

Deep Learning - Обзор

Содержание

- Программа курса
- Задачи CV
- Нейросетевая революция 2012
- Сверточные нейронные сети
- Задачи CV для сверточных нейронных сетей
- Инструменты
- CV текущее состояние (CVPR)

Программа курса

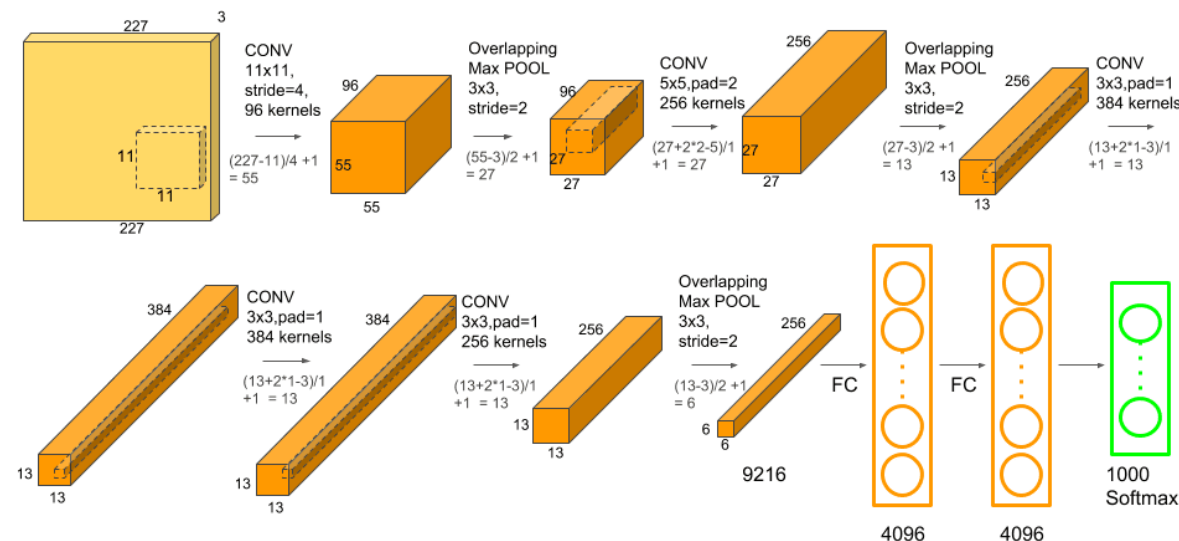
- Вводная лекция. Обзор задач CV, нефтяного инжиниринга. Применение свёрточных нейронных сетей, рекуррентных, обучение с подкреплением.
- ML-обучение. Переобучение, softmax, аугментация, дообучение. Суть машинного обучения, разметка...
- Feed forward, прямое, обратное распространение ошибки. Простейший пример.
- Свёрточные нейронные сети. Как работают.
- Задачи CV (классификация, детекторы, сегментация). Работа в группах по 2 (3-4 задачи)
- RNN
- Задачи RNN (временные ряды, CV)
- + Проект

