения методов самооценки в своей профессиональной деятельности, включая участие в тренингах, семинарах и проектах, направленных на развитие личной эффективности и профессионального роста
ОПК-2.	Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.1.	Знает основные математические модели и методы, используемые в естественных науках, включая статистическое моделирование, дифференциальные уравнения и численные методы, а также современные подходы к исследованию и анализу данных
		ОПК-2.2	Умеет разрабатывать и адаптировать математические модели для решения конкретных проблем в естественных науках, проводить их анализ и верификацию, а также интерпретировать полученные результаты в контексте научных исследований
		ОПК-2.3	Имеет практический опыт создания и исследования математических моделей в рамках научных проектов или

			исследований, включая участие в публикациях, конференциях или коллаборациях, где были разработаны и апробированы новые концепции и методы
ПК-3.	Способен решать задачи профессиональной деятельности, формулировать результат, увидеть следствия полученного результата	ПК-3.1.	Знает основные принципы и методы решения задач профессиональной деятельности, а также способы формулирования и представления результатов, включая анализ последствий и их значимость в контексте проекта
		ПК-3.2.	Умеет применять математические и компьютерные методы для решения конкретных задач, формулировать четкие и обоснованные результаты, а также анализировать их последствия для дальнейших действий и решений
		ПК-3.3.	Имеет практический опыт в решении профессиональных задач, включая участие в проектах, где были получены результаты и проанализированы их следствия, что способствовало принятию обоснованных решений
ПК-4.	Способен публично представлять собственные и известные научные результаты	ПК-4.1.	Знает основные принципы эффективного публичного выступления, методы визуализации данных и основные требования к научным презентациям, включая структуру и содержание
		ПК-4.2.	Умеет четко и логично формулировать свои научные результаты, адаптируя их для различных аудиторий, а также использовать визуальные средства для улучшения восприятия информации
		ПК-4.3.	Имеет практический опыт участия в научных конференциях, семинарах или других мероприятиях, где успешно представлял свои и известные научные результаты, получая обратную связь и взаимодействуя с аудиторией

### 3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостоятельная работа	
Лекции	Семинары (практические занятия)					
1	Базовые задачи и методы CV	8	8		39	Домашние задания, Тесты
2	Продвинутые задачи и методы CV	8	8		39	Домашние задания, Тесты
	<i>Зачет с оценкой</i>			4		
	<i>Итого:</i>	<i>16</i>	<i>16</i>	<i>4</i>	<i>78</i>	
	<i>Объем дисциплины (модуля) (в ак. ч.)</i>	<i>114</i>				
	<i>Объем дисциплины (модуля) (в зач. ед.)</i>	<i>3</i>				

### 4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Базовые задачи и методы CV	Изображение. Классические методы в CV. Современные архитектуры CNN. Построение эффективного пайплайна обучения сети. Vision Transformers, CLIP. Сегментация изображений, SAM. Современные детекторы объектов. Key-Point/Pose Estimation.
2	Продвинутые задачи и методы CV	Face Recognition/Re-Identification/Image Retrieval. Генерация. Multi-Object Tracking (MOT). Optical Character Recognition (OCR). SOTA Self-Supervised Learning. Ускорение и инференс моделей. Vision Language Models (VLMs). Action Recognition.

## 5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

### *Основная литература:*

1. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил. — (Серия «Библиотека программиста»). — ISBN 978-5-496-02536-2.
2. Richard Szeliski. Computer Vision: Algorithms and Applications. – Springer Science & Business Media, 2010.

### *Дополнительная литература:*

1. Траск Э. Грокаем глубокое обучение. — СПб.: Питер, 2019. — 352 с.: ил. — (Серия «Библиотека программиста»). — ISBN 978-5-4461-1334-7.
2. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил. – ISBN 978-5-97060-618-6.

## 6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	<a href="https://elibrary.ru/defaultx.asp">https://elibrary.ru/defaultx.asp</a>
2.	База данных для IT-специалистов	<a href="https://habr.com">https://habr.com</a>
3.	База данных ScienceDirect	<a href="https://www.sciencedirect.com">https://www.sciencedirect.com</a>
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	<a href="https://minobrnauki.gov.ru/">https://minobrnauki.gov.ru/</a>
5.	Федеральный портал «Российское образование»	<a href="https://www.edu.ru/">https://www.edu.ru/</a>
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	<a href="http://window.edu.ru/">http://window.edu.ru/</a>
7.	Единая коллекция цифровых образовательных ресурсов	<a href="http://school-collection.edu.ru/">http://school-collection.edu.ru/</a>
8.	Федеральный центр информационно - образовательных ресурсов	<a href="http://fcior.edu.ru/">http://fcior.edu.ru/</a>

