

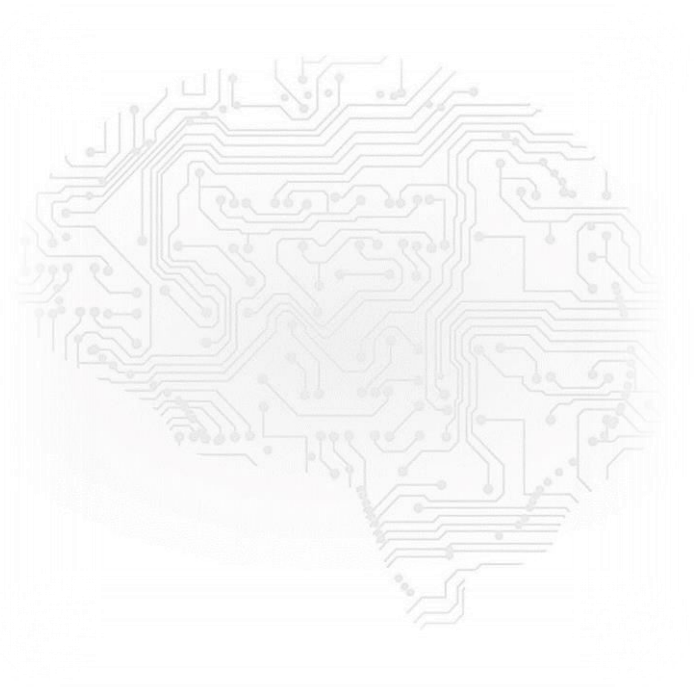
A faint background graphic of a human brain, where the internal structures are replaced by a complex network of circuit lines and nodes, symbolizing artificial intelligence or neural networks.

Встраиваемые системы машинного зрения нового поколения на основе глубоких нейронных сетей в реальных проектах для задач промышленности и безопасности

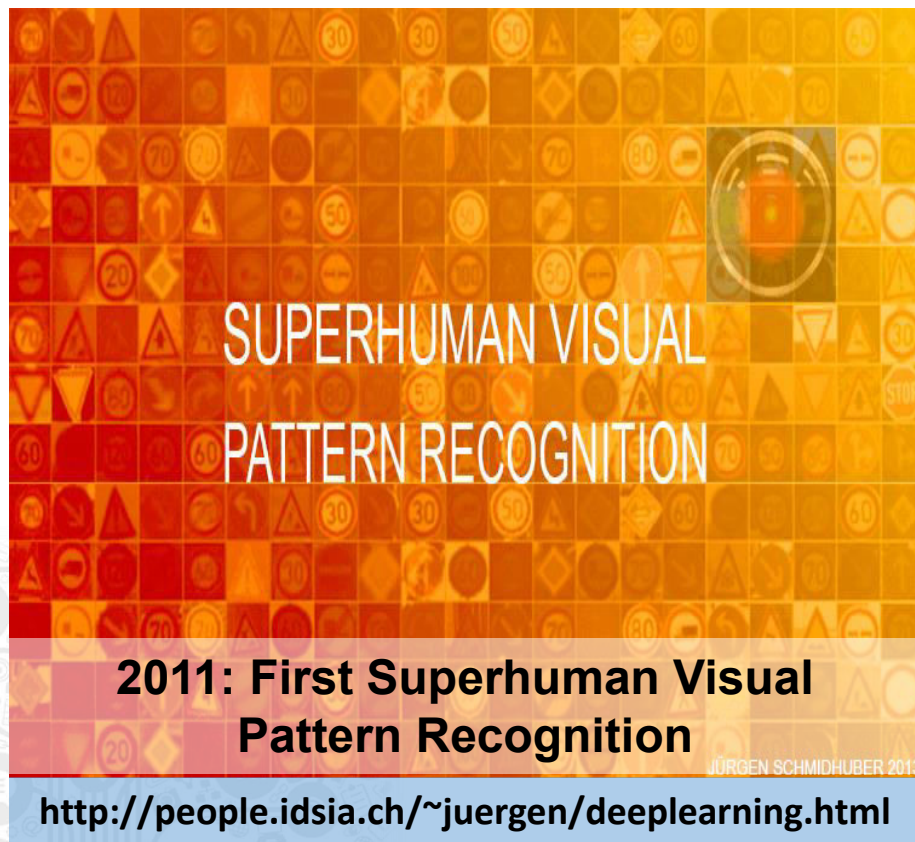
Борис Вишняков, начальник лаборатории анализа динамических сцен
ФГУП ГосНИИАС

