



СТАРТАПУ ИИ:

Классификация и детекция изображений

Обсуждаем: <https://t.me/devdvAI>, <https://t.me/devdvStartup>

Репозиторий: <https://github.com/akumidv/startup-khv-ai-study>

НАУКООБРАЗНО: МЕТОДЫ DATA MINING

ПРЕДСКАЗАНИЯ

PREDICTIVE

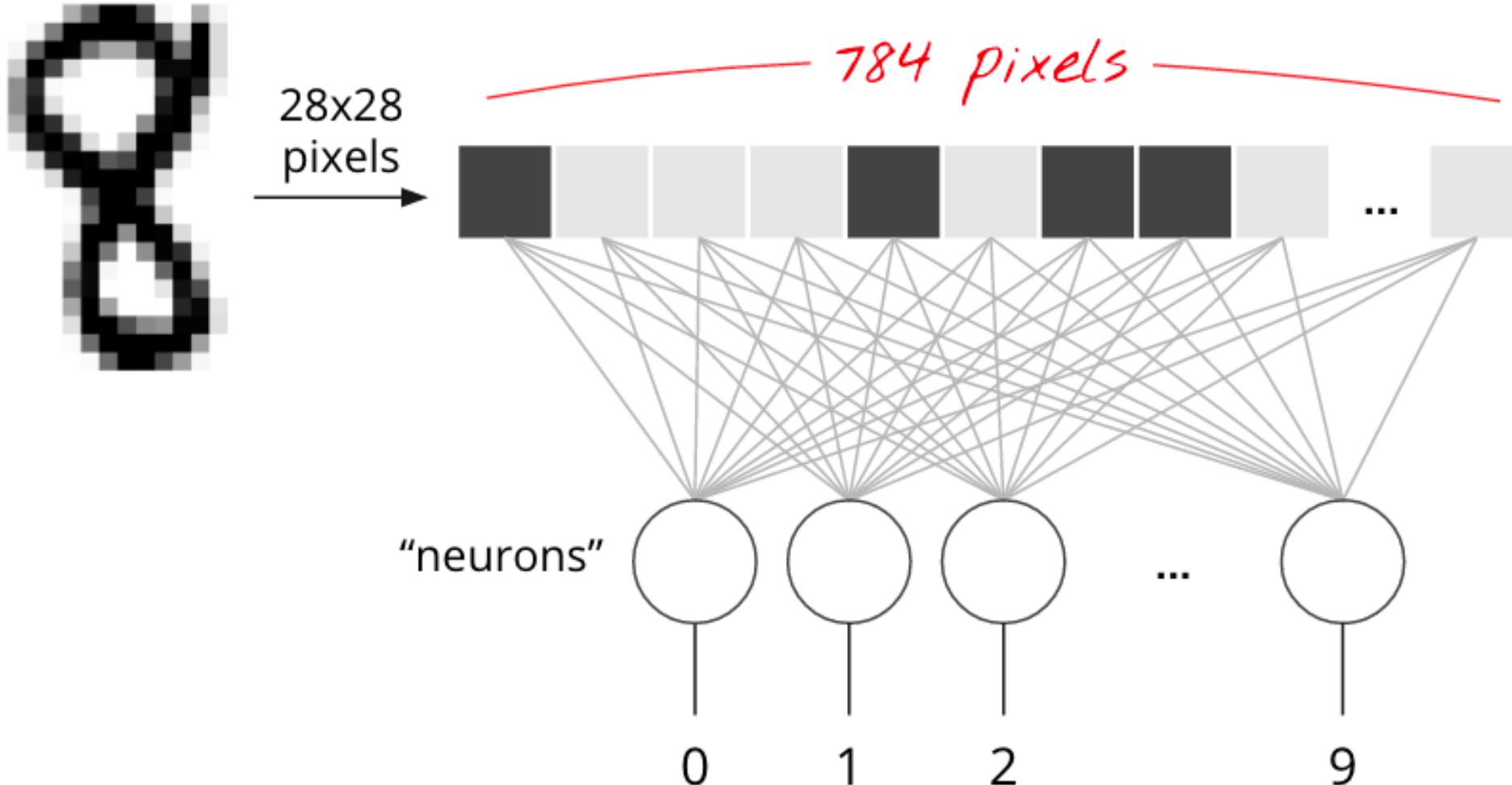
- КЛАССИФИКАЦИЯ
- БУДУЮЩАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
- ВОССТАНОВЛЕНИЕ УТРАЧЕННЫХ ДАННЫХ (ГЕНЕРАЦИЯ)
- ОПТИМИЗАЦИЯ УПРАВЛЕНИЯ ПО МОДЕЛИ

ОПИСАНИЕ

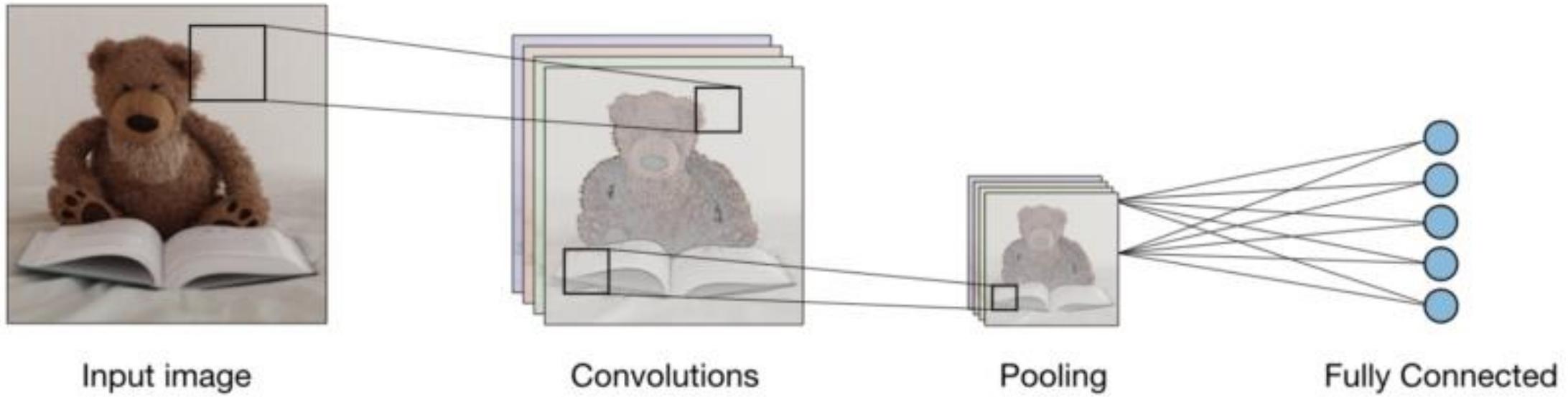
DESCRIPTIVE

- КЛАСТЕРИЗАЦИЯ
- РЕГРЕССИЯ (ЗАВИСИМЫЙ ПАРАМЕТР)
- АНОМАЛИИ (ИСКЛЮЧЕНИЯ)
- ИТОГ (СУММИРОВАНИЕ) – УНИКАЛЬНЫЕ ПРИЗНАКИ
- АССОЦИАЦИЯ – ЗАКОНОМЕРНОСТИ СВЯЗАННЫХ ДАННЫХ

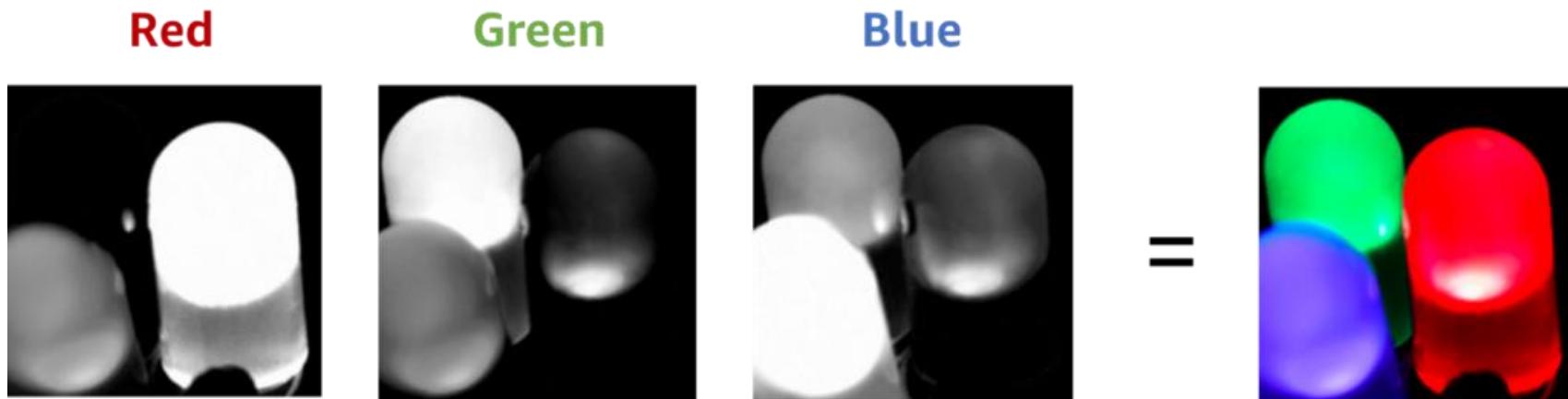
ВЕКТОР ИЗОБРАЖЕНИЯ



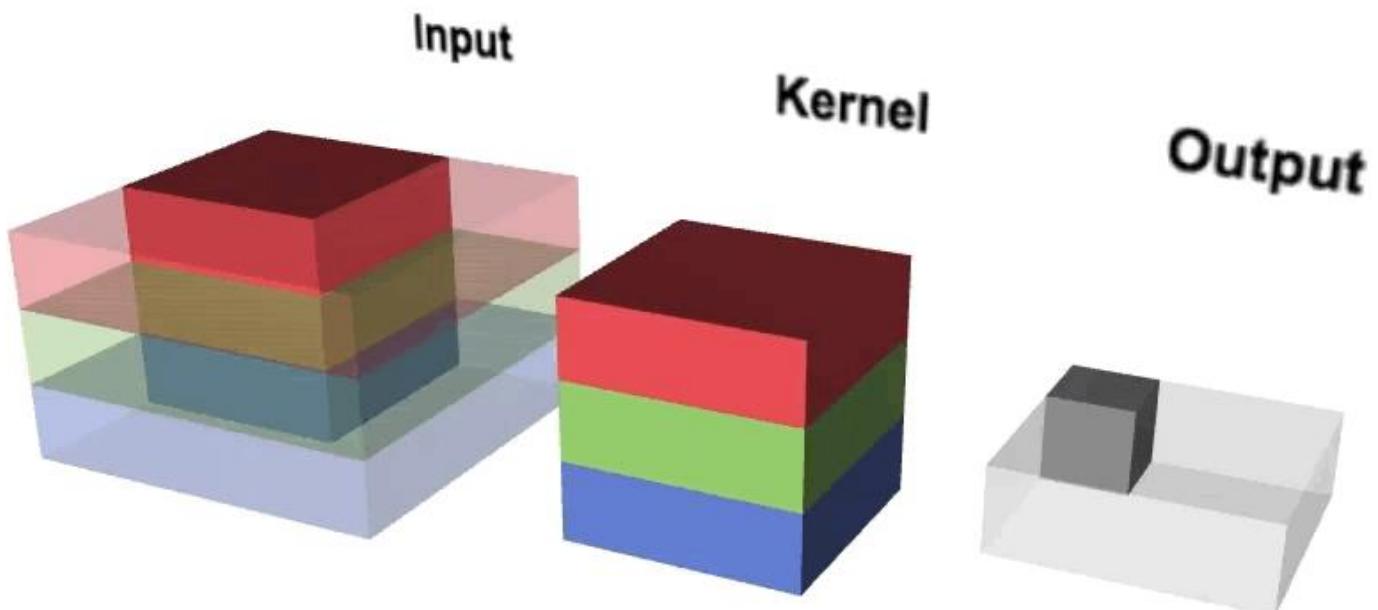
ТИПИЧНАЯ СХЕМА СЕТИ ДЛЯ ИЗОБРАЖЕНИЙ



СВЕРТОЧНАЯ СЕТЬ CONV2D



ФИЛЬТРАЦИЯ
ИЗОБРАЖЕНИЯ



COND2D В EXCEL

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																
2	Input				Kernel				Output							
3																
4	1	3	2	1		1	2	3								
5	1	3	3	1		0	1	0								
6	2	1	1	3		2	1	2								
7	3	2	3	3												
8																

$$(1*1)+(3*2)+(2*3)+(1*0)+(3*1) + (3*0)+(2*2)+(1*1)+(1*2) = 23$$

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																
2	Input				Kernel				Output							
3																
4	1	3	2	1	1	3	3	1	23	22						
5	1	3	3	1	0	1	0									
6	2	1	1	3	2	1	2									
7	3	2	3	3	2	3	3									
8																