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
<b>Операционные системы:</b>		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
<b>Браузеры:</b>		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
<b>Офисные приложения:</b>		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
<b>Программное обеспечение для планирования и учета времени:</b>		
Toggle app	зарубежное	свободно распространяемое
<b>Системы управления проектами:</b>		
Microsoft Imagine (Project)	зарубежное	лицензионное
<b>Системы управления базами данных:</b>		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
<b>Системы резервного копирования (backup):</b>		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
<b>Справочно-правовые системы:</b>		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
<b>Средства антивирусной защиты:</b>		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
<b>Среды разработки:</b>		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое

Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
<b>Пакеты программных средств и библиотек:</b>		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
<b>Системы управления библиографической информацией:</b>		
Zotero	зарубежное	свободно распространяемое
<b>Сервисы и службы:</b>		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

## 7. Методические и оценочные материалы

### Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Computer vision (Компьютерное зрение)» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, практические занятия, домашние задания, тесты, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

*Лекция* – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

*Участие в семинаре (практическом занятии)* – активная работа студента на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре студентам рекомендуется заранее ознакомиться с темой обсуждения, прочитать необходимые материалы и подготовить вопросы. Важно активно слушать и вовлекаться в дискуссию, высказывая свои мнения и аргументируя их. При ответах на вопросы преподавателя стоит быть уверенным, четким и логичным, опираясь на изученный материал. Также полезно поддерживать диалог с однокурсниками, чтобы обогатить обсуждение и расширить свои знания.

*Домашнее задание* – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

*Тест* – особая форма проверки знаний. Проводится после освоения одной или нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины (модуля).

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины (модуля), понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет

преподавателю быстро проверить уровень понимания студентами важных методологических категорий.

*Самостоятельная работа* – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

### **Система оценивания результатов обучения по дисциплине (модулю)**

#### **Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Computer vision (Компьютерное зрение)»**

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

**Промежуточная аттестация** по дисциплине (модулю) осуществляется в форме *зачета с оценкой*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Зачтено	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	Зачтено	
8	Отлично	Зачтено	
7	Хорошо	Зачтено	Студент обладает знаниями предмета почти в полном объеме рабочей
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	Зачтено	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	Зачтено	
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Computer vision (Компьютерное зрение)» оценивается следующим образом:

Активность	Вес	Количество	Описание
<b>Накопительная оценка</b>			
Домашние задания	60%	5	Набор задач по темам недели
Тесты		5	Письменная работа с набором задач, которые нужно решить за ограниченное время
<b>Промежуточная аттестация</b>			
Зачет с оценкой	40%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

**Формула расчёта итоговой оценки по дисциплине (модулю) «Computer vision (Компьютерное зрение)»:** « $0,6 \times$  накопительная оценка ( $0,85 \times$  среднее за домашние задания +  $0,15 \times$  среднее за тесты) +  $0,4 \times$  зачет с оценкой».

## **Текущий контроль успеваемости обучающихся по дисциплине (модулю)**

### **Примерные домашние задания**

#### **Домашнее задание 1.**

##### **Задание 1.**

1. Объясните, что такое цифровое изображение, и опишите основные параметры, характеризующие изображение (разрешение, глубина цвета, цветовое пространство).
2. Реализуйте на Python классический метод обнаружения границ на изображении (например, оператор Собеля или Кэнни) с использованием библиотеки OpenCV.
3. Проведите эксперимент: примените выбранный метод к нескольким изображениям с разной степенью шумов и прокомментируйте полученные результаты.

##### **Задание 2.**

1. Изучите архитектуру ResNet и опишите ключевую идею остаточных блоков (residual blocks) и их влияние на обучение глубоких сетей.
2. Постройте пайплайн обучения простой сверточной нейросети для задачи классификации изображений (например, CIFAR-10) с использованием PyTorch или TensorFlow. Включите этапы подготовки данных, настройки оптимизатора, функции потерь и метрик.
3. Проведите эксперимент с различными гиперпараметрами (например, размером батча, скоростью обучения) и опишите, как они влияют на обучение и качество модели.

##### **Задание 3.**

1. Опишите принцип работы Vision Transformer (ViT) и объясните, как происходит разбиение изображения на патчи и их обработка.
2. Ознакомьтесь с моделью CLIP (Contrastive Language-Image Pre-training) и кратко изложите, как она объединяет визуальные и текстовые представления.
3. Реализуйте на практике простой пример использования предобученной модели CLIP для поиска изображений по текстовому запросу с помощью библиотеки OpenAI CLIP или аналогичной.

#### **Домашнее задание 2.**

##### **Задание 1.**

1. Изучите основные методы сегментации изображений, такие как сегментация на основе порогов, метод активных контуров (snakes) и сегментация с использованием нейронных сетей (например, U-Net). Напишите краткое описание каждого метода.
2. Реализуйте алгоритм сегментации изображения с использованием модели SAM (Segment Anything Model) на наборе данных (например, COCO или ваш собственный набор изображений). Оцените качество сегментации с помощью метрик, таких как IoU (Intersection over Union).
3. Проведите эксперимент, изменяя параметры модели SAM, и проанализируйте, как это влияет на качество сегментации.

