ЛЕКЦИЯ 10

МЕТОДЫ И ТЕХНОЛОГИИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Основные понятия • определения • схемы • таблицы • графики • примеры

Преподаватель: проф.Ахмедиярова А.Т.

Структура и основные разделы

01 Понятие компьютерного зрения

02 Изображение как цифровые данные

03 Почему понимание изображения является сложной задачей

04 Предобработка изображений

05 Выделение границ изображения

06 От граничных элементов к контурам

07 Преобразование Хафа

08 Выделение областей изображения

09 Основные задачи современного компьютерного зрения

10 Классификация изображений

11 Детекция объектов

12 Сегментация

Понятие компьютерного зрения

Компьютерное зрение (Computer Vision) — направление искусственного интеллекта, связанное с получением, обработкой и интерпретацией визуальной информации. В учебнике В.Н. Бондарева и Ф.Г. Аде компьютерное зрение рассматривается как область, направленная на методы понимания изображений и формирование на их основе символического описания реальности.

Ключевая идея: Компьютерное зрение (Computer Vision) — направление искусственного интеллекта, связанное с получением, обработкой и интерпрет

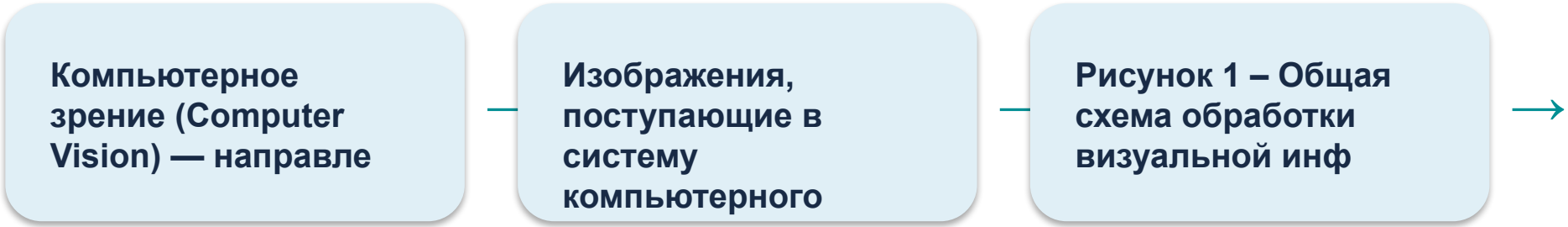
Понятие компьютерного зрения — продолжение

Изображения, поступающие в систему компьютерного зрения, могут быть различными: двумерными чёрно-белыми или цветными изображениями, стереоизображениями, диапазонными изображениями от датчиков глубины, последовательностями кадров и трёхмерными данными. Поэтому компьютерное зрение охватывает не одну отдельную задачу, а широкий класс методов анализа визуальной информации.

Понятие компьютерного зрения — продолжение

Рисунок 1 – Общая схема обработки визуальной информации

Схема: Понятие компьютерного зрения



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Изображение как цифровые данные

Для компьютера изображение первоначально представляет собой массив числовых значений. В чёрно-белом изображении каждому пикселю соответствует значение интенсивности. В цветном изображении пиксель обычно представлен несколькими компонентами, например каналами RGB.

Ключевая идея: Для компьютера изображение первоначально представляет собой массив числовых значений. В чёрно-белом изображении каждому пиксе

Изображение как цифровые данные — продолжение

Рисунок 2 – Упрощённое представление изображения как матрицы значений Если использовать 8 бит для одного канала интенсивности, значение пикселя может находиться в диапазоне от 0 до 255. При этом 0 соответствует минимальной интенсивности, а 255 — максимальной.

Изображение как цифровые данные — продолжение

Цветное изображение можно рассматривать как несколько матриц одинакового размера, каждая из которых описывает отдельный цветовой канал.

Сравнительная таблица: Изображение как цифровые данные

Изображение как цифровые данные	Для компьютера изображение первоначально представляет собой массив числовых значений. В чёрно-белом изображении каж
Почему понимание изображения является сложной задачей	Рисунок 2 – Упрощённое представление изображения как матрицы значений
Предобработка изображений	Если использовать 8 бит для одного канала интенсивности, значение пикселя может находиться в диапазоне от 0 до 255.