1

ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ



ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ



1988 год

Конволюционные (сверточные) сети

2006 год

Глубокое обучение

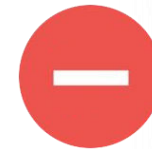
2011 год

Революция в машинном зрении –
первая публикация о глубоких
сверточных нейронных сетях в
распознавании изображений

ПЛЮСЫ И МИНУСЫ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ



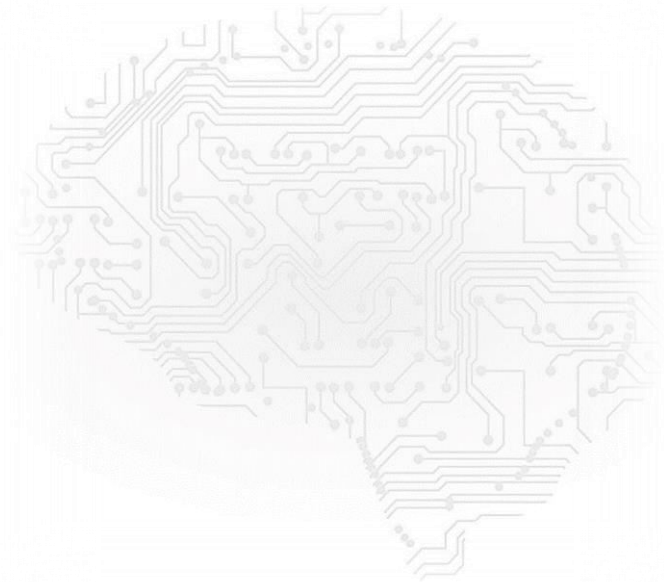
- Обучение без ограничений по сложности на сверхбольших объемах данных – за счет этого превосходит все методы прежних поколений по качеству
- Глубокие конволюционные нейросети учитывают специфику изображений как объекта распознавания
- Возможность обучения без учителя



- Для обучения требуются довольно большие объемы данных
- Для обучения требуется дорогое «железо» с большим количеством видеокарт
- Для работы также требуется мощные процессоры или видеокарты

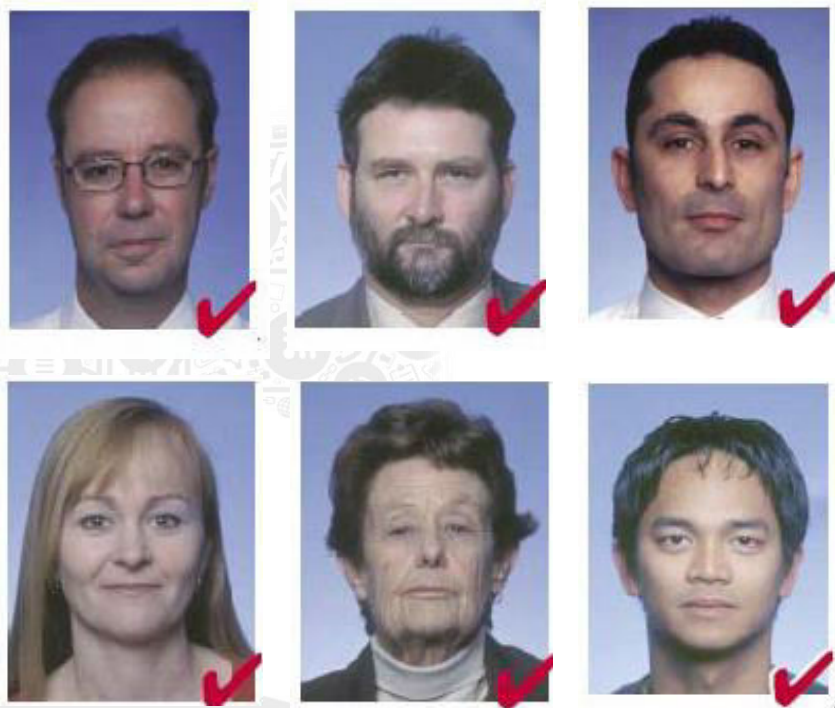
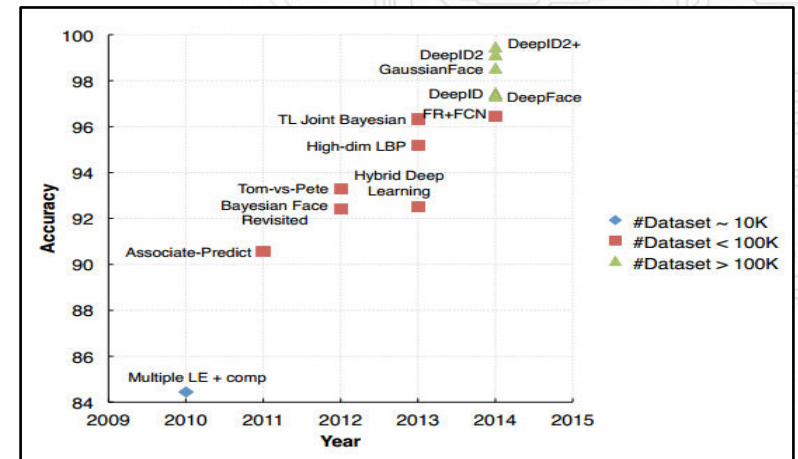
2

