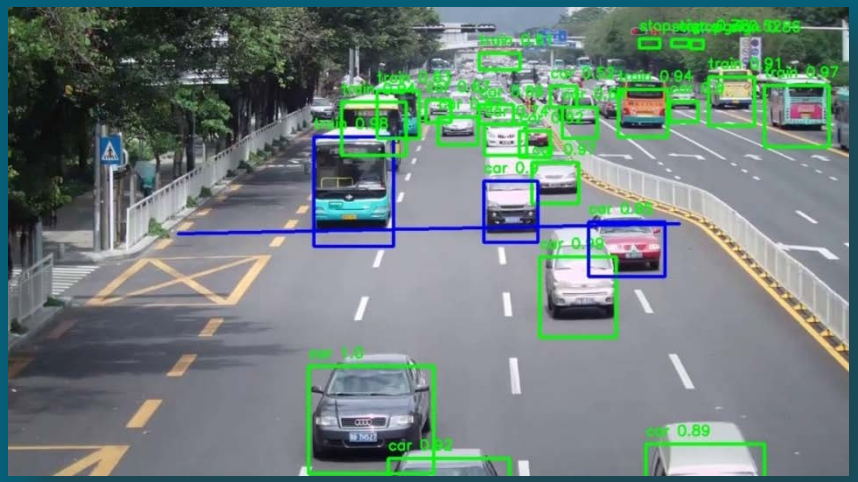
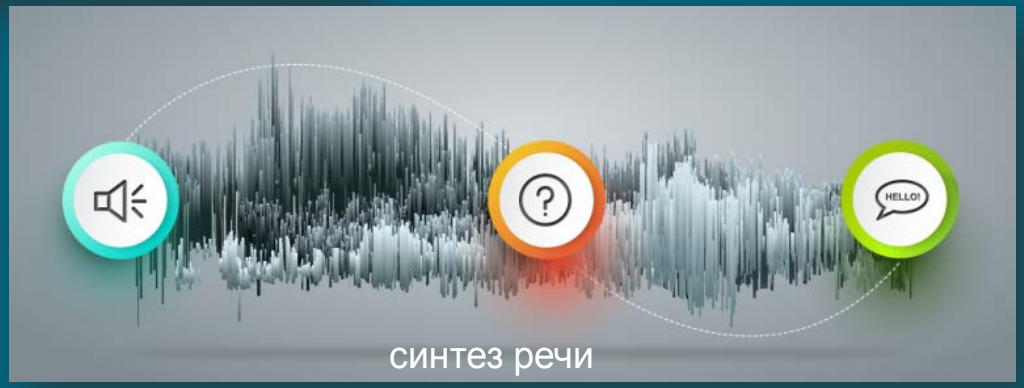
Применение FPGA для ускорения глубокого обучения

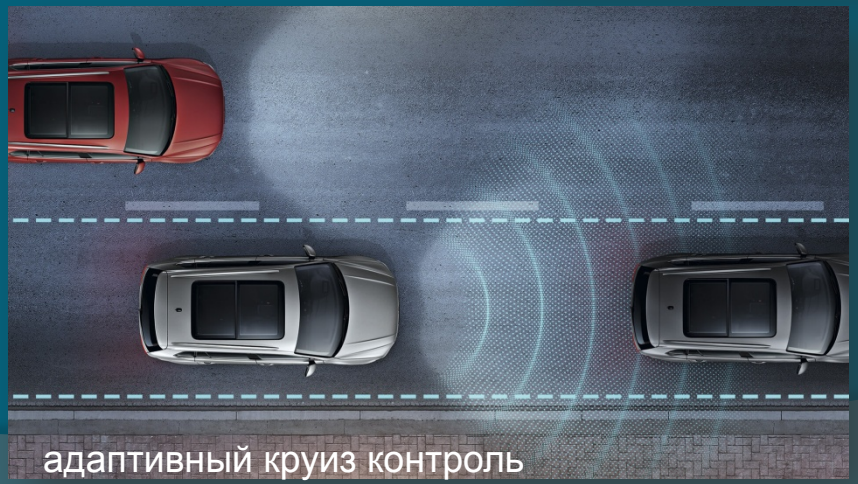
Работу выполнил
студент ГБПОУ «НАТК»
Морозов С.И.