Почему понимание изображения является сложной задачей

Учебник подчёркивает, что понимание изображения для человека кажется естественным, но для компьютера является сложной задачей. Одна из причин состоит в том, что двумерное изображение является проекцией трёхмерного мира. При проецировании часть информации, прежде всего глубина и пространственные отношения, теряется.

Ключевая идея: Учебник подчёркивает, что понимание изображения для человека кажется естественным, но для компьютера является сложной задачей

Почему понимание изображения является сложной задачей — продолжение

Другие сложности связаны с шумом датчиков, изменением освещения, тенями, отражающими свойствами поверхности, ракурсом, масштабом, частичным перекрытием объектов и необходимостью обрабатывать большие объёмы данных в реальном времени.

Предобработка изображений

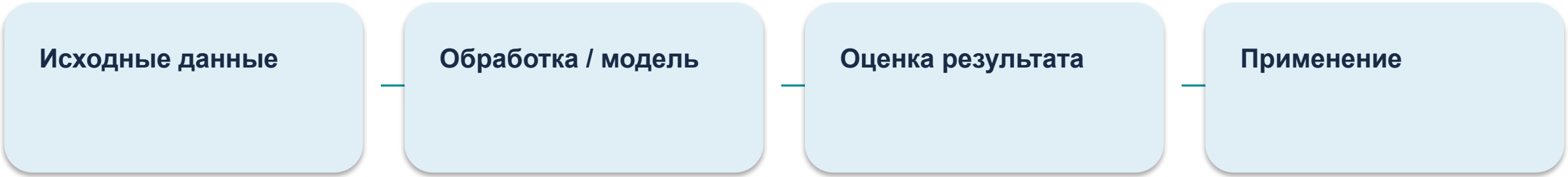
Перед распознаванием изображение часто подвергается предварительной обработке. Цель предобработки — уменьшить влияние шума, привести данные к ожидаемому формату и подготовить их к дальнейшему анализу. В классических системах могут применяться сглаживание, фильтрация, изменение контраста, нормализация яркости и геометрические преобразования.

Ключевая идея: Перед распознаванием изображение часто подвергается предварительной обработке. Цель предобработки — уменьшить влияние шума, п

Предобработка изображений — продолжение

В учебнике в качестве примера уменьшения шума рассматривается сглаживание изображения свёрткой с двумерным гауссовским окном.

Схема: Предобработка изображений



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Выделение границ изображения

Граница — область резкого изменения яркости или другого визуального признака. Выделение границ является одной из базовых операций классического компьютерного зрения. Она позволяет перейти от массива пикселей к более структурированному описанию формы объектов.

Ключевая идея: Граница — область резкого изменения яркости или другого визуального признака. Выделение границ является одной из базовых опер

Выделение границ изображения — продолжение

Рисунок 3 – Упрощённый пример выделенной границы объекта Классический подход заключается в анализе изменений интенсивности по соседним пикселям. После обнаружения отдельных граничных элементов возникает задача объединения их в связанные линии и контуры.

От граничных элементов к контурам

Если обнаружены отдельные точки границы, необходимо установить, какие из них принадлежат одной линии или одному объекту. Для этого применяются методы группировки и поиска. Один из классических методов — преобразование Хафа, позволяющее искать геометрические структуры, например прямые линии, через переход в пространство параметров.

Сравнительная таблица: От граничных элементов к контурам

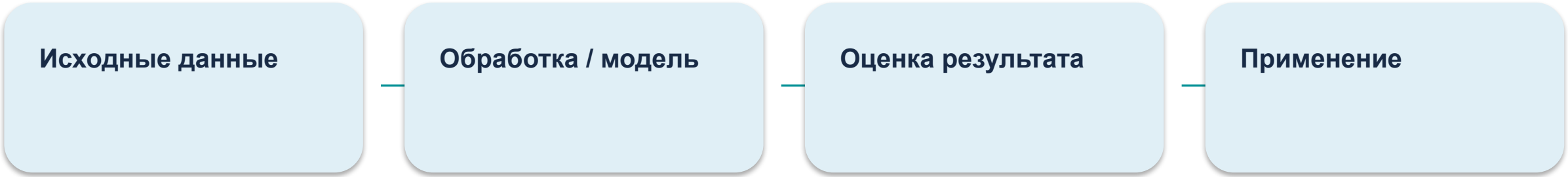
От граничных элементов к контурам	Если обнаружены отдельные точки границы, необходимо установить, какие из них принадлежат одной линии или одному объ