$f(x) = C5 * \$I\$4 + D5 * \$J\$4 + E5 * \$K\$4 + C6 * \$I\$5 + D6 * \$J\$5 + E6 * \$K\$5 + C7 * \$I\$6 + D7 * \$J\$6 + E7 * \$K\6

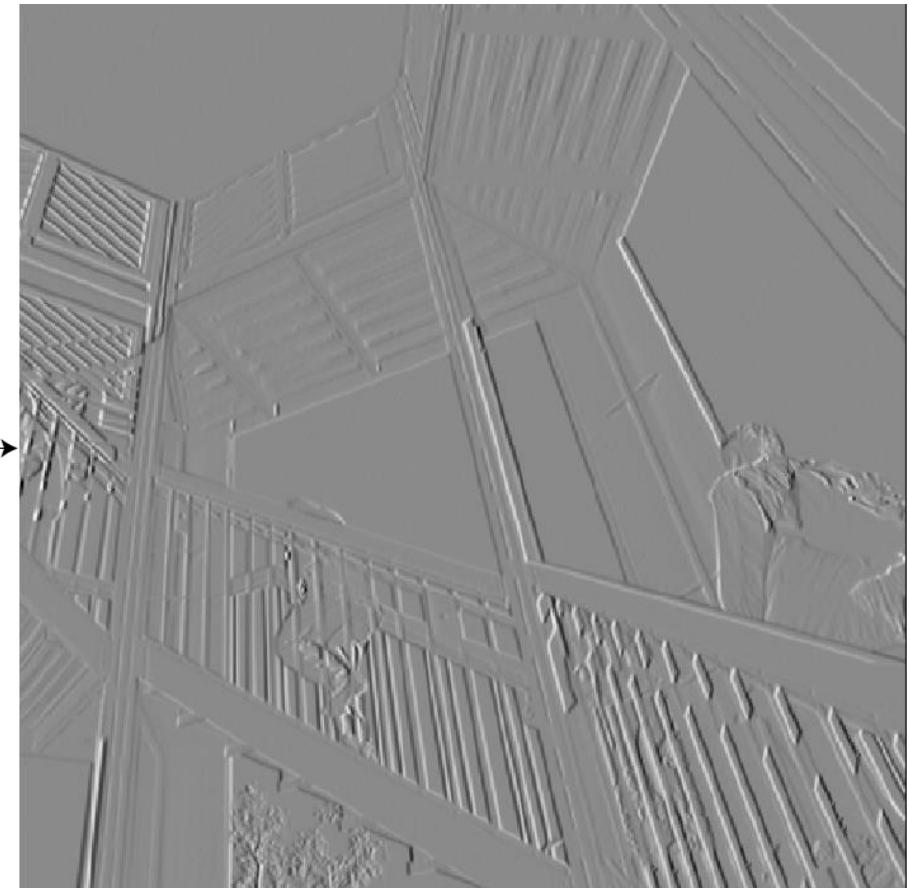
ГДЕ БЕРЕМ ЗНАЧЕНИЯ ЯДРА (KERNEL)?

ПРИМЕР ЗАДАННОГО ФИЛЬТРА СОБЕЛЯ



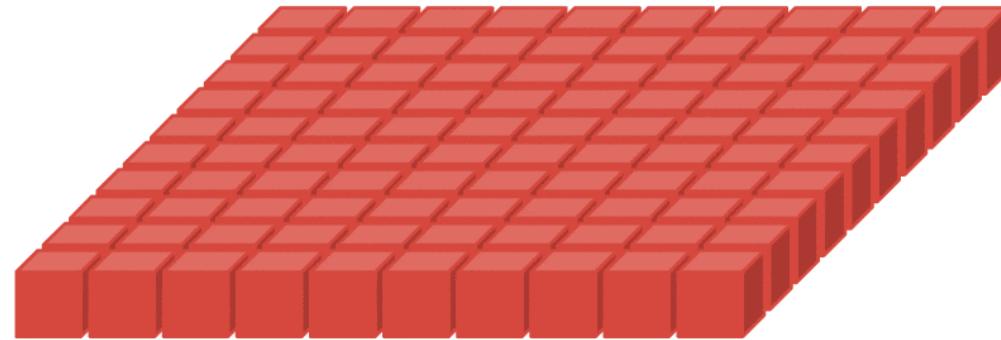
$$\begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}$$

