



# **Training Neural Networks on Embedded Devices Targeting Embedded Environments**

---

Prasanth Shaji, Deepak Venkataram

Master's Thesis  
Uppsala University

August 21, 2023



# Outline

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

## 1 Neural Networks in Embedded Devices

## 2 Embedded Linux

## 3 Neural Networks

## 4 Benchmark Applications

## 5 Discussion



UPPSALA  
UNIVERSITET

# Neural Network Applications on Embedded Devices

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# Embedded Devices





# Goal : Neural Network training on Scania C300

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# Goal : Neural Network training on Scania C300

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

Predictive vehicle maintenance using data from a fleet of connected trucks

Why Training on the C300?

- Repurpose existing C300 units to perform neural network related tasks
- Training on devices to reduce network bandwidth usage, ensure data privacy, etc.

*Scope Repurpose Scania C300, train neural network model, assess training performance*



# Goal : Neural Network training on Scania C300

Predictive vehicle maintenance using data from a fleet of connected trucks

## Why Training on the C300?

- Repurpose existing C300 units to perform neural network related tasks
- Training on devices to reduce network bandwidth usage, ensure data privacy, etc.

*Scope Repurpose Scania C300, train neural network model, assess training performance*



# Goal : Neural Network training on Scania C300

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

Predictive vehicle maintenance using data from a fleet of connected trucks

## Why Training on the C300?

- Repurpose existing C300 units to perform neural network related tasks
- Training on devices to reduce network bandwidth usage, ensure data privacy, etc.

**Scope** Repurpose Scania C300, train neural network model, assess training performance



# ARM Boards : Examples

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

SoC	Application Domain	Architecture	Processor Core
nRF51822	Ultra low power, BLE	ARM v6-M	ARM Cortex M0
AM2732	Automotive	ARM v7-R	ARM Cortex R5
AM3358	Industrial / IoT	ARM v7-A	ARM Cortex A8
i.MX6S	Multimedia applications	ARM v7-A	ARM Cortex A9
BCM2711	Educational / IoT	ARM v8-A	ARM Cortex-A72
Snapdragon - Kryo 240	Smartphones	ARM v8-A	ARM Cortex-A73, ARM Cortex-A53
Apple A16 Bionic	iPhone 14 Pro	ARM v8.6-A	APL1W10
Apple M1	Macbooks / Mac Mini / iPads	ARM v8.5-A	APL1102



# ARM Boards : Outline

## ARM | Silicon Vendor | System Maker

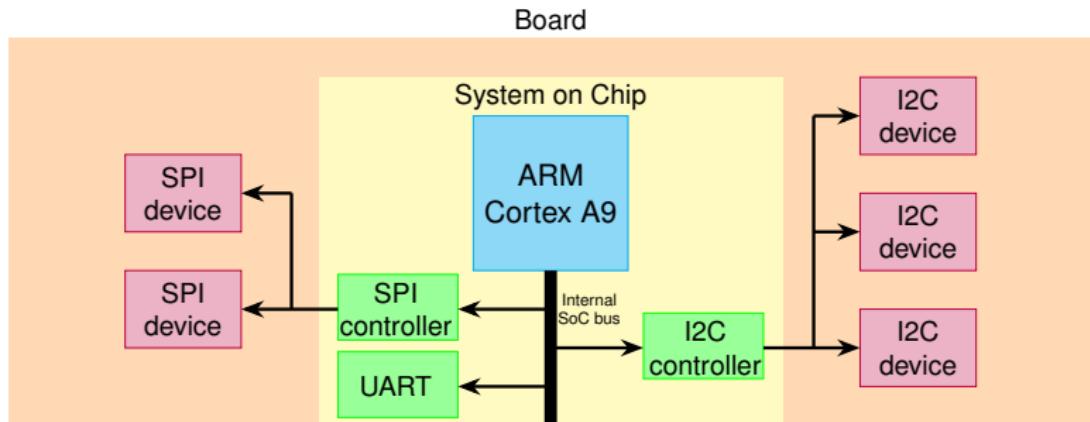
Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# Embedded Linux

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

Kernel, Bootloaders, Device Trees  
Cross Compilation, Toolchains, Root Filesystems

