

академия  
больших  
данных

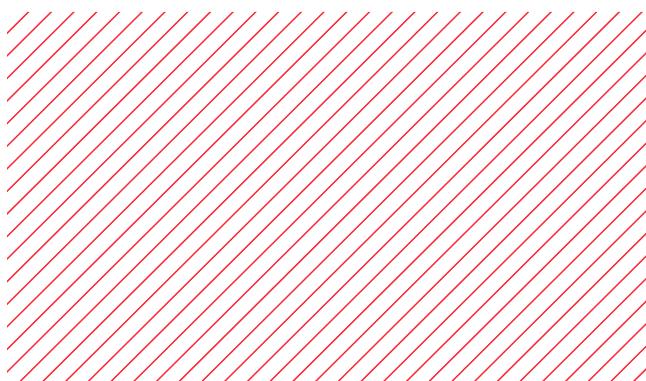


mail.ru  
group

# Как вообще делать задачи

Даниил Лысухин

Программист-исследователь в команде машинного  
зрения





# Схема любой задачи

---

- Данные
- Модель
- Обучение
- Оптимизация
- Развёртывание



# Данные

---

- Собрать
- Прочистить разметку
  - Ручками
  - Эвристиками
  - Обученной моделью
- \*Сбалансировать

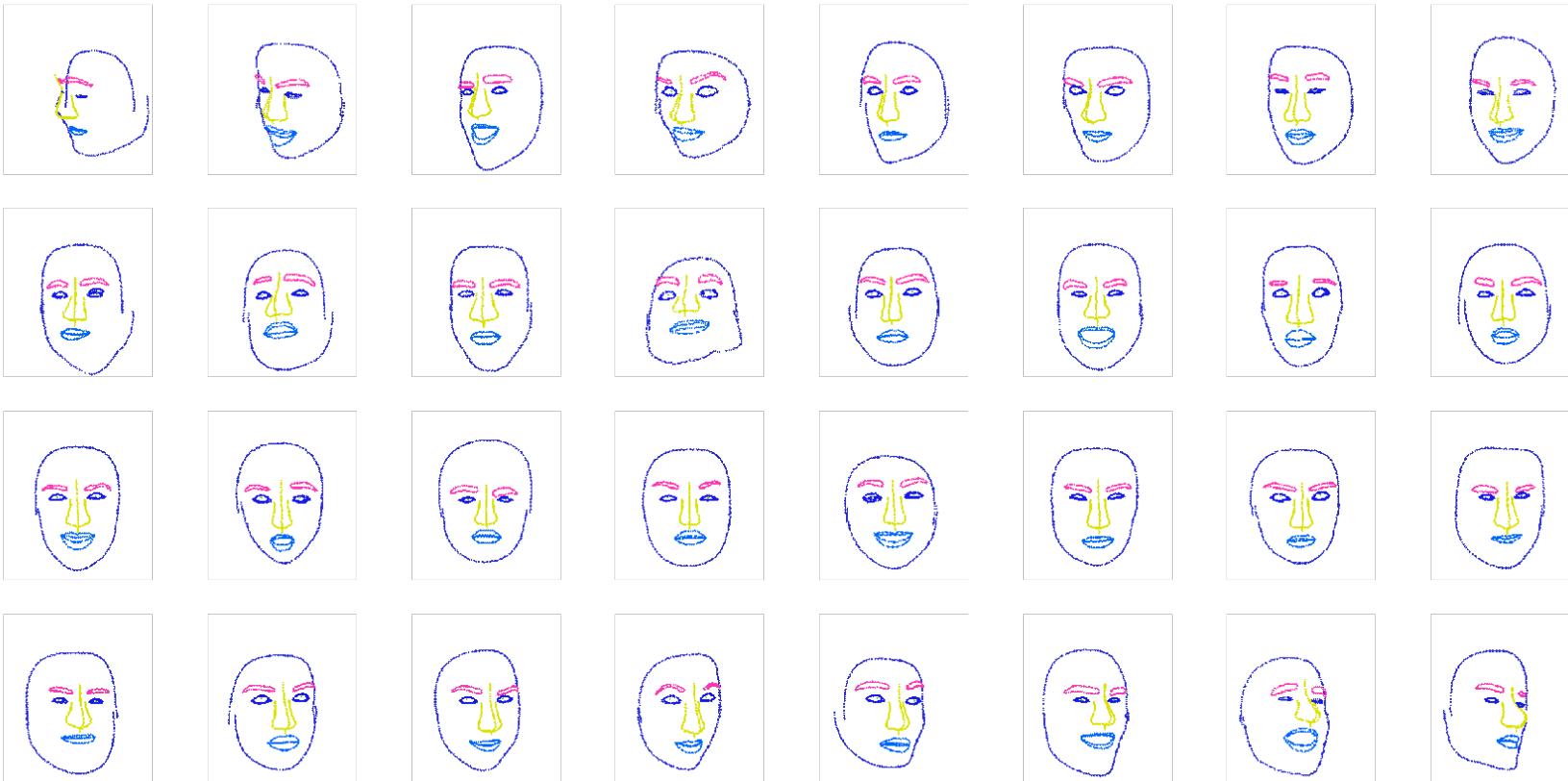
# Чистка

---



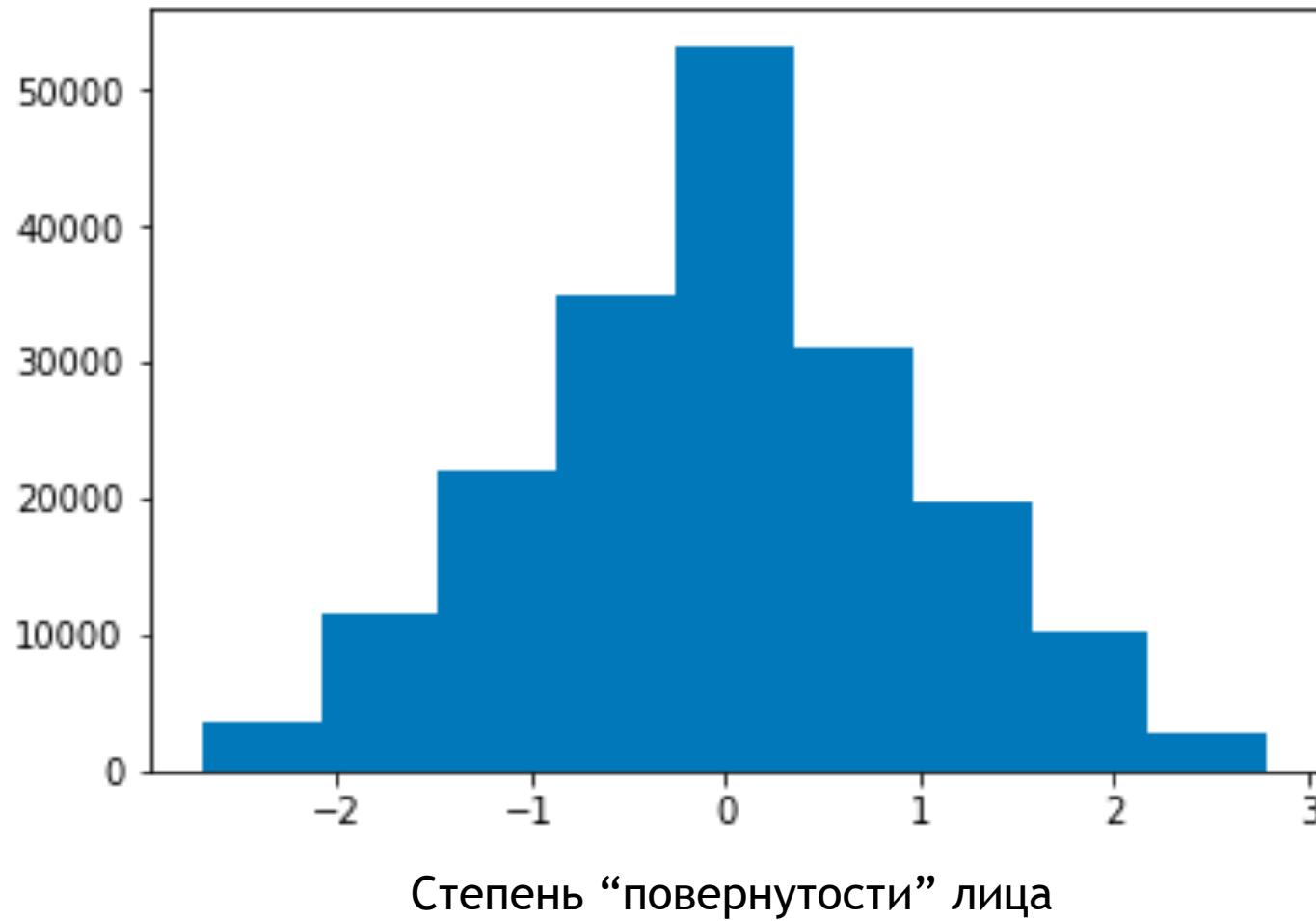
# Несбалансированность

---



# Несбалансированность

---





# Модель

---

- Подобрать архитектуру / Loss
  - Пролистать статьи!