Horizontal Sobel kernel



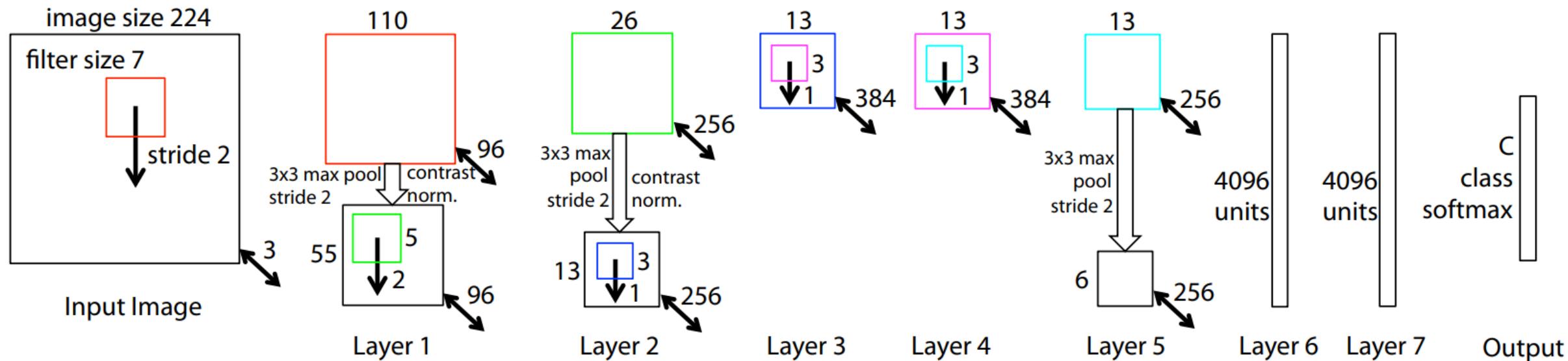
MAX POOLING – СУММИРОВАНИЕ МАКСИМУМОВ

УПЛОТНЯЕМ ЗНАЧЕНИЯ ПРЕДЫДУЩЕГО СЛОЯ

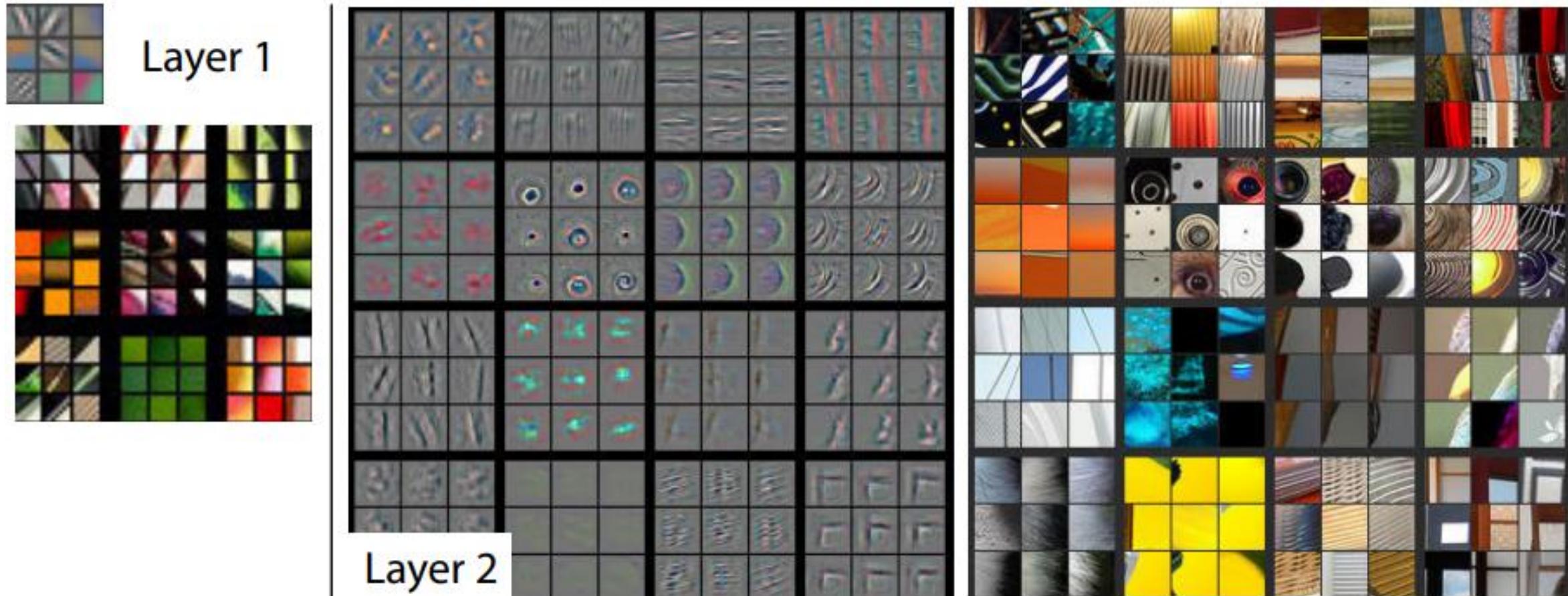


ВИЗУАЛЬНОЕ ПОНИМАНИЕ СВЕРТОЧНОЙ СЕТИ

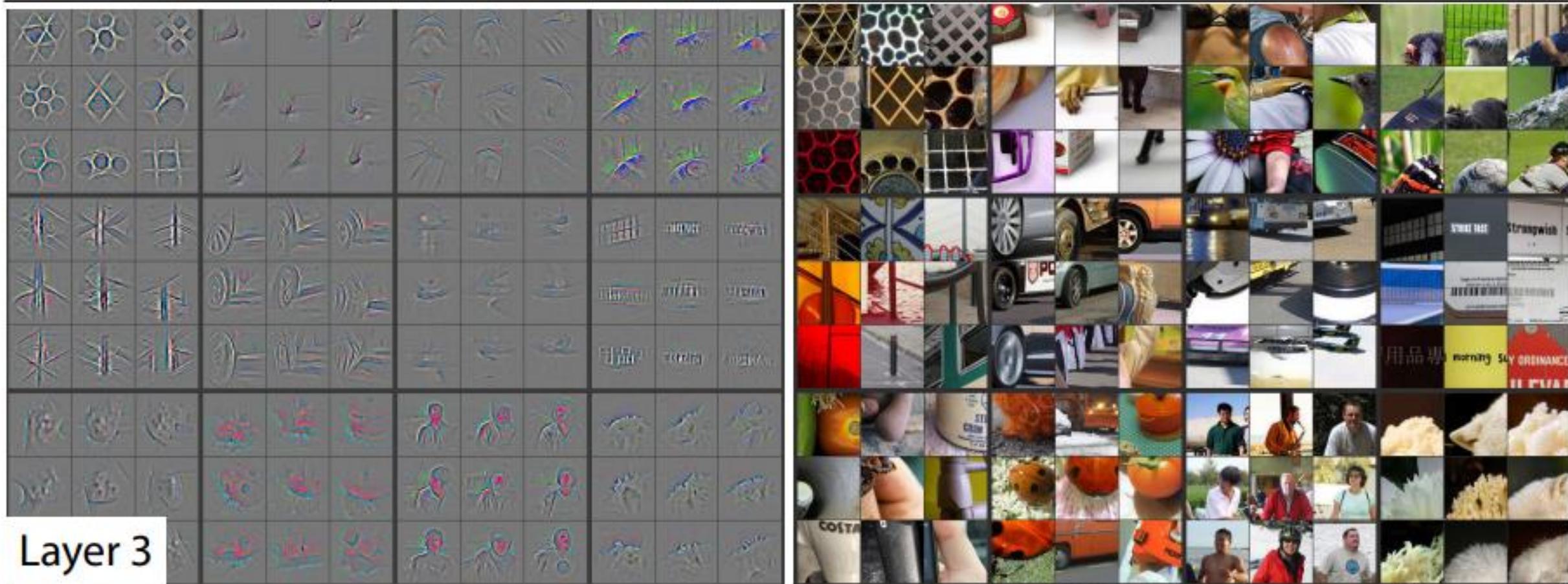
IMAGENET



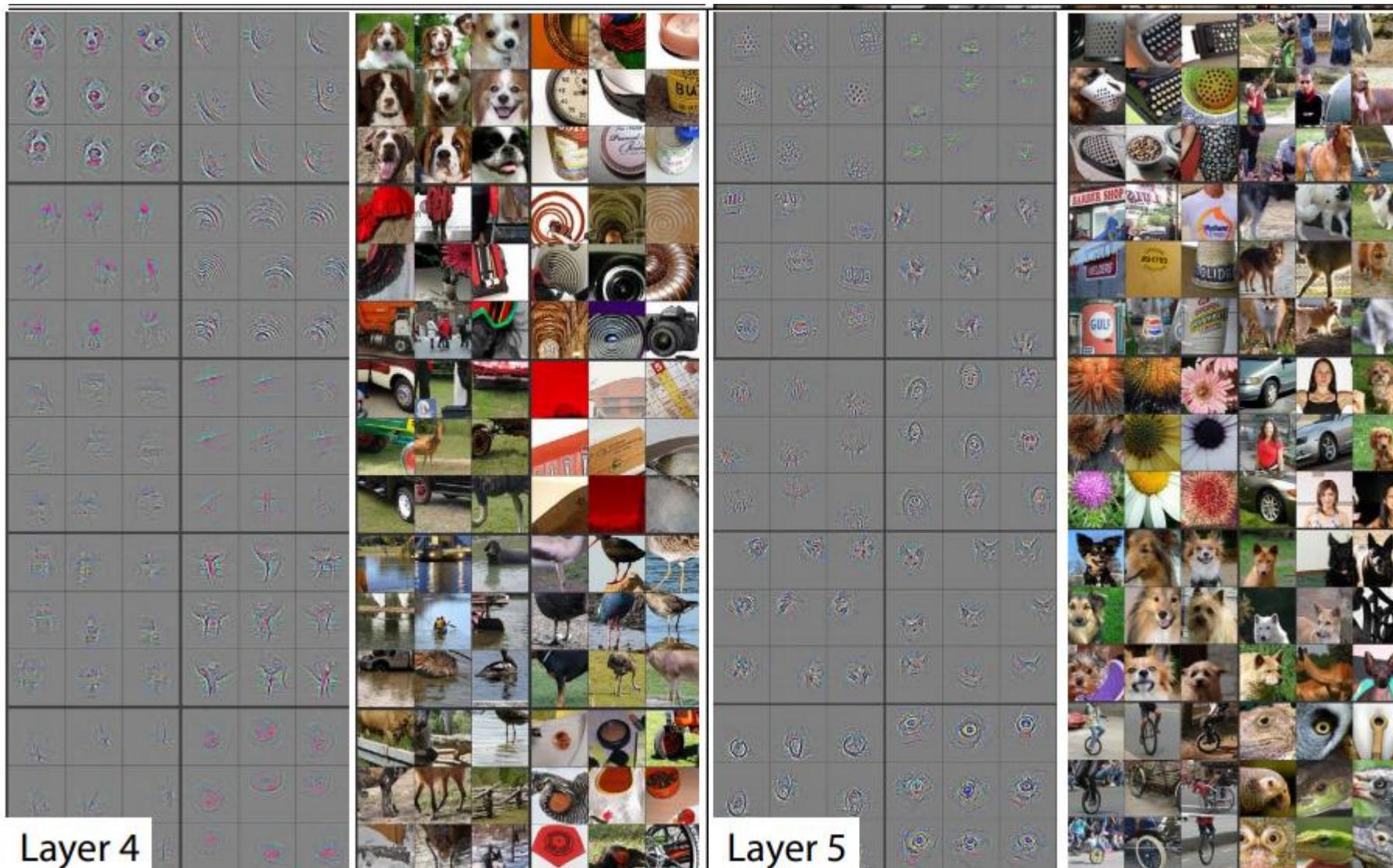
ВИЗУАЛЬНОЕ ПОНИМАНИЕ СВЕРТОЧНОЙ СЕТИ



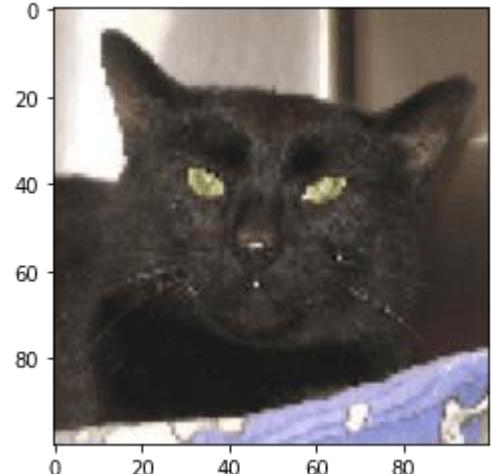
ВИЗУАЛЬНОЕ ПОНИМАНИЕ СВЕРТОЧНОЙ СЕТИ



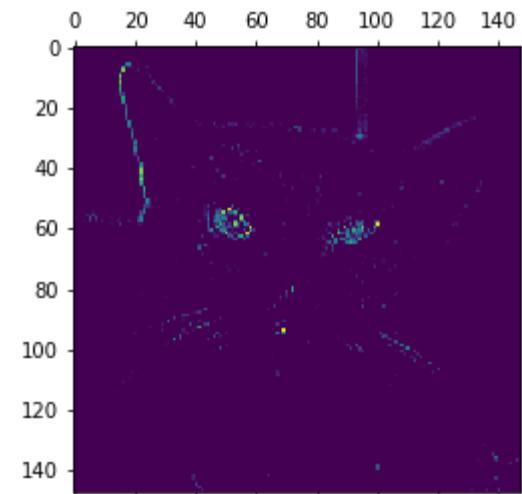
ВИЗУАЛЬНОЕ ПОНИМАНИЕ СВЕРТОЧНОЙ СЕТИ



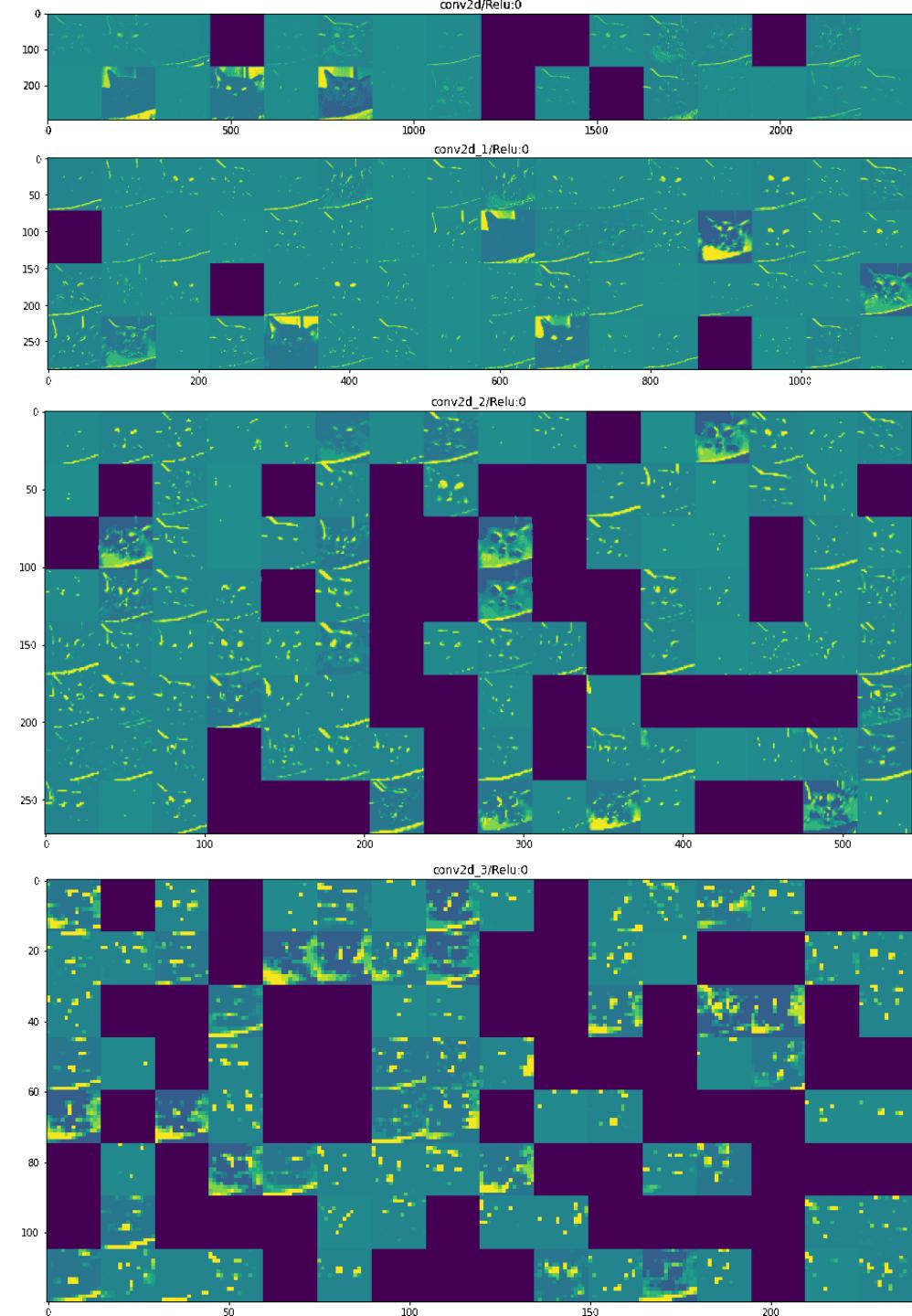
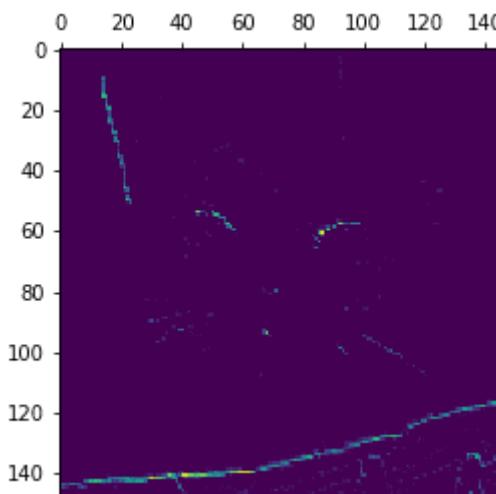
НА КОШКЕ



