

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«24» июня 2025 г.
Протокол № 2

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

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Машинное обучение

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2025

**Москва
2025**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	8
6. Материально-техническое обеспечение	8
7. Методические и оценочные материалы	10

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 Машинное обучение, 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 Машинное обучение и входит в обязательную часть Блока 1.

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): формирование комплексного понимания принципов и технологий обработки изображений и видео для эффективного решения практических задач в области искусственного интеллекта.

Задачи изучения дисциплины (модуля):

- изучить основы цифрового представления изображений и традиционных алгоритмов их анализа для подготовки к работе с современными технологиями;
- освоить структуры глубоких нейронных сетей, предназначенных для распознавания объектов, разделения сцен и локализации элементов на изображениях;
- понять адаптацию трансформерных моделей к задачам компьютерного зрения, таким как категоризация, обнаружение и разбиение изображений;
- ознакомиться с современными стратегиями обнаружения объектов без и с использованием опорных точек для повышения точности и эффективности;
- разработать навыки проектирования экспериментов и декомпозиции сложных задач на подзадачи для создания эффективных решений в области зрения.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- принципы получения и хранения изображения в компьютере, классические необучаемые методы обработки изображений;
- как устроены архитектуры бэкбонов нейронных сетей для задач классификации, сегментации, детекции;
- как трансформерные архитектуры адаптированы для области CV в задачах классификации/детекции/сегментации;
- актуальные подходы в задаче детектирования объектов: anchor-based, anchor-free;
- устройство и принципы работы архитектур нейросетей для задач сегментации и pose estimation;
- устройство и принципы работы фундаментальных моделей в CV и VLM;
- 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.

уметь:

- 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.

владеть:

- навыками проведения и оценки экспериментов по обучению сетей сообразных поставленной задаче;
- навыками декомпозирования сложносоставных задач на более мелкие и созданий 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)
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
ОПК-3.	Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.1.	Знает основные принципы программирования, архитектуры программного обеспечения и современные языки программирования, а также особенности отечественных информационных технологий и сетевых ресурсов
		ОПК-3.2.	Умеет разрабатывать прикладные программные средства, используя современные инструменты и технологии, а также интегрировать их с сетевыми ресурсами для решения конкретных задач

		ОПК-3.3.	Имеет практический опыт разработки программных средств, используемых при построении математических моделей в естественных науках
ПК-2.	Способен математически корректно ставить естественнонаучные и прикладные задачи	ПК-2.1.	Знает основные методы и подходы к математическому моделированию, а также теоретические основы естественных и прикладных наук, необходимые для корректной формулировки задач
		ПК-2.2.	Умеет анализировать практические ситуации и формулировать на их основе математические модели, включая выбор адекватных методов решения и формулировку условий задачи
		ПК-2.3.	Имеет практический опыт в разработке и решении математических задач в рамках проектов или научных исследований, где были успешно поставлены и решены естественнонаучные и прикладные задачи
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

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостоя тельная работа	
Лекции	Семинары (практичес кие занятия)					
1	CV до эпохи нейронных сетей	8	4		14	Домашние задания, Тесты
2	Задача классификации в CV	12	6		19	Домашние задания, Тесты
3	Базовые задачи в CV	16	8		25	Домашние задания, Тесты, Соревнование
4	Специализированные задачи в CV	12	6		19	Домашние задания, Тесты
5	Large Models в CV	12	6		19	Домашние задания, Тесты, Соревнование
	<i>Экзамен</i>			4		
	Итого:	60	30	4	96	
	Объем дисциплины (модуля) (в ак. ч.)	190				
	Объем дисциплины (модуля) (в зач. ед.)	5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	CV до эпохи нейронных сетей	Введение в дисциплину. Изображение. Базовые операции Классические методы обработки изображений
2	Задача классификации в CV	Архитектуры CNN Эффективный пайплайн обучения сети Трансформеры в CV: ViT, Deit, Swin
3	Базовые задачи в CV	Keypoint Detection Сегментация изображений, SAM Детекция объектов: anchor-based Детекция объектов: anchor-free
4	Специализированные задачи в CV	Face Recognition/Re-Identification/Image Retrieval Multi-Object Tracking (MOT) Optical Character Recognition (OCR)
5	Large Models в CV	Self-Supervised Learning Vision Language Models (VLMs) Overview

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. Компьютерное зрение. Современные методы и перспективы развития : монография / ред. Р. Дэвис, М. Терк ; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2022. - 692 с. - ISBN 978-5-93700-148-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2109506>.