Преобразование Хафа

Преобразование Хафа используется для обнаружения объектов или контуров, которые можно описать параметрической моделью. Наиболее известный пример — поиск прямых линий на изображении. Каждая граничная точка голосует за множество возможных параметров линии. В параметрическом пространстве области с большим числом голосов соответствуют вероятным линиям исходного изображения.

Ключевая идея: Преобразование Хафа используется для обнаружения объектов или контуров, которые можно описать параметрической моделью. Наибол

Схема: Преобразование Хафа



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Выделение областей изображения

Альтернативой анализу контуров является выделение областей. В этом случае алгоритм объединяет соседние пиксели, обладающие сходными характеристиками, например яркостью, цветом или текстурой. Один из классических принципов — наращивание областей. Начальная точка или небольшой набор точек постепенно расширяется за счёт соседних пикселей, удовлетворяющих заданному критерию сходства.

Ключевая идея: Альтернативой анализу контуров является выделение областей. В этом случае алгоритм объединяет соседние пиксели, обладающие сх

Выделение областей изображения — продолжение

Такие методы применяются и к трёхмерным данным, например при сегментации облаков точек.

Основные задачи современного компьютерного зрения

Рисунок 4 – Классификация, детекция и сегментация

Классификация изображений

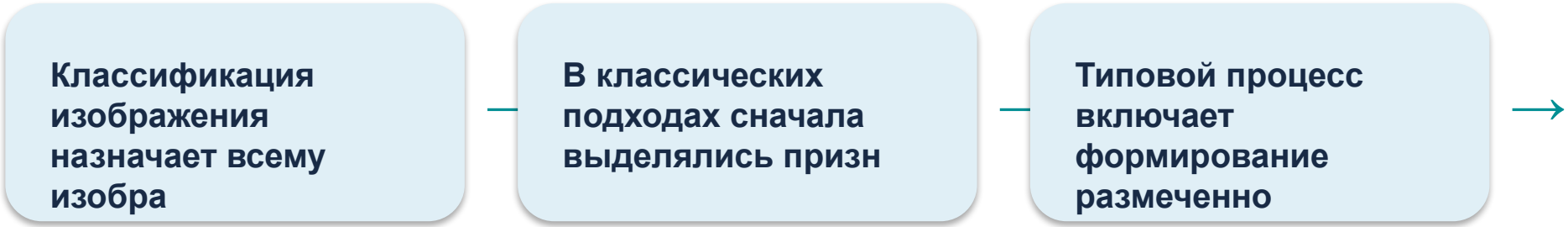
Классификация изображения назначает всему изображению одну или несколько категорий. Например, система может определить, изображён ли на фотографии автомобиль, человек или другое состояние объекта. В классических подходах сначала выделялись признаки изображения, после чего применялся классификатор.

Ключевая идея: Классификация изображения назначает всему изображению одну или несколько категорий. Например, система может определить, изобр

Классификация изображений — продолжение

Современные глубокие нейронные сети способны обучать визуальные признаки непосредственно из данных. Типовой процесс включает формирование размеченного набора изображений, разделение на train/validation/test, нормализацию, обучение модели и оценку качества на независимых данных.

Схема: Классификация изображений



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Сравнительная таблица: Классификация изображений

Классификация изображений	Классификация изображения назначает всему изображению одну или несколько категорий. Например, система может определ
Детекция объектов	В классических подходах сначала выделялись признаки изображения, после чего применялся классификатор. Современные г
Сегментация	Типовой процесс включает формирование размеченного набора изображений, разделение на train/validation/test, нормали

Детекция объектов

Детекция объектов является более сложной задачей, чем классификация. Необходимо определить не только класс объекта, но и его положение на изображении. Обычно результат представляется ограничивающей рамкой (bounding box) и меткой класса. Например, система видеонаблюдения может обнаружить на одном кадре несколько автомобилей и людей.