Нейросетевая революция 2012

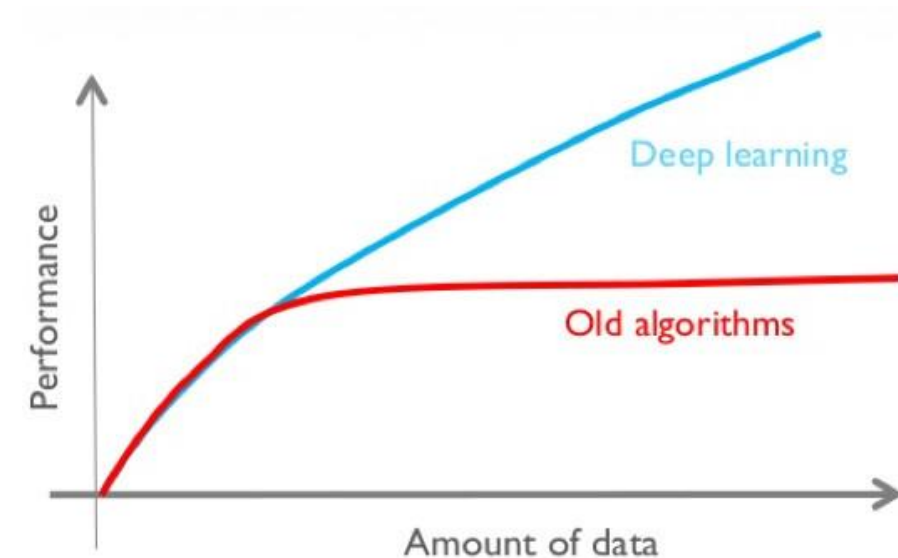
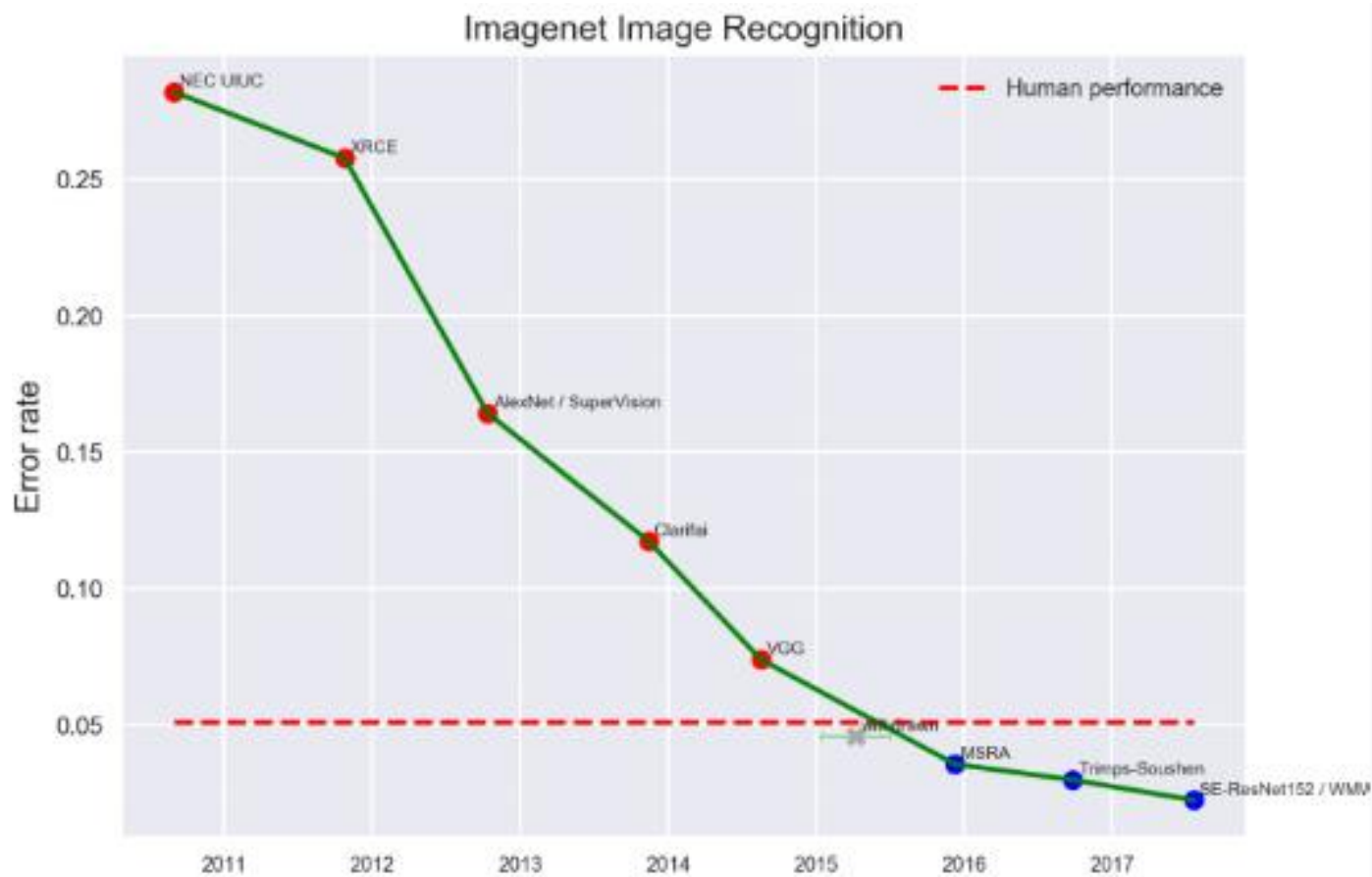


- 1М изображений
- 1К классов объектов

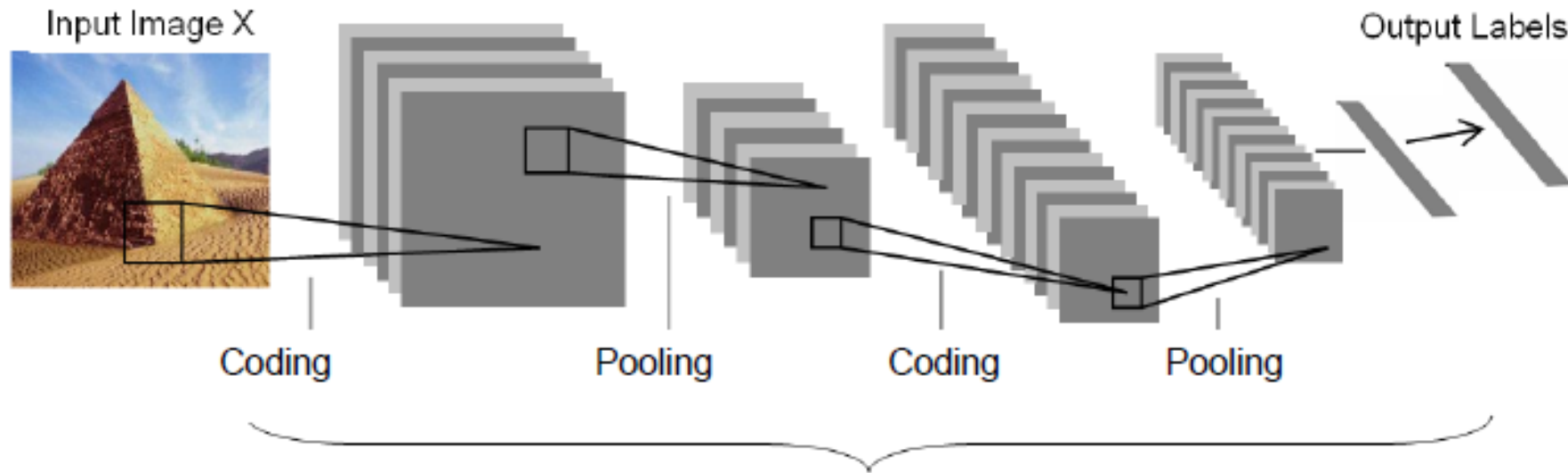
- top-5 ошибка 15.3%
- Улучшение на 10.8%



