

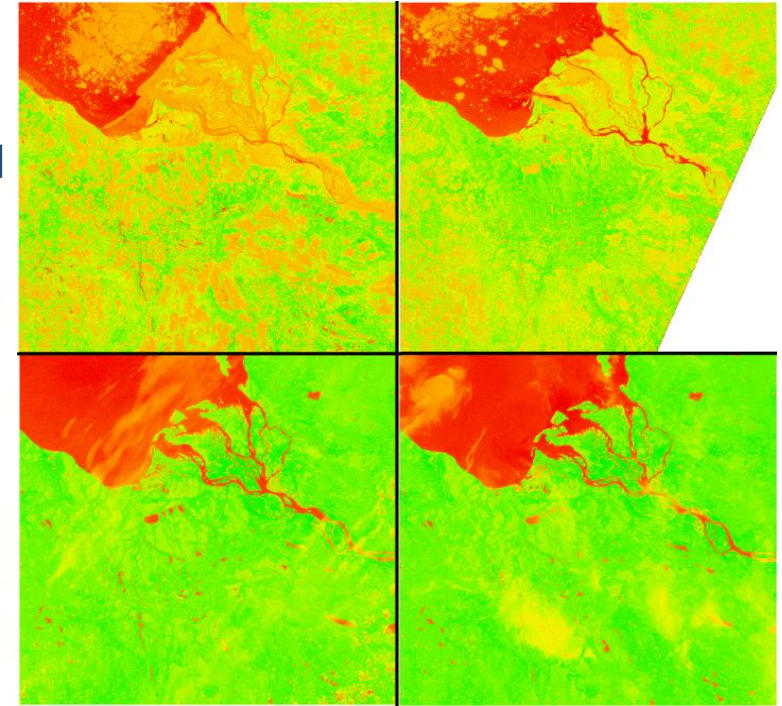
Нейронные сети для компьютерного зрения

Докладчик: Ян Пустынный,
инженер отдела картографии и геоинформационных систем
ГАУ АО «Управление ИКТ АО», г. Архангельск

16 апреля 2019

ГАУ АО «Управление ИКТ АО»

- 1) Работа со спутниковыми и БПЛА снимками
- 2) Геометрические и географические преобразования
- 3) Дешифрирование и классификация снимков



Содержание

- 1) Понятие и сферы применения компьютерного зрения
- 2) Реализация компьютерного зрения
- 3) Нейронные сети и компьютерное зрение
- 4) Применение нейронных сетей для нетривиальных задач компьютерного зрения

Часть 1

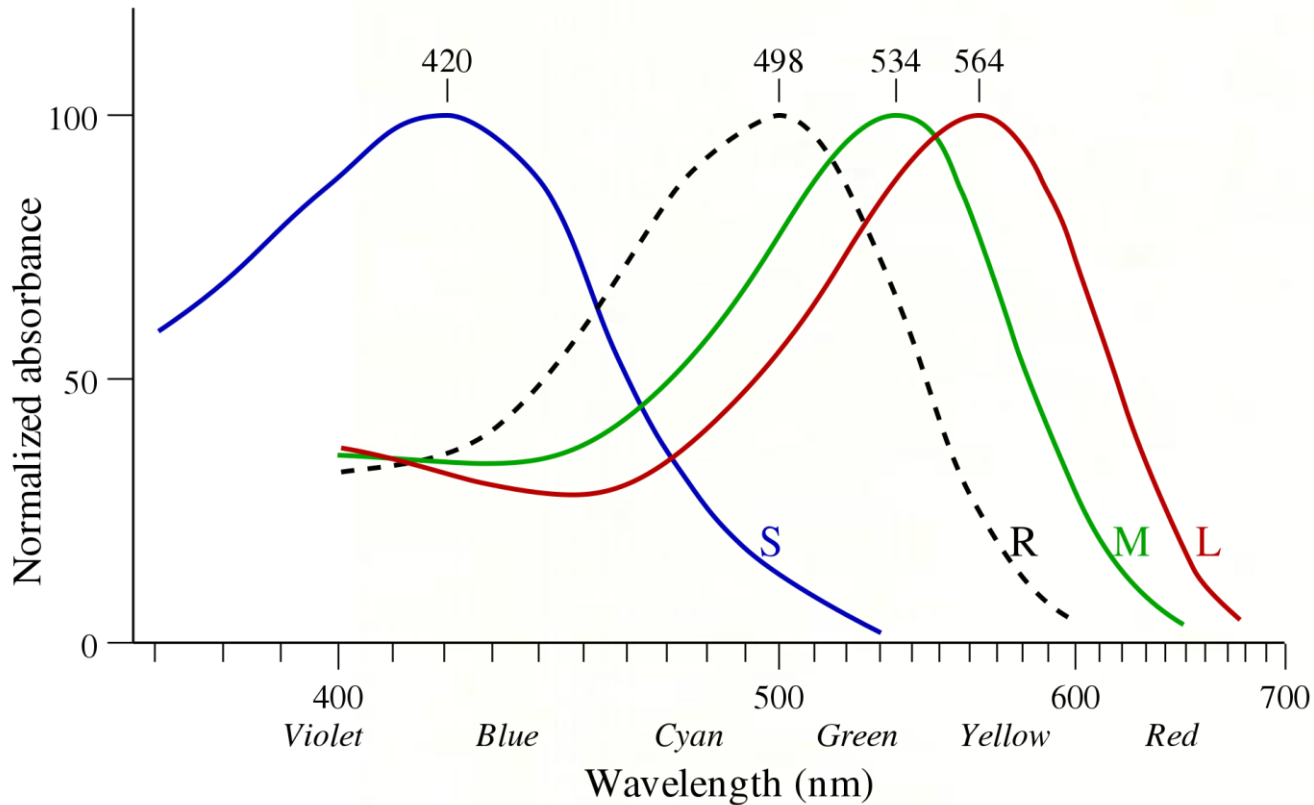
Понятие и сферы применения компьютерного зрения

Что такое компьютерное зрение?

