



Конференция
для IT-специалистов,
интересующихся темой
Artificial Intelligence

Искусственный интеллект глазами разработчика



НовГУ им.
Ярослава Мудрого



Искусственный интеллект глазами разработчика

29 мая 2021



Мастерская инструментов
разработки
mir.dev

AI на обычном микроконтроллере? Легко!

Воркшоп на удалённом
“железе”, который можно
повторить дома

Александр Белоусов

Руководитель проектов, компания МИР

Антон Волков

Магистрант, НовГУ им.Ярослав Мудрого



НовГУ имени
Ярослава Мудрого

Содержание

- Машинное обучение на микроконтроллерах
- Обучение нейронной сети для распознавания рукописных цифр
- Формирование прошивки с нейронной сетью
- Поставка прошивки на удаленную плату с микроконтроллером
- Распознавание изображения на удаленной плате
- Подведение итогов

“Мелкая революция”



Микроконтроллер – небольшой компьютер, предназначенный для выполнения одной задачи или программы в устройстве

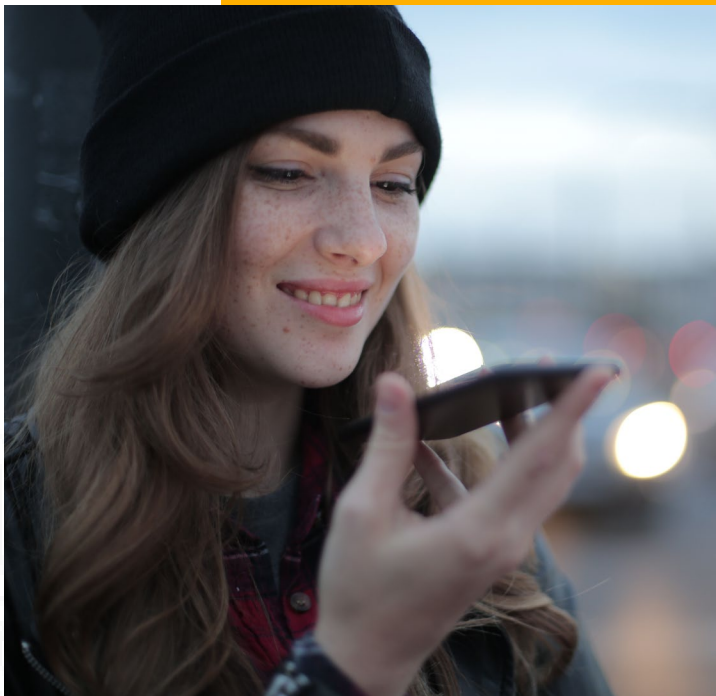
- доступны
- потребляют мало энергии
- мгновенный результат здесь и сейчас

Встроенные (embedded) системы есть практически в каждом электронном устройстве

**270
млрд**

микроконтроллеров
в мире, ежегодный
прирост около
30 млрд

ML на микроконтроллерах?



Раньше:

- сложно, дорого, специальные вычислители
- решаем задачи в облаке – издержки на передачу данных

“Привет, Siri” - “Окей, Google”

- AI на микропроцессоре

Сейчас:

- 270 млрд микроконтроллеров могут решать задачи ML
- Tiny Machine Learning (TinyML)

Микровселенная возможностей



Tiny Machine Learning (TinyML)

Новейшая технология встроенного программного обеспечения, которая перемещает оборудование в сферу, где машины могут автоматически учиться и расти в процессе использования

- TensorFlow Lite для микроконтроллеров
- Arm CMSIS-NN
- CoreML — библиотека Apple для устройств iOS.
- PyTorch Mobile — мобильная версия библиотеки PyTorch от Facebook

TinyML - можем делать больше, располагая меньшим

Tensorflow Lite



[tensorflow.org/
lite/guide](https://tensorflow.org/lite/guide)

Готовые решения:

- **Обнаружение объектов** — используется для распознавания нескольких объектов на изображении, с поддержкой до 80 различных элементов
- **Умные ответы** — генерирует интеллектуальные ответы, подобные тем, которые вы получаете при взаимодействии с диалоговым ИИ или чат-ботом
- **Рекомендации** — предлагает системы индивидуальных рекомендаций, основанных на поведении пользователей

