

Extending AlFES with Depthwise Convolution for On-Device Training on Microcontrollers

Author: Zico Krassenburg
z.i.k.krassenburg@student.tudelft.nl

Responsible professor: Marco Zuñiga Zamalloa

Supervisor: Hao Liu

1. Introduction

- AI is increasingly used in embedded systems like smart doorbells and wildlife monitoring cameras.
- Microcontrollers have a lot of restrictions like computational power, power usage
- MobileNet is a neural network architecture specifically designed for these resource constrained devices. It enables tasks such as image classification and object detection while requiring significantly fewer computations than traditional convolutional neural networks.
- A key component of MobileNet is depthwise convolution (DepthConv), which reduces the number of parameters, computational operations, memory usage, and energy consumption.
- AlFES (Artificial Intelligence for Embedded Systems) is an open-source framework that enables machine learning directly on microcontrollers and even supports on-device training but currently lacks the depthconv operator.

2. Research question

How can we extend AlFES to add DepthConv operators?

3. Methodology

Analyze AlFES Conv2d layer Implement DepthConv2d operator Test on Arduino Nano

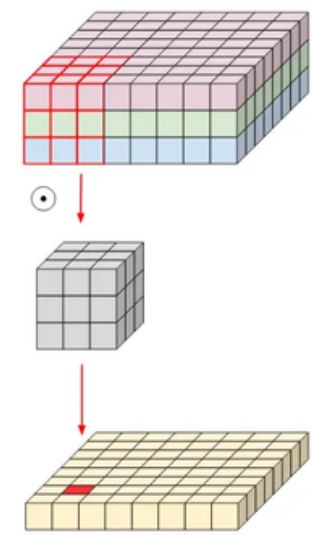


Figure 1: Convolution

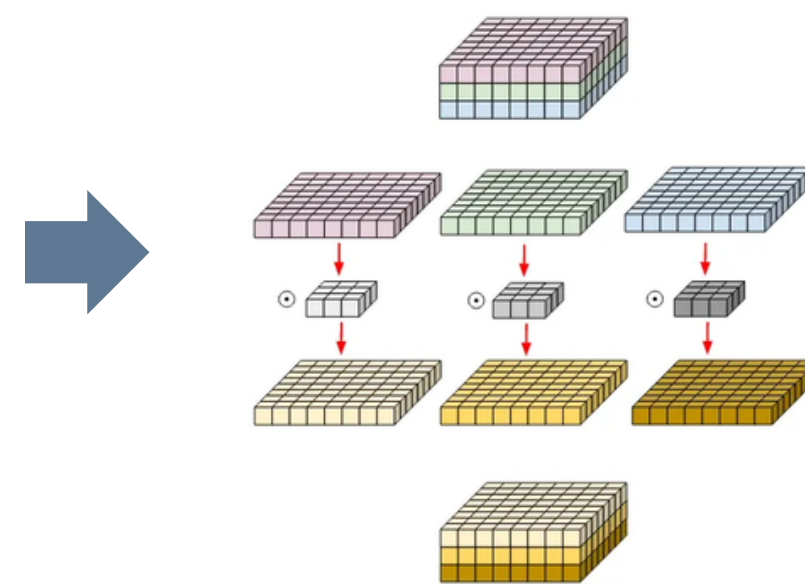


Figure 2: Depthwise Convolution

In order to test the implementation i made use of a synthetic dataset consisting of 8x8 images with 3 input channels containing horizontal, vertical and diagonal lines

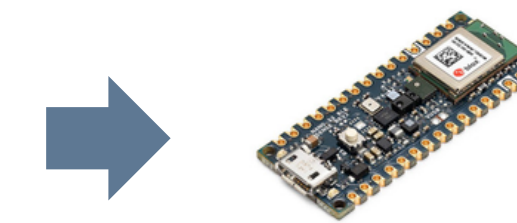


Figure 3: Arduino nano 33 BLE



Figure 4: Example image of the synthetic dataset

Both experiments where run for 250 epochs using the adam optimizer of AlFES with learning rate of 0.01