Нейросетевая революция 2012



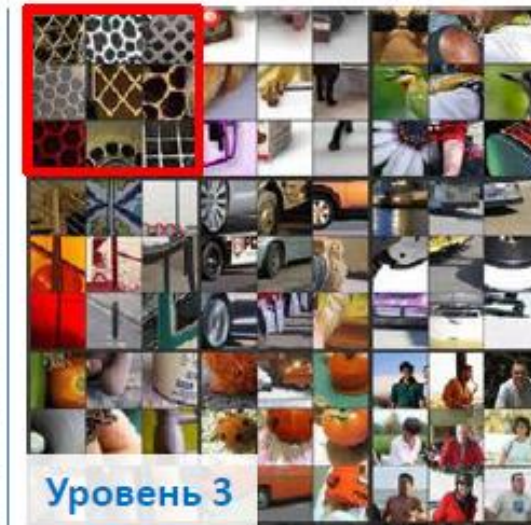
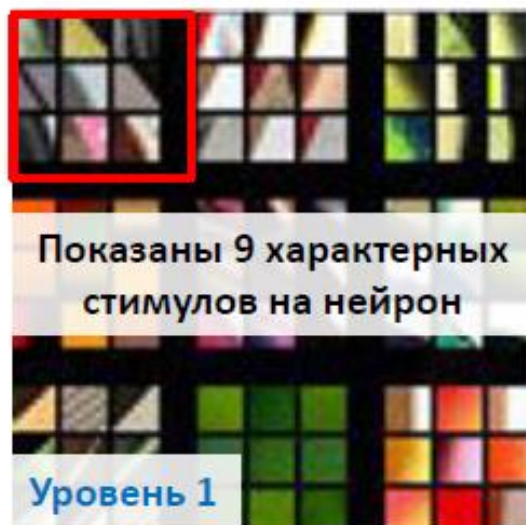
Сверточные нейронные сети



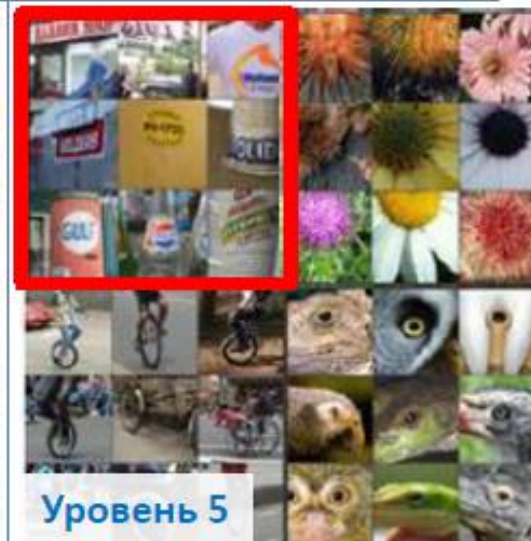
Архитектура "Coding + Pooling" = "Кодирование + Объединение"

- Иерархическое выделение **локальных** признаков при обучении и/или самообучении на сверхбольших выборках изображений
- Много десятков уровней (кодирование + объединение данных)
- **Инвариантность к сдвигу** и масштабу изображения (**свертка**)
- На верхних уровнях – собственно классификатор (полносвязный)

Сверточные нейронные сети

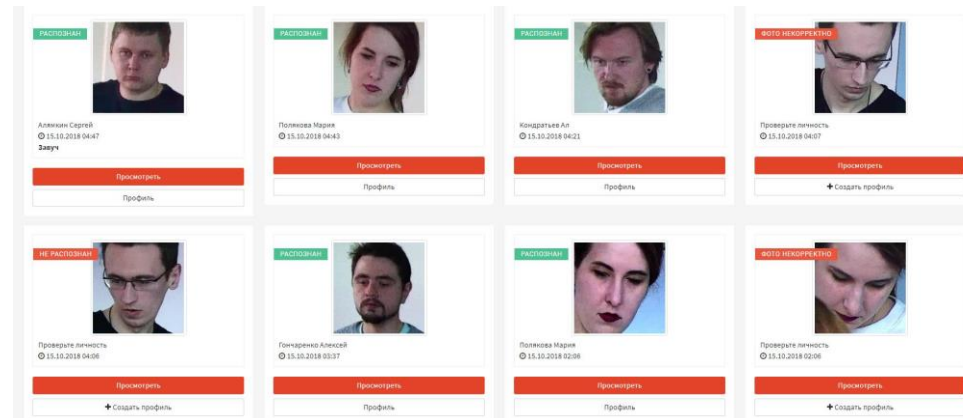
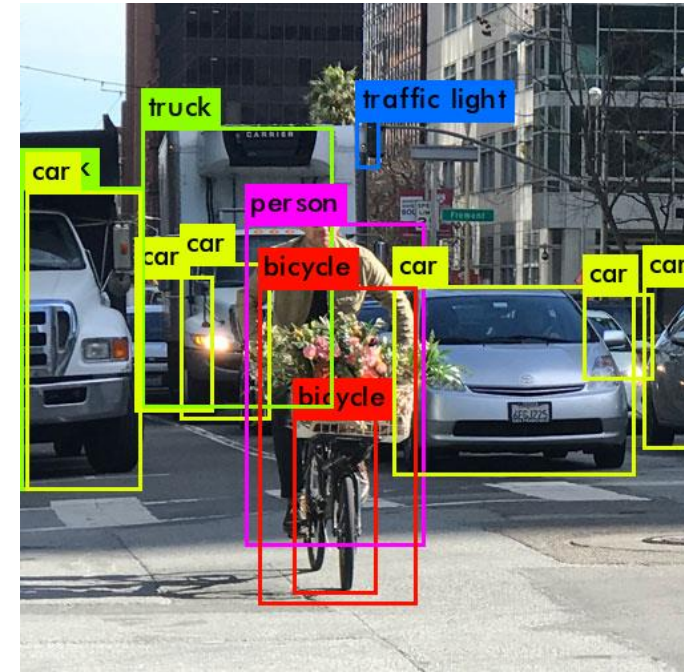