Operating System for Embedded Devices  
- Customized for Specific Hardware



# Embedded Linux

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

Kernel, Bootloaders, Device Trees  
Cross Compilation, Toolchains, Root Filesystems

Operating System for Embedded Devices  
- Customized for Specific Hardware



# Build Systems

Neural  
Networks in  
Embedded  
Devices

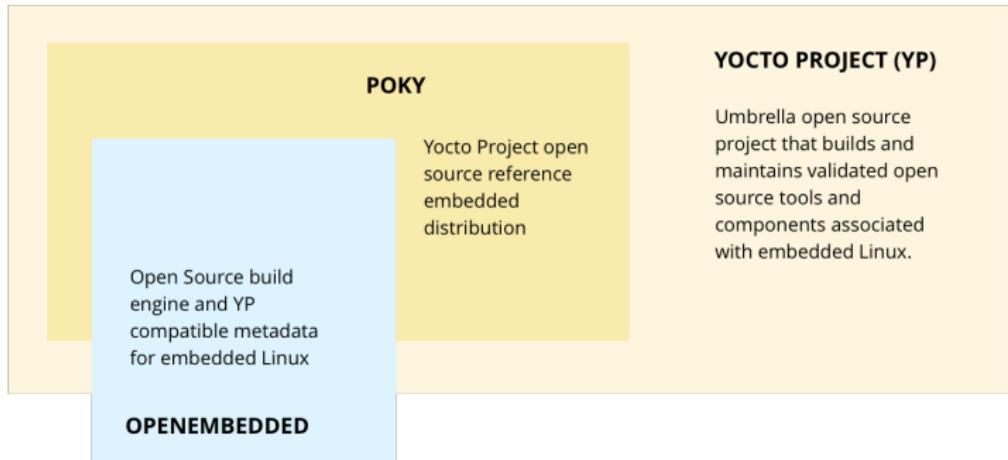
Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

## Yocto Project





# The Yocto Project

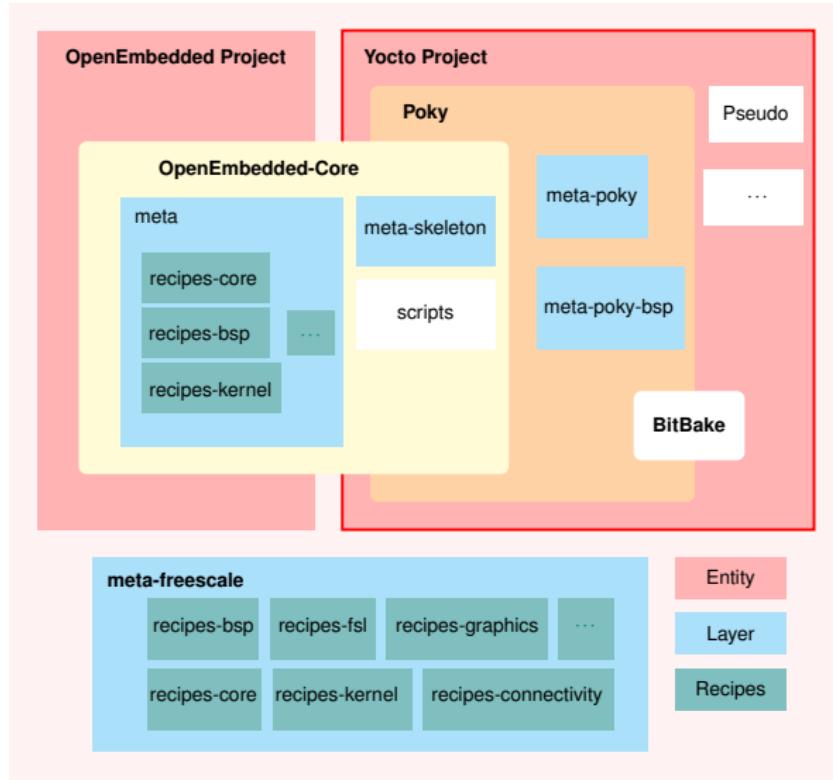
Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