1^Й СЛОЙ, 1^Й ЭЛЕМЕНТ



1^Й СЛОЙ, ЗОЙ ЭЛЕМЕНТ



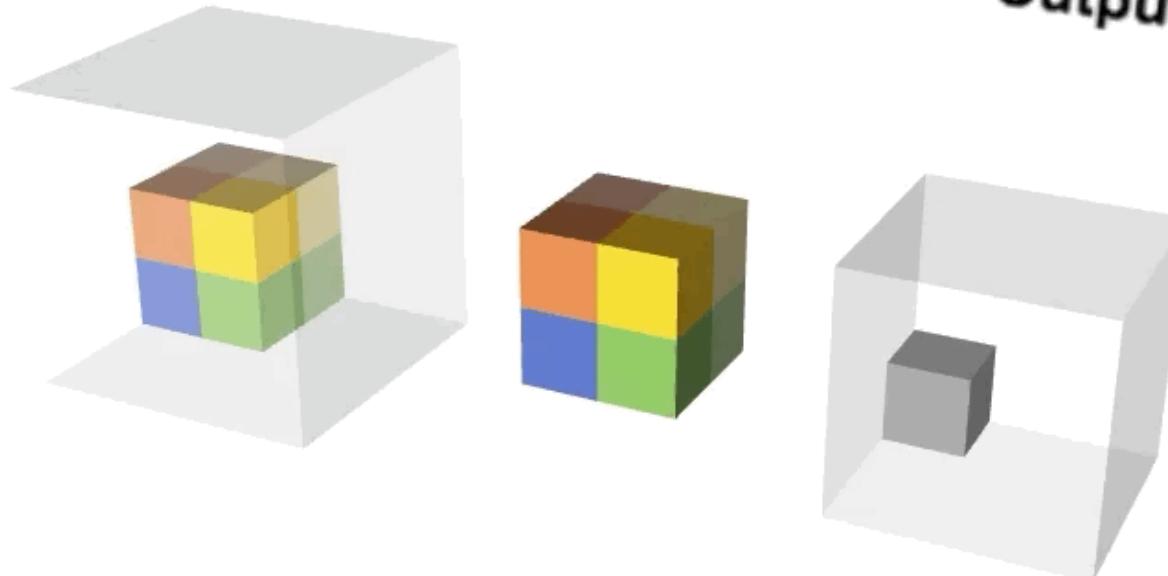
CONV3D – ПОИСК ПАТТЕРНОВ 3D ИЛИ +ВРЕМЯ



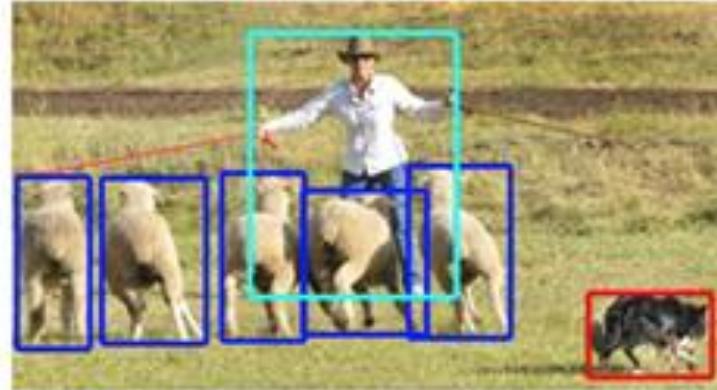
Input

Kernel

Output



ИЗОБРАЖЕНИЯ



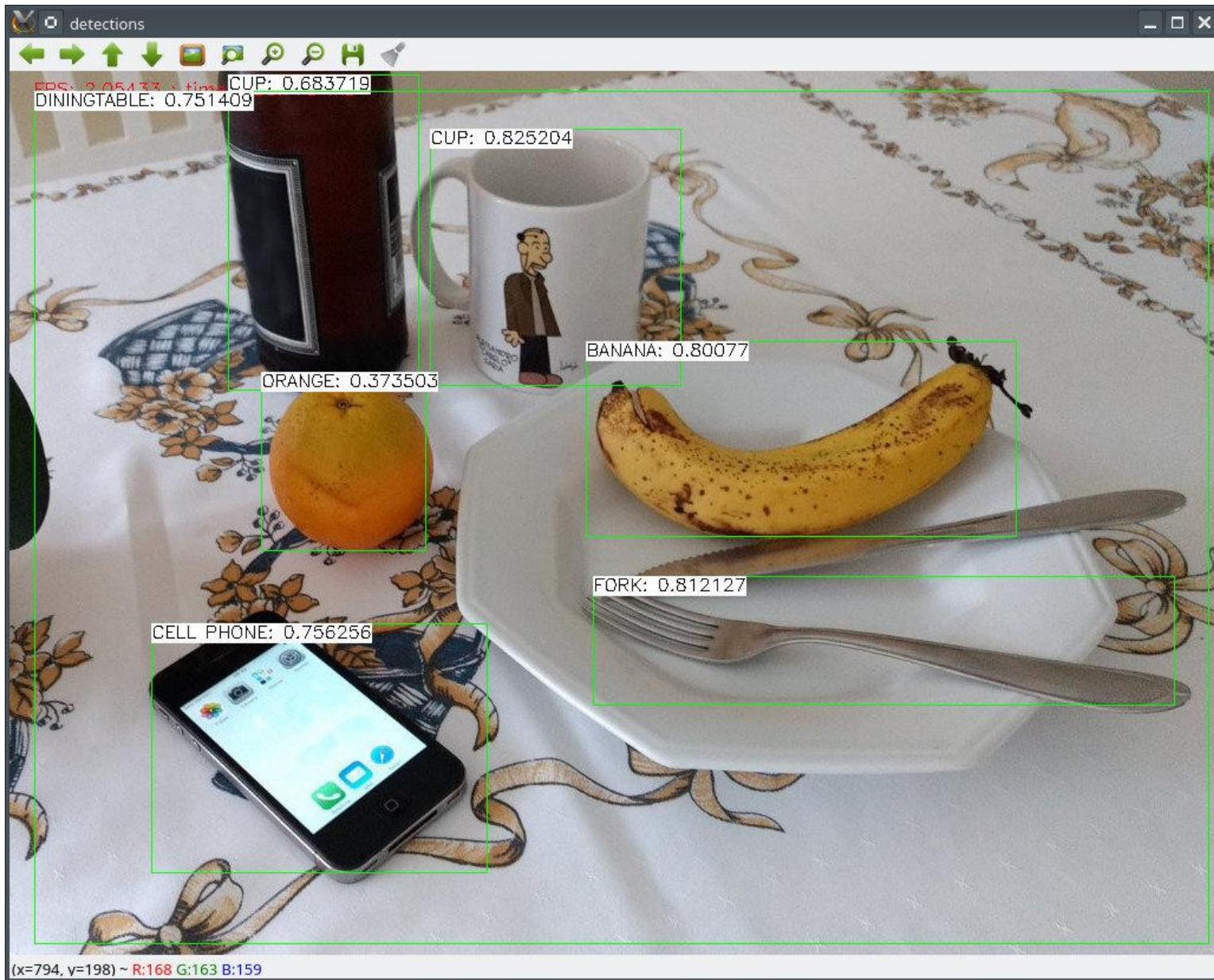
- **ДЕТЕКЦИЯ БРАКА НА КОНВЕЙЕРЕ ИЛИ ПОЛОТНЕ**
Факт брака, уровень допустимости
- **ПОДСЧЕТ ЭЛЕМЕНТОВ И НАХОЖДЕНИЯ В ЗОНЕ**
Количество выпускаемой продукции, объема и количества материалов, факта и времени нахождения на территории (например уборка)

- **ОПРЕДЕЛЕНИЕ КОМПЛЕКТНОСТИ**
Наличие защитных элементов спецодежды, всех элементов объекта
- **КЛАССИФИКАЦИЯ ГОТОВНОСТИ/ЗРЕЛОСТИ**
Сельхозпродукция, здоровье животных