Компьютерное зрение — теория и технологии создания программных и аппаратных средств, которые способны извлекать полезную информацию из изображений.

Данное понятие связано с двумя другими: обработка изображений и машинное зрение. Машинное зрение сосредотачивается на применении, в основном промышленном, например, автономные роботы и системы визуальной проверки и измерений. Обработка изображений или анализ изображений, в основном сосредоточены на работе с двухмерными изображениями, то есть как преобразовать одно изображение в другое.

Зрение



Человеческий глаз воспринимает длины волн λ от 400 нм (фиолетовый) до 700 нм (красный).

Свет от объекта попадает на фоторецепторы, состоящие из колбочек и палочек. Затем, полученные сигналы переходят по зрительному нерву в мозг и обрабатываются

История развития компьютерного зрения

Персептрон был впервые смоделирован на универсальной ЭВМ IBM-740 в 1958 году. Аппаратный вариант персептрона -- Mark I был изготовлен в 1960 году и предназначался для распознавания зрительных образов. Его рецепторное поле состояло из 400 точек (матрица 20x20 элементов), и он был способен решать ряд несложных задач, в частности распознавать печатные буквы.

Исследования в области синтеза систем компьютерного зрения бурно развивались на протяжении 60-х годов по мере того, как расширялось использование вычислительных машин и становилась очевидной потребность в более быстрой и эффективной связи человека с ЭВМ. К началу 60-х годов задачи компьютерного зрения в основном охватывали область космических исследований, требовавших обработки большого количества цифровой информации. Запущенные в то время космические летательные аппараты передавали на Землю тысячи телевизионных изображений Земли, обратной стороны поверхности Луны. Полученные тогда цифровые изображения требовали удаления различного рода искажений, в частности оптических. Кроме того, эти изображения нуждались в обработке и анализе с целью решения различного рода навигационных задач -- таких как определение места площадок, пригодных для посадки спускаемых космических аппаратов. Позднее, в 70-е годы, наряду с ростом разрешающей способности изображений в видимом спектре эта информация стала дополняться полученной в тепловизионном и других спектральных диапазонах частот, что позволяло проводить более глубокий анализ данных. Начали развиваться различные подходы к распознаванию объектов на изображении, например структурные, признаковые и текстурные.



Применения компьютерного зрения

Видеонаблюдение

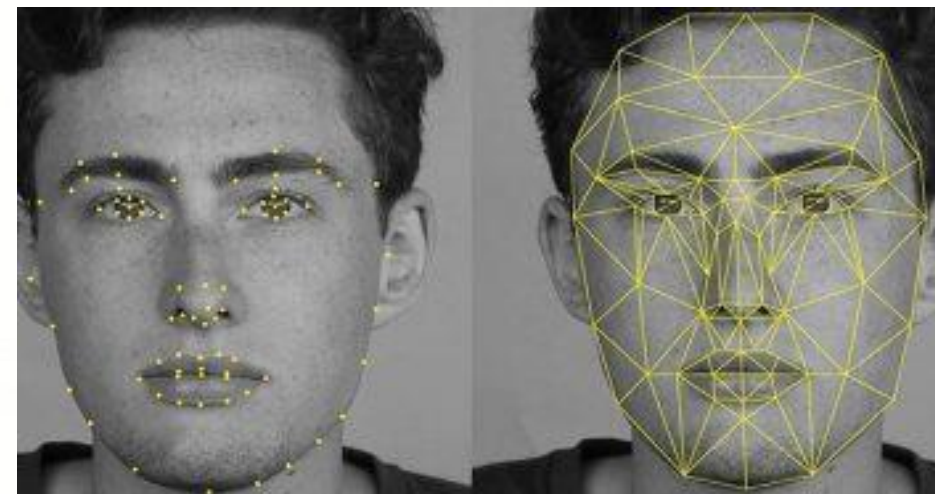
- Детектирование движения и запись видео. (видеорегистраторы)
- Охрана определённых зон. Охрана предметов. (музеи)
- Детектирование оставленных предметов/машин. (вокзалы)
- Анализ траекторий движения. (рекомендации маркетингу).
- Сопровождение объекта используя несколько камер.
- Детектирование поскользнувшихся и упавших людей.
- Обнаружение воровства на кассе. (POS monitor)



Применения компьютерного зрения

Биометрия

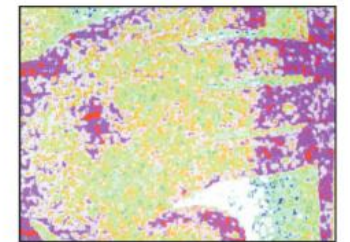
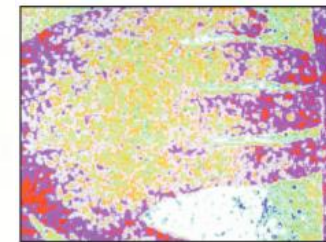
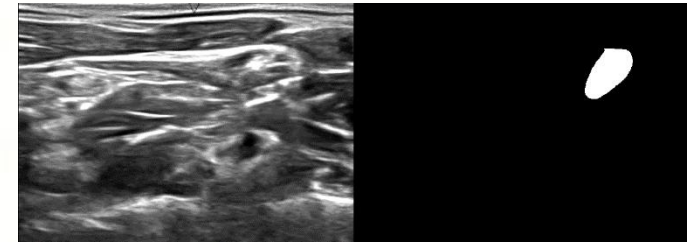
- Распознавание отпечатков пальцев
- Распознавание лица (2D, 3D)
- Идентификация по изображению роговицы