НОВЫЕ ПОДХОДЫ К РЕШЕНИЮ СТАРЫХ ЗАДАЧ



РЫВОК В КАЧЕСТВЕ РАСПОЗНАВАНИЯ ЛИЦ

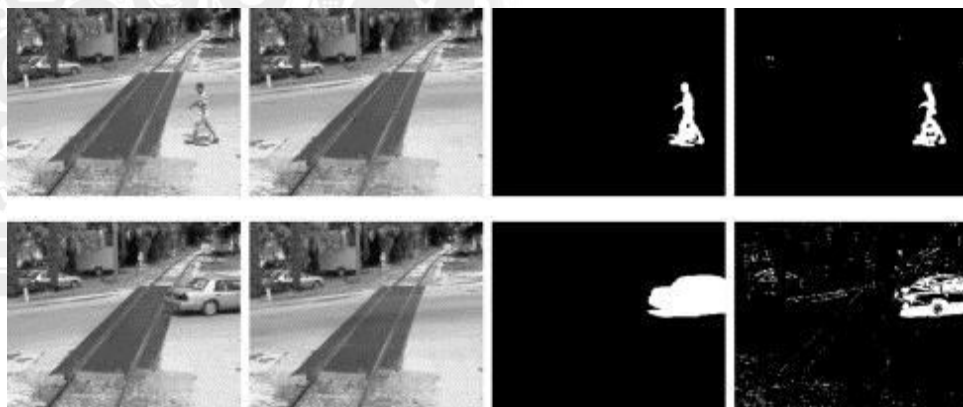
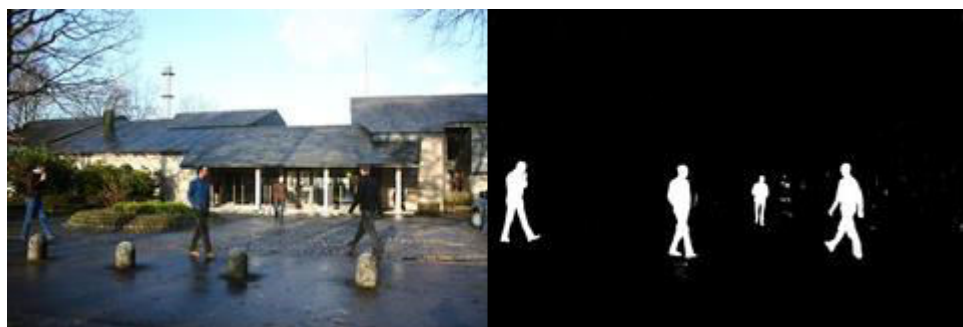
Появление нового класса алгоритмов на основе глубоких конволюционных нейронных сетей, позволяющего вплотную подойти к созданию системы автоматического распознавания лиц с качеством, превышающим уровень тренированного человека-оператора