- Реализовать
- \*Использовать pretrained



# Модель: vanilla regression

---

```
model = models.resnet18(pretrained=True)
model.fc = nn.Linear(model.fc.in_features, 2 * NUM PTS, bias=True)
```



# SOTA

---

- Modern Facial Landmark Estimation and Tracking (<https://youtu.be/0yC0hhRpb7Y>)
- Practical Face Landmarks Detector (<https://arxiv.org/pdf/1902.10859.pdf>)
- Look At Boundary: A Boundary-Aware Face Alingment Algorithm (<https://arxiv.org/pdf/1805.10483.pdf>)
- Style-Aggregated Network for Facial Landmark Detection (<https://arxiv.org/pdf/1803.04108.pdf>)
- Wing Loss for Robust Facial Landmark Localisation with CNNs (<https://arxiv.org/pdf/1711.06753.pdf>)
- Adaptive Wing Loss for Robust Face Alingment via Heatmap Regression (<https://arxiv.org/pdf/1904.07399.pdf>)
- An intriguing failing of CNNs and the CoordConv solution (<https://arxiv.org/pdf/1807.03247.pdf>)

# Идея: wing loss

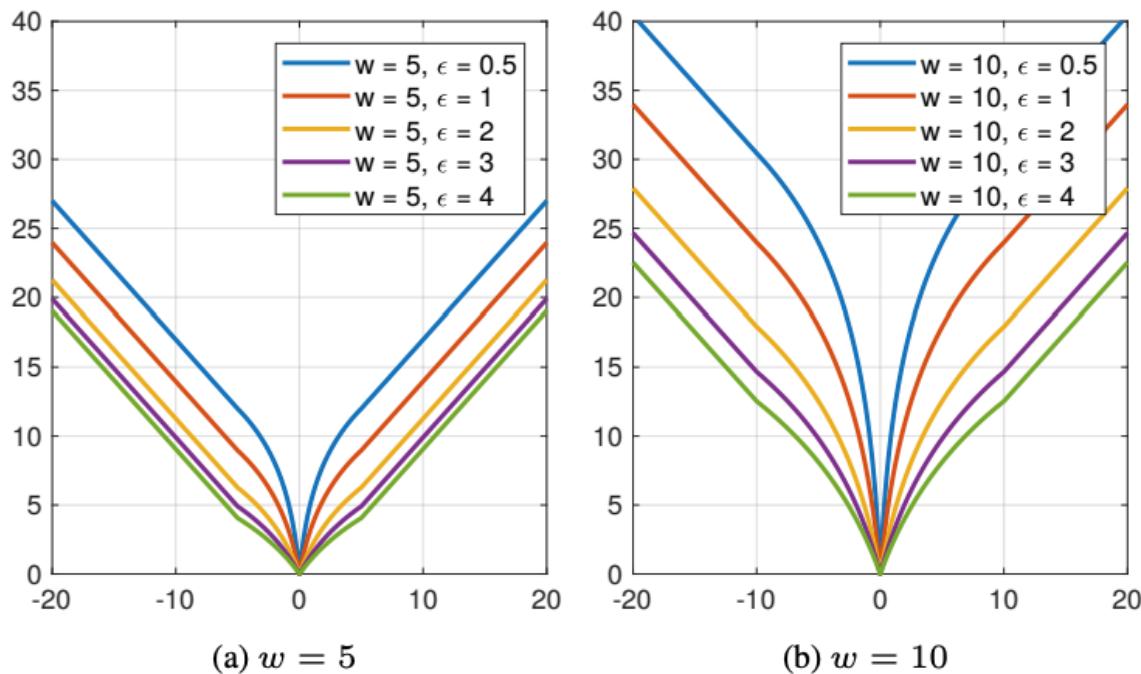


Figure 1. Our Wing loss function (Eq. 5) plotted with different parameter settings, where  $w$  limits the range of the non-linear part and  $\epsilon$  controls the curvature. By design, we amplify the impact of the samples with small and medium range errors to the network training.