Ключевая идея: Детекция объектов является более сложной задачей, чем классификация. Необходимо определить не только класс объекта, но и его

Детекция объектов — продолжение

Для каждого объекта модель должна определить координаты рамки и категорию.

Сегментация

Сегментация выполняет классификацию на уровне пикселей. При семантической сегментации каждому пикселю присваивается класс. При instance segmentation дополнительно различаются отдельные экземпляры объектов одного класса.

Распознавание объектов

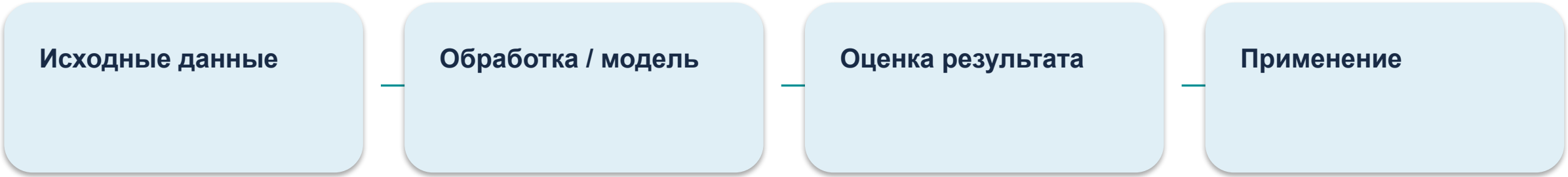
В классической компьютерной системе распознавание объектов может основываться на глобальных или локальных признаках. Учебник Бондарева и Аде рассматривает 2D-методы с глобальными признаками, локальными признаками и графами отношений, а также пример 3D-распознавания. Глобальные признаки описывают объект целиком.

Ключевая идея: В классической компьютерной системе распознавание объектов может основываться на глобальных или локальных признаках. Учебник

Распознавание объектов — продолжение

Локальные признаки описывают отдельные характерные части — углы, отверстия, участки контура. Локальный подход полезен, когда объект частично перекрыт: для распознавания может быть достаточно видимой части, содержащей характерные признаки.

Схема: Распознавание объектов



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Стереозрение и глубина

Для восстановления глубины можно использовать изображения, полученные двумя камерами с известным взаимным расположением. Одна и та же точка трёхмерной сцены отображается в двух изображениях с некоторым смещением — диспаритетом. На основе этого смещения и параметров камер можно оценить глубину. Для удалённых точек диспаритет становится малым.

Ключевая идея: Для восстановления глубины можно использовать изображения, полученные двумя камерами с известным взаимным расположением. Одна

Стереозрение и глубина — продолжение

Чтобы получить трёхмерную структуру сцены, требуется найти соответствующие точки на изображениях обеих камер.

Облака точек

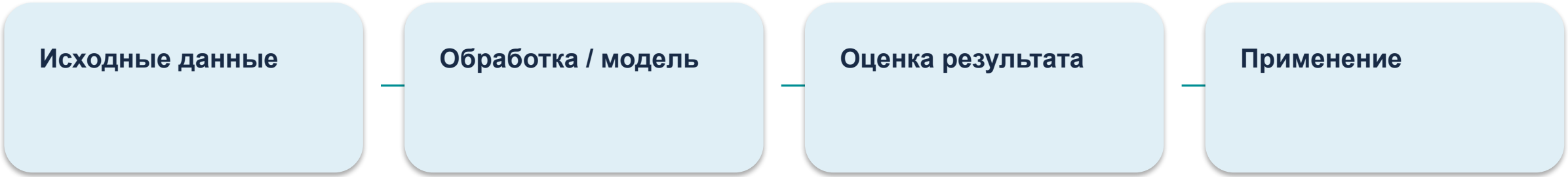
Датчики глубины могут формировать набор трёхмерных точек. Такой набор называют облаком точек. Каждая точка имеет координаты x, y, z, но первоначально связи между точками могут отсутствовать. Далее система должна объединить точки в поверхности или области.

