

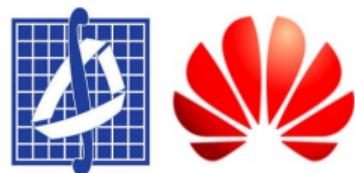
# Нейронные сети

## Лекция 6. Введение в современные нейронные сети

Бабин Д.Н., Иванов И.Е., Петюшко А.А.

МаТИС

22 октября 2021г.



- ① Историческая справка
- ② Нейросетевой подход
- ③ Постановка основных задач компьютерного зрения
- ④ Примеры прикладных задач

# Нейросетевой подход

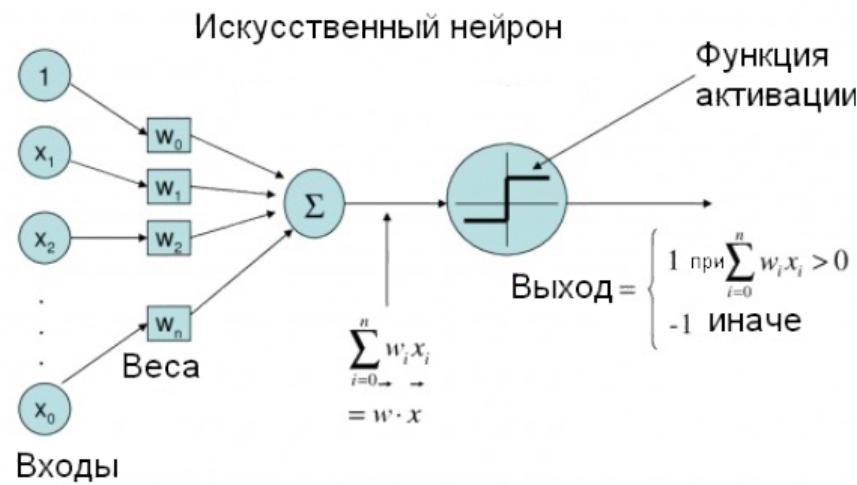
- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети

# Нейросетевой подход

- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети
- Простейшая нейронная сеть — однослойный перцептрон — состоит из линейного слоя и следующей за ним нелинейной функции

# Нейросетевой подход

- На данный момент подавляющее большинство алгоритмов при работе с компьютерным зрением использует нейронные сети
- Простейшая нейронная сеть — однослойный перцептрон — состоит из линейного слоя и следующей за ним нелинейной функции
- В 1943 г. Маккалок и Питтс формализуют понятие нейрона



- 1949 — Хебб предлагает первый алгоритм обучения перцептрана

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрана и неокогнитрана (прототип сверточных сетей)

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрона и неокогнитрона (прототип сверточных сетей)
- 1989 — Лекун представляет первую работу по объединению свёрточных нейросетей и метода обратного распространения ошибки

- 1949 — Хебб предлагает первый алгоритм обучения перцептрана
- 1958 — Розенблattt изобретает однослойный перцептрон и демонстрирует его способность решать задачу классификации
- 1974 — Галушкин и Вербос предлагают метод обратного распространения ошибки (впоследствии улучшен Хинтоном в 1986)
- 1975–1980 — Фукусима предлагает мощную модель когнитрона и неокогнитрона (прототип сверточных сетей)
- 1989 — Лекун представляет первую работу по объединению свёрточных нейросетей и метода обратного распространения ошибки
- 2012 — Крижевский и Хинтон публикуют первую успешную глубокую классификационную нейросеть AlexNet