На какие элементы изображения реагируют нейроны разных уровней: **чем выше слой сети, тем выше уровень абстракции**

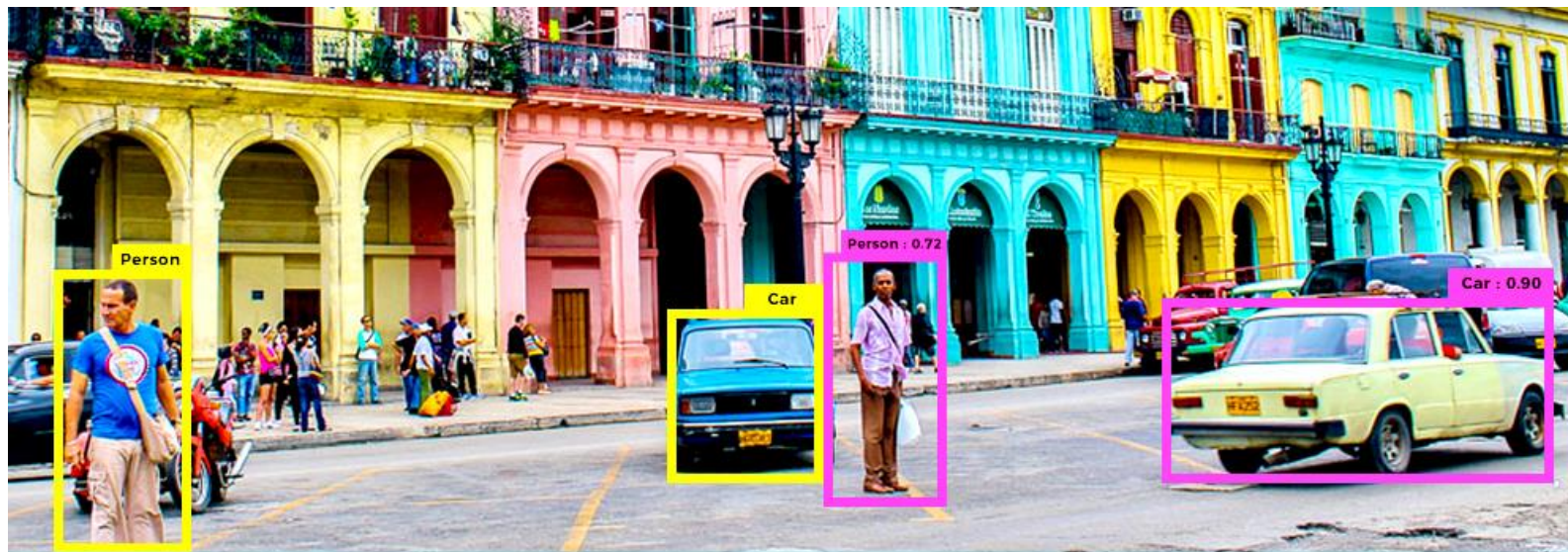
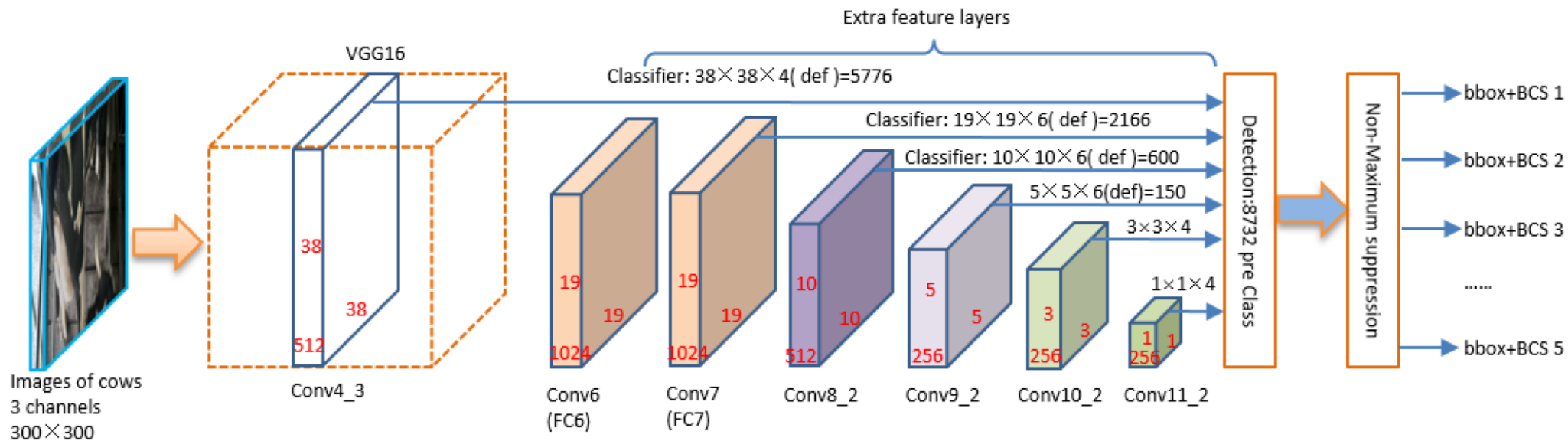