Ключевая идея: Датчики глубины могут формировать набор трёхмерных точек. Такой набор называют облаком точек. Каждая точка имеет координаты x

Облака точек — продолжение

В приведённом в учебнике примере облака точек сегментируются на плоские участки, после чего строятся гипотезы объектов и оценивается возможность их захвата манипулятором робота.

Схема: Облака точек



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Свёрточные нейронные сети CNN

Свёрточные нейронные сети являются основным классом глубоких моделей для анализа изображений. Свёрточный слой применяет небольшие обучаемые фильтры к локальным участкам изображения и формирует карты признаков. Идея локального анализа хорошо соответствует структуре изображений: важны не только значения отдельных пикселей, но и локальные комбинации — границы, текстуры, формы.

Ключевая идея: Свёрточные нейронные сети являются основным классом глубоких моделей для анализа изображений. Свёрточный слой применяет небол

Свёрточные нейронные сети CNN — продолжение

В глубоких CNN ранние слои формируют простые признаки, а последующие — более сложные визуальные представления.

Свёрточные нейронные сети CNN — продолжение

Пример классификации изображений в TensorFlow/Keras import tensorflow as tf from tensorflow import keras from tensorflow.keras import layers model = keras.Sequential([layers.Input(shape=(128, 128, 3)), layers.Conv2D(32, 3, activation='relu'), layers.MaxPooling2D(), layers.Conv2D(64, 3, activation='relu'), layers.MaxPooling2D(), layers.Flatten(), layers.Dense(64, activation='relu'), layers.Dense(3, activation='softmax')]) model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy']) model.fit(train_ds, validation_data=val_ds, epochs=15) В примере сеть получает цветное изображение 128×128 пикселей и формирует вероятности трёх классов.

Свёрточные нейронные сети CNN — продолжение

Это учебная схема; реальная архитектура и параметры должны определяться экспериментально.

Подготовка изображений для обучения

Аугментация может включать допустимые повороты, отражения, сдвиги и изменения масштаба. Она должна сохранять смысл класса. Например, не каждое геометрическое преобразование допустимо в техническом или медицинском изображении.

Сравнительная таблица: Подготовка изображений для обучения

Подготовка изображений для обучения	Аугментация может включать допустимые повороты, отражения, сдвиги и изменения масштаба. Она должна сохранять смысл

Transfer Learning

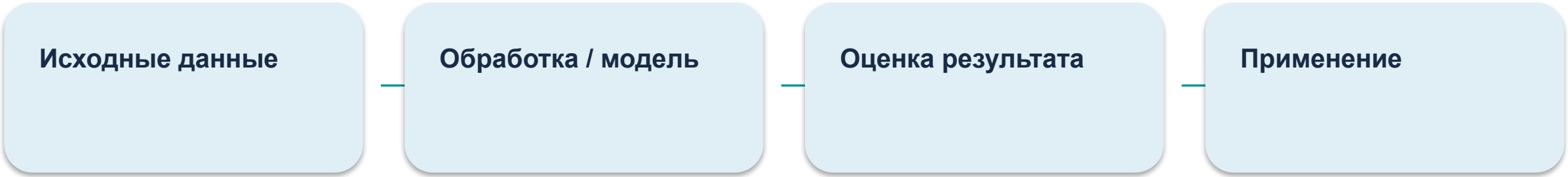
Для многих задач не требуется обучать глубокую сеть с нуля. Можно использовать модель, предварительно обученную на большом наборе изображений, и адаптировать её к новой задаче. Этот подход называется transfer learning. Преимущества — уменьшение требований к объёму размеченных данных и сокращение вычислительных затрат.

Ключевая идея: Для многих задач не требуется обучать глубокую сеть с нуля. Можно использовать модель, предварительно обученную на большом на

Transfer Learning — продолжение

Однако необходимо проверить, насколько исходная область данных близка к новой задаче.

Схема: Transfer Learning



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Оценка классификации изображений

Для классификации изображений применяются уже знакомые метрики accuracy, precision, recall, F1-score и confusion matrix. При дисбалансе классов общая accuracy может быть недостаточной. В многоклассовой задаче важно анализировать ошибки отдельно по каждому классу.