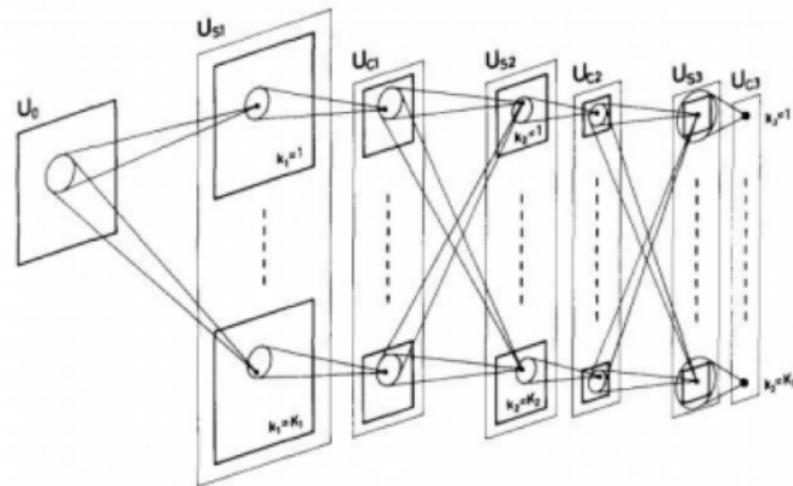
# Неокогнитрон

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры
- Два вида нейронов:
  - Простые (Simple), отвечающие за локальные признаки
  - Сложные (Complex), отвечающие за компенсацию искажения
  - Организованы в каскадную структуру S-C-S-C-S-C...
  - В сверточной сети S=свертка, C=субдискретизация

# Неокогнитрон

- Фукусима предложил практически современный метод построения архитектуры нейросетей, заимствованный им из модели первичной зрительной коры
- Два вида нейронов:
  - Простые (Simple), отвечающие за локальные признаки
  - Сложные (Complex), отвечающие за компенсацию искажения
  - Организованы в каскадную структуру SCSCSC...
  - В сверточной сети S=свертка, C=субдискретизация



- Главный минус: не было предложено метода обратного распространения ошибки для обучения

# Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году

---

<sup>1</sup><https://arxiv.org>

<sup>2</sup><https://github.com>

# Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:

---

<sup>1</sup><https://arxiv.org>

<sup>2</sup><https://github.com>

# Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
  - Появились большие наборы данных для обучения ( $\sim 10^6$  изображений)

---

<sup>1</sup><https://arxiv.org>

<sup>2</sup><https://github.com>

# Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
  - Появились большие наборы данных для обучения ( $\sim 10^6$  изображений)
  - Появились мощные видеокарты для обучения нейронных сетей

---

<sup>1</sup><https://arxiv.org>

<sup>2</sup><https://github.com>

# Нейронные сети — почему сейчас?

- Несмотря на то, что весь математический аппарат был готов уже к середине 80-х годов прошлого века, революция в компьютерном зрении произошла только в 2012 году
- Основные причины начала новой эры в нейронных сетях:
  - Появились большие наборы данных для обучения ( $\sim 10^6$  изображений)
  - Появились мощные видеокарты для обучения нейронных сетей
  - Открытость исследований (arxiv<sup>1</sup>, github<sup>2</sup>)

---

<sup>1</sup><https://arxiv.org>

<sup>2</sup><https://github.com>

# Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.

---

<sup>3</sup><https://colab.research.google.com>

<sup>4</sup><https://aws.amazon.com/ru/ec2/>

# Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.
- Для обучения промышленных нейросетей (например, Mask R-CNN) необходим объем памяти GPU от 12-16 ГБ

---

<sup>3</sup><https://colab.research.google.com>

<sup>4</sup><https://aws.amazon.com/ru/ec2/>

# Ресурсы для работы с нейросетями

- На данный момент обучение нейросетей немыслимо без использования **GPU** (Graphical Processing Unit) от NVIDIA, **TPU** (Tensor Processing Unit) от Google, AI Chip (Da Vinci) от Huawei и пр.
- Для обучения промышленных нейросетей (например, Mask R-CNN) необходим объем памяти GPU от 12-16 ГБ
  - В силу этого подходящие карты на данный момент: Titan X / V / RTX, GTX 1080 Ti, RTX 2080 Ti / 3080 / 3090, P100, V100, A100
  - Для любого проекта необходимо обеспечить доступ к данному оборудованию (например, удаленно через ssh)
  - Либо воспользоваться облачными сервисами (например, через Google Colaboratory<sup>3</sup> либо Amazon Web Services<sup>4</sup>)