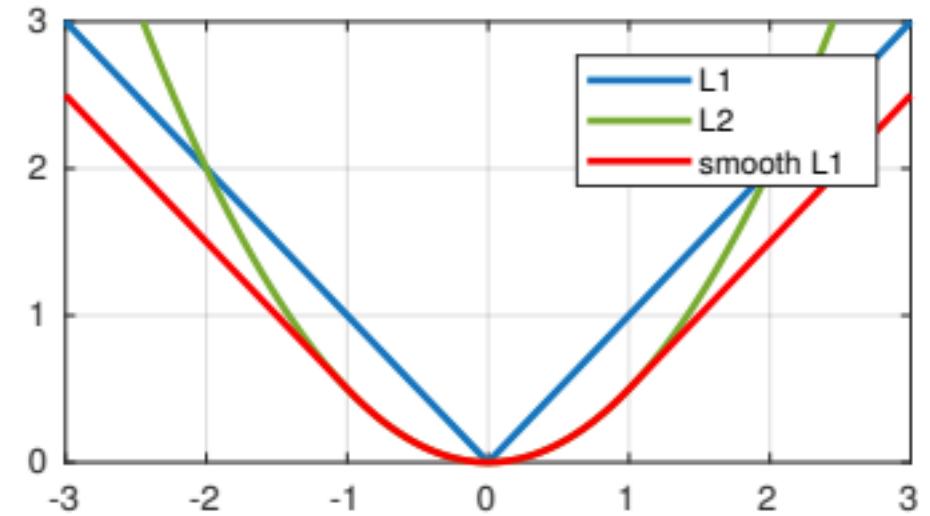


Figure 3. Plots of the L1, L2 and smooth L1 loss functions.