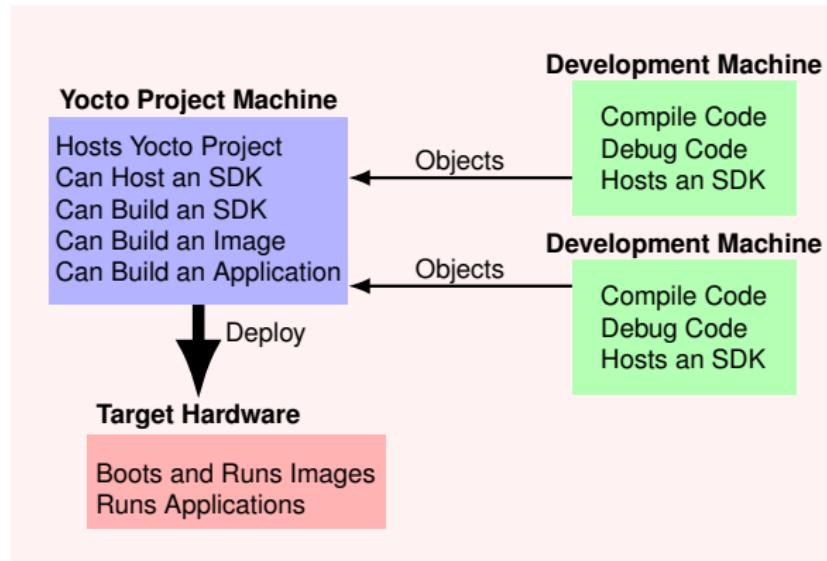
Benchmark  
Applications

Discussion





# Yocto Project : Development





# Software Versions

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion



# Scania C300 to MCIMX6Q-SDB

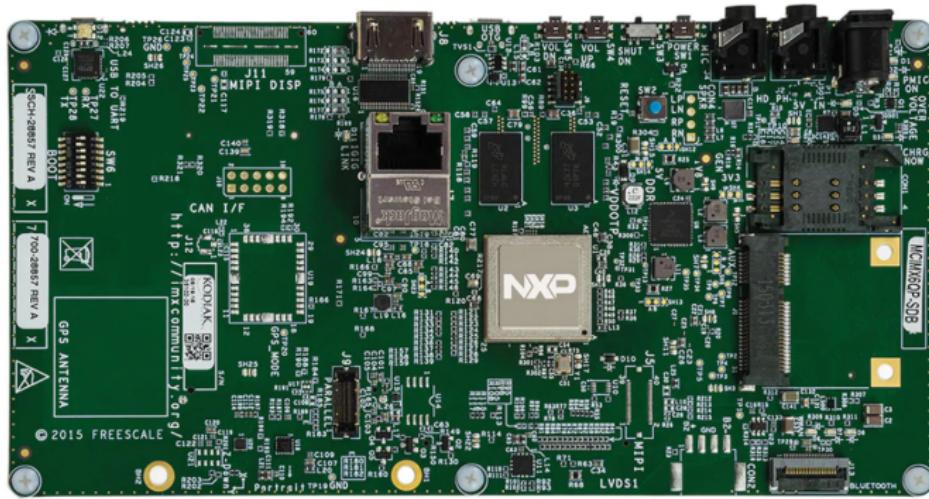
# Neural Networks in Embedded Devices

# Embedded Linux

# Neural Networks

## Benchmark Applications

## Discussion





# Handwritten Digit Recognition

## Neural Network



MNIST



# Neural Network Training

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

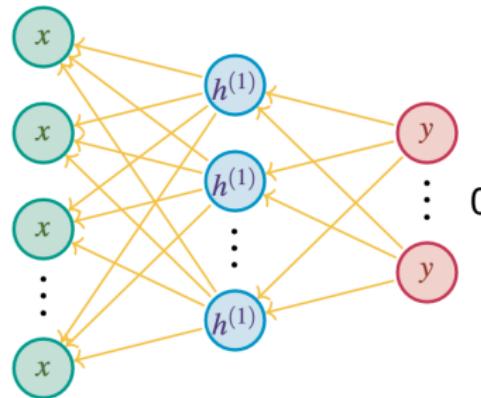
Neural  
Networks

Benchmark  
Applications

Discussion



label = 0





# Learning Algorithm

---

**Algorithm 1** Mini Batch Gradient Descent with learning rate  $\eta$  and the Mean Squared Error (MSE) cost function

---

**Require:** initial weights  $w^{(0)}$ , number of epochs  $E$ , batch size  $B$ , training data with  $T$  entries

**Ensure:** final weights  $w^{(E*T)}$

```