---

<sup>3</sup><https://colab.research.google.com>

<sup>4</sup><https://aws.amazon.com/ru/ec2/>

# Задача классификации

- **Задача:** отнести входное изображение к одному из классов
  - По фотографии человека определить пол, возраст и т. д.
  - По фотографии автомобиля определить марку и модель

# Задача классификации

- **Задача:** отнести входное изображение к одному из классов
  - По фотографии человека определить пол, возраст и т. д.
  - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
  - Решается с помощью сверточных нейросетей (СНС)

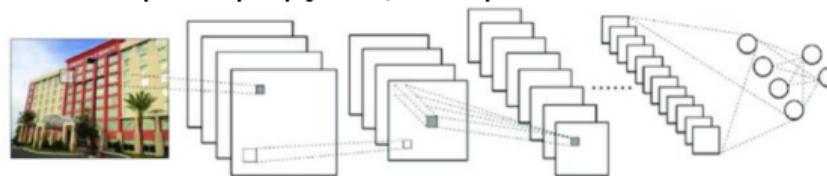
# Задача классификации

- **Задача:** отнести входное изображение к одному из классов
  - По фотографии человека определить пол, возраст и т. д.
  - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
  - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
  - Не любое изображение можно описать одним классом

# Задача классификации

- **Задача:** отнести входное изображение к одному из классов
  - По фотографии человека определить пол, возраст и т. д.
  - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
  - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
  - Не любое изображение можно описать одним классом

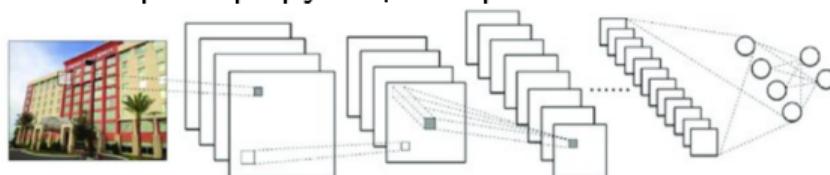
Пример функционирования СНС



# Задача классификации

- **Задача:** отнести входное изображение к одному из классов
  - По фотографии человека определить пол, возраст и т. д.
  - По фотографии автомобиля определить марку и модель
- Самая простая задача компьютерного зрения
  - Решается с помощью сверточных нейросетей (СНС)
- **Проблема:** двусмысленность изображений
  - Не любое изображение можно описать одним классом

Пример функционирования СНС



Основные представители:

- 2012 — AlexNet
- 2014 — VGGNet
- 2014 — Inception
- 2015 — ResNet
- 2016 — ResNeXt
- 2017 — MobileNet
- 2017 — (P)NASNet
- 2019 — EfficientNet
- 2020 — ViT

# Проблемная классификация

Сено, лошадь или женщина?



# Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником

# Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект

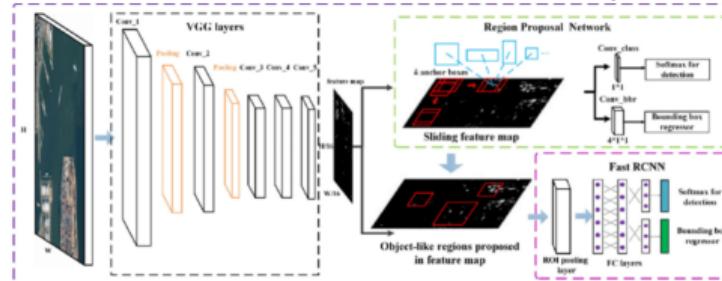
# Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
  - Faster R-CNN, YOLO, SSD

# Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
  - Faster R-CNN, YOLO, SSD

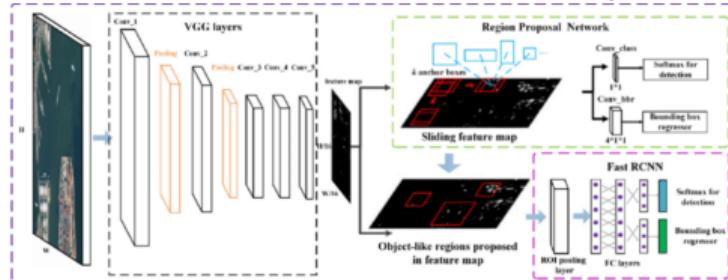
Пример функционирования детектора (Faster R-CNN)