НОВОЕ ПОКОЛЕНИЕ СИСТЕМ ВИДЕОАНАЛИТИКИ

Распознавание объектов вместо моделирования фона (background/foreground)

Прошрое поколение

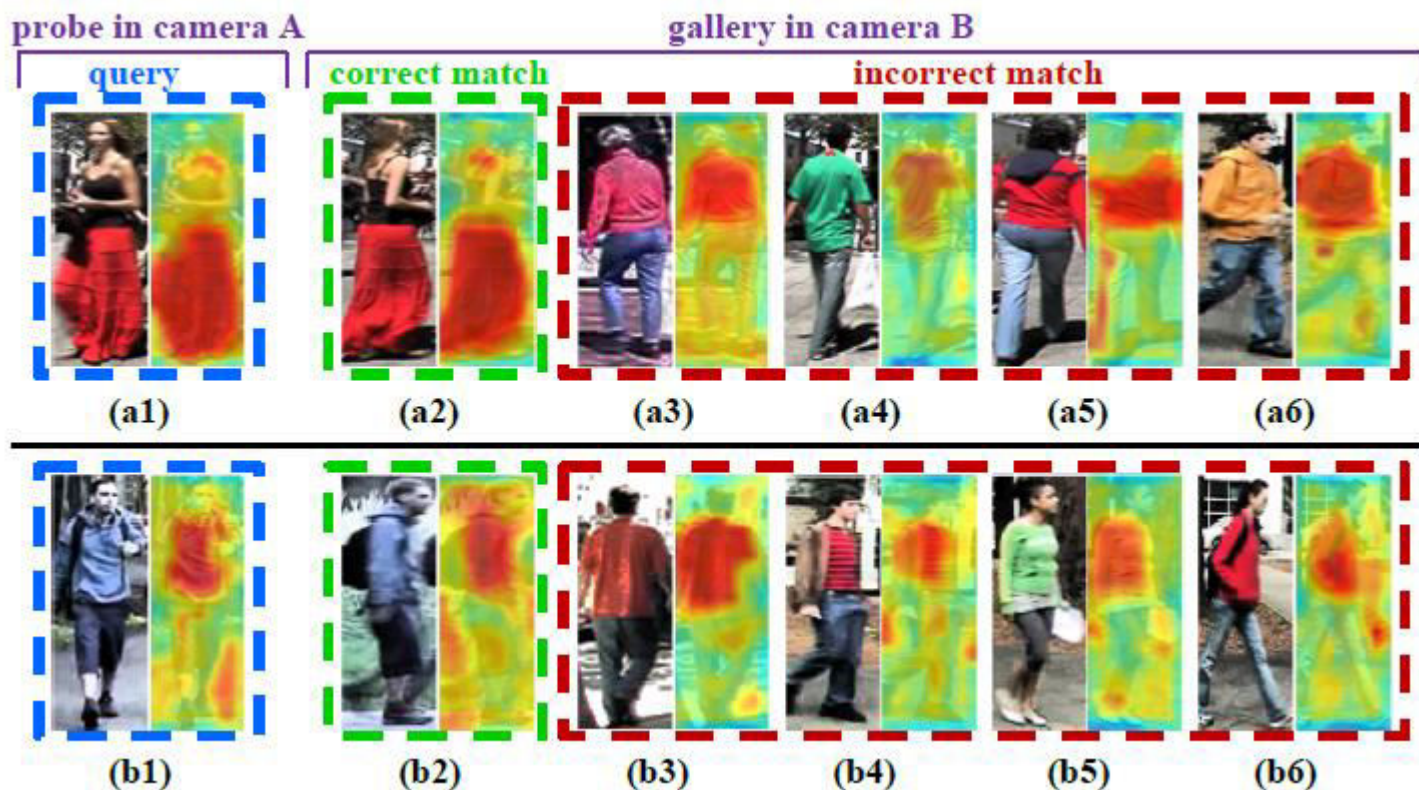


Новое поколение



РЕИДЕНТИФИКАЦИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Генерация аналога биометрического шаблона для лиц – Deep ID для объекта (человека, автомобиля и пр.) с последующим сравнением шаблона с шаблонами других объектов