for  $e = 0 \rightarrow E - 1$  do
    for  $b = 0 \rightarrow T/B$  do
        for  $t = b * B \rightarrow (b + 1) * B$  do
            estimate  $\nabla \mathcal{L}(w^{(t)})$ 
            compute  $\Delta w^{(b)} += -\nabla \mathcal{L}(w^{(t)})$ 
        end for
         $w^{(e+1)} := w^{(e)} + \eta \Delta w^{(e)}$ 
    end for
end for
return  $w^{(T)}$ 
```

▷  $\mathcal{L}$  here is MSE

---



# Neural Network Inference

Neural  
Networks in  
Embedded  
Devices

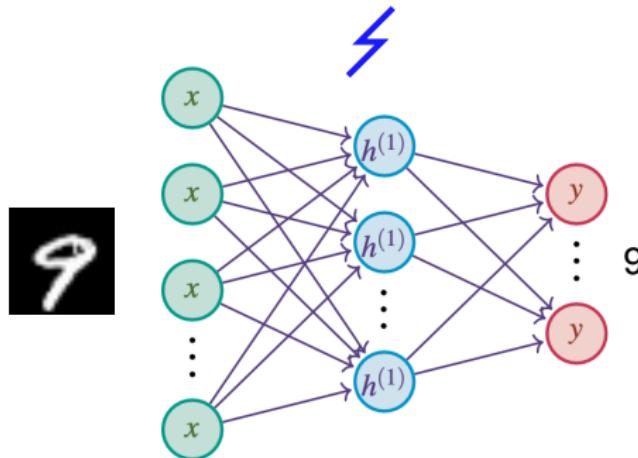
Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

9





# Traditional Paradigm

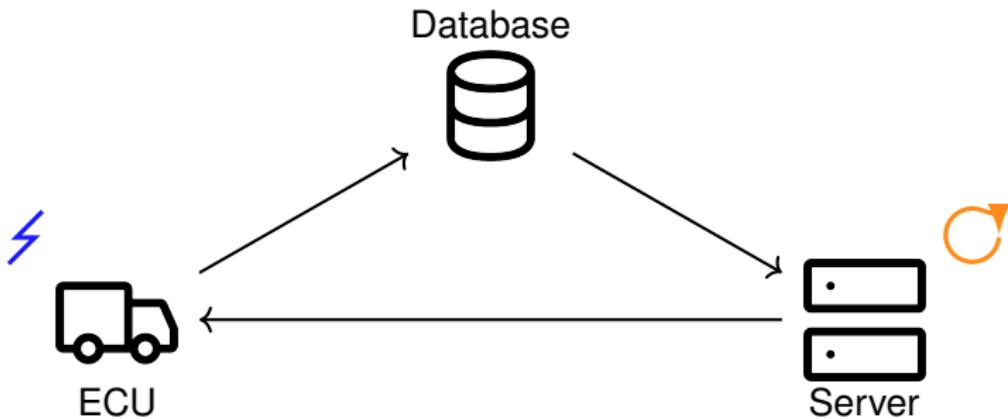
Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# Federated Learning

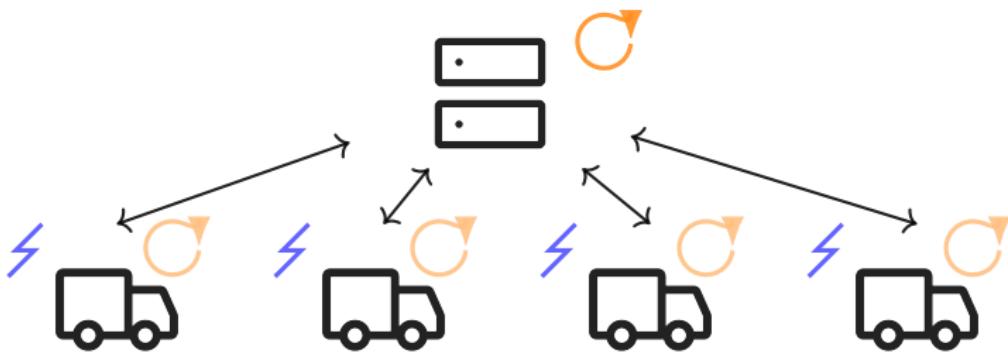
Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# C based HDR-NN