компьютерное зрение



синтез речи



адаптивный круиз контроль

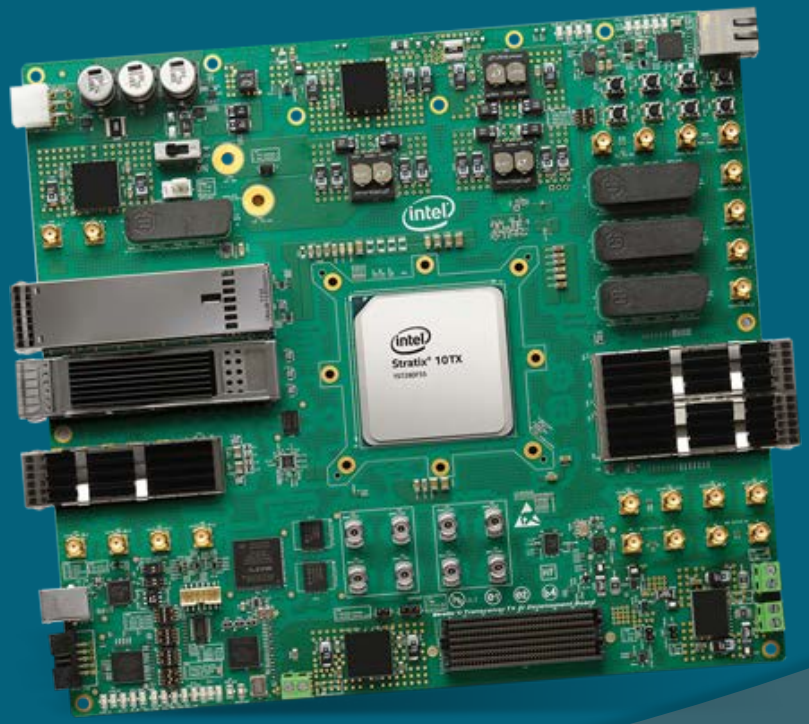
Цели и задачи

Целью является изучить принципы ускорения приложений глубинного обучения с помощью FPGA на примере приложения компьютерного зрения.

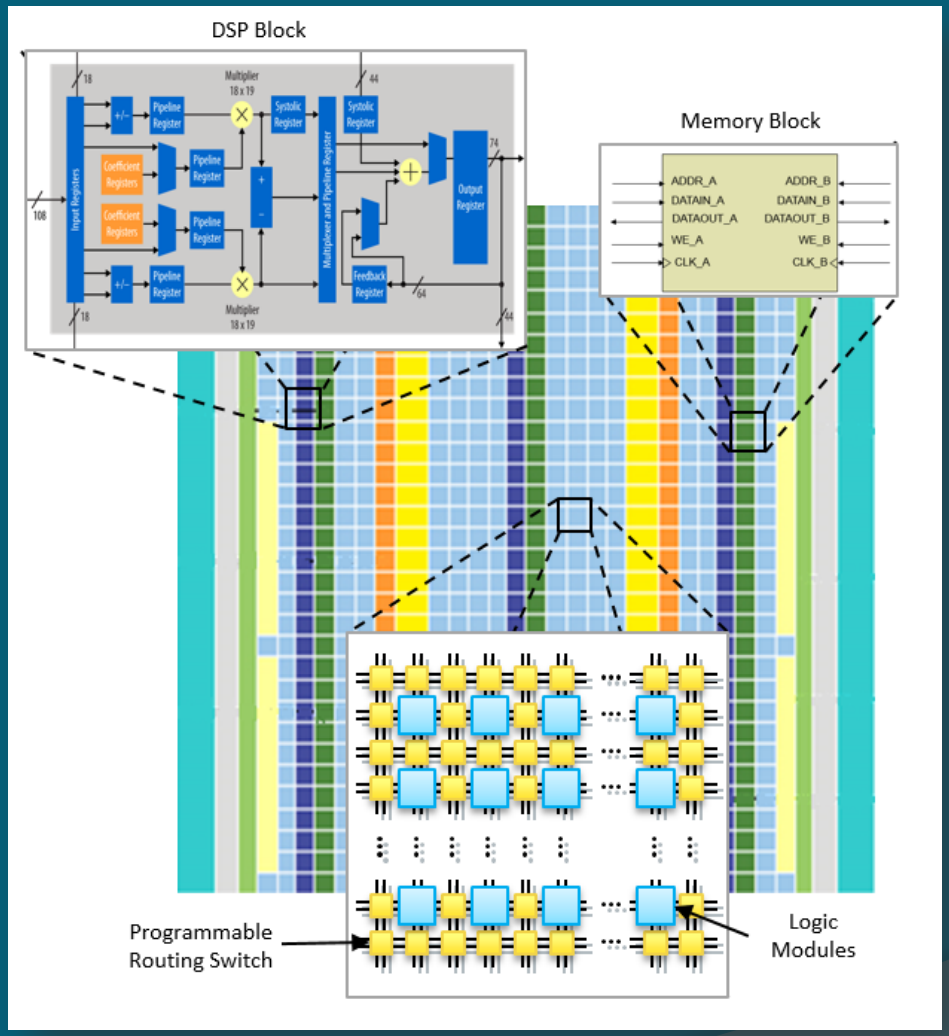
- Задачи:
- Изучение FPGA с точки зрения разработчика программного обеспечения
- Углубленное изучение конволюционных нейронных сетей их выполнение и построение
- Узнать как развернуть приложение компьютерного зрения, которое использует глубинное обучение для извлечения шаблонов из данных
- Использование Intel® Distribution of OpenVINO™ toolkit для конволюционной нейронной сети и обработки данных на основе вычислительных процессоров Intel® и массивов FPGA
- Как использование FPGA и стека ускорения для вычислительных модулей на базе процессоров Intel® Xeon® переводит приложения для облачных вычислений и центров обработки данных на более высокий уровень эффективности

FPGA

Field-Programmable Gate Array (Программируемая логическая интегральная схема). Используется для создания конфигурируемых цифровых электронных схем.