2. Коул, А. Искусственный интеллект и компьютерное зрение. Реальные проекты на Python, Keras и TensorFlow : практическое руководство / А. Коул, С. Ганджу, М. Казам. - Санкт-Петербург : Питер, 2023. - 624 с. - (Бестселлеры O'Reilly). - ISBN 978-5-4461-1840-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123884>.

Дополнительная литература:

1. Прохоренок, Н. А. OpenCV и Java. Обработка изображений и компьютерное зрение : практическое руководство / Н. А. Прохоренок. - Санкт-Петербург : БХВ-Петербург, 2018. - 321 с. - (Профессиональное программирование). - ISBN 978-5-9775-3955-5. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1856805>.

2. Корк, П. Машинное зрение. Основы и алгоритмы с примерами на Matlab : практическое руководство / П. Корк ; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2023. - 586 с. - ISBN 978-5-93700-222-8. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2204248>.

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;
- механическими калькуляторами;
- 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 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	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

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
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		

КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
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
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
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, семинары, домашние задания, тесты, соревнования, 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.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

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.

Соревнование – организованное мероприятие, в рамках которого участники соперничают друг с другом для достижения определенной цели, демонстрируя свои навыки, знания или способности в заданной области.

Бонусные баллы — это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

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 **экзамена**, 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	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	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
9	Отлично	
8	Отлично	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		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.
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей 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.
6	Хорошо	
5	Удовлетворительно	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	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

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

Активность	Вес	Описание
Накопительная оценка		
Домашние задания	60%	Набор задач по темам недели
Тесты		Письменная работа с набором задач, которые нужно решить за ограниченное время
Соревнования		Мероприятие, в рамках которого участники соперничают друг с другом для достижения определенной цели, демонстрируя свои навыки, знания или способности в заданной области
Промежуточная аттестация		
Экзамен	40%	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Итоговая оценка рассчитывается по накопительной при условии, если средний балл студента составляет 4 и более баллов, по формуле: « $0,55 \times$ среднее за домашние задания + $0,2 \times$ среднее за тесты + $0,25 \times$ среднее за соревнования».

Если студент не выполняет условие для получения оценки по накопительной системе, ему необходимо сдать экзамен. В данном случае формула расчёта итоговой оценки по дисциплине (модулю) «Computer vision (Компьютерное зрение)»: « $0,6 \times$ накопительная оценка ($0,55 \times$ среднее за домашние задания + $0,2 \times$ среднее за тесты + $0,25 \times$ среднее за соревнования) + $0,4 \times$ экзамен».

В случае, если средний балл студента составляет 4 и более баллов, но он хочет улучшить оценку, итоговая оценка по дисциплине (модулю) выставляется по формуле: « $0,7 \times$ накопительная оценка ($0,55 \times$ среднее за домашние задания + $0,2 \times$ среднее за тесты + $0,25 \times$ среднее за соревнования) + $0,3 \times$ экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1.

Ваша задача — изучить репозиторий [ultralytics](#) и обучить сеть YOLO для детекции объектов на датасете [SkyFusion](#). Это специализированный датасет для обнаружения малых объектов (транспорт, корабли и самолёты) на спутниковых снимках. В процессе выполнения задания вы пройдёте через несколько этапов, за каждый из которых получите баллы:

- **2 балла** — конвертация train и valid частей датасета в формат разметки YOLO;
- **4 балла** — обучение сети (например, YOLOv8n) с default-параметрами;
- **6 баллов** — добавить [ClearML](#)-логирование эксперимента;
- **8 баллов** — любая простая модификация конфига модели (например, yolov8.yaml);
- **10 баллов** — побить результат mAP:0,5 > 0,62 на test-сет, который уже приложен для вас.

Ваш результат должен превышать указанные значения метрик, чтобы получить соответствующее количество баллов.

Инструкция

1. **Начните** с [ultralytics quickstart](#) или обратитесь к ноутбуку с практикой.
2. **Обязательно ознакомьтесь** с примечаниями.
3. **Измените код по своему усмотрению**, чтобы получить требуемое количество баллов.
4. **Запустите код** и убедитесь, что результаты корректно отображены. Рекомендуется сохранить логи.
5. **Создайте папку** с именем в формате «ФИО_HW_номер_дз». В эту папку загрузите всю папку ultralytics, в которой уже лежит обученная модель, все скрипты и скрины.