Ключевая идея: Для классификации изображений применяются уже знакомые метрики accuracy, precision, recall, F1-score и confusion matrix. При

Оценка классификации изображений — продолжение

Например, если система хорошо распознаёт нормальные изображения, но пропускает редкий опасный дефект, высокая общая точность не означает пригодность модели.

Оценка детекции объектов

Для детекции необходимо оценивать не только правильность класса, но и точность локализации объекта. Часто используется мера IoU — Intersection over Union, то есть отношение площади пересечения, предсказанной и эталонной рамок к площади их объединения. На основе порога IoU определяют корректность обнаружения, после чего рассчитывают precision и recall детектора.

Ключевая идея: Для детекции необходимо оценивать не только правильность класса, но и точность локализации объекта. Часто используется мера I

Оценка детекции объектов — продолжение

Для сравнения детекторов часто используется средняя точность по классам, однако конкретная схема оценки зависит от принятого протокола.

Распознавание лиц

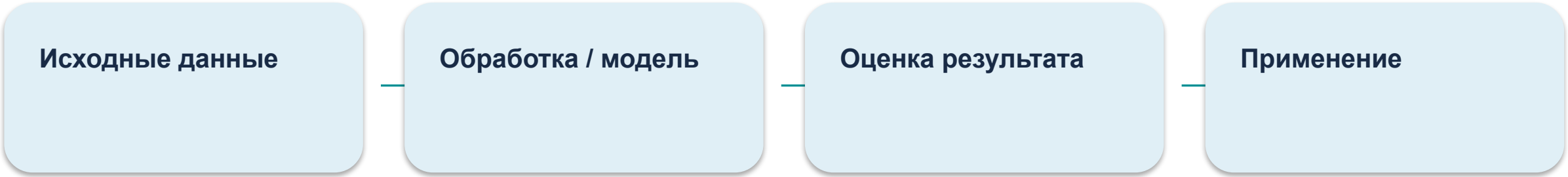
Распознавание лиц включает несколько различных подзадач: обнаружение лица на изображении, выравнивание, формирование числового представления лица и сравнение с известными образцами. Распознавание лиц связано с биометрическими и персональными данными.

Ключевая идея: Распознавание лиц включает несколько различных подзадач: обнаружение лица на изображении, выравнивание, формирование числовог

Распознавание лиц — продолжение

Поэтому при разработке таких систем необходимо учитывать требования конфиденциальности, права доступа, защиту шаблонов и риск ошибочной идентификации.

Схема: Распознавание лиц



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Сравнительная таблица: Распознавание лиц

Распознавание лиц	Распознавание лиц включает несколько различных подзадач: обнаружение лица на изображении, выравнивание, формировани
Видеопоток и отслеживание объектов	Распознавание лиц связано с биометрическими и персональными данными. Поэтому при разработке таких систем необходимо

Видеопоток и отслеживание объектов

Видео добавляет временное измерение. Система должна не только обнаруживать объект в отдельном кадре, но и связывать его положения между кадрами. Эта задача называется tracking. Пример: в системе видеонаблюдения один и тот же автомобиль последовательно появляется на множестве кадров. Детектор находит объект, а алгоритм отслеживания поддерживает его идентификатор во времени.

Ключевая идея: Видео добавляет временное измерение. Система должна не только обнаруживать объект в отдельном кадре, но и связывать его полож

Компьютерное зрение в информационной безопасности

При использовании видеонаблюдения и биометрии необходимо учитывать требования законодательства и политики организации. Техническая возможность распознавания объекта не означает автоматической допустимости его применения.

Компьютерное зрение в промышленности и робототехнике

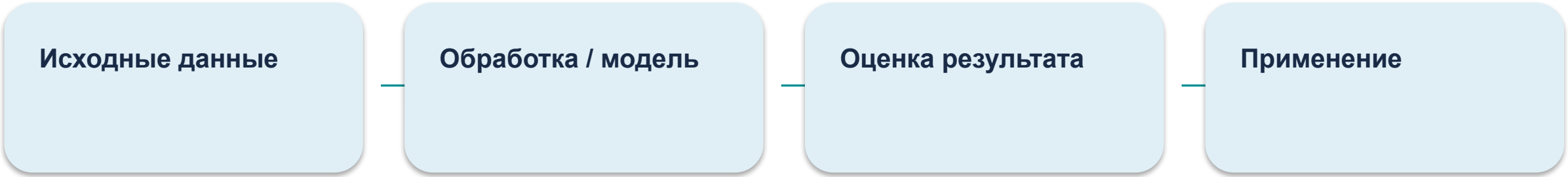
В основном компьютерное зрение связывает с системами технического зрения роботов. В робототехнике визуальная система не только распознаёт объект, но и оценивает его положение, ориентацию и возможность взаимодействия.