КОНТРОЛЬ ТЕХНИКИ БЕЗОПАСНОСТИ НА ПРОИЗВОДСТВЕ

Обучение на СИЗ по видео-
и модельным данным



Распознавание силуэта
человека с помощью
глубоких нейронных сетей



Классификация
"есть СИЗ" – "нет СИЗ"



АНАЛИЗ БОЛЬШИХ ДАННЫХ В ЗАДАЧАХ ПРОМЫШЛЕННОСТИ

Поиск полезных
ископаемых



Транспортные и
логистические задачи

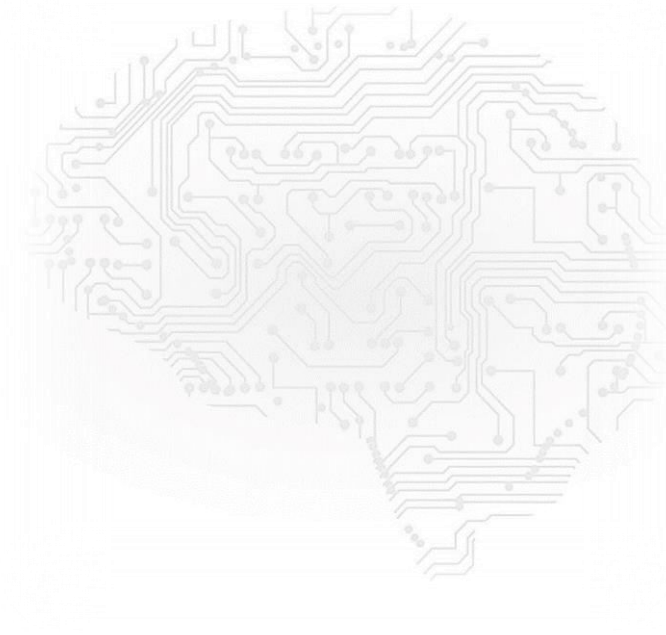


Предсказание
неисправностей



3

ТЕКУЩИЕ РЕШЕНИЯ SoC



КЛАССИЧЕСКИЕ РЕШЕНИЯ

Ambarella S2

- ✓ ARM 2 Cores A9
- ✓ FPU
- ✓ Neon
- ✓ ISP
- ✓ Max 1-4(?) GB RAM

TI DM388

- ✓ ARM 1 Cores A8
- ✓ FPU
- ✓ Neon
- ✗ ISP
- ✓ Max 1 GB RAM

HiSilicon 3519

- ✓ ARM 2 Cores A7+A17
- ✓ FPU
- ✓ Neon
- ✓ ISP
- ✓ Max 1 GB RAM



- Достаточное количество оперативной памяти для загрузки в память нейронной сети
- FPU, Neon



- Очень небольшая shared-память между ISP и CPU
- Экономия производителей камер на FLASH и RAM
- Закрытый ISP в тулчейнах производителей камер
- Очень слабый CPU (2012-2014 гг.)

ПЛИС (FPGA)

NVIDIA Tegra Jetson TX2

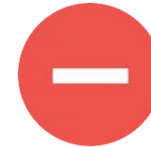
- ✓ ARM 4 Cores A57
- ✓ NVIDIA Pascal, 256 ядер CUDA
- ✓ Neon
- ✓ 8 GB RAM



- Достаточное количество оперативной памяти для загрузки в память нейронной сети
- Достаточная мощность видеопроцессора
- Достаточная мощность процессора

Intel (Altera) Stratix 10

- ✓ ARM 4 Cores A53
- ✓ FPGA с более 5.5M элементами
- ✓ Neon
- ✓ Max 16 GB DRAM



- Стоимость
- Пока не применяются в камерах

НЕЙРОННЫЕ СЕТИ НА БОРТУ КАМЕР

Распознавание лиц

- ✗ Нейросетевой детектор лиц
 - Детектор лиц на базе Виолы-Джонса
- ✗ Распознавание в реальном времени
- ✓ Распознавание в близком к реальному времени