Задачи CNN

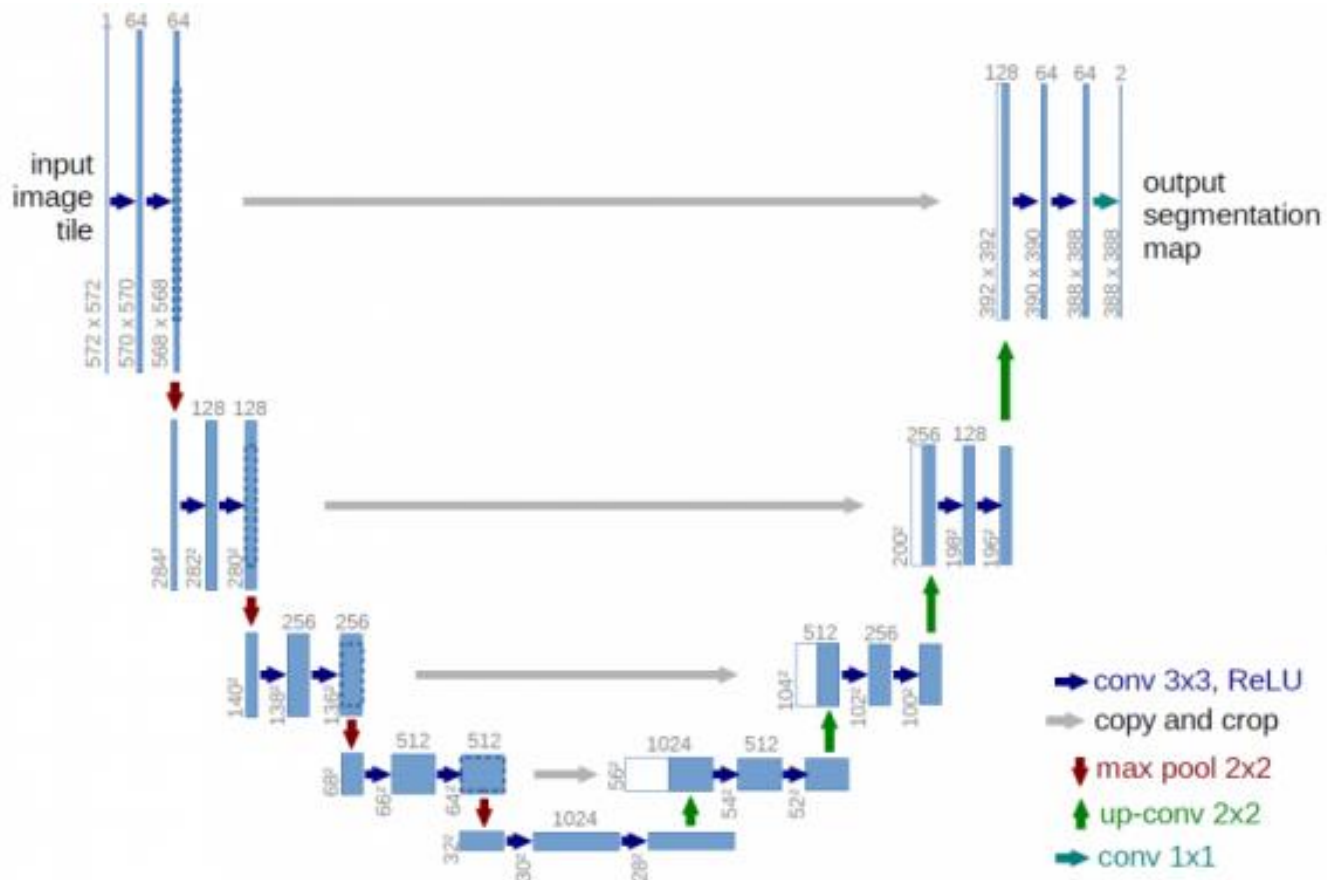
- Классификация изображений
- Детектирование объектов
- Сегментация изображений
- Распознавание текста на изображениях
- Metric Learning
- ...