Примечание

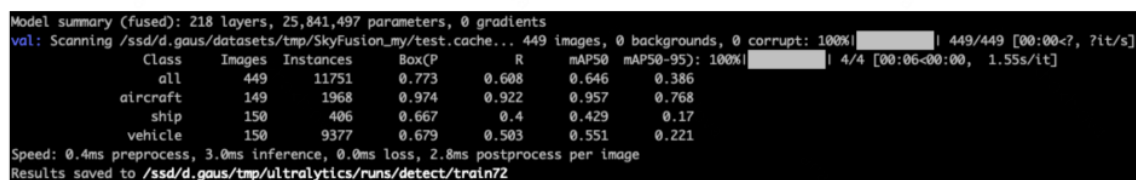
1. Тестовый набор данных уже представлен в формате YOLO и не требует дополнительной конвертации.
2. Для оценки результатов на тестовой выборке используйте скрипт `model.val(data="SkyFusion.yaml", split='test')`, где в конфигурационном файле `SkyFusion.yaml` уже определён путь к тестовым данным:
`path: /ssd/datasets/tmp/SkyFusion/`

...

`test: test/`

Это указывает на расположение тестовых изображений в директории `/ssd/datasets/tmp/SkyFusion/test`.

3. Для подтверждения интеграции с ClearML включите в итоговый архив файл `clearml.txt`, содержащий ссылку на ваш эксперимент с графиками. Убедитесь, что эксперимент настроен на публичный доступ и виден внешним пользователям.
4. В библиотеке `ultralytics` уже реализована интеграция с ClearML, поэтому дополнительная модификация кода для логирования не потребуется.
5. Для получения максимальных 10 баллов (при достижении $mAP:0,5 > 0,62$) необходимо приложить скриншот результатов теста, например:



```
Model summary (fused): 218 layers, 25,841,497 parameters, 0 gradients
val: Scanning /ssd/d.gaus/datasets/tmp/SkyFusion_my/test.cache... 449 images, 0 backgrounds, 0 corrupt: 100%|██████████| 449/449 [00:00<7, ?it/s]
      Class  Images  Instances  Box(P)   R      mAP50  mAP50-95: 100%|██████████| 4/4 [00:06<00:00, 1.55s/it]
      all      449     11751    0.773   0.608   0.646   0.386
  aircraft    149     1968    0.974   0.922   0.957   0.768
      ship     150      406    0.667   0.4     0.429   0.17
      vehicle  150     9377    0.679   0.503   0.551   0.221
Speed: 0.4ms preprocess, 3.0ms inference, 0.0ms loss, 2.8ms postprocess per image
Results saved to /ssd/d.gaus/tmp/ultralytics/runs/detect/train72
```

В скриншоте должна быть отражена строка "all classes" и значение метрики `mAP50` (как в примере, где достигнуто значение 0,646).

6. Строго запрещается включать тестовые данные (папку `test`) в обучающую выборку. Это является нарушением и будет проверяться при оценке работы.

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.

1. Какова основная функция фотодиодов в камере?
- Увеличение яркости изображения
 - Преобразование света в электрический сигнал
 - Сохранение изображения на носителе
 - Управление экспозицией камеры

Ответ: b.

2. Что такое дебайеризация?
- Процесс преобразования черно-белого изображения в цветное
 - Устранение цветных артефактов на изображениях
 - Преобразование данных с сенсора в полноцветное изображение
 - Процесс выравнивания изображений

Ответ: c.

3. Каково основное назначение цветовой модели HSV?
- Отображение цвета в терминах яркости
 - Создание черно-белых изображений
 - Облегчение работы с цветами, учитывая восприятие человека
 - Хранение изображения в трех других первичных цветах: циан, маджента, желтый

Ответ: c.

4. Какой цвет соответствует комбинации всех трех первичных цветов в модели RGB?
- Черный
 - Белый
 - Серый
 - Красный

Ответ: b.

5. Что такое попиксельная операция?
- Операция, применяющаяся ко всему изображению целиком
 - Операция, применяемая к отдельным пикселям изображения
 - Процесс пересчета значения пикселя в новой цветовой модели
 - Улучшение контраста изображения

Ответ: b.

6. Какая операция используется для повышения контраста изображения?
- Выравнивание гистограммы
 - Бинаризация
 - Интерполяция
 - Дебайеризация

Ответ: a.

7. Как называется процесс изменения разрешения изображения?
- Модуляция
 - Интерполяция
 - Гомография
 - Коррекция

Ответ: b.

8. Что такое алиасинг в контексте цифровой обработки изображений?
- Возникновение нежелательных артефактов в результате дискретизации
 - Процесс сглаживания изображения
 - Увеличение разрешения изображения
 - Уменьшение разрешения изображения

Ответ: a.

9. Какой из следующих методов чаще всего используется для борьбы с алиасингом в обработке изображений?