Применения компьютерного зрения

Медицинские технологии

- Поиск аномалий на медицинских изображениях
- Помощь в проведении анализов
- Классификация рентгеновских снимков
- Детектирование эпилептических приступов
- Автоматическая диагностика болезней глаз



Применения компьютерного зрения

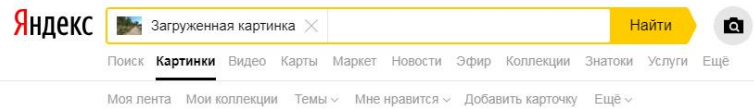
Улучшение качества изображений

- Удаление шумов.
- Улучшение цветопередачи.
- Улучшение контраста.
- Просветление тёмных фрагментов.
- Удаление нежелательных объектов.
- Удаление эффекта “красных глаз”



Применения компьютерного зрения

Поиск похожих изображений



Исходная картинка
4160x3120

Таких же картинок не найдено

Кажется, на картинке

- уличный пейзаж
- участок
- земельный участок
- фото верхний рогачик херсонская обл
- земельные участки

Похожие картинки



Все Картинки Карты Покупки Ещё Настройки Инструменты

Результатов: примерно 2 (0,95 сек.)



Размер изображения:
4160 x 3120

Изображения других размеров не найдены.

Запрос, который может быть связан с изображением: **dirt road**

Jason Aldean - Dirt Road Anthem - YouTube

<https://www.youtube.com/watch?v=Lb9q1ScC4cg> ▾ Перевести эту страницу
17 июн. 2011 г. - Check out the official video for "Dirt Road Anthem" from My Kinda Party, available here:
<http://smarturl.it/aldean-mykindaparty> Subscribe to ...

Определение DIRT ROAD в кембриджском словаре английского ...

<https://dictionary.cambridge.org/ru/словарь/английский/dirt-road> ▾
dirt road: Определение dirt road: a road in the countryside that is made from soil and not paved: .
Узнать больше.

Похожие изображения



Пожаловаться на картинку

Грунтовая дорога

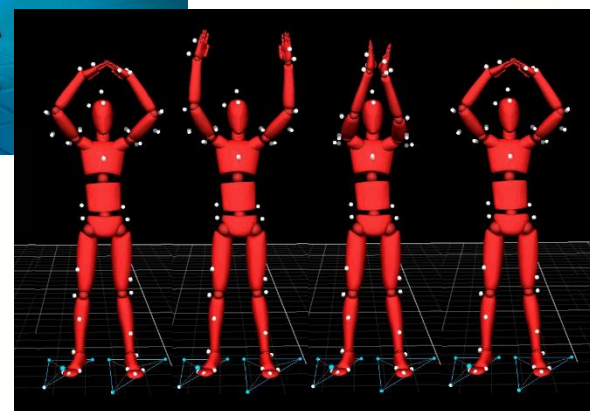
Грунтовая дорога — автомобильная или тракторная дорога, построенная из природного грунта или грунта с добавками других материалов. Согласно российской строительной терминологии это дороги, имеющие дорожные одежды низшего типа по классификации СНиПТ 2.05.02-08 или с отсутствующей одеждой. Википедия

Оставьте отзыв

Применения компьютерного зрения

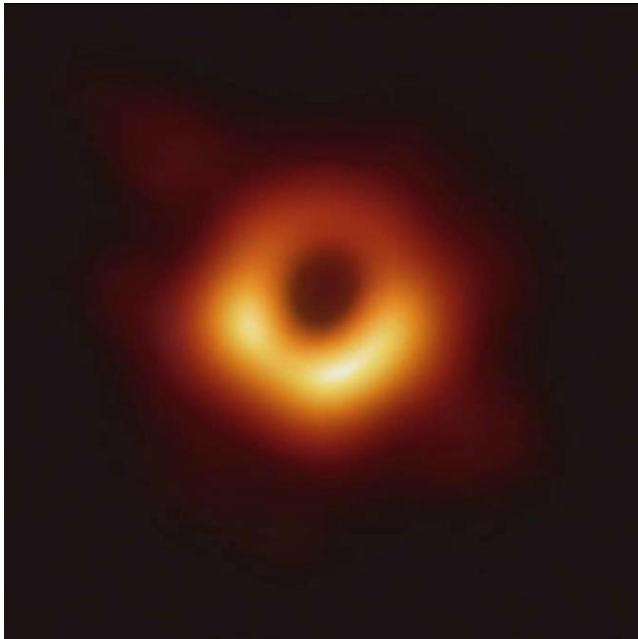
Технологии виртуальной и дополненной реальности

- VR
- AR
- Xbox Kinect



Применения компьютерного зрения

Снимок черной дыры в центре галактики M87



Для получения снимка был разработан специальный алгоритм компьютерного зрения с целью сопоставления данных с разных телескопов в радиодиапазоне и их обработки

Часть 2

Реализация систем компьютерного зрения

Реализация систем компьютерного зрения

Получение изображений:

цифровые изображения получают от одного или нескольких датчиков изображения, которые помимо различных типов светочувствительных камер включают датчики расстояния, радары, ультразвуковые камеры и т. д. В зависимости от типа датчика, получающиеся данные могут быть обычным 2D изображением, 3D изображением или последовательностью изображений. Значения пикселей обычно соответствуют интенсивности света в одной или нескольких спектральных полосах (цветные или изображения в оттенках серого), но могут быть связаны с различными физическими измерениями, такими как глубина, поглощение или отражение звуковых или электромагнитных волн, или ядерным магнитным резонансом.