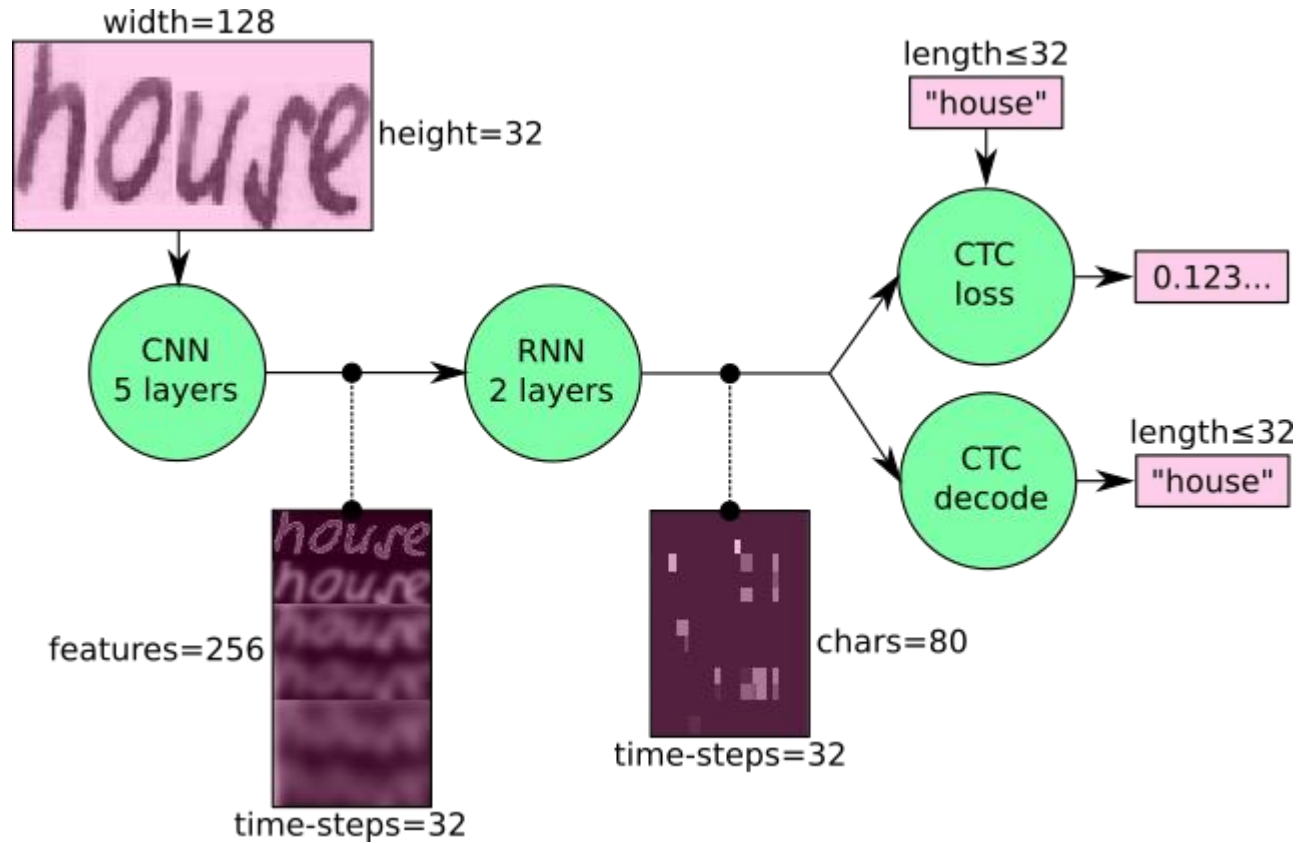
Задачи CNN – Детектирование объектов



Задачи CNN – Сегментация объектов



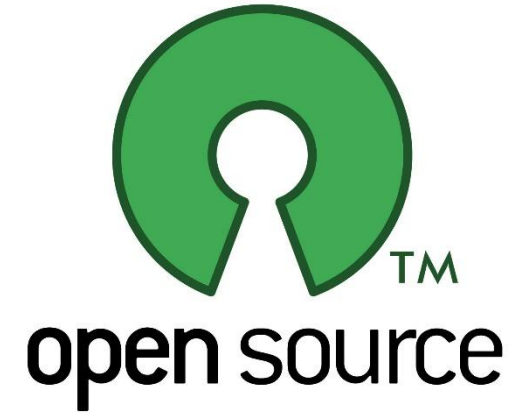
Задачи CNN – Optical Character Recognition