KEYPOINTS

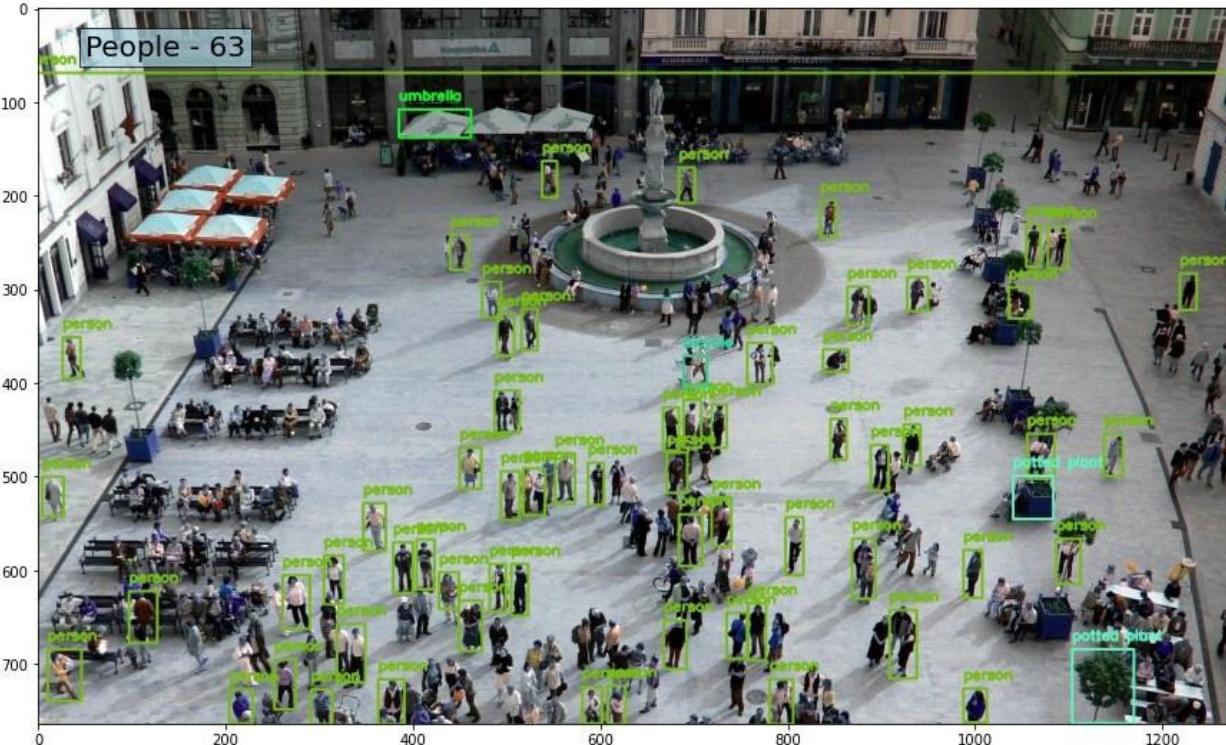


ПРИМЕР ДЕТЕКЦИИ ИЗОБРАЖЕНИЙ

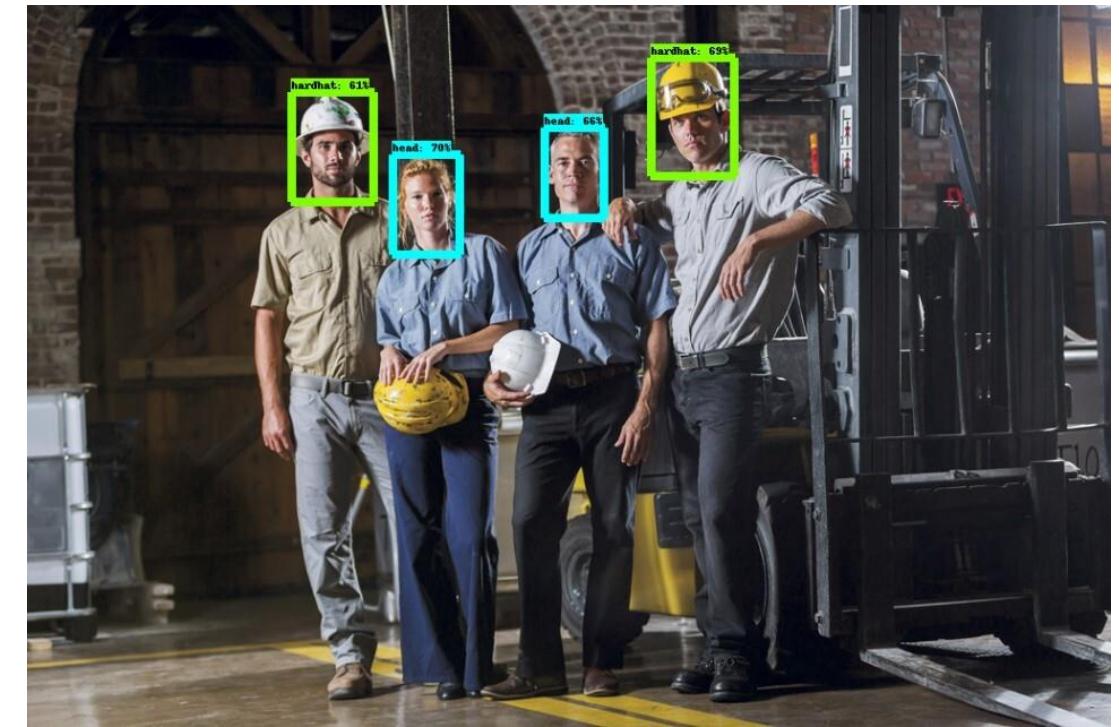


ПРИМЕРЫ ЗАДАЧ

ДЕТЕКЦИЯ И ПОДСЧЁТ



ДЕТЕКЦИЯ И КЛАССИФИКАЦИЯ



ДЕТЕКЦИЯ И РЕГРЕССИЯ ПАРАМЕТРОВ СТЫКОВ

06.02.19 – 27 ВАГОНОВ С УГЛЕМ



09.03.21 – 11 ЦИСТЕРН С НЕФТЬЮ



Defect Detector

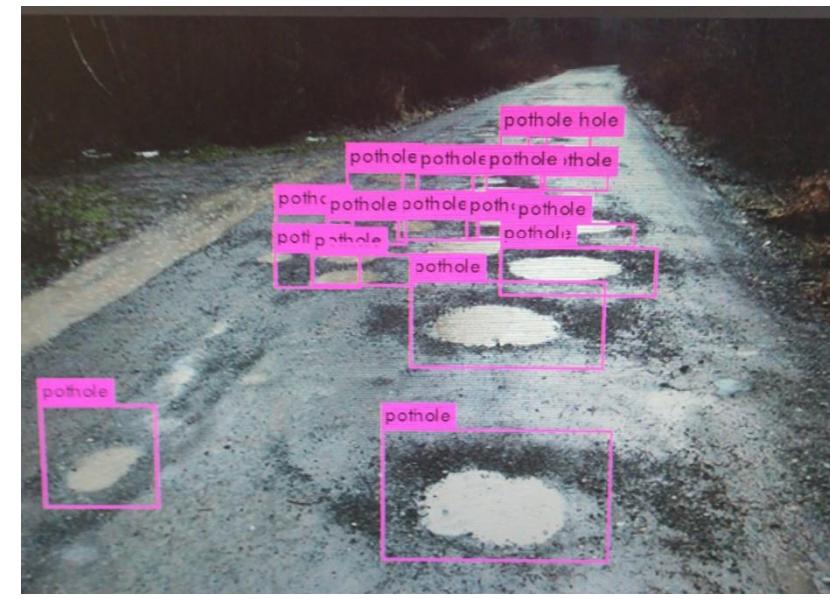
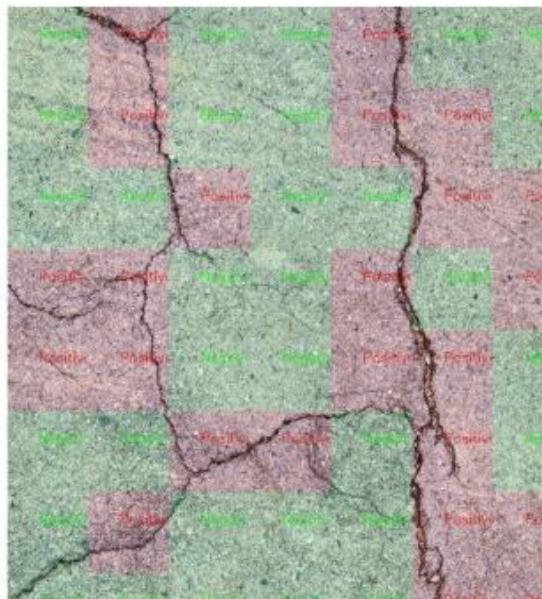
Open video... Open CSV...

Мин. зазор, мм 0 обе нити

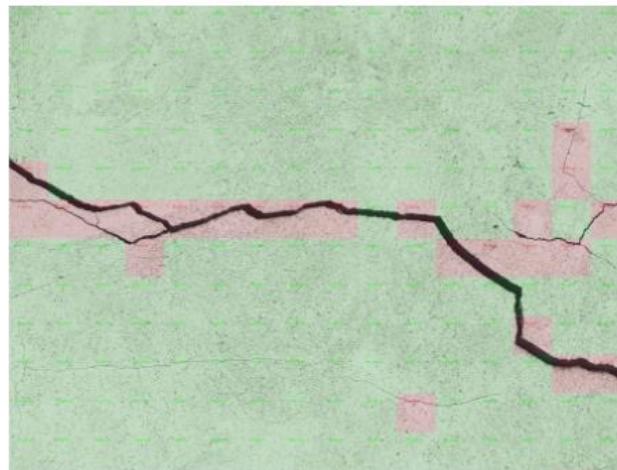
(Л) 200 км 53 м: зазор 13 мм (кадр 86)
(Л) 200 км 78 м: зазор 20 мм (кадр 156)
(Л) 200 км 103 м: зазор 22 мм (кадр 225)
(Л) 200 км 128 м: зазор 14 мм (кадр 294)
(Л) 200 км 153 м: зазор 28 мм (кадр 364)
(Л) 200 км 178 м: зазор 26 мм (кадр 433)
(Л) 200 км 203 м: зазор 23 мм (кадр 503)
(Л) 200 км 228 м: зазор 26 мм (кадр 573)
(Л) 200 км 253 м: зазор 28 мм (кадр 642)
(Л) 200 км 278 м: зазор 26 мм (кадр 712)
(Л) 200 км 304 м: зазор 20 мм (кадр 782)
(Л) 200 км 329 м: зазор 20 мм (кадр 851)
(Л) 200 км 354 м: зазор 14 мм (кадр 921)
(Л) 200 км 379 м: зазор 24 мм (кадр 991)
(Л) 200 км 404 м: зазор 24 мм (кадр 1060)
(Л) 200 км 429 м: зазор 25 мм (кадр 1130)
(Л) 200 км 454 м: зазор 7 мм (кадр 1199)
(Л) 200 км 479 м: зазор 15 мм (кадр 1269)
(Л) 200 км 504 м: зазор 21 мм (кадр 1339)
(Л) 200 км 529 м: зазор 21 мм (кадр 1408)
(Л) 200 км 554 м: зазор 27 мм (кадр 1478)
(Л) 200 км 579 м: зазор 23 мм (кадр 1548)
(Л) 200 км 604 м: зазор 25 мм (кадр 1617)
(Л) 200 км 629 м: зазор 44 мм (кадр 1686)
(Л) 200 км 654 м: зазор 17 мм (кадр 1756)
(Л) 200 км 680 м: зазор 22 мм (кадр 1826)
(Л) 200 км 705 м: зазор 18 мм (кадр 1896)
(Л) 200 км 730 м: зазор 8 мм (кадр 1966)
(Л) 200 км 755 м: зазор 16 мм (кадр 2035)
(Л) 200 км 780 м: зазор 7 мм (кадр 2104)
(Л) 200 км 805 м: зазор 6 мм (кадр 2174)
(Л) 200 км 830 м: зазор 19 мм (кадр 2244)
(Л) 200 км 855 м: зазор 14 мм (кадр 2313)
(Л) 200 км 880 м: зазор 25 мм (кадр 2383)
(Л) 200 км 905 м: зазор 7 мм (кадр 2453)
(Л) 200 км 930 м: зазор 12 мм (кадр 2522)
(Л) 200 км 955 м: зазор 7 мм (кадр 2592)
(Л) 200 км 980 м: зазор 7 мм (кадр 2661)
(Л) 201 км 2 м: зазор 5 мм (кадр 2731)
(Л) 201 км 27 м: зазор 7 мм (кадр 2801)
(Л) 201 км 64 м: зазор 11 мм (кадр 2905)

ДЕТЕКЦИЯ И СЕГМЕНТНАЯ КЛАССИФИКАЦИЯ

АСФАЛЬТ И
ГРАВИЙ



БЕТОН



СЕГМЕНТАЦИЯ



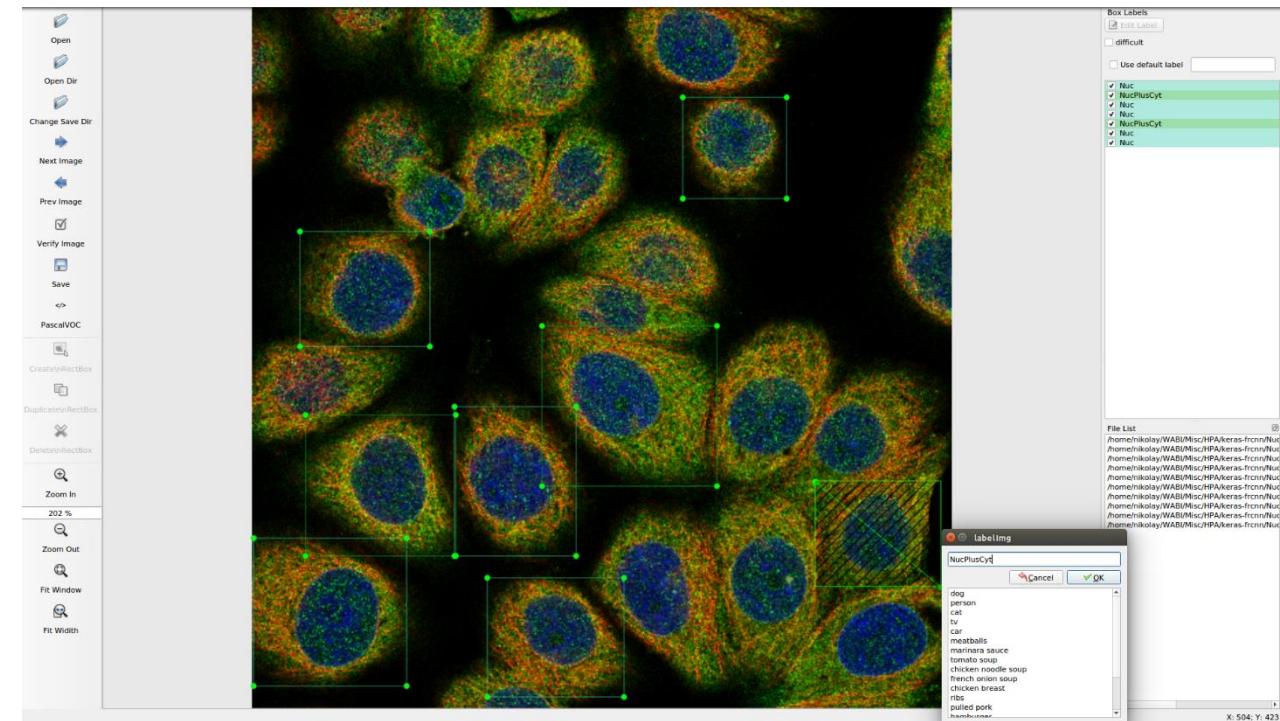
ПРОЦЕДУРА СОЗДАНИЯ СИСТЕМЫ ДЕТЕКЦИИ

СБОР ДАННЫХ

ОБРАБОТКА И РАЗМЕТКА ДАННЫХ

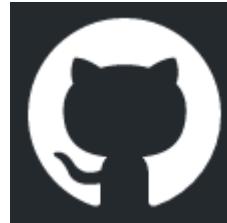
СОЗДАНИЕ И ОБУЧЕНИЕ МОДЕЛИ

ПЕРЕНОС В ПРОИЗВОДСТВО И ИНТЕГРАЦИЯ



ИСТОЧНИКИ ДАТАСЕТОВ

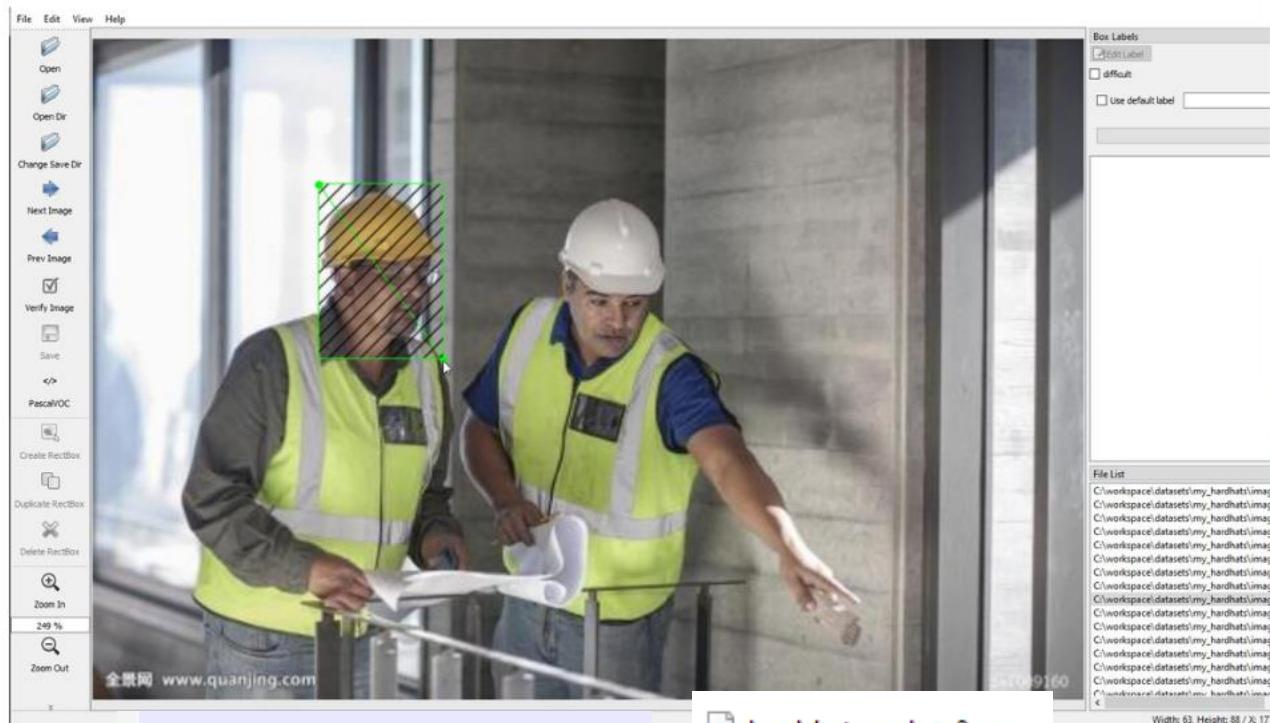
kaggle





- **Object segmentation**
- **Recognition in context**
- **Superpixel stuff segmentation**
- **330K images (>200K labeled)**
- **1.5 million object instances**
- **80 object categories**
- **91 stuff categories**
- **5 captions per image**
- **250,000 people with keypoints**