Реализация систем компьютерного зрения

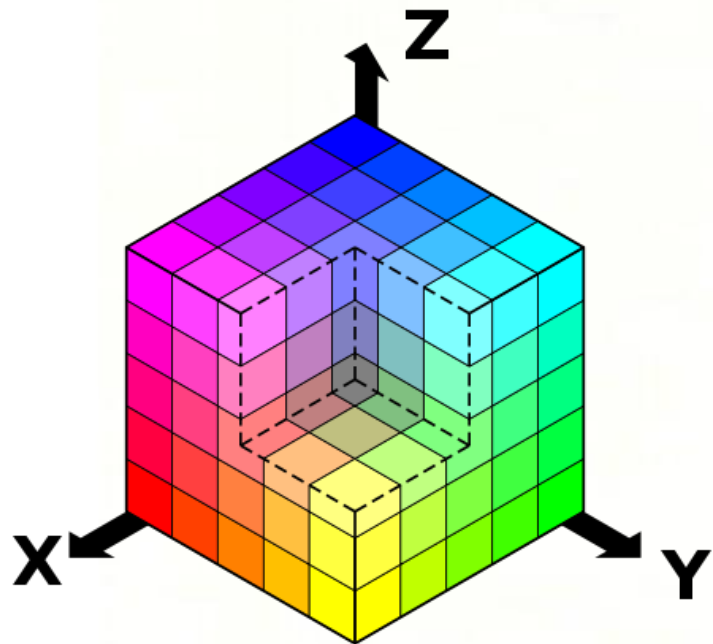
Предварительная обработка:

перед тем, как методы компьютерного зрения могут быть применены к данным с тем, чтобы извлечь определённую долю информации, необходимо обработать данные, с тем чтобы они удовлетворяли некоторым условиям, в зависимости от используемого метода. Примерами являются:

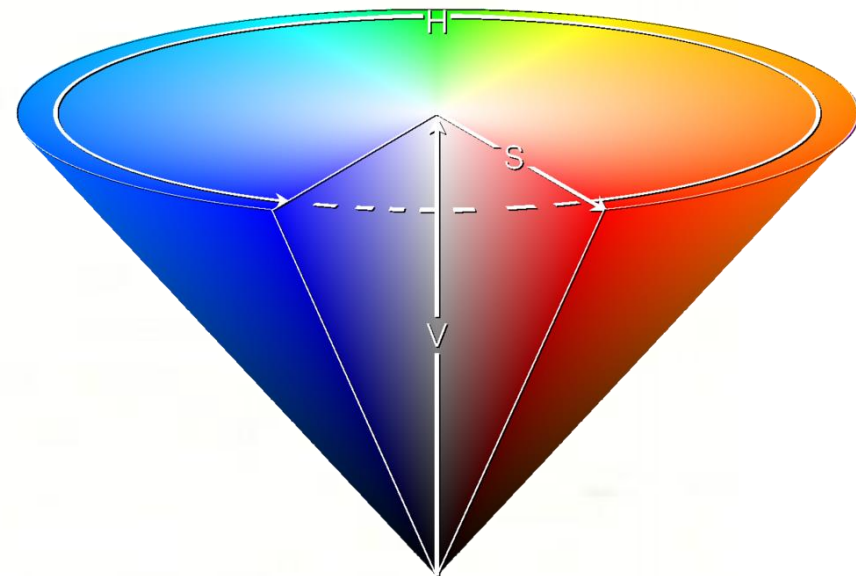
1. Уменьшение размера изображения
2. Удаление шума с тем, чтобы удалить искажения, вносимые датчиком
3. Улучшение контрастности, для того, чтобы нужная информация могла быть обнаружена
4. Масштабирование для лучшего различения структур на изображении

Цветовые модели

RGB



HSV



Цветовые модели



Реализация систем компьютерного зрения

Выделение деталей:

детали изображения различного уровня сложности выделяются из видеоданных. Типичными примерами таких деталей являются:

Линии, границы и кромки

Локализованные точки интереса, такие как углы, капли или точки: более сложные детали могут относиться к структуре, форме или движению.



Реализация систем компьютерного зрения

Выделение деталей:

Сначала на картинке выявляют потенциально значимые места — то есть предполагаемые объекты или их границы. Это можно сделать несколькими способами. Рассмотрим, к примеру, алгоритм Difference of Gaussians (DoG, разность гауссиан).

Он подразумевает, что исходную картинку несколько раз подвергают размытию по Гауссу, каждый раз используя разный радиус размытия. Затем результаты сравнивают друг с другом.

Этот способ позволяет выявить на изображении наиболее контрастные фрагменты — к примеру, яркие пятна или изломы линий.

После того как значимые места найдены, их описывают в числах. Запись фрагмента картинки в числовом виде называется дескриптором. С помощью дескрипторов можно быстро, полно и точно сравнить фрагменты изображения, не используя сами фрагменты. Существуют разные алгоритмы получения дескрипторов — например, SIFT, SURF, HOG и многие другие.

Реализация систем компьютерного зрения

Детектирование/Сегментация:

на определённом этапе обработки принимается решение о том, какие точки или участки изображения являются важными для дальнейшей обработки.