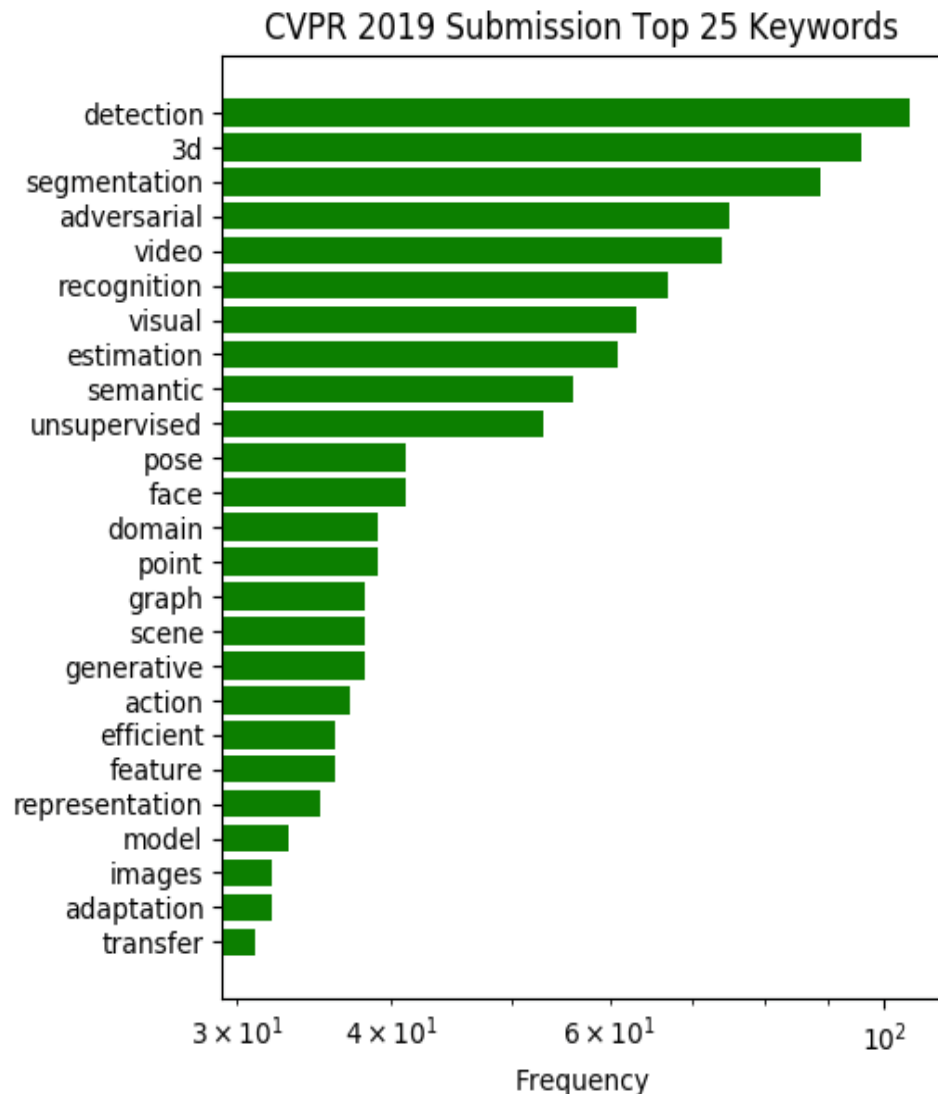
Инструменты CV



Catalyst



Текущее состояние и тренды CV

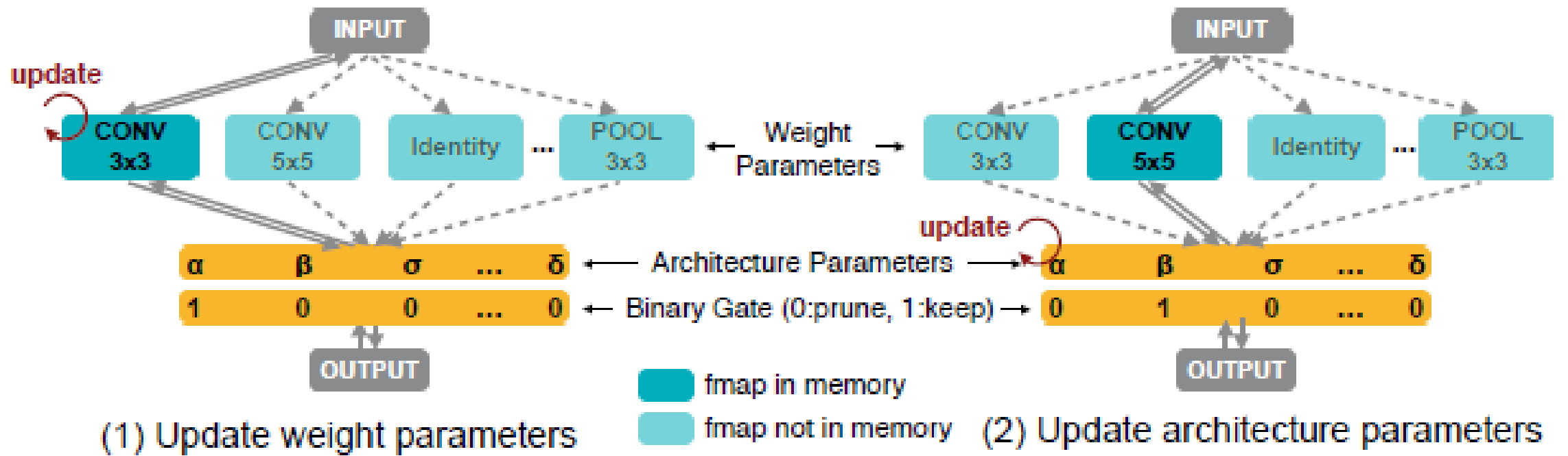


<https://towardsdatascience.com/latest-computer-vision-trends-from-cvpr-2019-c07806dd570b>

Тренды:

- Автоматический поиск архитектур (NAS)
- Оптимизация Inference
- Semi-supervised learning
- GAN
- Graph Neural Networks

NN optimization: Proxyless Neural Architecture Search (NAS)