Видеоаналитика

- ✗ Нейросетевой детектор людей и машин
- ✗ Нейросетевой трекер движущихся объектов
- ✗ Реидентификация
- ✓ Нейросетевой детектор (как распознаватель) оставленных предметов

Обнаружение СИЗ

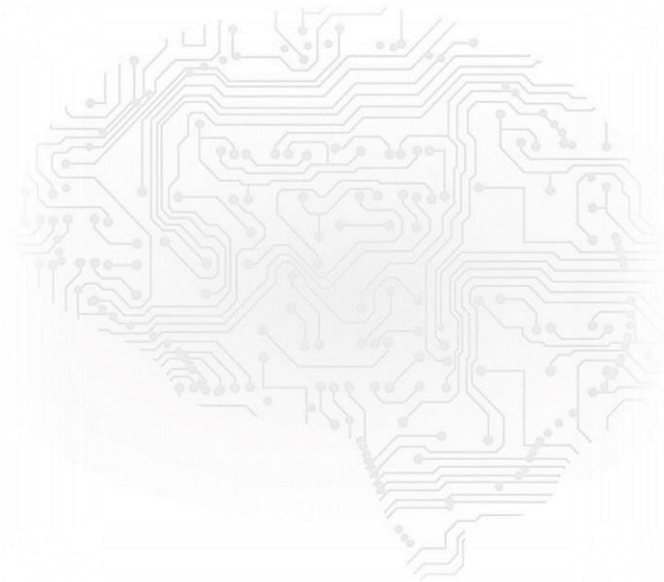
- ✗ Нейросетевой детектор людей и машин
 - Детектор на базе моделей фона
- ✓ Нейросетевой классификатор СИЗ

Выводы:

1. Нейросетевые детекторы не доступны на текущем уровне развития мобильных процессоров
2. Нейросетевые классификаторы/распознаватели могут быть использованы уже сейчас