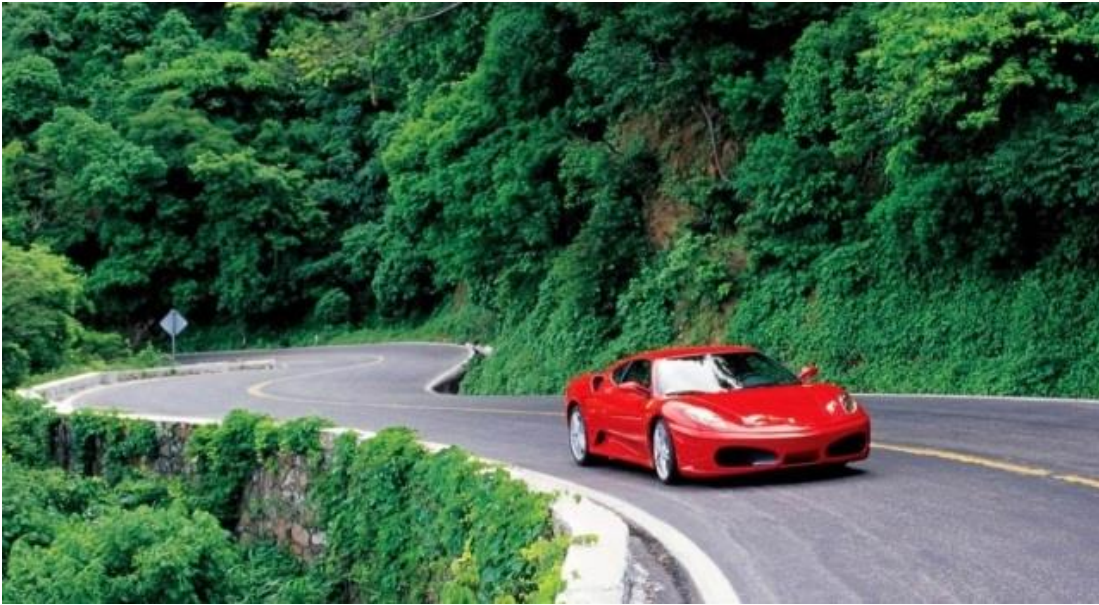
Примерами являются:

Выделение определённого набора интересующих точек

Сегментация одного или нескольких участков изображения, которые содержат характерный объект

Реализация систем компьютерного зрения

Детектирование/Сегментация:



Реализация систем компьютерного зрения

Высокоуровневая обработка:

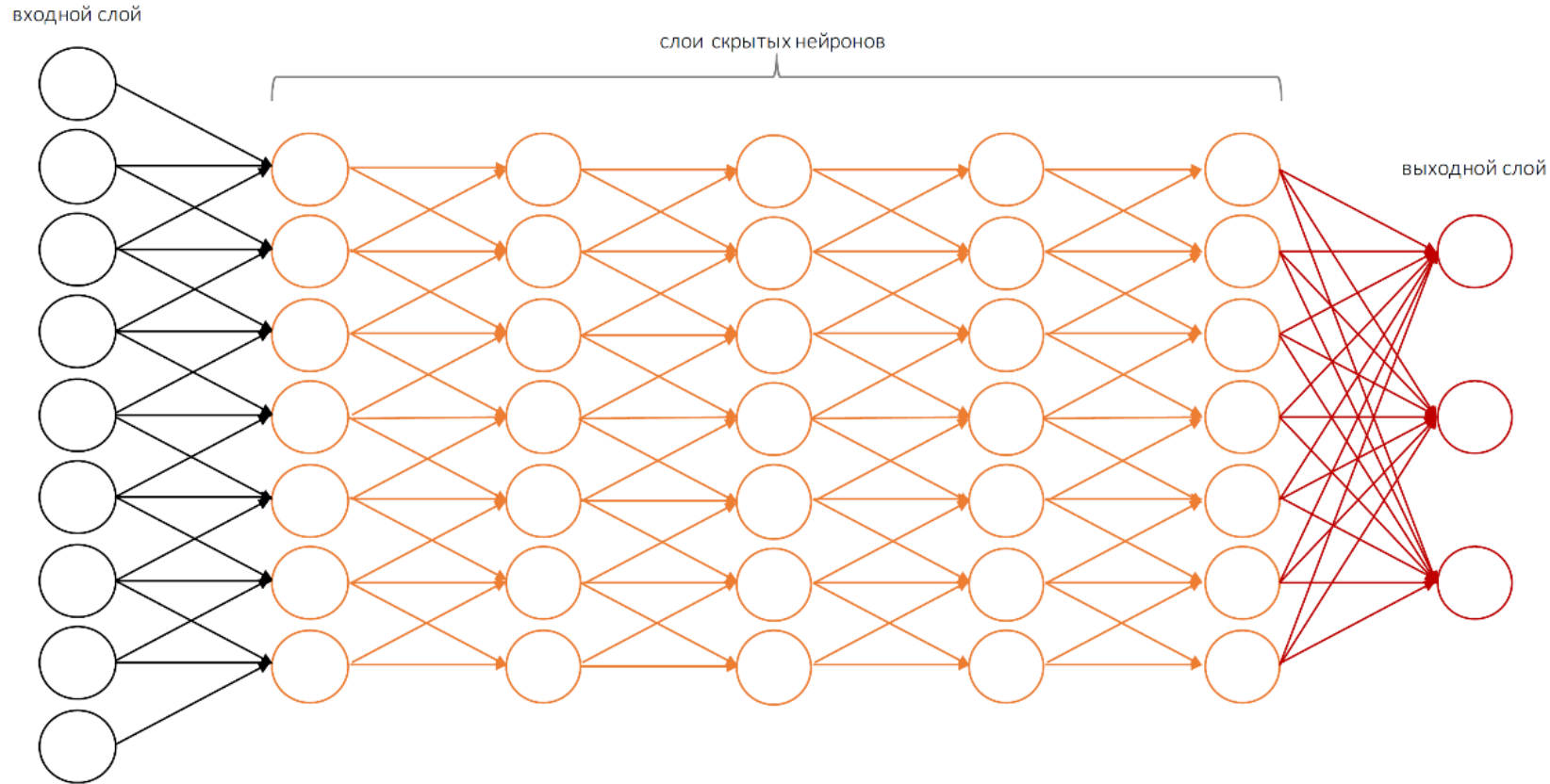
на этом шаге входные данные обычно представляют небольшой набор данных, например набор точек или участок изображения, в котором предположительно находится определённый объект. Примерами являются:

- Проверка того, что данные удовлетворяют условиям, зависящим от метода и применения
- Оценка характерных параметров, таких как положение или размер объекта
 - Классификация обнаруженного объекта по различным категориям

Часть 3

Нейронные сети и компьютерное зрение

Перцептрон

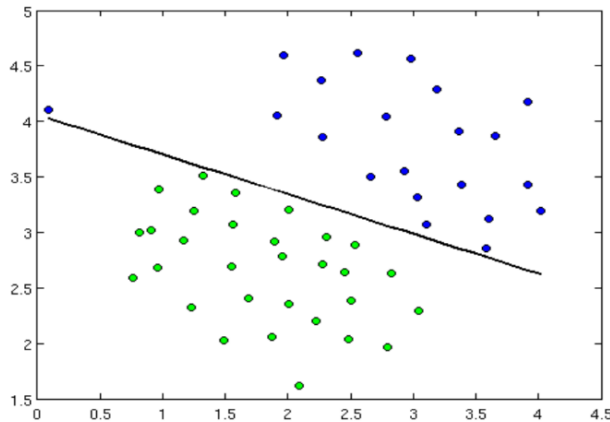


Для выполнения сложных операций, нужно большее количество скрытых слоев

Перцептрон

Перцептрон представляет собой линейный классификатор, то есть алгоритм, который классифицирует объект путем разделения двух категорий прямой. Объектом обычно является вектор-функция x , взятая с весом w и смещенная на b : $y = w * x + b$.

На выходе перцептрон выдает результат y , основанный на нескольких вещественных входных объектах путем формирования линейной комбинации с использованием весовых коэффициентов (иногда с последующим пропуском результата через нелинейную функцию активации)



$$y = \varphi\left(\sum_{i=1}^n w_i x_i + b\right) = \varphi(\mathbf{w}^T \mathbf{x} + b)$$

где w – вектор весовых коэффициентов,
 x – вектор входных объектов,
 b – смещение,
 φ – функция нелинейной активации

Перцептрон в Keras

```
import numpy as np

from keras.datasets import mnist
from keras.models import Sequential
from keras.layers.core import Dense
from keras.utils import np_utils

from keras.models import Sequential
from keras.layers.core import Dense

nb_classes = 10

# Обучающая и боевая выборки перетасовывается и разделяются
(X_train, y_train), (X_test, y_test) = mnist.load_data()

X_train = X_train.reshape(60000, 784)
X_test = X_test.reshape(10000, 784)
X_train = X_train.astype("float32")
X_test = X_test.astype("float32")
X_train /= 255
X_test /= 255
print(X_train.shape[0], 'train samples')
print(X_test.shape[0], 'test samples')

# конвертируем векторы классов в бинарные матрицы
Y_train = np_utils.to_categorical(y_train, nb_classes)
Y_test = np_utils.to_categorical(y_test, nb_classes)

model.compile(optimizer='adam', loss='categorical_crossentropy',
              metrics=['accuracy'])

model.fit(X_train, Y_train, batch_size=128, nb_epoch=5, verbose=2,
          validation_data=(X_test, Y_test))
```


Сверточная нейронная сеть

Предложена Яном
Лекуном в 1988 году.



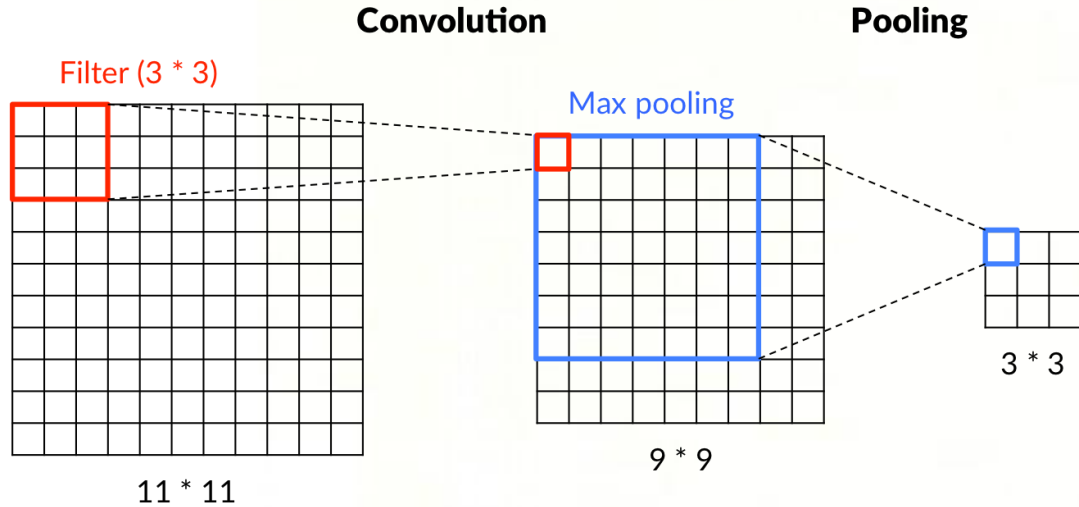
Состоит из различных типов слоев, идущих последовательно:

- Слой свертки
- Слой активации
- Слой субдискретизации (пулинга)

Сверточная нейронная сеть

Слой свёртки

Особенностью свёрточного слоя является сравнительно небольшое количество параметров, устанавливаемое при обучении.