##### **Задание 2.**

1. Изучите архитектуры современных детекторов объектов, такие как YOLO, SSD и Faster R-CNN. Напишите краткое сравнение их особенностей и областей применения.

2. Реализуйте детектор объектов (например, YOLOv5) на наборе данных (например, Pascal VOC или COCO). Оцените производительность модели с помощью метрик, таких как mAP (mean Average Precision).

3. Изучите методы оценки позы (Key-Point Estimation), такие как OpenPose или PoseNet. Реализуйте один из этих методов на наборе данных и визуализируйте результаты, отметив ключевые точки на изображениях.

### **Задание 3.**

1. Изучите методы распознавания лиц (Face Recognition) и переидентификации (Re-Identification). Напишите краткое описание подходов, таких как FaceNet и DeepFace, а также их применения.

2. Реализуйте систему распознавания лиц с использованием предобученной модели (например, FaceNet) на наборе изображений. Оцените точность распознавания на тестовом наборе данных.

3. Изучите методы многократного отслеживания объектов (MOT), такие как SORT (Simple Online and Realtime Tracking) и Deep SORT. Реализуйте один из методов на видео с несколькими движущимися объектами и визуализируйте результаты отслеживания.

## **Домашнее задание 3.**

### **Задание 1.**

1. Изучите основные этапы OCR: предобработка изображения, сегментация символов, распознавание и постобработка. Опишите типичные проблемы OCR и методы их решения.

2. Реализуйте простую систему OCR на базе открытой библиотеки (например, Tesseract) и протестируйте её на наборе сканированных документов или фотографий с текстом.

3. Изучите методы распознавания действий (Action Recognition) в видео, например, 3D-CNN или модели на основе трансформеров. Опишите ключевые особенности и задачи, которые решают эти методы.

### **Задание 2.**

1. Ознакомьтесь с принципами самоконтролируемого обучения (Self-Supervised Learning) и опишите, как оно применяется в компьютерном зрении (например, методы contrastive learning: SimCLR, MoCo).

2. Реализуйте тренировку простой модели самоконтролируемого обучения на наборе изображений (например, CIFAR-10) с использованием одного из популярных методов.

3. Изучите архитектуру и применение Vision Language Models (VLMs), таких как CLIP или Flamingo. Опишите, как они связывают визуальные и текстовые данные.

### **Задание 3.**

1. Изучите методы оптимизации моделей для ускорения инференса: квантование, праунинг, knowledge distillation. Опишите принципы каждого метода.

2. Выберите предобученную модель (например, ResNet) и проведите эксперимент по её оптимизации с помощью одного из методов ускорения. Сравните скорость и точность до и после оптимизации.

3. Исследуйте библиотеки и инструменты для ускорения инференса (TensorRT, ONNX Runtime, OpenVINO) и опишите их основные возможности.

## Примерные задания по тестам

### Тест 1.

#### Вопрос 1.

Какова основная функция фотодиодов в камере?

- A. Увеличение яркости изображения
- B. Преобразование света в электрический сигнал
- C. Сохранение изображения на носителе
- D. Управление экспозицией камеры

Ответ: B.

#### Вопрос 2.

Что такое дебайеризация?

- A. Процесс преобразования черно-белого изображения в цветное
- B. Устранение цветных артефактов на изображениях
- C. Преобразование данных с сенсора в полноцветное изображение
- D. Процесс выравнивания изображений

Ответ: C.

#### Вопрос 3.

Каково основное назначение цветовой модели HSV?

- A. Отображение цвета в терминах яркости
- B. Создание черно-белых изображений
- C. Облегчение работы с цветами, учитывая восприятие человека
- D. Хранение изображения в трех других первичных цветах: циан, маджента, желтый

Ответ: D.

#### Вопрос 4.

Какой цвет соответствует комбинации всех трех первичных цветов в модели RGB?

- A. Черный
- B. Белый
- C. Серый
- D. Красный

Ответ: B.

#### Вопрос 5.

Что такое попиксельная операция?

- A. Операция, применяющаяся ко всему изображению целиком
- B. Операция, применяемая к отдельным пикселям изображения
- C. Процесс пересчета значения пикселя в новой цветовой модели
- D. Улучшение контраста изображения

Ответ: B.

### Тест 2.

#### Вопрос 1.

Какого типа детектирования не существует?

- A. Two-stage
- B. Three-stage
- C. Anchor-based
- D. Anchor-free

Ответ: B.

#### Вопрос 2.

Какая из этих метрик/операций не используется при подсчете AP?

- A. IoU
- B. Precision

C. Recall

D. NMS

Ответ: D.