# Задача детекции

- **Задача:** по входному изображению требуется найти объекты заранее заданных классов и обвести их прямоугольником
- Сложнее, чем классификация, так как необходимо не только распознать класс, но и локализовать объект
- Существуют алгоритмы, которые решают эту задачу в т. ч. в реальном времени
  - Faster R-CNN, YOLO, SSD

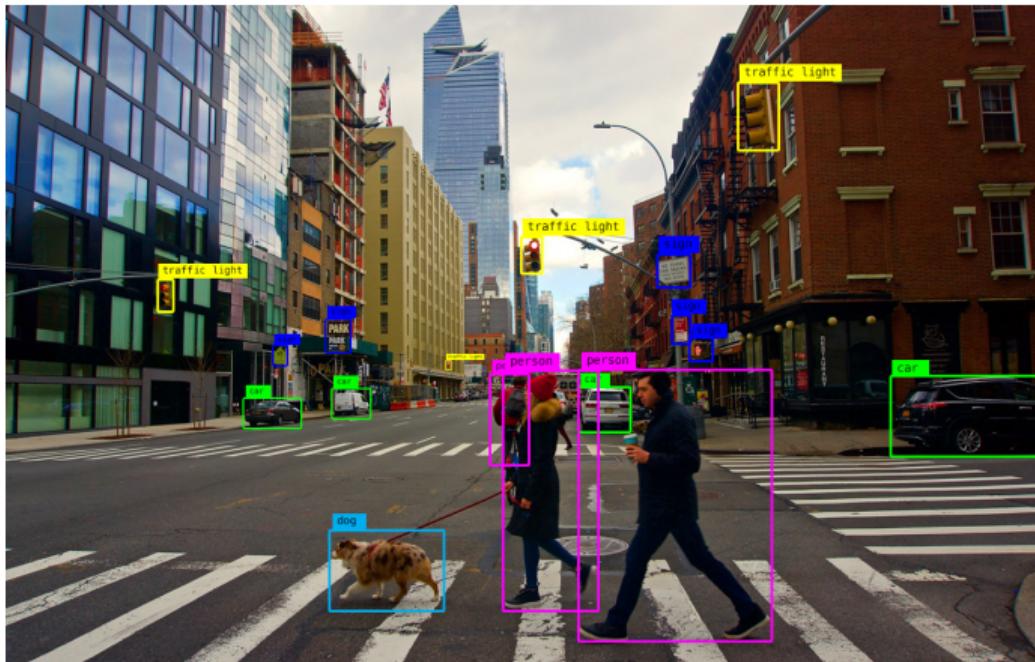
Пример функционирования детектора (Faster R-CNN)



Основные представители:

- 2013 — Overfeat
- 2014 — R-CNN
- 2015 — Faster R-CNN
- 2015 — YOLO
- 2015 — SSD
- 2016 — R-FCN
- 2017 — FPN
- 2017 — RetinaNet
- 2019 — CenterNet
- 2019 — EfficientDet
- 2020 — DETR

# Пример работы детектора



# Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта

# Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку

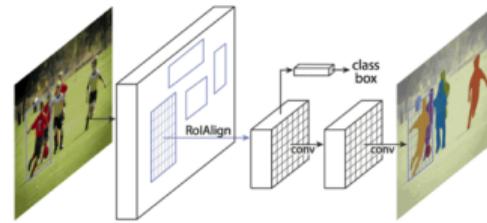
# Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

# Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

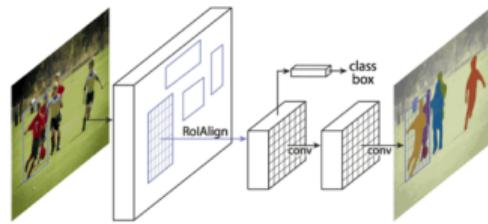
Пример функционирования решения-комбайна (Mask R-CNN)