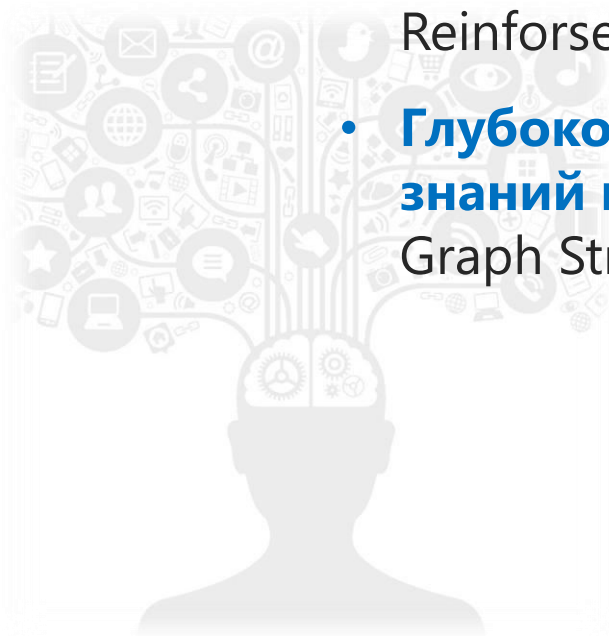
5

БОНУС: ВТОРОЙ ВИТОК НЕЙРОСЕТЕВОЙ РЕВОЛЮЦИИ

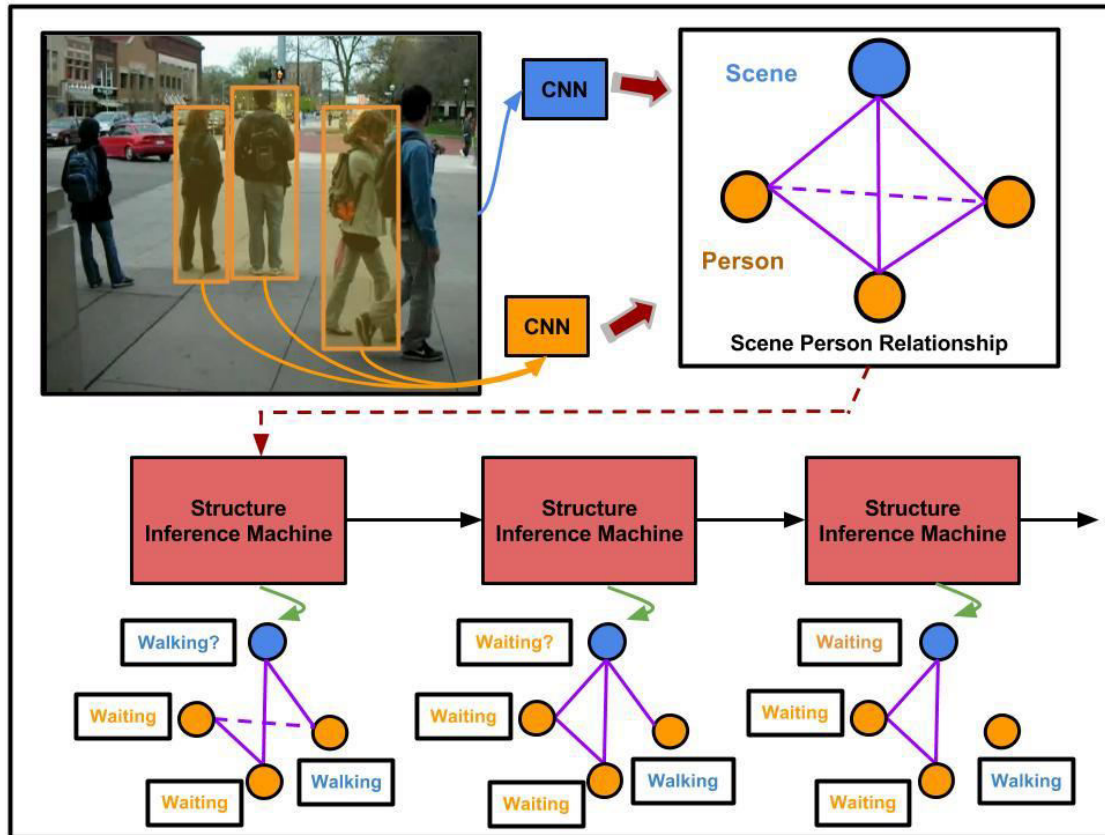


ТЕКУЩЕЕ ПОЛОЖЕНИЕ ДЕЛ

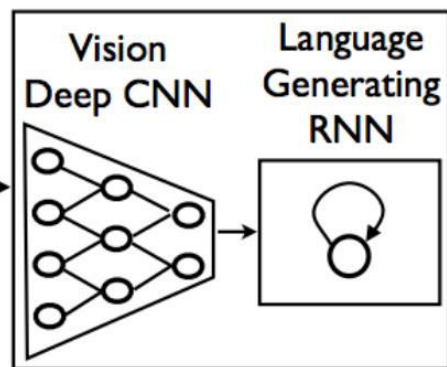
- **Глубокие соревнующиеся сети для имитации данных**
GAN, Domain Transfer Learning, Zero-Shot Learning
- **Интерпретация динамической визуальной информации на естественном языке**
Action Detection and Prediction, Video Annotation, Video and Language Understanding, Text-to-Video, VQA
- **Обучение глубоких сетей как активных агентов**
Reinforcement Learning, Lifelong Learning
- **Глубокое обучение с использованием структурных моделей, баз знаний и программ логического вывода**
Graph Structured CNN, Deep Visual Reasoning