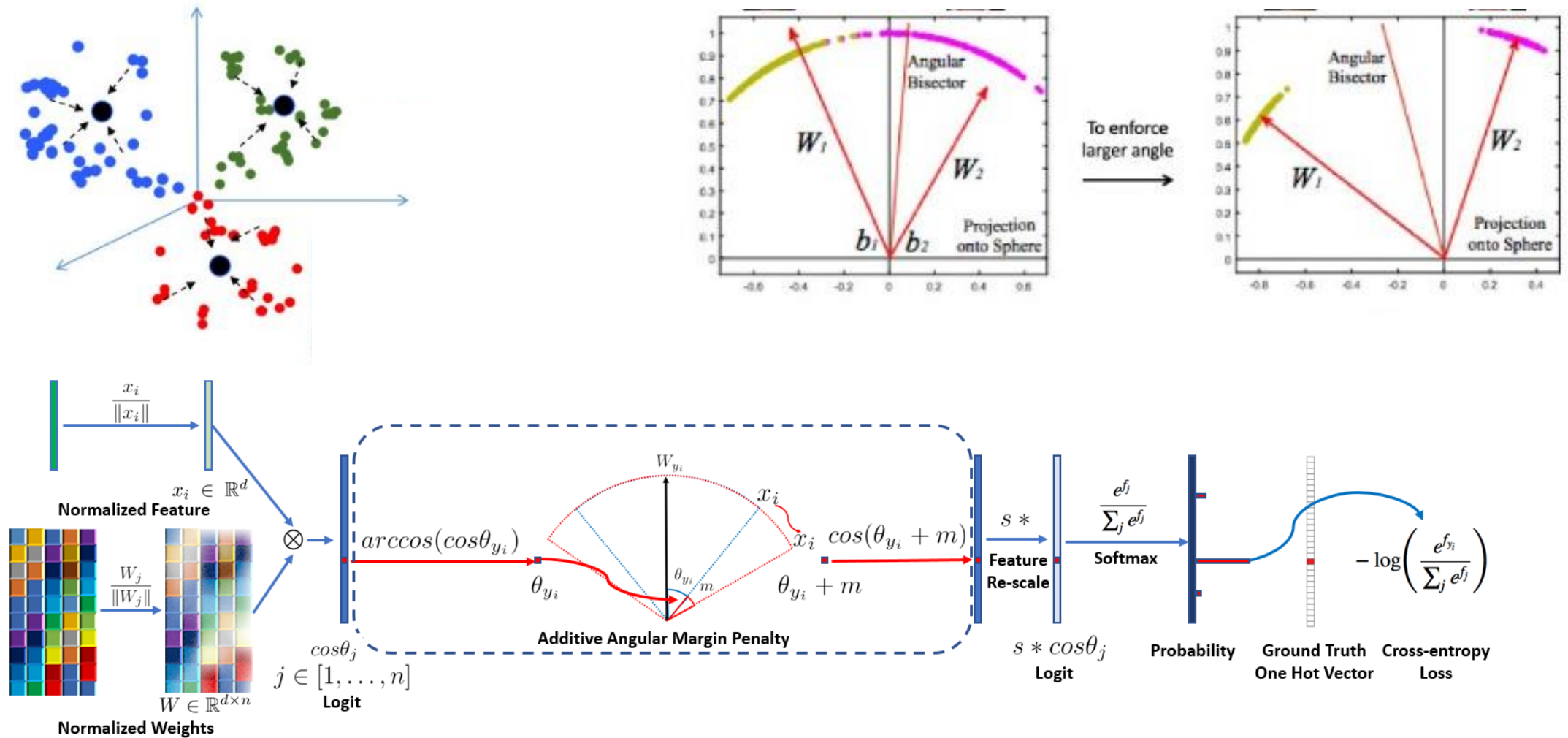


NN optimization: Proxyless NAS

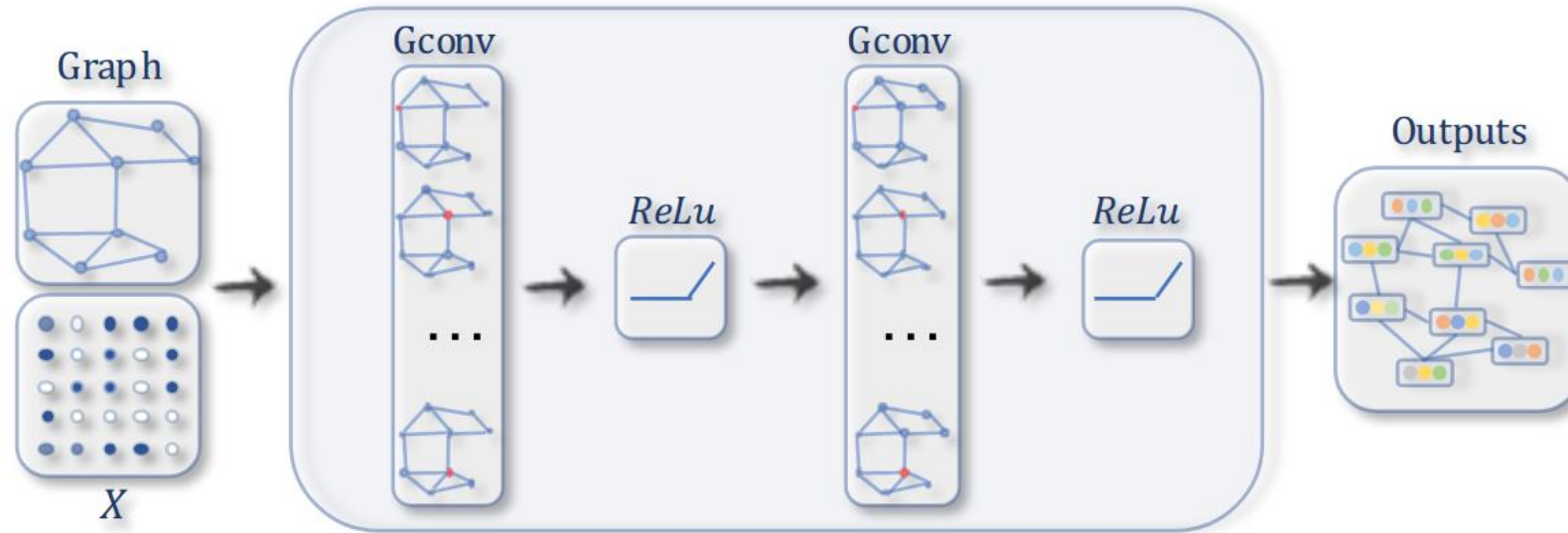
Model	Top-1	Top-5	Mobile Latency	Hardware -aware	No Proxy	No Repeat	Search cost (GPU hours)
MobileNetV1 [16]	70.6	89.5	113ms	-	-	✗	Manual
MobileNetV2 [30]	72.0	91.0	75ms	-	-	✗	Manual
NASNet-A [38]	74.0	91.3	183ms	✗	✗	✗	48,000
AmoebaNet-A [29]	74.5	92.0	190ms	✗	✗	✗	75,600
MnasNet [31]	74.0	91.8	76ms	✓	✗	✗	40,000
MnasNet (our impl.)	74.0	91.8	79ms	✓	✗	✗	40,000
Proxyless-G (mobile)	71.8	90.3	83ms	✗	✓	✓	200
Proxyless-G + LL	74.2	91.7	79ms	✓	✓	✓	200
Proxyless-R (mobile)	74.6	92.2	78ms	✓	✓	✓	200

Table 2: ProxylessNAS achieves state-of-the art accuracy (%) on ImageNet (under mobile latency constraint $\leq 80ms$) with $200\times$ less search cost in GPU hours. “LL” indicates latency regularization loss. Details of MnasNet’s search cost are provided in appendix C.

Face recognition: Arc face



Convolution GNN



Варианты задач:

- Node-level - регрессия и классификация узлов
- Edge-level - предсказание наличия связи по двум узлам
- Graph-level - классификация графов

Варианты обучения:

- Semi-supervised для Node-level: По размеченным узлам восстановить метки остальных
- Supervised для Graph-level: ConvGNN + pooling + readout для классификации графов
- Unsupervised для Embedding: Обучение GAE для эмбедингов, по которым можно восстановить *матрицу смежности*