РАЗМЕТКА

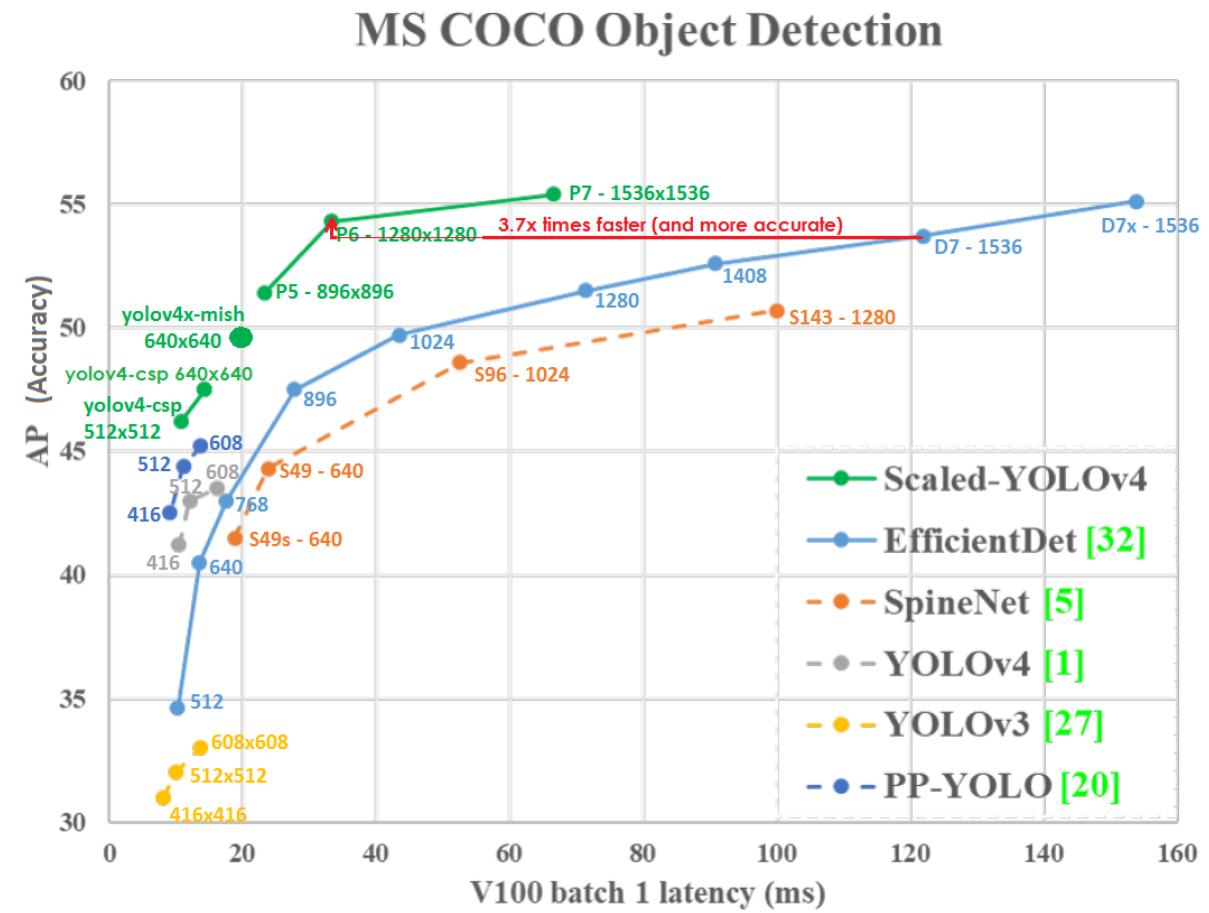
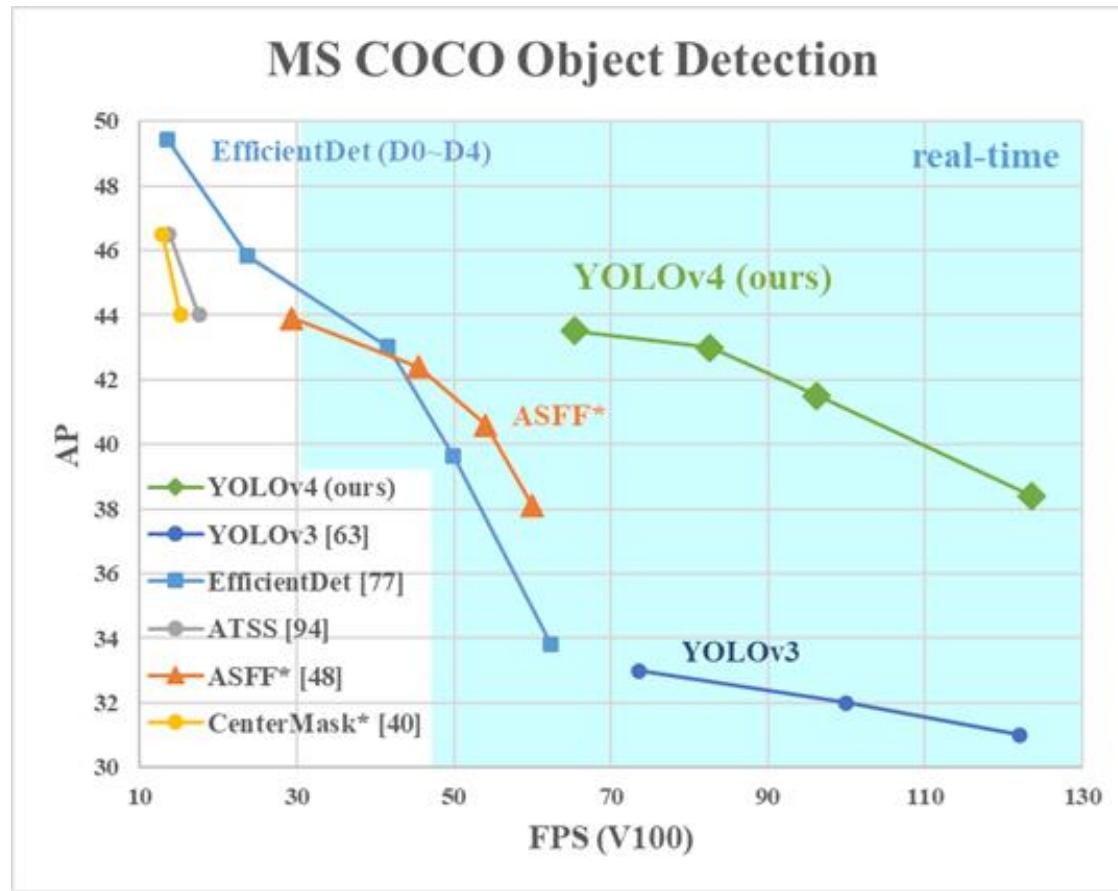


```
<annotation>
  <folder>images</folder>
  <filename>hard_hat_workers0.png</filename>
  <size>
    <width>416</width>
    <height>416</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>helmet</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <occluded>0</occluded>
    <difficult>0</difficult>
    <bndbox>
      <xmin>357</xmin>
      <ymin>116</ymin>
      <xmax>404</xmax>
      <ymax>191</ymax>
    </bndbox>
  </object>
</annotation>
```