Инструментарий



- Сервис all-hw.com - сервис удаленного доступа к платам
- Сервис Google Colaboratory colab.research.google.com - обучение модели нейронной сети
- Приложение STM32CubeMX с пакетом X-CUBE-AI Pack - формирование прошивки с нейронной сетью
- Приложение STM32CubeIDE - отправка прошивки на плату и удаленная отладка

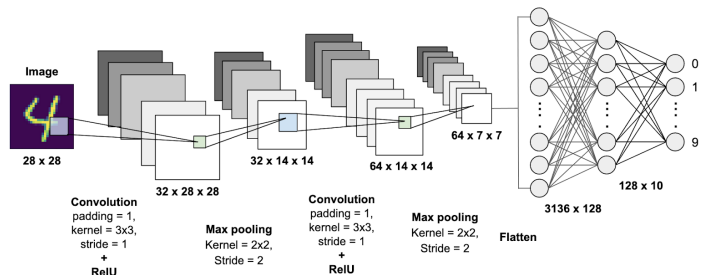
Для демонстрации будем использовать плату

STM32L562E Discovery в сервисе <https://all-hw.com/>

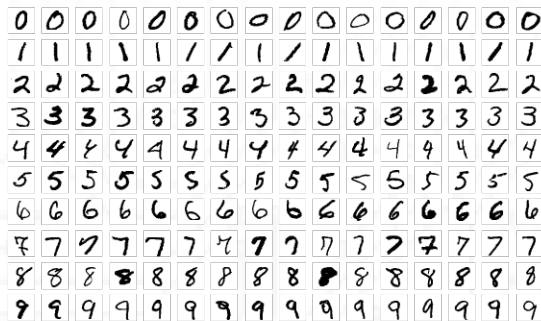
Обучение нейронной сети. Этапы



- Строим нейронную сеть



- Обучаем на наборе MNIST



- Преобразуем модель в формате tflite

```
converter = tf.lite.TFLiteConverter.from_keras_model(model)
```

Отправка прошивки на плату



Framework	Runtime	Application Base
1. Keras 2. TF Lite	1. TF Lite Micro 2. X Cube AI	1. Validation 2. App template 3. Not selected

Основа для прошивки



1

Show graph

Analyze

Validate on desktop

Validate on target

2

Validation on target

Use communication port: ☒ Automatic ☐ Manual

Automatic compilation and download

☒ Enabled Communication

3

C0	0	.	.	.	.	.	.	.	.
C1	.	0	.	.	.	.	.	.	.
C2	.	.	0	.	.	.	.	.	.
C3	.	.	.	0	.	.	.	.	.
C4	.	.	.	.	1	.	.	.	.
C5	.	.	.	.	.	0	.	.	.
C6	.	.	.	.	.	.	0	.	.
C7	.	.	.	.	.	.	.	0	.
C8	.	.	.	.	.	.	.	.	0
C9	.	.	.	.	.	.	.	.	0

Evaluation report (summary)

Mode	acc	rmse	mae	l2r
X-cross #1	100.00%	0.0000000000	0.0000000000	0.000000

Что добавили в свою прошивку



UART с глобальным прерыванием

Инициализация и запуск
нейросети

Переопределены callback'и UART
для принятия набора байтов

Главный цикл:

- Проверка входных данных
- Валидация нейросети с входными данными
- Вывод результатов

```
HELLO!  
  
Received array of 785 elements!  
  
Array has been converted!  
  
Prediction results:  
Out[0]: 0.000000  
Out[1]: 0.000000  
Out[2]: 0.000000  
Out[3]: 0.000000  
Out[4]: 1.000000  
Out[5]: 0.000000  
Out[6]: 0.000000  
Out[7]: 0.000000  
Out[8]: 0.000000  
Out[9]: 0.000000  
Predicted digit: 4
```

Подведение итогов



- 1 Познакомились с сервисом для удаленной работы с платами <https://all-hw.com/>
- 2 С помощью TensorFlow и Google Collab обучили нейронную сеть для распознавания рукописных цифр
- 3 Использовали обученную нейронную сеть для создания прошивки удаленной платы с помощью STM32CubeMX и пакета X-CUBE-AI Pack
- 4 Поставили созданную прошивку с нейронной сетью на удаленную плату и запустили нейронную сеть глубокого обучения на микроконтроллере
- 5 Провели распознавание рукописных цифр на удаленном микроконтроллере



Искусственный интеллект глазами разработчика

29 мая 2021



Мастерская инструментов
разработки
mir.dev

Вопросы?



НовГУ имени
Ярослава Мудрого