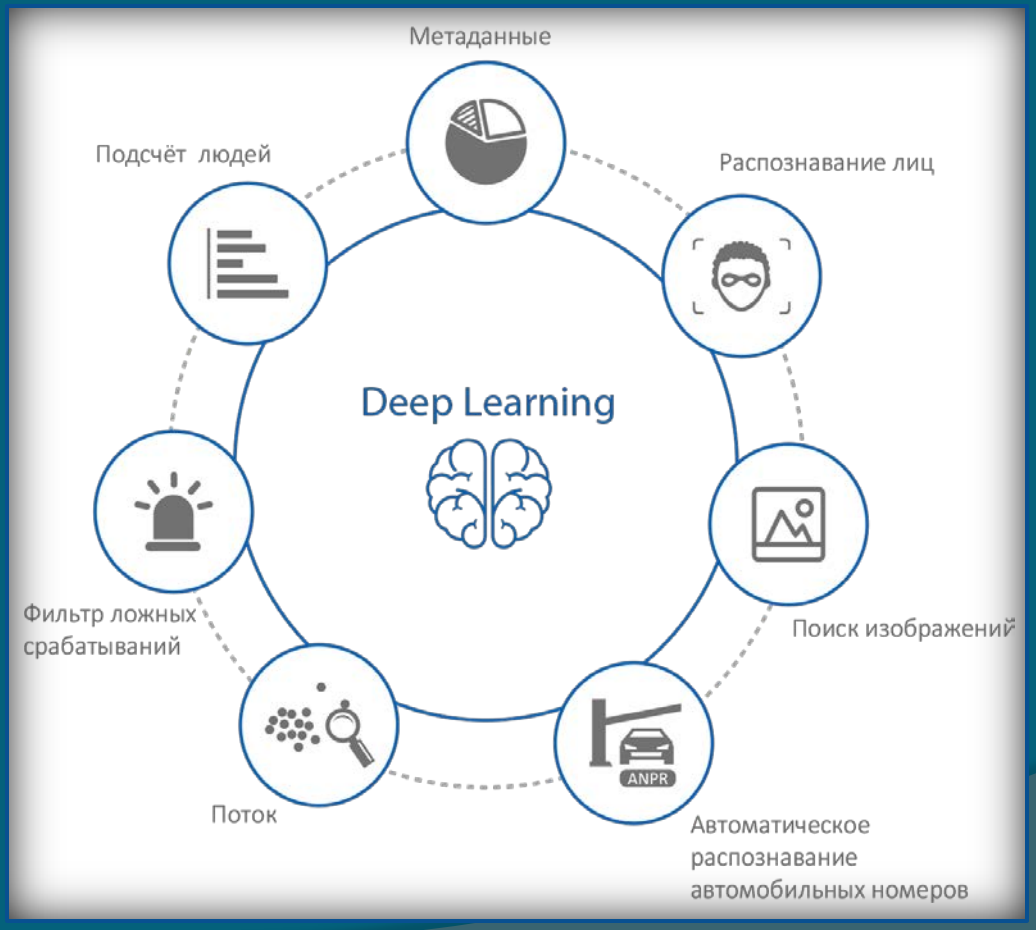


Архитектура FPGA

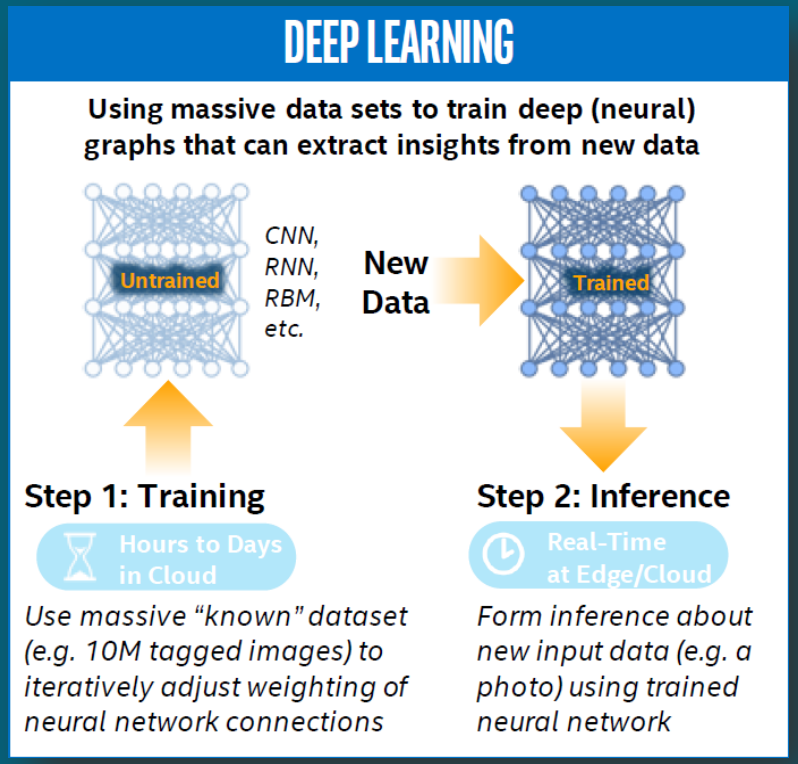


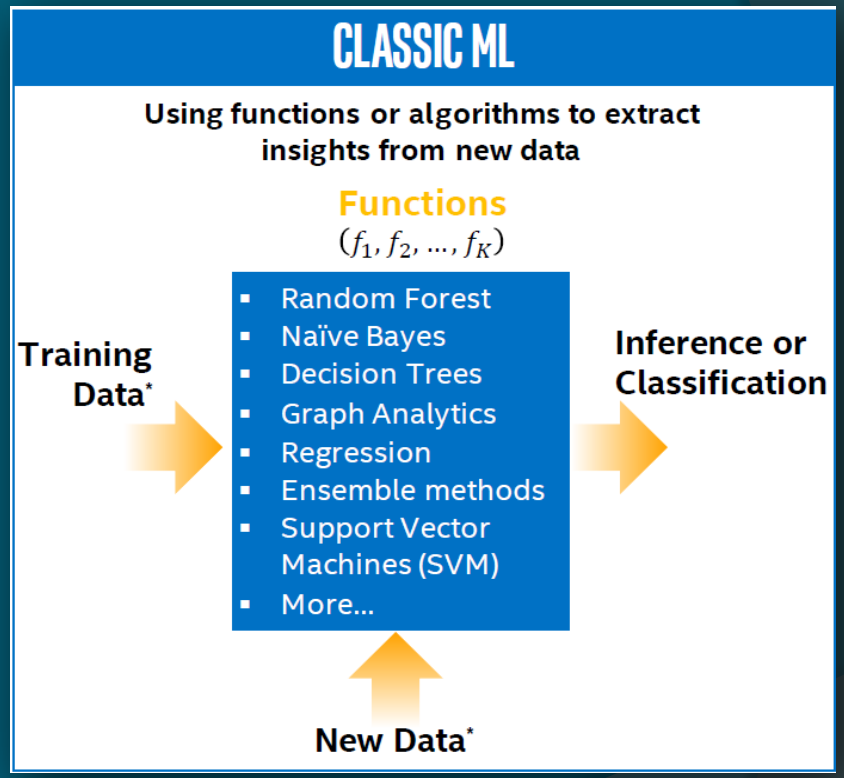
Глубокое обучение

Совокупность методов обучения на основе представлений, а не специализированных алгоритмов под конкретные задачи.