Ключевая идея: В основном компьютерное зрение связывает с системами технического зрения роботов. В робототехнике визуальная система не тольк

Компьютерное зрение в промышленности и робототехнике — продолжение

Например, робот-манипулятор получает данные от нескольких датчиков глубины, объединяет трёхмерные точки, сегментирует поверхности, формирует гипотезы объектов и выбирает подходящую область для захвата.

Схема: Компьютерное зрение в промышленности и робототехнике



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Типичные ошибки при разработке системы компьютерного зрения

Оценивать модель только на обучающей выборке. Смешивать почти одинаковые изображения одного объекта между train и test. Игнорировать изменение освещения, ракурса и качества камеры. Применять аугментацию, которая меняет смысл класса. Оценивать многоклассовую модель только общей accuracy. Использовать изображения с персональными или биометрическими данными без необходимых прав и мер защиты.

Ключевая идея: Оценивать модель только на обучающей выборке.

Типичные ошибки при разработке системы компьютерного зрения — продолжение

Считать детекцию и классификацию одной и той же задачей. Предполагать, что модель, обученная на одном типе камер или объектов, автоматически сохранит качество в другой среде. Пример полного конвейера: классификация изображений Пусть требуется классифицировать изображения на три категории. Сначала формируется размеченный набор данных и проверяется корректность меток.

Типичные ошибки при разработке системы компьютерного зрения — продолжение

Затем изображения приводятся к одинаковому размеру, значения пикселей нормализуются, а данные разделяются на train, validation и test. В качестве базовой модели можно использовать небольшую CNN. После обучения анализируются loss и accuracy на train и validation, затем выполняется независимая проверка на test.

Типичные ошибки при разработке системы компьютерного зрения — продолжение

Если модель переобучается, применяются аугментация, регуляризация, Dropout, ранняя остановка или transfer learning. В заключение строится confusion matrix и анализируются типичные ошибочные изображения. Пример полного конвейера: детекция объектов В задаче детекции кроме класса требуется разметка местоположения объектов. Для каждого обучающего изображения задаются рамки и категории.

Типичные ошибки при разработке системы компьютерного зрения — продолжение

После обучения детектор выдаёт набор рамок, вероятностей и классов. При проверке необходимо анализировать ложные обнаружения, пропущенные объекты и точность рамок. В практической системе также оцениваются скорость обработки кадров и устойчивость к изменению условий съёмки.

Основные выводы

Компьютерное зрение преобразует цифровые изображения и видео в содержательную информацию о сцене. Изображение является числовым массивом, а цветное изображение — многоканальным тензором. Классические этапы компьютерного зрения включают предобработку, выделение границ, контуров и областей. Преобразование Хафа относится к классическим методам поиска параметрических геометрических структур.

Ключевая идея: Компьютерное зрение преобразует цифровые изображения и видео в содержательную информацию о сцене.

Основные выводы — продолжение

Основные современные задачи — классификация, детекция, сегментация, распознавание и отслеживание объектов. 3D-зрение использует стереоданные, изображения глубины и облака точек. CNN позволяют автоматически обучать иерархические визуальные признаки. При оценке CV-моделей необходимо учитывать качество разметки, независимость test и особенности конкретной задачи.

Основные выводы — продолжение

Применение биометрических и визуальных данных требует мер безопасности, конфиденциальности и правовой обоснованности.

Схема: Основные выводы

Компьютерное зрение преобразует цифровые изображ

Изображение является числовым массивом, а цветно

Классические этапы компьютерного зрения включают

Преобразование Хафа относится к классическим мет

Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Как изображение представляется в памяти компьютера?