Для входного изображения 100×100 пикселей и фильтра с ядром 3×3 пикселя, количество параметров будет равно 167.

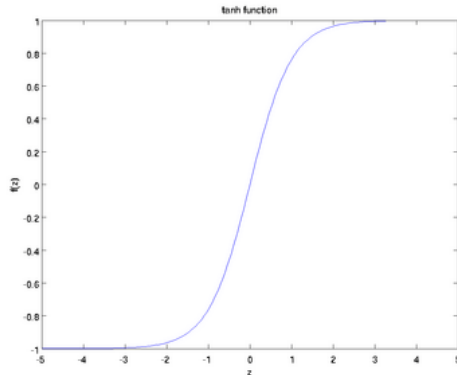
Сверточная нейронная сеть

Слой активации

Скалярный результат каждой свёртки попадает на функцию активации, которая представляет собой некую нелинейную функцию.

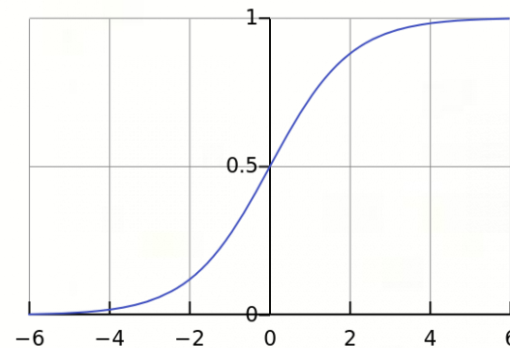
Функции активации

Тангенциальная



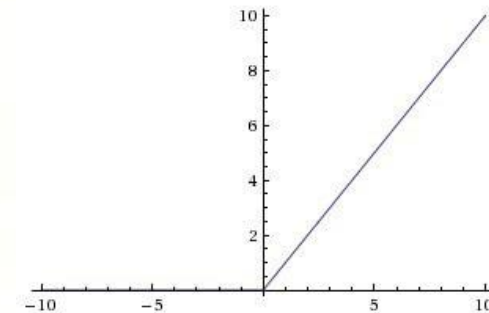
$$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$$