- a. Предварительное увеличение разрешения
- b. Размытие изображения с помощью фильтра Гаусса
- c. Удаление шумов с помощью морфологических операций
- d. Использование более сложных методов интерполяции

Ответ: b.

10. Какой метод используется для преобразования цветного изображения в двухцветное?

- a. Выравнивание гистограммы
- b. Бинаризация
- c. Дебайеризация
- d. Морфологическая операция

Ответ: b.

Тест 2.

1. Какого типа детектирования не существует?

- a. Two-stage
- b. Three-stage
- c. Anchor-based
- d. Anchor-free

Ответ: b.

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

- a. IoU
- b. Precision
- c. Recall
- d. NMS

Ответ: d.

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.

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.

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.

6. Что из перечисленных ниже трюков не привело к улучшению качества детектирования у Mask R-CNN?

- a. Использование RoIAlign
- b. Использование ResNeXt50 + FPN
- c. Использование bipartite лосса и NMS-free инференса

- d. Обучение сразу на две задачи: детекцию и сегментацию

Ответ: с.

7. Какой из перечисленных вариантов правильно описывает принцип инференса в головах у SSD?

- a. На первых головах детектируются самые маленькие объекты, на последних – самые большие
- b. На первых головах детектируются самые большие объекты, на последних – самые маленькие
- c. На средних головах детектируются самые большие объекты, на первых и последних – самые маленькие
- d. Любые объекты детектируются на любых головах

Ответ: а.

8. Что предсказывают локализационные выходы anchor-based детекторов?

- a. Координаты центров боксов, w , h
- b. Поправки к расположению и размеру якорей
- c. Координаты левого верхнего и правого нижнего углов боксов
- d. Нормализованные координаты центров боксов, w , h

Ответ: b.

9. Какой из этих детекторов не является anchor-free?

- a. FCOS
- b. CenterNet
- c. CornerNet
- d. RetinaNet

Ответ: d.

10. Какого из этих элементов нет в типичном построении современного детектора?

- a. Бэкбона
- b. Голов(ы)
- c. Шеи
- d. Корня

Ответ: d.

Тест 3.

1. Какой подвид задачи Keypoint Detection не является особо распространенным?

- a. Human Pose Estimation
- b. Facial Landmarks Detection
- c. Vehicle Keypoint Detection
- d. Hand Pose Estimation

Ответ: с.

2. Какую из этих метрик не используют для оценки Keypoint Detection?

- a. OKS
- b. PCKh-0.5
- c. PCJ
- d. AP
- e. AR

Ответ: e.

3. Какой тип предсказания не используют для НРЕ?

- a. Предсказание тепловых карт точек
- b. Регрессия координат точек
- c. Детекция отдельных частей тела и выбор точек по боксам
- d. Одновременное предсказание тепловых карт точек и направлений их соединений

Ответ: с.

4. Какие есть отличительные особенности практически у всех CNN методов для НРЕ?
- Использование Attention
 - Каскадное построение сети
 - Комбинация heatmaps и регрессии
 - Комбинация глобального и локального контекста
 - Несколько голов для предсказаний

Ответ: b, d.

5. Какого свойства **нет** у Stacked Hourglass сети?
- Переагрегация признаков с помощью SE блока
 - Промежуточная супервизия
 - Стекинг архитектурно идентичных модулей
 - Использование L2 лосса
 - Сжатие и разжатие feature maps с использованием skip connections

Ответ: a.

6. Какого из этих шагов нет в пайплайне распознавания лиц?
- Детекция лиц
 - Детекция ключевых точек лица
 - Трекинг лиц
 - Проверка Liveness
 - Нормализация лиц
 - Матчинг дескрипторов ключевых точек

Ответ: f.

7. **Каких** типов задач не бывает в распознавании лиц?
- Аутентификация
 - Верификация
 - Идентификация
 - Классификация

Ответ: a, d.

8. **Какие** из этих проблем характерны для датасетов для FaceRec-a?
- Несбалансированные по расам/полу/возрасту распределения с длинными хвостами
 - Грязная разметка: один и тот же ID в разных классах, разные ID в одном классе
 - Все датасеты собраны азиатами и не применимы для других доменов
 - Обучение и тест пересекаются по ID
 - Датасеты состоят только из знаменитостей, получается слишком мало уникальных ID

Ответ: a, b, d.

9. Что мы называем центроидой класса в Softmax Loss-e?
- Столбцы матрицы весов FC слоя
 - Строки матрицы весов FC слоя
 - Эмбединг, получающийся перед FC слоем
 - Средний эмбединг посчитанный по данным одного класса

Ответ: b.