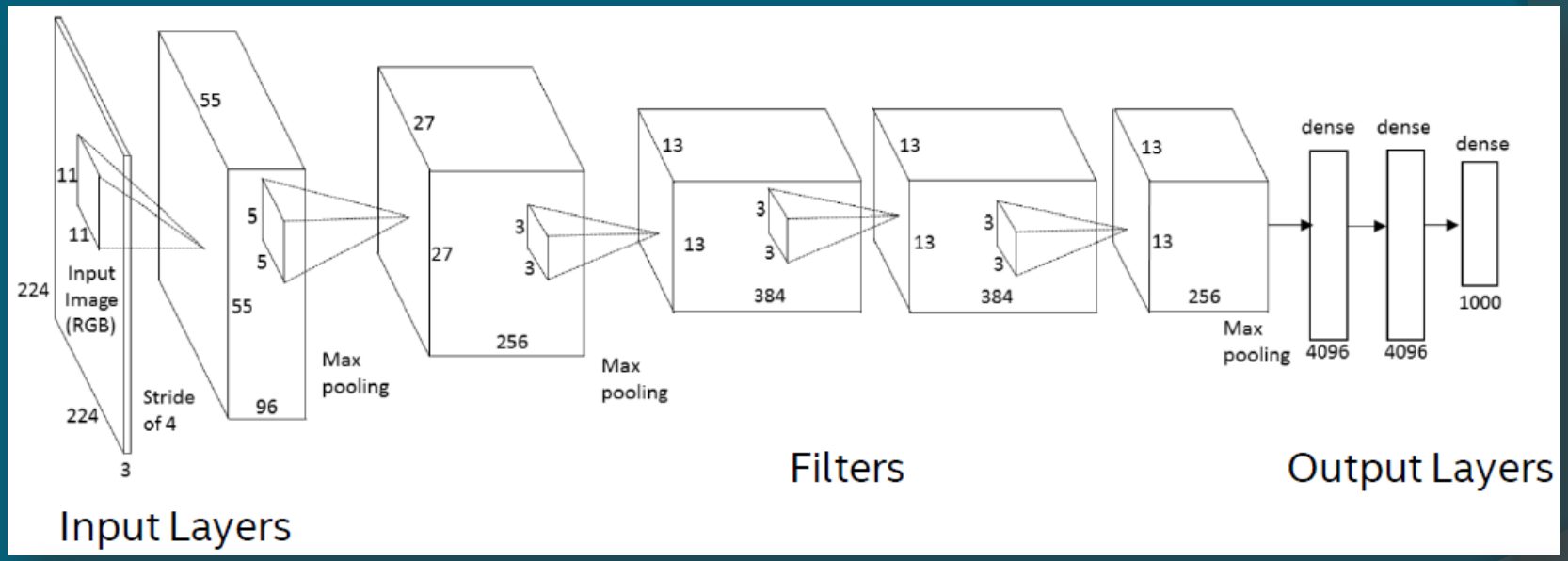
Глубокое обучение





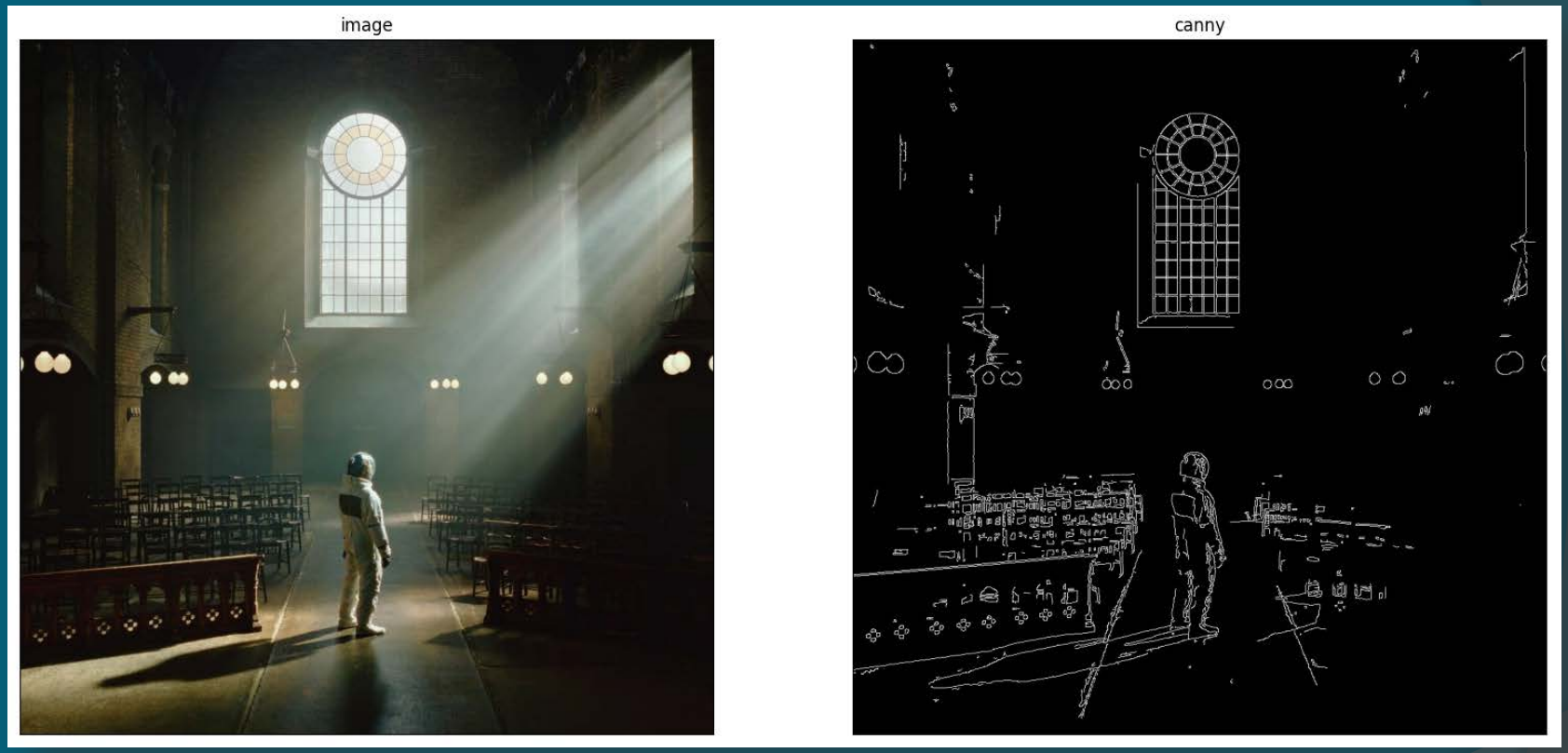
Свёрточная нейронная сеть

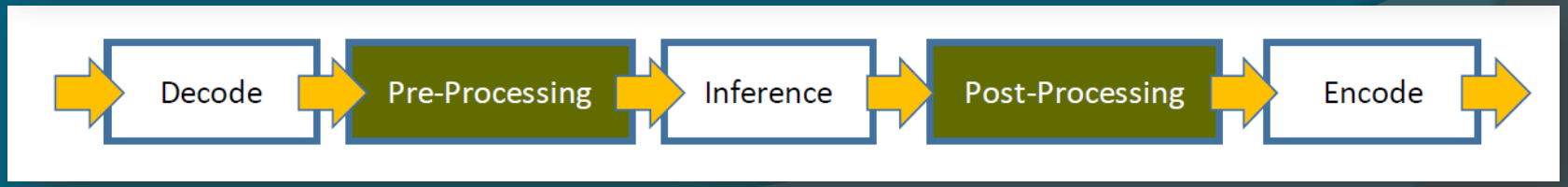
Свёртка уменьшает объем данных и количество вычислений