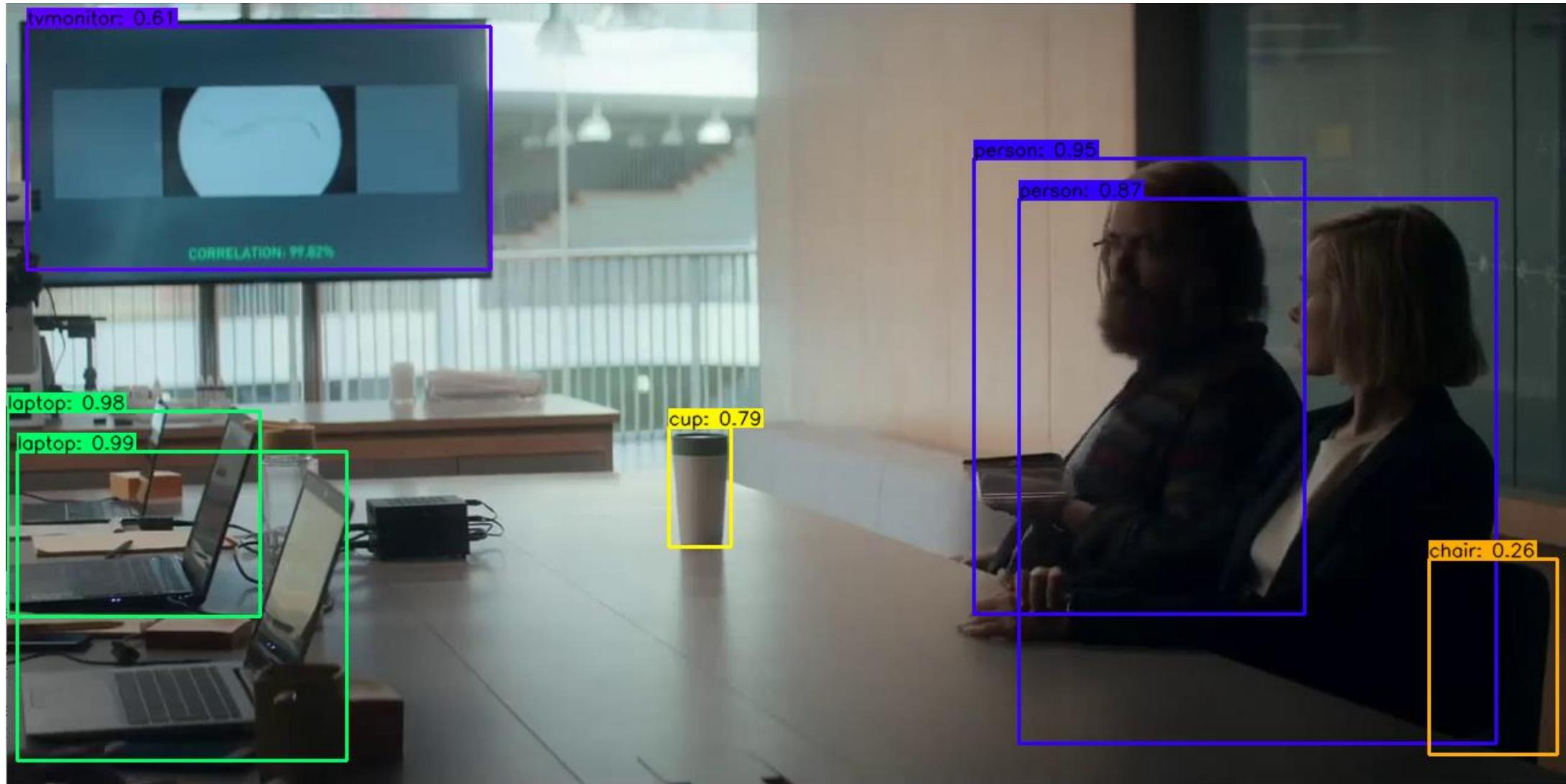


YOLO – АРХИТЕКТУРА CNN ДЛЯ ДЕТЕКЦИИ

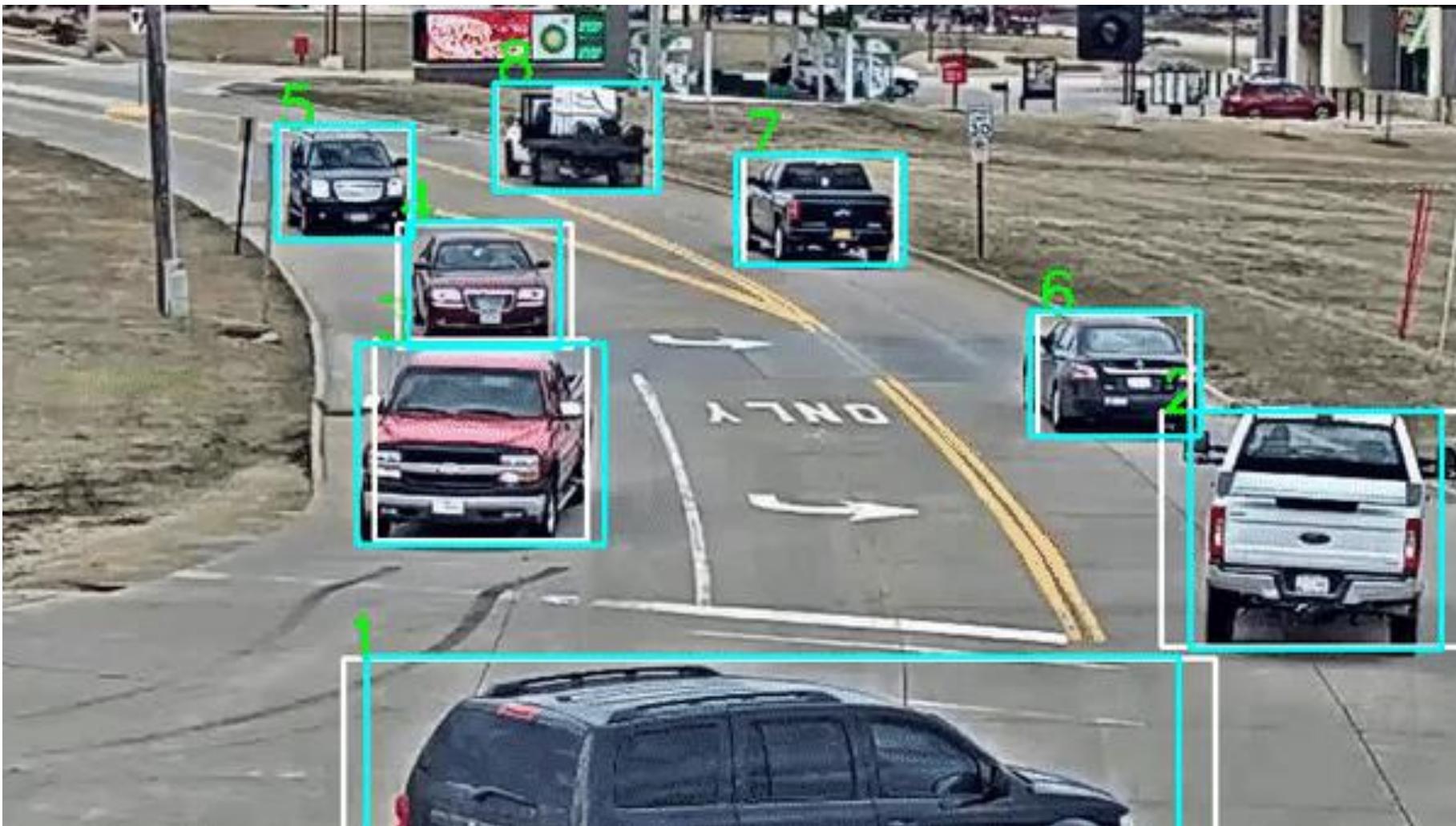
YOU ONLY LOOK ONCE



YOLO



ТРЕКИНГ

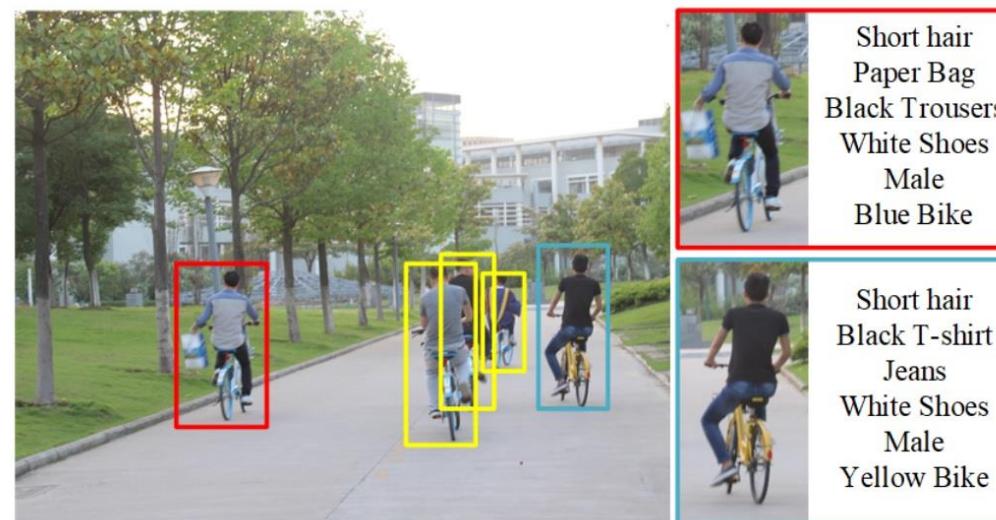


ОСНОВНЫЕ ПОДОХОДЫ ТРЕКИНГА

- Element ReIdentification

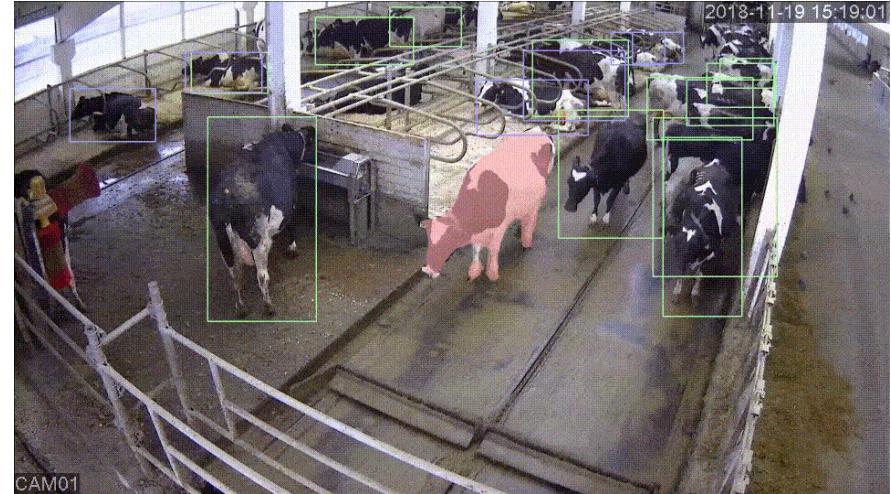


- По набору классификаций



ОСНОВНЫЕ ПОДОХОДЫ ТРЕКИНГА

- По уникальному объекту (лицу, рисунку)



- По близости в кадре ключевых точек (скелету или близость позы, центр масс, расчетной скорости и направления, размеру, координат проекции на землю)



КЛАССИЧЕСКАЯ МАТЕМАТИКА

Простой и популярный: модель движения + алгоритм целераспределения

SORT: Алгоритм целераспределения (венгерский) и фильтр Калмана

DeepSORT: + близость между ReID описаниям

ВОССТАНОВЛЕНИЕ IMAGE GPT: 64x64 IMAGES

ВОССТАНОВЛЕНИЕ ПОЛОВИНЫ ИЗОБРАЖЕНИЯ

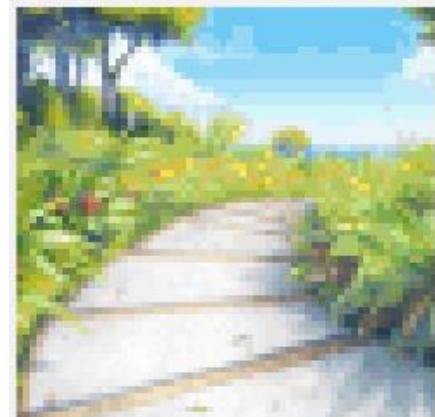
MODEL INPUT



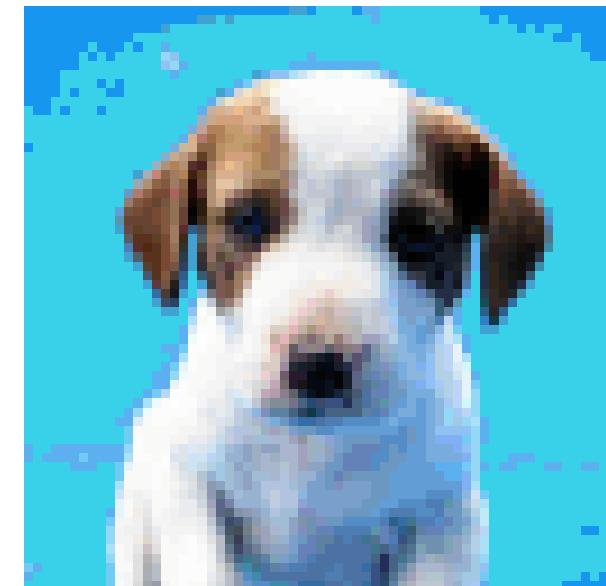
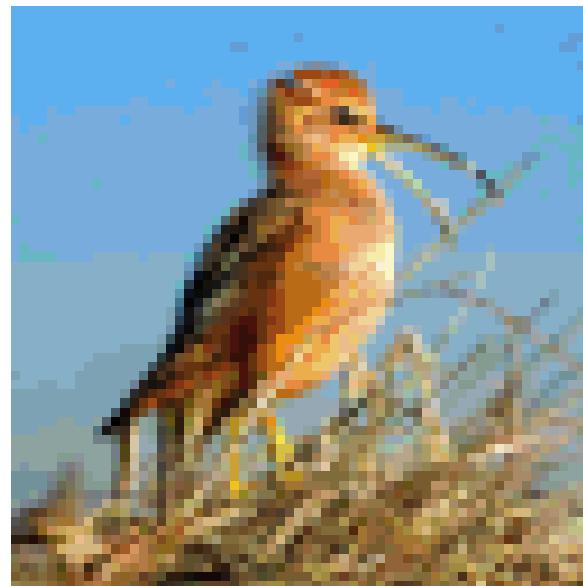
COMPLETIONS →