# Задача сегментации

- **Задача:** по входному изображению найти объекты заранее заданных классов и выдать пиксельную маску объекта
- Сложнее, чем детекция — нужно точно определять границы объекта, а не просто заключить в рамку
- На данный момент существуют алгоритмы-комбайны (решающие одновременно задачи классификации, детекции и сегментации)

Пример функционирования решения-комбайна (Mask R-CNN)



Основные представители:

- 2014 — FCN
- 2015 — U-Net
- 2015 — SegNet
- 2016 — RefineNet
- 2016 — PSPNet
- 2017 — Mask R-CNN
- 2018 — DeepLabv3+
- 2020 — ResNeSt
- 2020 — DetectoRS



# Пример работы решения-комбайна



# Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра

# Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео

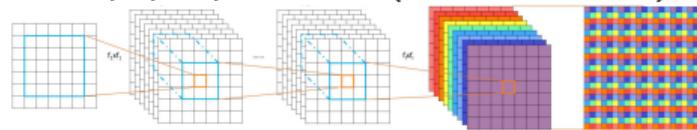
# Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

# Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

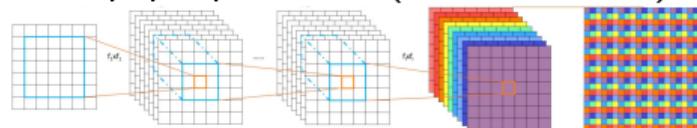
Пример функционирования алгоритма сверхразрешения (ESPCN, 2016)



# Задачи улучшения изображений

- **Задача сверхразрешения:** увеличить пространственное разрешение фотографии / видео-кадра
- **Метод решения:** применить нейронную сеть, заранее обученную на множестве картинок / видео
- **Цель:** улучшить классические алгоритмы интерполяции

Пример функционирования алгоритма сверхразрешения (ESPCN, 2016)



Метод сверхразрешения	PSNR
Бикубическая интерполяция	24.32 dB
Бикубическая интерполяция + GIMP	24.69 dB
Бикубическая интерполяция + Photoshop	24.72 dB
<i>Нейро сверхразрешение</i>	<i>27.20 dB</i>