10. Какие из перечисленных лоссов явно работают со сближением/отталкиванием дескрипторов?
- Softmax Loss
 - Contrastive Loss
 - Triplet Loss
 - CosFace

Ответ: b, c.

Примерное задание к соревнованию

1. Формат соревнования

Участники (индивидуально или в мини-группах) получают датасет изображений для задачи классификации.

Цель — построить модель, показывающую наилучшее качество на скрытой тестовой выборке.

Соревнование проходит в течение одного занятия:

- первая часть — разработка и обучение модели;
- финальная часть — отправка предсказаний и формирование рейтинга.

Победитель определяется по значению основной метрики (например, Accurasy или F1-score).

2. Задание

Этап 1. Базовая обработка изображений

1. Проанализировать данные:

- размер изображений,
- количество классов,
- баланс классов.

2. Выполнить предобработку:

- ресайз,
- нормализацию,
- базовые аугментации (по возможности).

Этап 2. Построение моделей

Необходимо реализовать и сравнить два подхода:

1. Базовый подход

Один из вариантов:

- простая CNN (2–3 сверточных слоя),
- либо классификация на основе заранее извлечённых признаков (по желанию).

2. Улучшенный подход

- более глубокая CNN

или

- использование предобученной модели (transfer learning).

Этап 3. Выбор модели

- Разделить данные на обучающую и валидационную выборки.
- Сравнить модели по валидационной метрике.
- Выбрать лучшую модель.
- Сделать финальное предсказание для тестовой выборки.

3. Ограничения

• Запрещено использовать готовые решения с внешним обучением на тестовых данных.

- Разрешено использовать предобученные CNN.
- Тестовая выборка используется только один раз для финальной отправки.

4. Метрика

Основная метрика объявляется в начале занятия:

- Accurasy

или

- F1-score.

Именно по этой метрике формируется итоговый рейтинг.

5. Критерии оценивания

Оценивание включает:

Соревновательный результат

- Значение метрики на тестовой выборке.

- Позиция в итоговом рейтинге.

Качество подхода

- Корректность предобработки изображений.
- Реализация базовой CNN.
- Реализация улучшенной модели.
- Корректное разделение данных на обучение и валидацию.
- Наличие сравнения моделей.
- Обоснование выбора финальной модели.
- Использование корректной терминологии (CNN, аугментация, переобучение и др.).

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? А. Accuracy Б. F1-score В. mAP Г. IoU	В	ОПК-3
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-2
5.	Как называется математическая модель, используемая для описания процессов в компьютерном зрении, таких как фильтрация изображений с помощью свёрточных ядер?	свёрточная нейронная сеть	ОПК-2
6.	Как называется библиотека для создания прикладных программных средств в компьютерном зрении, позволяющая обучать и инференсить нейронные сети на Python?	PyTorch	ОПК-3
7.	Как называется математически корректная постановка задачи в CV, заключающаяся в предсказании класса объекта на изображении с использованием вероятностных моделей?	задача классификации	ПК-2
8.	Как называется результат решения задачи pose estimation в CV, выраженный в виде набора ключевых точек на теле человека для анализа движений?	скелетная модель	ПК-3

9.	Укажите название метода оптимизации, часто применяемого для обучения моделей в компьютерном зрении, минимизирующего функцию потерь с использованием градиентов.	стохастический градиентный спуск	ОПК-2
10.	Укажите название фреймворка, используемого для реализации пайплайнов обработки изображений и видео в CV, с поддержкой GPU-ускорения.	OpenCV	ОПК-3
11.	Укажите название метрики оценки качества детекции объектов, измеряющей среднюю точность на разных уровнях пересечения предсказаний и истинных областей.	mean average precision	ПК-2
12.	Как называется следствие применения фундаментальных моделей в CV, таких как CLIP, для решения задач VQA или Image Captioning?	мультимодальное понимание	ПК-3
13.	Как называется математическая операция, лежащая в основе трансформерных архитектур в CV, позволяющая модели учитывать зависимости между всеми элементами последовательности?	механизм внимания	ОПК-2
14.	Как называется тип архитектуры нейросетей, адаптированный для задач сегментации изображений, такой как U-Net или Mask R-CNN?	encoder-decoder	ОПК-3
15.	Как называется подход к постановке задачи сегментации, где каждый пиксель присваивается к классу, с учётом контекста соседних пикселей?	семантическая сегментация	ПК-2
16.	Укажите название типа экспериментов в CV, оценивающих точность модели на тестовом наборе данных после обучения на видеопотоке.	валидация на видео	ПК-3