**Вопрос 3.**

Что такое Average Precision?

A. Это площадь под Precision-Recall кривой

B. Это Precision при score=0.5

C. Это средний precision при скорях 0.5..0.95

D. Это среднее геометрическое между Precision и Recall для всех возможных скоров детектора

Ответ: A.

**Вопрос 4.**

Как изменятся показатели метрик при увеличении  $th\_iou$  с 0.5 до 0.95

A. AP уменьшится

B. AP увеличится

C. Максимальный Recall увеличится

D. Максимальный Precision увеличится

Ответ: A.

**Вопрос 5.**

Какой из этих лоссов обычно не используется в задачах детектирования?

A. GIoU Loss

B. CIoU Loss

C. BIoU Loss

D. DIoU Loss

Ответ: C.

**Тест 3.**

**Вопрос 1.**

Выберите **ложные** утверждения о задаче OCR:

A. Задача OCR предназначена для преобразования текстовой информации в изображения.

B. OCR может работать как с печатным, так и с рукописным текстом.

C. Scene Text Recognition - это подзадача OCR, связанная с распознаванием текста на документах

D. Все датасеты OCR содержат расшифрованные тексты.

E. Метрики OCR базируются на расстоянии Левенштейна.

F. Расстояние Левенштейна можно использовать только для сравнения слов одинаковой длины.

G. В задаче OCR обычно рассматриваются две метрики - Character Error Rate и Word Error Rate.

Ответ: A, C, D, F.

**Вопрос 2.**

Выберите **ложное** утверждение о пайплайне OCR:

A. Извлечение ключевой информации зависит от бизнес-задачи.

B. Выравнивание текста перед детекцией необходимо только при работе с печатными документами.

C. Детекция текста — это процесс нахождения и локализации текстовых областей на изображении.

D. Распознавание текста может использовать как CTC-based, так и Attention-based подходы.

E. Для извлечения ключевой информации можно использовать графовые сети и языковые модели.

F. Подходы локализации текста делятся на regression-based и segmentation-based.  
Ответ: B.

### Вопрос 3.

Выберите верные утверждения о детекции текста:

- A. В OCR можно детектировать как отдельные символы, так и слова или строки.
- B. Регрессионные методы лучше подходят для текста произвольной формы.
- C. Сегментационные методы детекции текста определяют текстовые области, предсказывая маски для отдельных букв или слов.
- D. Регрессионные методы детекции текста используют детекторы объектов (anchor-based/anchor-free).
- E. Сложность детекции текста обусловлена только низким качеством изображений.
- F. Сегментационные методы требуют сложной постобработки для извлечения текстовых областей.

Ответ: A, C, D, F.

### Вопрос 4.

Выберите **ложные** утверждения про DBNet:

- A. DBNet использует карту бинаризации для предсказанных хитмапов чтобы выделить текстовые области.
- B. DBNet — это объектный детектор, который использует anchor-based подход.
- C. Бинаризация в DBNet осуществляется с фиксированным порогом, заданным вручную.
- D. DBNet аппроксимирует бинаризацию с помощью непрерывной функции.
- E. В основе DBNet лежит архитектура FPN, которая помогает учитывать контекст на разных уровнях.
- F. DBNet лучше всего подходит для детекции текста на документах, но неэффективна для детекции текста “in the wild”.

Ответ: B, C, F.

### Вопрос 5.

Выберите верные утверждение о модели CRNN:

- A. CRNN объединяет сверточные и рекуррентные нейросети для распознавания текста.
- B. CRNN использует объектный детектор для локализации текста перед распознаванием.
- C. CRNN использует механизм внимания (attention) при предсказании текста.
- D. Выходные векторы CNN в CRNN имеют фиксированную длину, которая соответствует количеству символов в тексте.
- E. CRNN использует сверточные слои для классификации символов.
- F. CRNN является end-to-end решением.
- G. CRNN предсказывает сразу готовую строку текста, для которой не требуется этапа декодирования.
- H. CRNN предсказывает дублирующиеся символы из-за того, что receptive field эмбедингов накладываются друг на друга.

Ответ: A, H.

### Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Какую метрику для оценки качества работы модели классификации картинок стоит выбрать, если данные сильно не сбалансированы по классам? A. Accuracy	B	УК-6

	B. F1-score C. mAP D. IoU		
2.	Какой механизм в сетях ResNet позволил углублять все дальнейшие сети?	Skip connection / Residual Block	ОПК-2
3.	Какой тип нейросетевых моделей позволяет решать задачи VQA – Vision Question Answering и Image Captioning?	VLM / Vision Language Models/ Multimodal Models	ПК-3
4.	Какой первый one-stage подход был предложен для детектирования объектов? Его же последующие версии являются на данный момент самыми используемыми CNN детекторами в CV.	YOLO	ПК-4