Чем чёрно-белое изображение отличается от цветного с точки зрения структуры данных? Почему восстановление трёхмерной сцены по одному изображению является сложной задачей? Какие факторы затрудняют анализ изображений? Для чего выполняется предобработка изображения? Что такое граница изображения и какова цель выделения контуров?

Ключевая идея: Чем чёрно-белое изображение отличается от цветного с точки зрения структуры данных?

Схема: Как изображение представляется в памяти компьютера?

Чем чёрно-белое изображение отличается от цветно

Почему восстановление трёхмерной сцены по одному

Какие факторы затрудняют анализ изображений? Для

Что такое граница изображения и какова цель выде

Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Чем выделение областей отличается от анализа границ?

Чем классификация изображения отличается от детекции объектов?

Что понимается под распознаванием объекта?

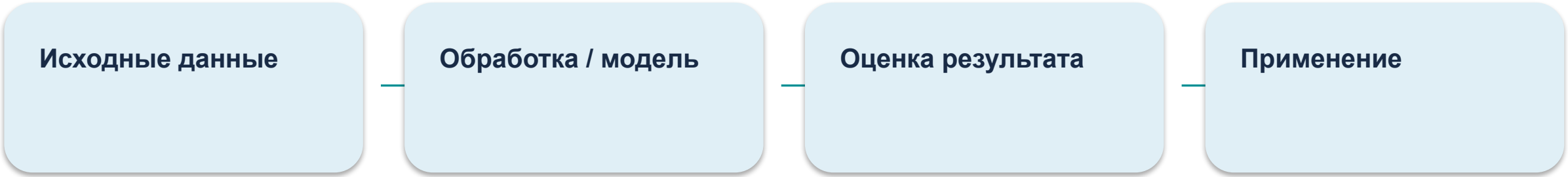
Какие применения компьютерного зрения возможны в информационной безопасности?

Чем классические методы CV отличаются от глубокого обучения?

Какие ошибки при разделении train/test могут привести к завышенной оценке качества? Почему условия съёмки важны для обобщения модели? Какие риски необходимо учитывать при работе с биометрическими изображениями?

Ключевая идея: Какие ошибки при разделении train/test могут привести к завышенной оценке качества? Почему условия съёмки важны для обобщения

Схема: Чем классические методы CV отличаются от глубокого обучения?



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Список источников

1. Бондарев В.Н., Аде Ф.Г. Искусственный интеллект: учебное пособие. Севастополь, 2002. Глава 10 «Компьютерное зрение»: основные понятия, системы компьютерного зрения, выделение границ и областей, преобразование Хафа, 2D- и 3D-методы распознавания. 2. Беликова Е.К., Гудилина И.С., Круглова М.А., Саратовская Л.Б., Ямпольская А.В.

Ключевая идея: 1. Бондарев В.Н., Аде Ф.Г. Искусственный интеллект: учебное пособие. Севастополь, 2002. Глава 10 «Компьютерное зрение»: основ

Список источников — продолжение

English Reader in Artificial Intelligence = Хрестоматия по искусственному интеллекту на английском языке. Учебное пособие. Москва, 2024. – 171 с. 3. Бессмертный И.А., Авдюшина А.Е., Кугаевских А.В., Королева Ю.А., Рущенко Н.Г. Системы искусственного интеллекта. Методическое пособие. Санкт-Петербург: Университет ИТМО, 2024. – 281 с. 4. Андреянов Н.В., Евдокимова Т.С., Павлов А.Д., и др.

Список источников — продолжение

Методы искусственного интеллекта. Учебное пособие. Казань: КНИТУ-КАИ, 2024. – 392 с.

График: иллюстрация зависимости



Иллюстративные значения (не результаты эксперимента).

Итоги и вопросы для обсуждения

1. Объясните понятие «Понятие компьютерного зрения» и приведите пример.

2. Объясните понятие «Изображение как цифровые данные» и приведите пример.

3. Объясните понятие «Почему понимание изображения является сложной задачей» и приведите пример.

4. Объясните понятие «Предобработка изображений» и приведите пример.

5. Объясните понятие «Выделение границ изображения» и приведите пример.

6. Объясните понятие «От граничных элементов к контурам» и приведите пример.

7. Объясните понятие «Преобразование Хафа» и приведите пример.