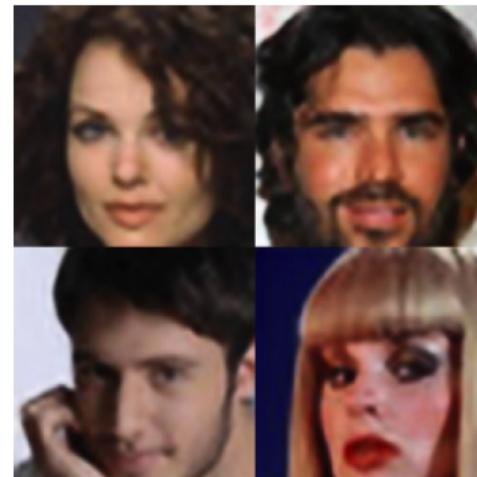
# Задачи улучшения изображений

## Сверхразрешение для лиц — пример

Бикубическая интерполяция: PSNR = 24.32



НейроСверх-разрешение: PSNR = 27.20



Малоразмерный вход

# Задачи улучшения изображений

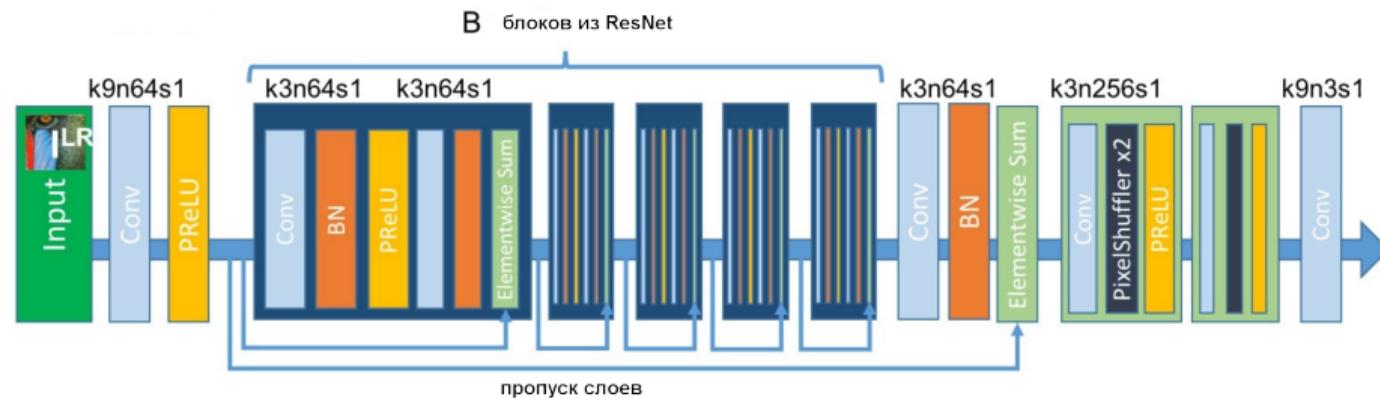
- **Задача:** улучшить качество дефектных изображений
  - Устранить смаз (motion blur)
  - Устранить размытие (defocus blur)
  - Удалить артефакты, являющиеся результатом попытки устраниить дефекты зашитым в устройстве алгоритмом (image reconstruction)

# Задачи улучшения изображений

- **Задача:** улучшить качество дефектных изображений

- Устранить смаз (motion blur)
- Устранить размытие (defocus blur)
- Удалить артефакты, являющиеся результатом попытки устраниить дефекты зашитым в устройстве алгоритмом (image reconstruction)

Пример функционирования алгоритма деблюра (SRGAN, 2016)



# Задачи улучшения изображений

Исходное  
размытие



Деблор для лиц — пример  
Классический  
Wiener



Коммерческое  
решение



Нейросетевое  
решение



## Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер



## Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер
- Зачастую итоговое решение состоит из следующих модулей:
  - Детектирование автомобиля
  - Распознавание марки и модели (опционально)
  - Трекинг автомобиля
  - Детектирование номера
  - Распознавание номера
  - Объединение результатов с разных кадров / поиск ключевых кадров

## Распознавание автомобильных номеров

- **Задача:** по видеопотоку детектировать транспортное средство и распознать госномер
- Зачастую итоговое решение состоит из следующих модулей:
  - Детектирование автомобиля
  - Распознавание марки и модели (опционально)
  - Трекинг автомобиля
  - Детектирование номера
  - Распознавание номера
  - Объединение результатов с разных кадров / поиск ключевых кадров



# Примеры прикладных задач

## Ре-идентификация пешеходов

- **Задача:** идентифицировать пешеходов на камерах наблюдения



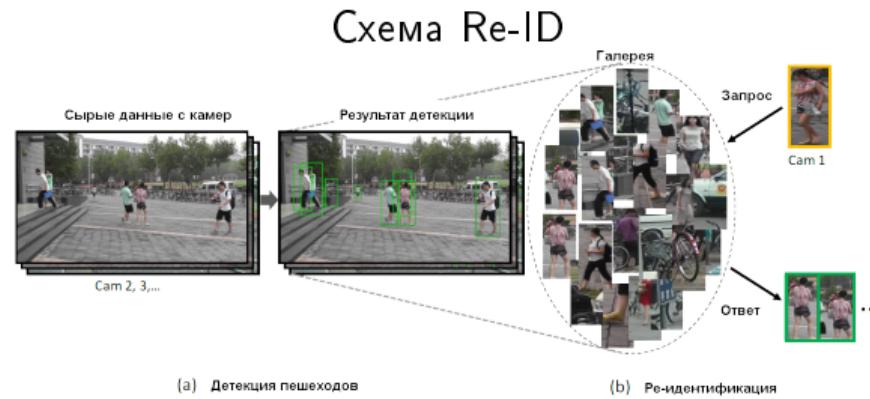
## Ре-идентификация пешеходов

- **Задача:** идентифицировать пешеходов на камерах наблюдения
- **Метод решения:**
  - Детектирование пешеходов
  - Трекинг пешеходов
  - Извлечение признаков
  - Поиск в пространстве признаков