Layer 1 neurons	Layer 2 neurons	Fully Connected Calculations	Convolutional Calculations
224 x 224 x 3 = 150,528	55 x 55 x 96 = 290,400	43 Billion	97 Million

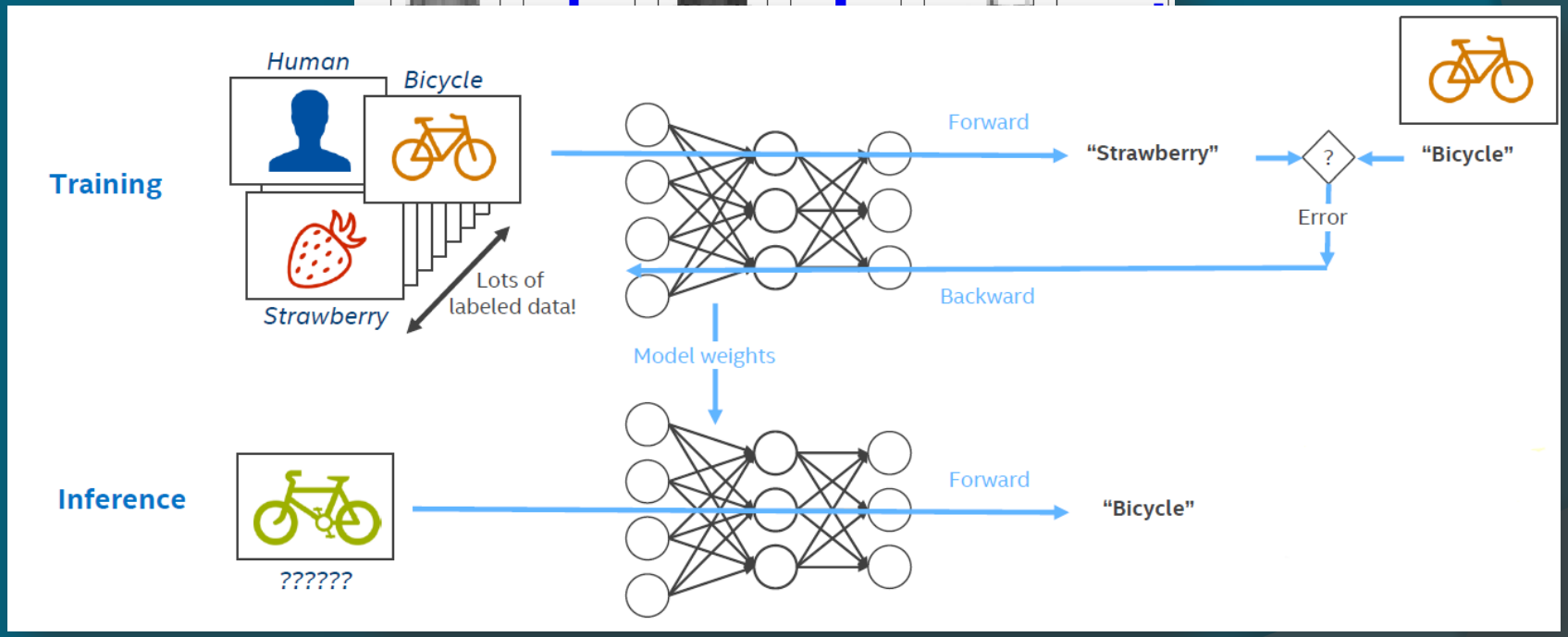
Pre-processing

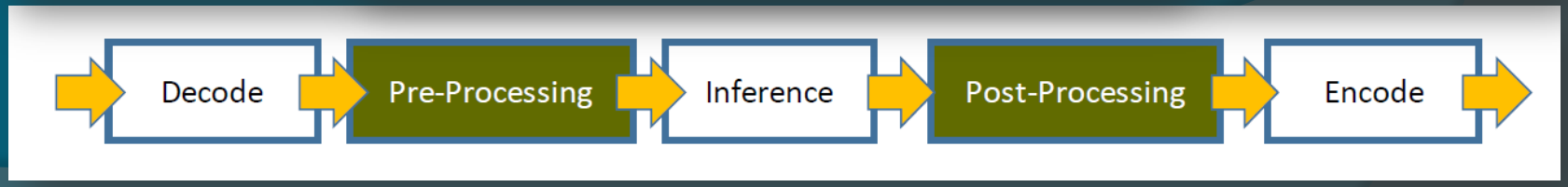




inference

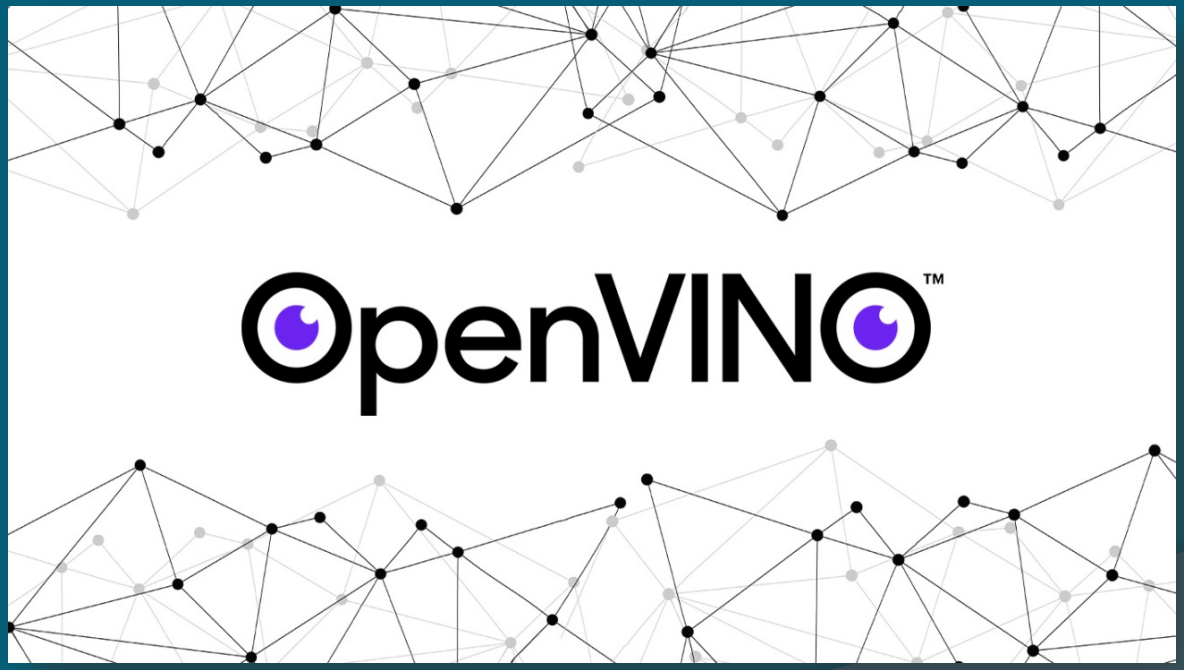
Вычислительно интенсивный процесс выполнения обученной сети на входе данных (видео / изображения) чтобы найти образец