ПОНИМАНИЕ ПОВЕДЕНИЯ

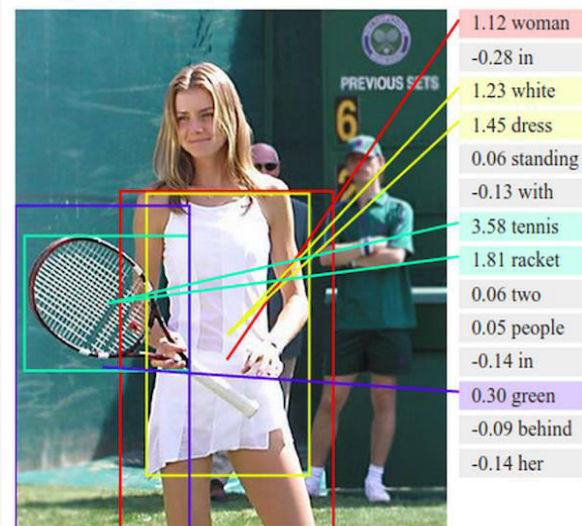
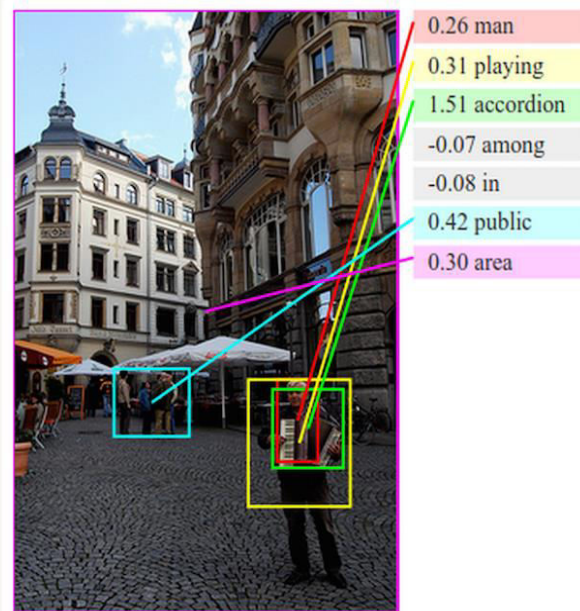
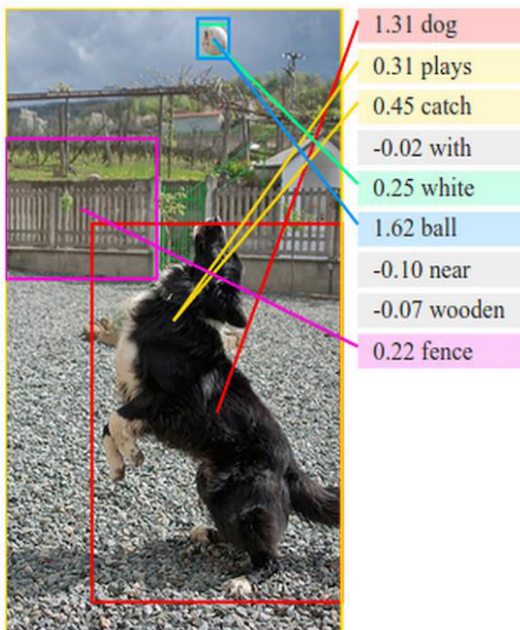
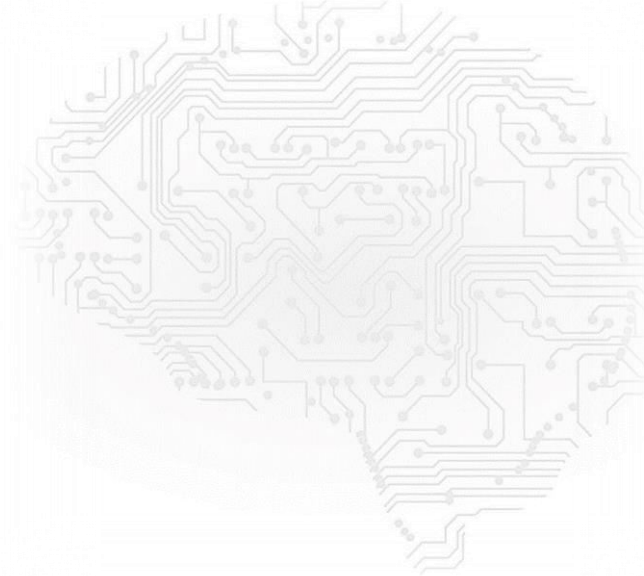


ПОНИМАНИЕ СЦЕНЫ



A group of people shopping at an outdoor market.

There are many vegetables at the fruit stand.



GANs – GENERATIVE ADVERSARIAL NETWORKS

Summer ↔ **Winter**

summer $\rightarrow$ winter

winter $\rightarrow$ summer

Monet  Photos

Monet $\rightarrow$ photo

photo $\rightarrow$ Monet

Zebras ↔ Horses

zebra $\rightarrow$ horse

horse $\rightarrow$ zebra