[github.com/paperbound/hdrnn-arm](https://github.com/paperbound/hdrnn-arm)

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

```
/*-HDR-Neural-Network-*/
typedef struct
{
    float *bias;
    float **weights;
    float **nabla_w;
} Neuron;

typedef struct LayerT
{
    int size;
    int incidents;
    Neuron **neurons;
    float **activations;
    float **z_values;
    float **nabla_b;
    struct LayerT *next;
    struct LayerT *previous;
} Layer; // Network.layers.except.for.input

typedef struct
{
    Layer **layers;
    int depth;
} Network; // HDRNN
```

```
static void feed_forward(Network *network, float *image)
{
    // First layer is the image
    float *activations = image;

    Layer *layer = network->layers;
    while (layer != NULL)
    {
        for (int i = 0; i < layer->size; i++)
        {
            Neuron *neuron = &layer->neurons[i];
            float zvalue = 0;
            for (int j = 0; j < layer->incidents; j++)
            {
                zvalue += neuron->weights[j] * activations[j];
            }
            layer->z_values[i] = zvalue + neuron->bias;
            layer->activations[i] = sigmoid(layer->z_values[i]);
        }
        activations = layer->activations;
        layer = layer->next;
    }
}
```



# HDR-NN configurability

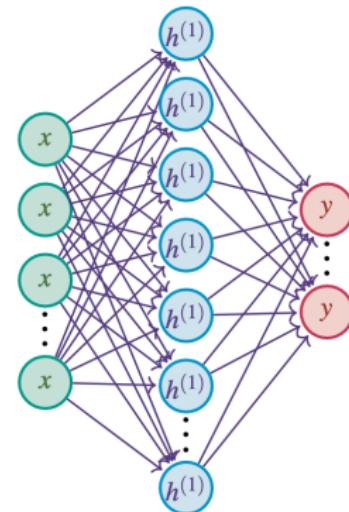
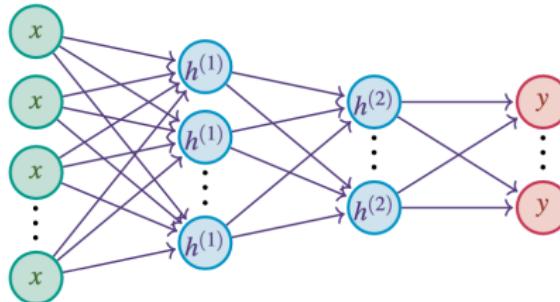
Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion





# HDR-NN configurability

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

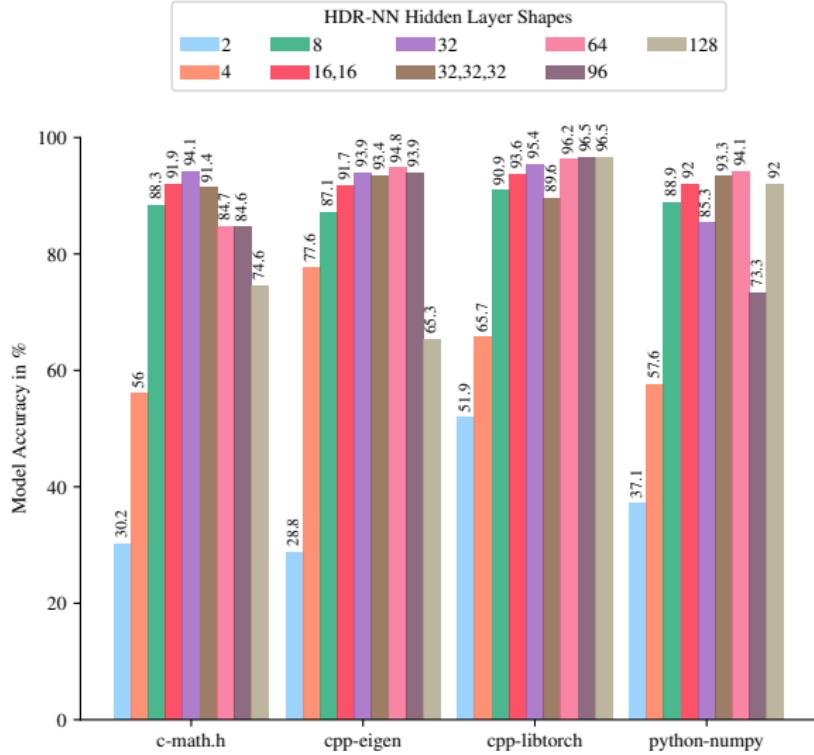
```
→ ./bin/hdrnn
nothing to do
Usage: ./bin/hdrnn

      <command> [<args>]

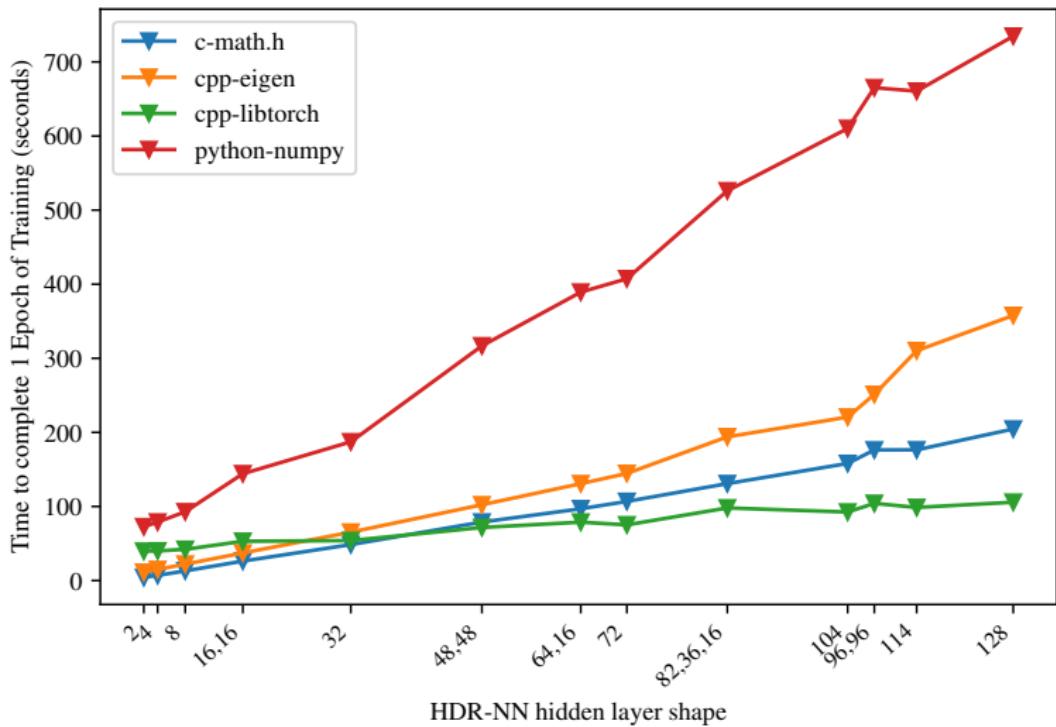
Commands:
  infer [-i, --image IMAGE_PATH] [-n, --net c-math.nn]
  train [-s, --shape 32] [-e, --epochs 30]
        [-q, --quiet] [-bs, --batch_size=10]
        [-lr, --learning_rate 3] [-n, --net c-math.nn]
```



# Accuracy



# Training Time





# Development Journey

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

- Linux, software packages for MCIMX6Q
- Implementing back propagation in C
- Implementing HDR-NN in C, C++ / Eigen
- Compiling libtorch for armv7hl



# Software Development on Embedded Devices

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

- Embedded Platforms are fragmented in both *hardware* and *software*
- Challenging to advance software
- Research Approaches : Tiny ML, Tiny Training Engine



# Future Work

---

Neural  
Networks in  
Embedded  
Devices

Embedded  
Linux

Neural  
Networks

Benchmark  
Applications

Discussion

- Reverse Engineering Efforts on the C300
- Performance engineering the c based implementation
- Adding more target devices
- Adding more implementations (TTE)