# Примеры прикладных задач

## Ре-идентификация пешеходов

- **Задача:** идентифицировать пешеходов на камерах наблюдения
- **Метод решения:**
  - Детектирование пешеходов
  - Трекинг пешеходов
  - Извлечение признаков
  - Поиск в пространстве признаков



## Распознавание лиц

- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
  - **Верификация** — принадлежат ли две входные картинки одному человеку?
  - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход

## Распознавание лиц

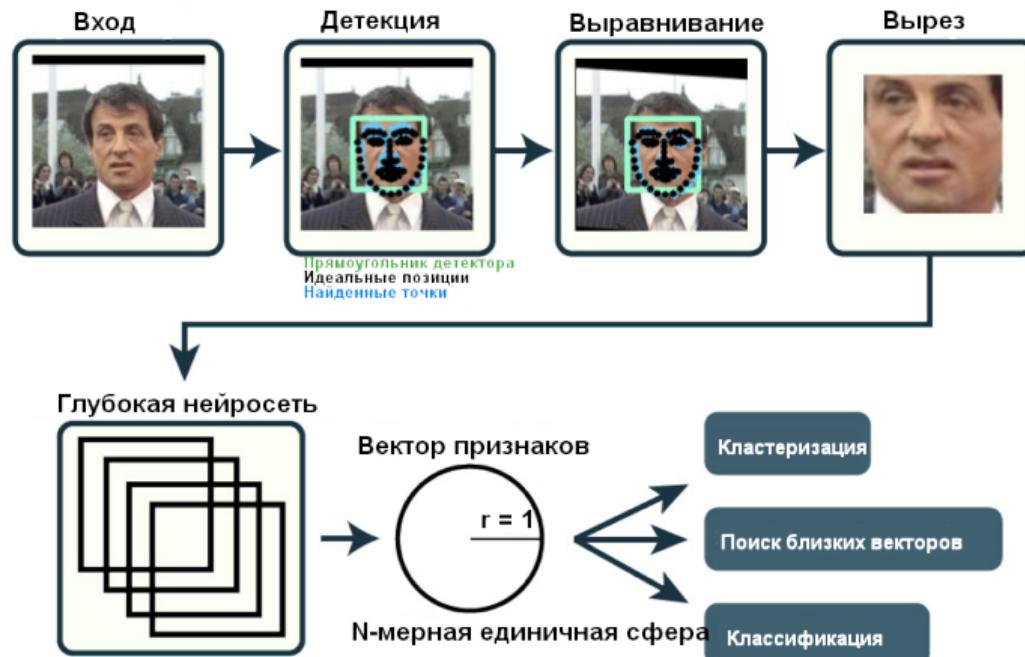
- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
  - **Верификация** — принадлежат ли две входные картинки одному человеку?
  - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход
- **Метод решения:** с помощью каскада СНС

## Распознавание лиц

- **Задача:** по входному изображению лица распознать человека. Есть два подтипа задач:
  - **Верификация** — принадлежат ли две входные картинки одному человеку?
  - **Идентификация** — нахождение в базе данных человека, фотография которого подана на вход
- **Метод решения:** с помощью каскада СНС
- Главная особенность — нахождение  $N$ -мерного вектора признаков, соответствующего данному человеку

# Примеры прикладных задач

## Типичная система распознавания лиц:



## Синтез (создание) лиц

- **Задача:** сгенерировать набор синтетических лиц, неотличимых от реальных фотографий
  - Зачастую нужно для расширения обучающей выборки для задачи распознавания лиц
  - Необходимо генерировать как можно больше разных личностей

## Синтез (создание) лиц

- **Задача:** сгенерировать набор синтетических лиц, неотличимых от реальных фотографий
  - Зачастую нужно для расширения обучающей выборки для задачи распознавания лиц
  - Необходимо генерировать как можно больше разных личностей
- **Метод решения:** с помощью специальных генеративных состязательных нейросетей (GAN)

# Примеры прикладных задач

Какое из этих лиц — синтетическое<sup>5</sup>?



<sup>5</sup><https://thispersondoesnotexist.com>

Спасибо за внимание!