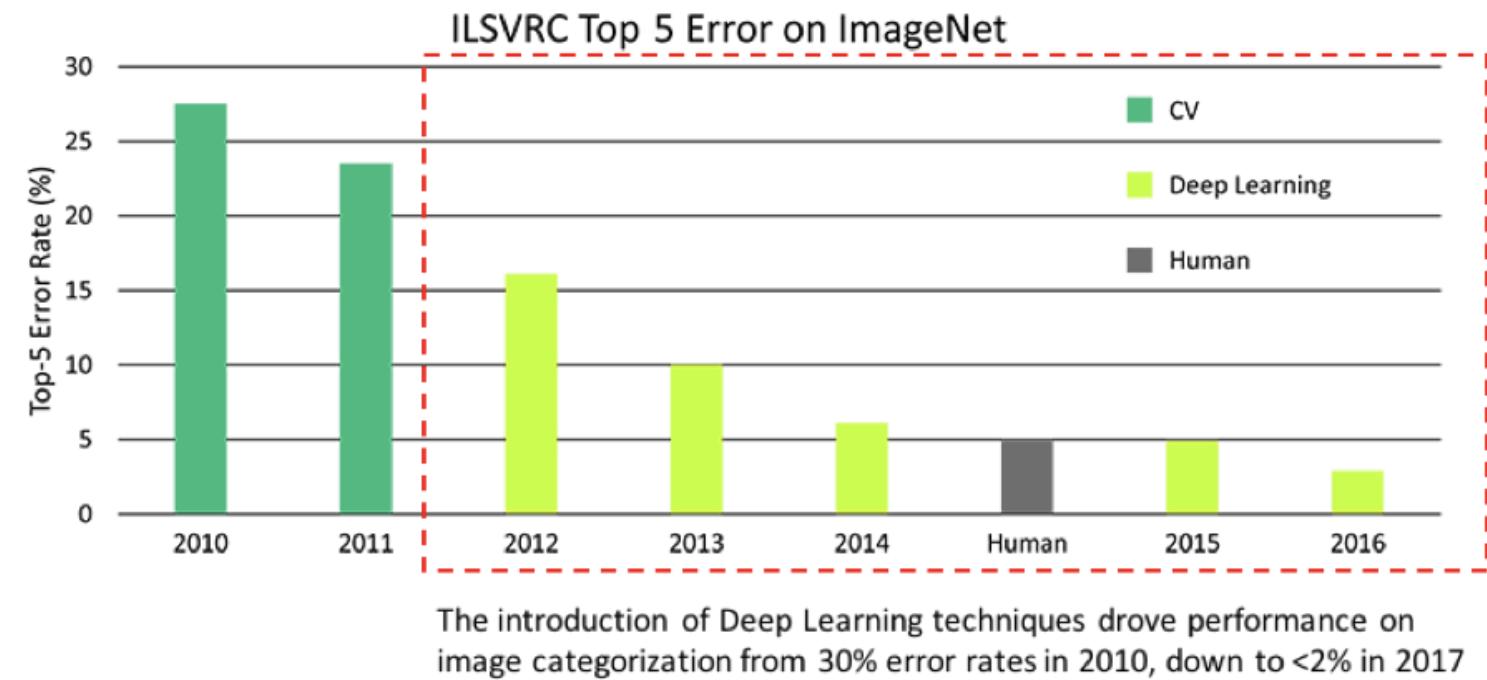
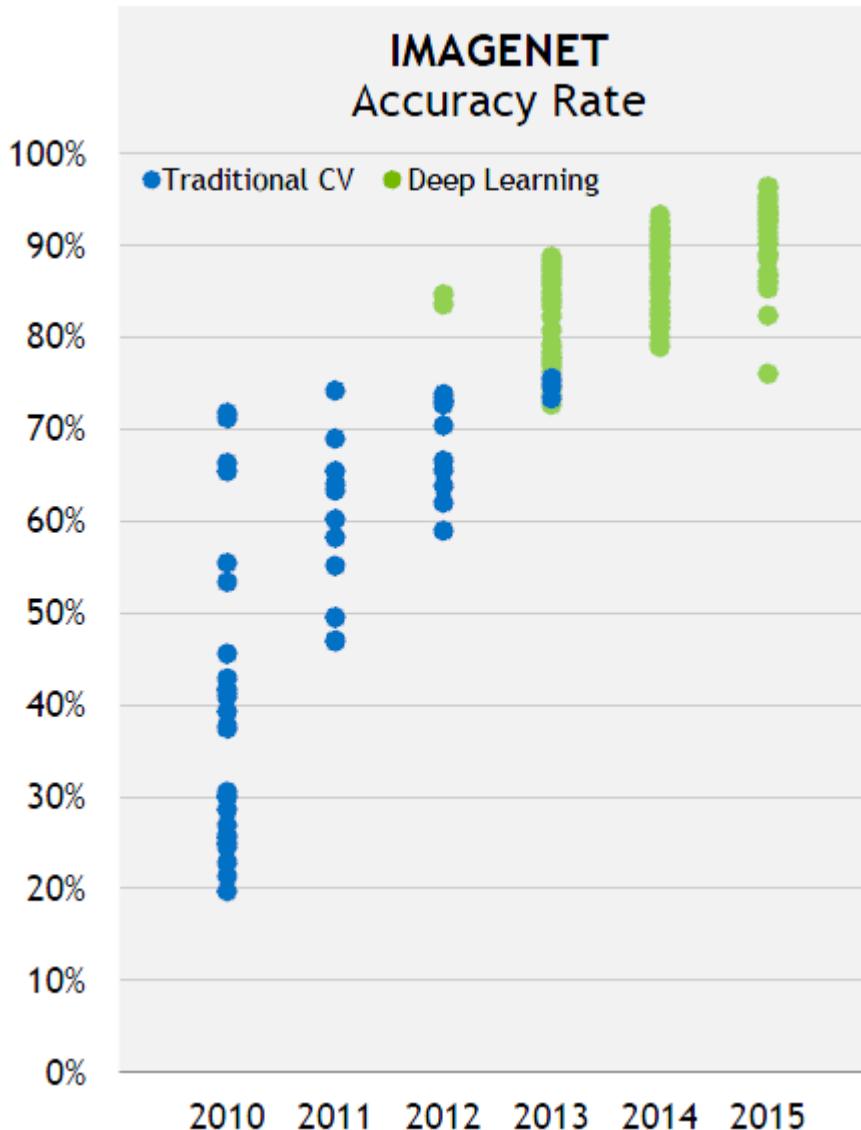
ORIGINAL



ГЕНЕРАЦИЯ ИЗОБРАЖЕНИЯ ПО ОПИСАНИЮ



АЛГОРИТМЫ: СОРЕВНОВАНИЯ ILSVRC



The introduction of Deep Learning techniques drove performance on image categorization from 30% error rates in 2010, down to <2% in 2017

СРАВНИТЕ

2020 Г.

NVIDIA 3090

36.84 TFLOPS

\$2 200



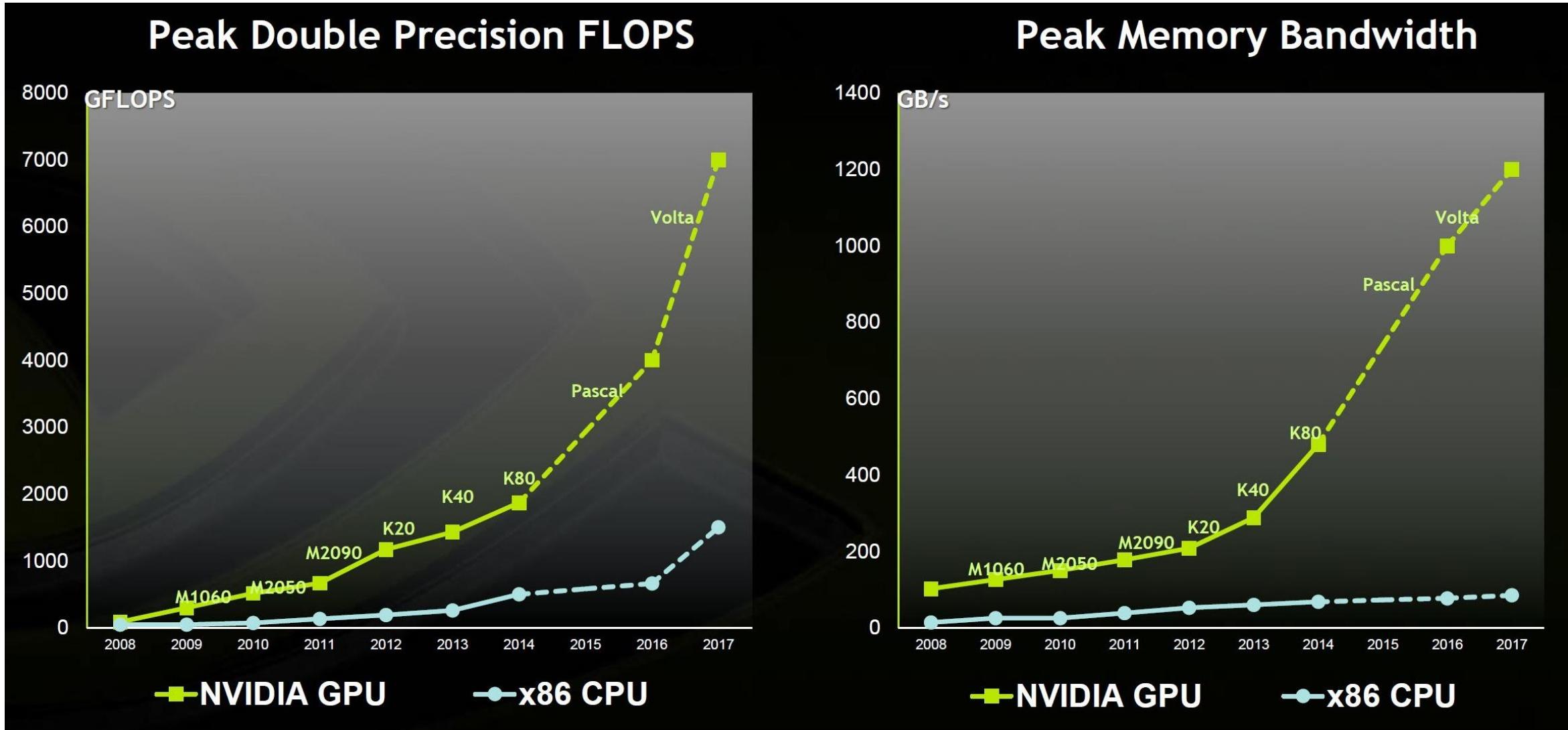
2002-2004 САМЫЙ МОЩНЫЙ
EARTH SIMULATOR (JAPAN)

35.86 TFLOPS

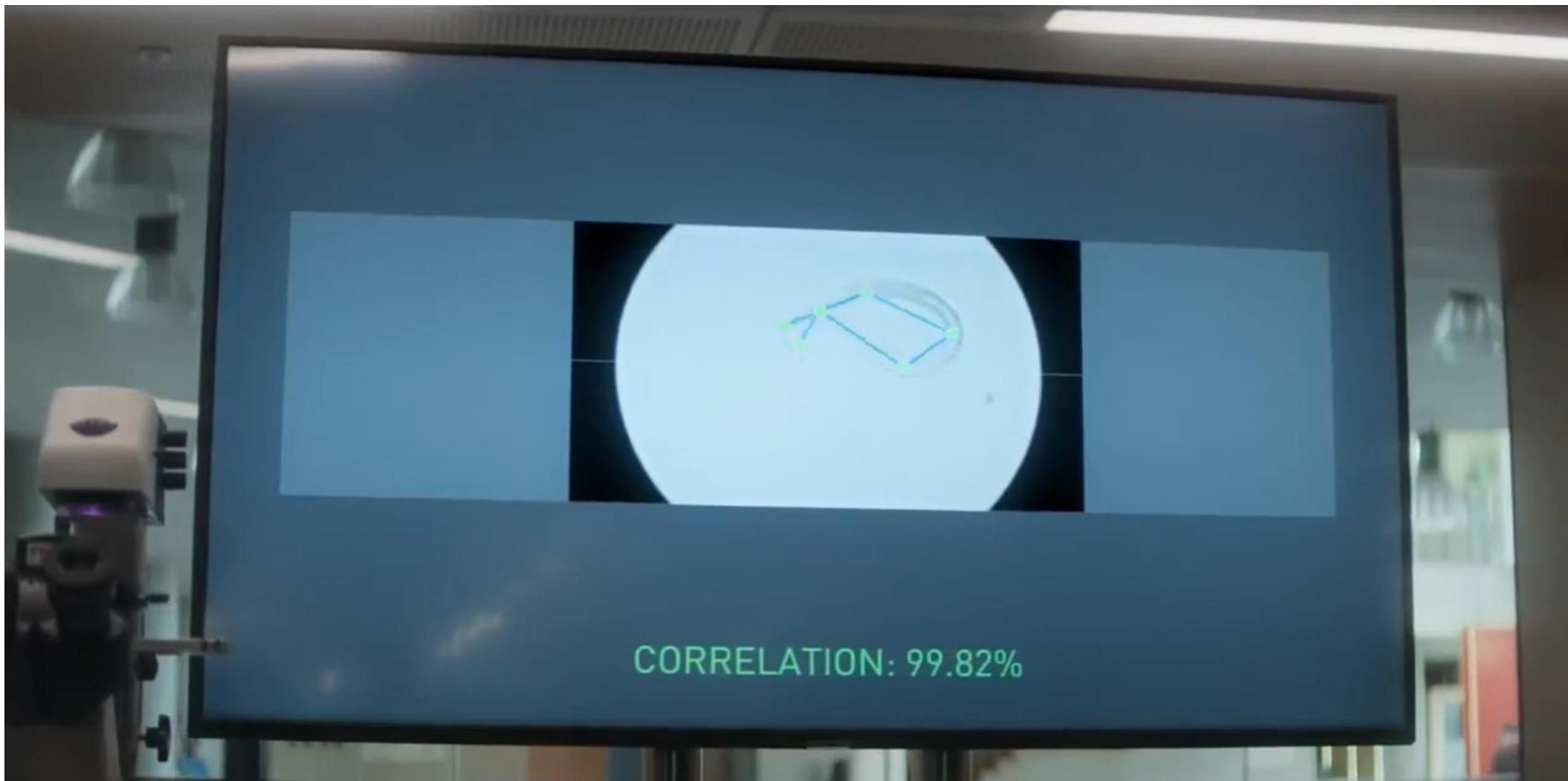
\$500 000 000



ПОЧЕМУ СТАЛО ВОЗМОЖНЫМ?



DEVS – УЖЕ ВОЗМОЖНО ПРЕДСКАЗАТЬ 10СЕК?



КОНТАКТЫ

ОБСУЖДАЕМ

<https://t.me/devdvAI>

<https://t.me/devdvStartup>



РЕПОЗИТОРИЙ

<https://github.com/akumidv/startup-khv-ai-study>

АНДРЕЙ КУМИНОВ

+7 914 770 5846

<https://facebook.com/akuminov>

<https://vk.com/akumidv>