Сигмоидальная



$$A = \frac{1}{1+e^{-x}}$$

ReLU

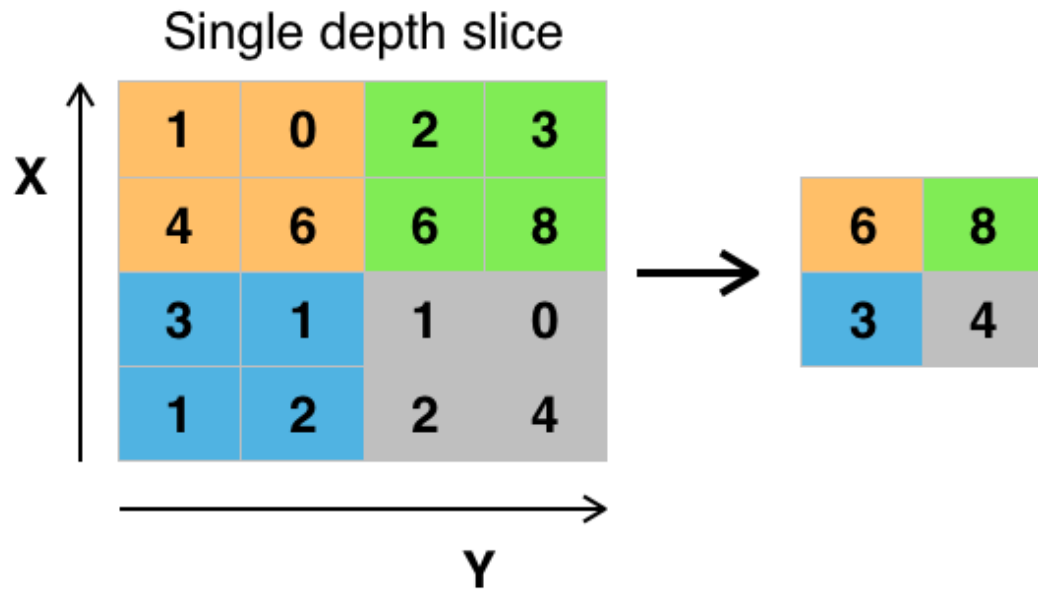


$$f(s) = \max(0, s)$$

Сверточная нейронная сеть

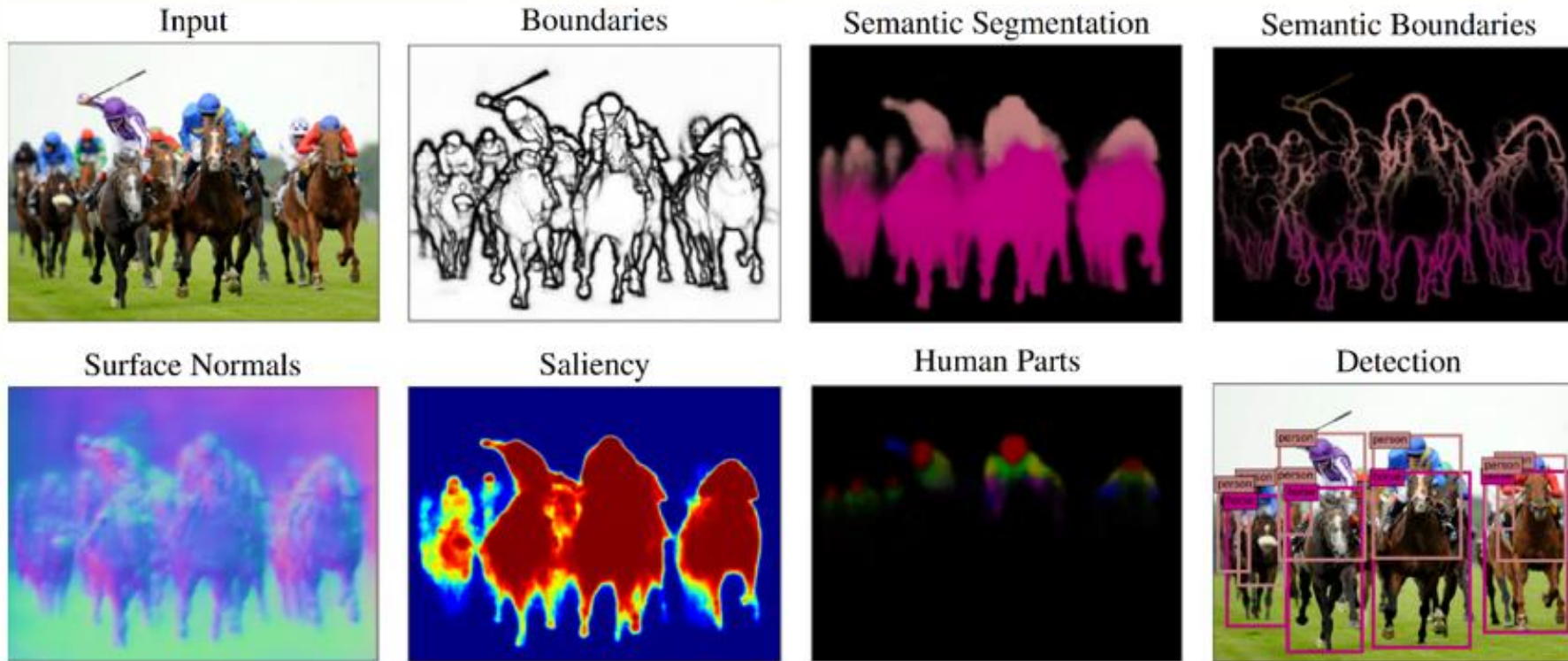
Слой субдискретизации (пулинга)

представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера 2×2) уплотняется до одного пикселя, проходя нелинейное преобразование. Наиболее употребительна при этом функция максимума.



Сверточная нейронная сеть (CNN)

Классические задачи



Сверточная нейронная сеть для распознавания рукописных цифр MNIST

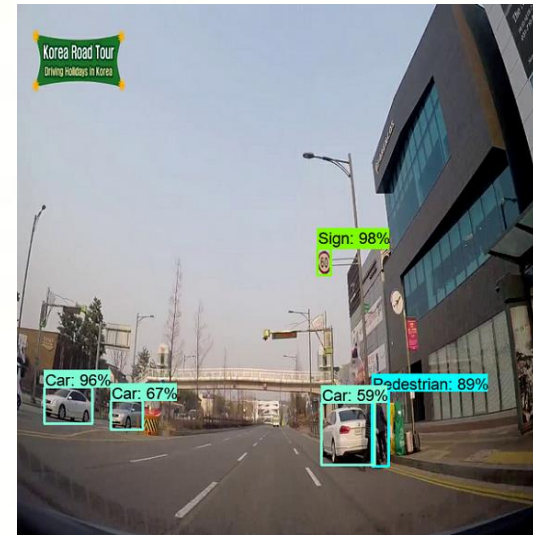
https://www.asozykin.ru/deep_learning/2017/05/08/CNN-for-MNIST.html

Часть 4

Применение нейронных сетей для нетривиальных задач компьютерного зрения

Система помощи водителю ADAS

- обнаружение и классификация различных объектов
- семантическая сегментация дорожных сцен.
- Распознавание дорожной разметки.



Распознавание достопримечательностей

По словам специалистов Google, распознавание достопримечательностей стоит особняком среди актуальных проблем машинного обучения. Например, даже набора данных Google-Landmarks с 2 млн изображений недостаточно, чтобы описать множество менее известных памятников. Это связано с тем, что достопримечательности, как правило, представляют из себя неподвижные предметы, и их внутриклассовая вариация крайне мала. Это усложняет поиск ключевых особенностей распознаваемых объектов, так как их зачастую просто нет. Разнообразие можно воссоздавать искусственно, делая снимки памятника с различных ракурсов, однако такой подход ведёт к разрастанию баз данных. С подобной проблемой также приходится сталкиваться разработчикам систем распознавания произведений искусства.



Список литературы

1. Marvin Minsky and Seymour Papen, Perceptions An Introduction to Computational Geometry <https://drive.google.com/file/d/1UsoYSWypNjRth-Xs81FsoyqWDSdnhjIB/view>
2. Понс, Форсайт, Компьютерное зрение
3. Гонсалес Р. Вудс Р., Цифровая Обработка Изображения
4. А. Потапов, Системы компьютерного зрения: современные задачи и методы, <https://docplayer.ru/26522327-Sistemy-kompyuternogo-zreniya-sovremennye-zadachi-i-metody.html>
5. Richard Szeliski, Computer Vision: Algorithms and Applications <http://szeliski.org/Book/>
6. M. Nielsen, "Neural Networks and Deep Learning", <http://neuralnetworksanddeeplearning.com/>
7. Andrew Ng, Machine learning course <https://www.coursera.org/learn/machine-learning>

Спасибо за внимание!