# Идея: heatmap regression

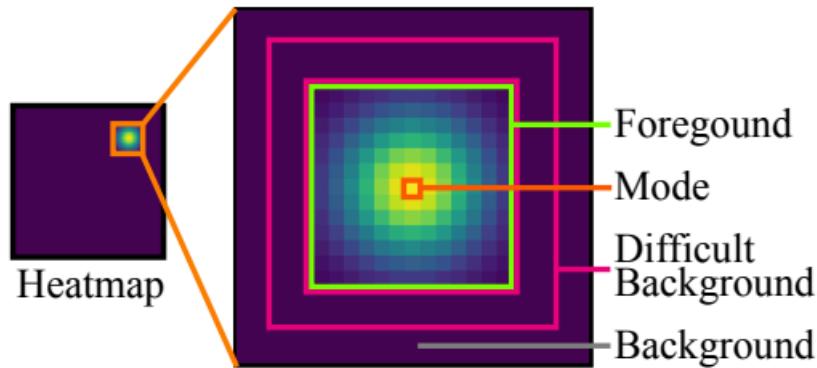
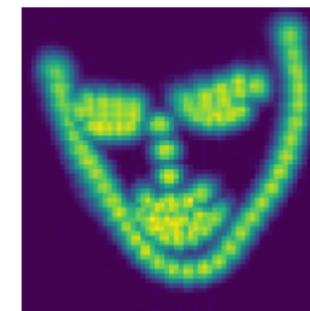
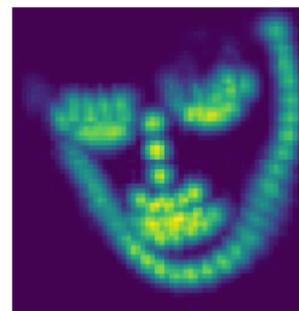
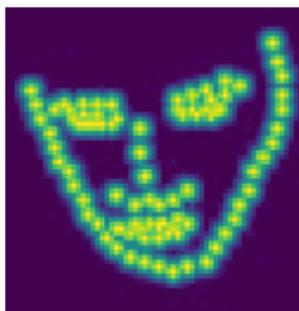
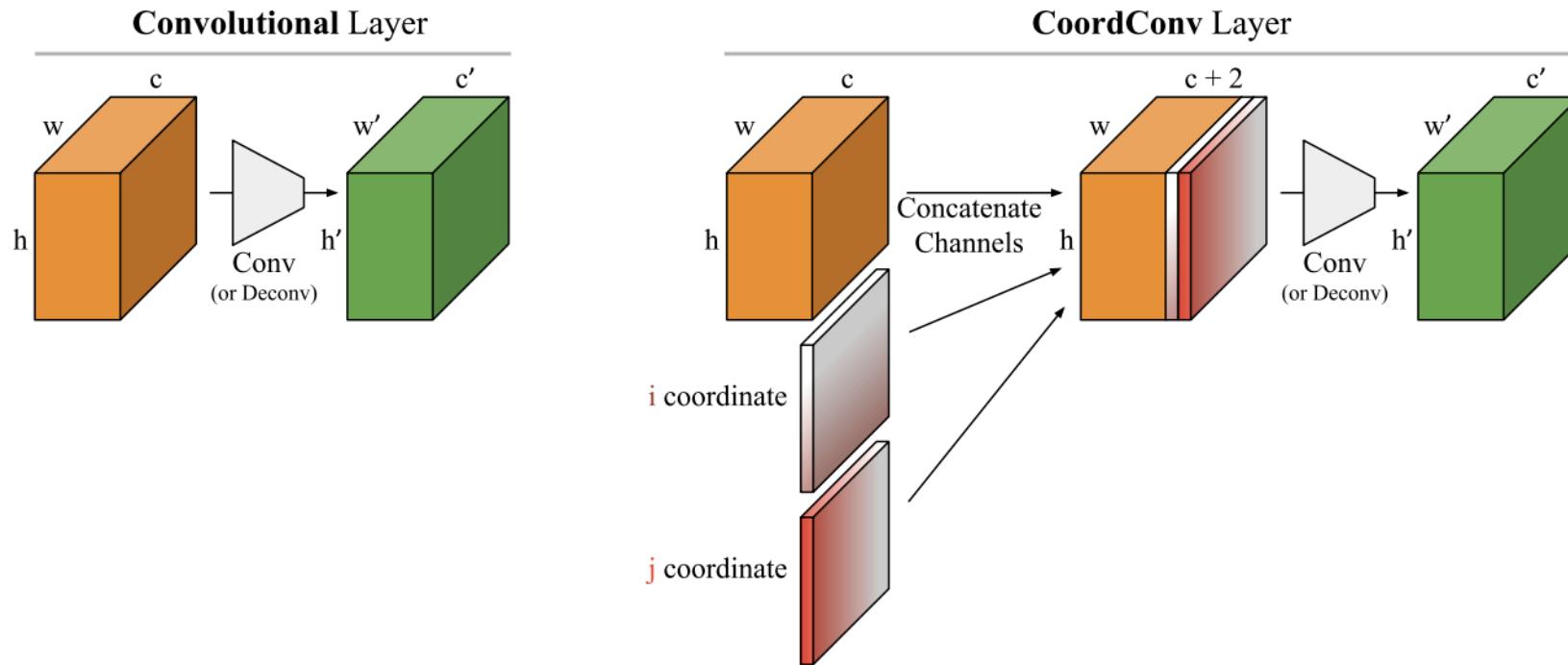


Figure 1: **Pixel type definitions.** (Best viewed in color).



# Идея: CoordConv





# Обучение

---

- Бейзлайн
  - Подмножество выборки
  - Гиперпараметры “чтобы работало” (оптимизатор, lr\_policy, ...)
  - Облегченная архитектура
- Усложнение (цикл)
  - Вся выборка
  - Тюнинг параметров модели
  - Аугментации
- ...



# Аугментации

---

- Для повышения разнообразия данных
  - Геометрия:
    - Отражения
    - Повороты на небольшие углы
    - Малые перспективные преобразования
  - Цвет:
    - Контраст и яркость
    - Hue, Saturation, Value
  - Optional:
    - Шум, размытие, ...
- Не выходить из распределения!



# Оптимизация

---

- Fuse BN
- Половинная / иная точность (fp16, TensorRT, ...)
- \*Distillation



# Развертывание

---

- Модель “в готовом виде” (TorchScript, caffe2, protobuf, ...)
- Production code
- Интеграционные тесты