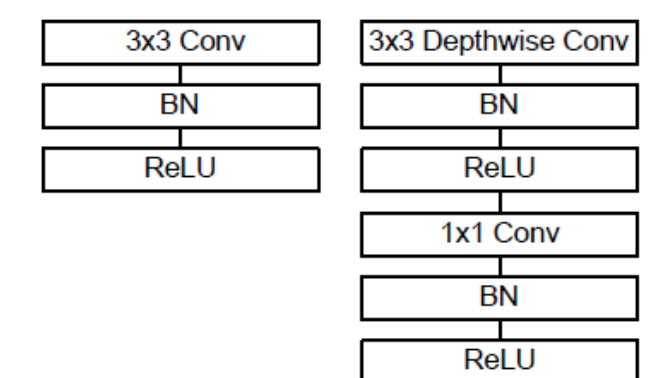


Figure 5: Architecture for testing based of a MobilenetV1 block. The convolution block(left) and the depthwise seperable block(right)

4. Results

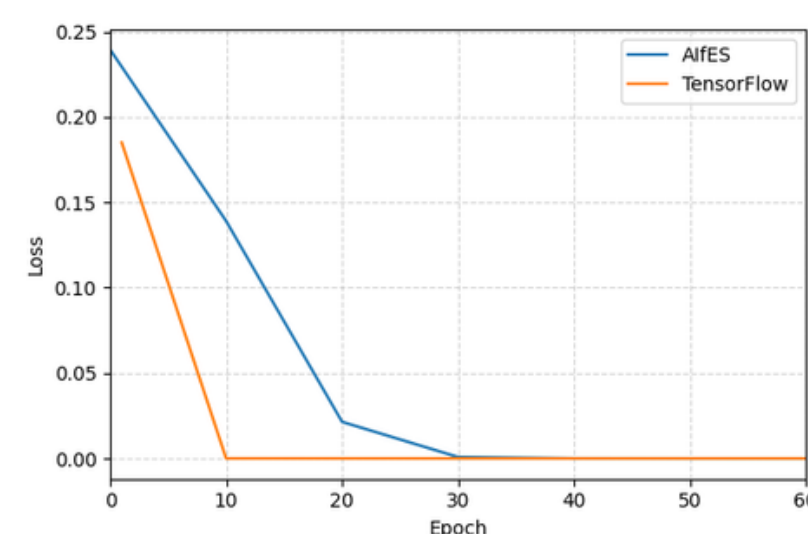


Figure 6: loss vs epoch of AlFES and tensorflow setup

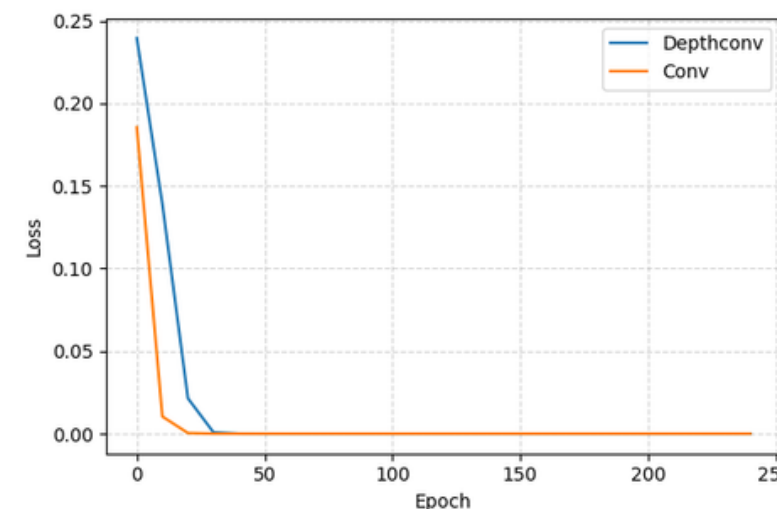


Figure 7: loss vs epoch of Conv architercture and Depthconv architecture

Architecture	Parameters	Training Memory
Conv2D	2948 B	50108 B
DepthConv + PW Conv	2300 B	60572 B

Table 1: Memory usage of Conv architercture and Depthconv architecture

For the memory comparison the parameters consist of the trainable weights and biases while training memory consists of intermediate results and gradients

5. Conclusion

- Result indicate that training works as the loss decreases over epochs
- The tests show that the Depthconv layer was able to learn the classification task
- The Depthconv layer can be used to create mobilenet architectures in AlFES

6. Future work

- AiFES is still missing more advanced techniques used by later versions of Mobilenet like branching
- The depthconv layer can be tested on a bigger dataset and a bigger neural network on a more capable microcontroller