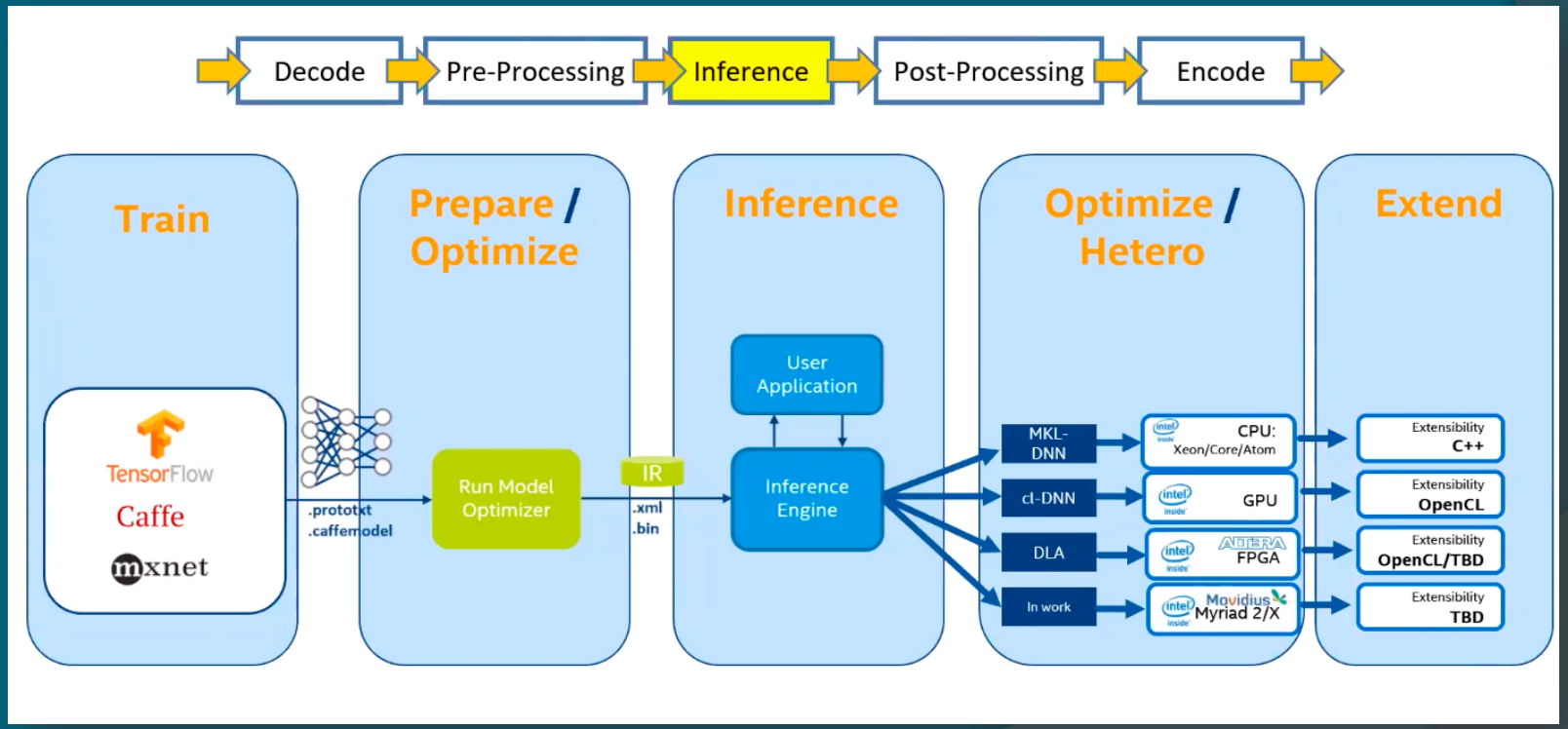
OPENVINO

OpenVINO – это бесплатный набор инструментов, основанный на сверточных нейронных сетях (CNN), инструментарий расширяет рабочие нагрузки на аппаратное обеспечение Intel® и максимизирует производительность.

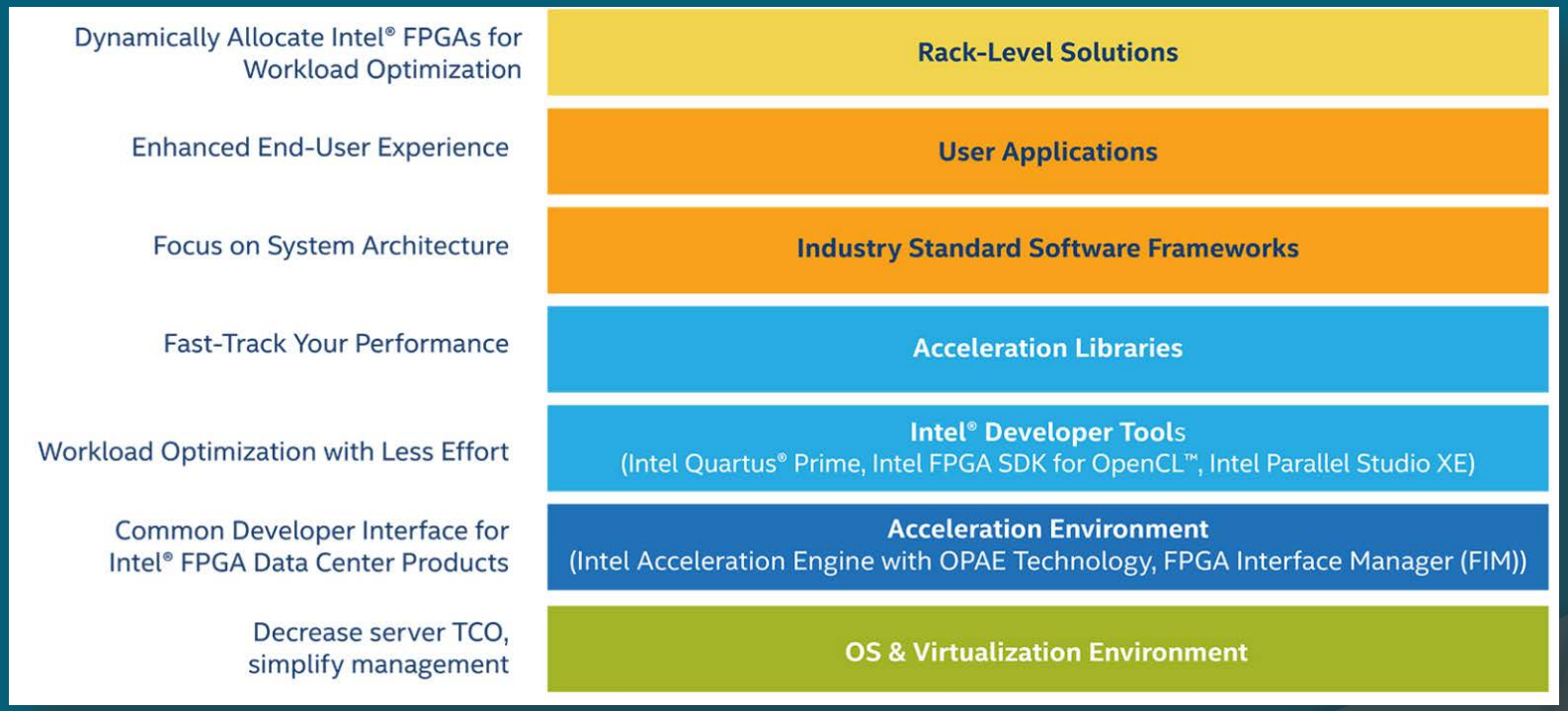


OpenVINO

- Оптимизатор модели
- Механизм логического вывода
- Model Zoo и Model Downloader
- Предсобранные OpenCV и OpenVX



FPGA Acceleration Stack



Перспективы

Искусственный интеллект — это одна из самых быстроразвивающихся технологий. Нейронные сети глубокого обучения – наиболее актуальное направление развития этой технологии.

В дальнейшем, мною планируется применить полученные в ходе исследования навыки на практике, реализовав алгоритм компьютерного зрения на отдельном чипе с применением аппаратного ускорения с использованием FPGA