Real-Time Traffic Sign Recognition on Microcontrollers

Aykut Emre Celen I Supervisors: Qing Wang, Ran Zhu

EEMCS, Delft University of Technology

Introduction

- Real-time traffic sign recognition on microcontrollers introduces challenges due to **limited memory and processing capacity**.
- This study investigates the trade-offs between **model size, classification accuracy, and inference latency** within hardware constraints.
- We present **AykoNet**, an efficient network architecture for traffic sign recognition specifically optimized for **Raspberry Pi Pico**.

Figure 1: Raspberry Pi Pico

Research Question

How can we create an optimal TinyML [1] model for real-time traffic sign recognition on microcontrollers?

Background

- MobileNetV1 [2] is an efficient CNN that introduces **depthwise separable convolutions** for mobile devices.

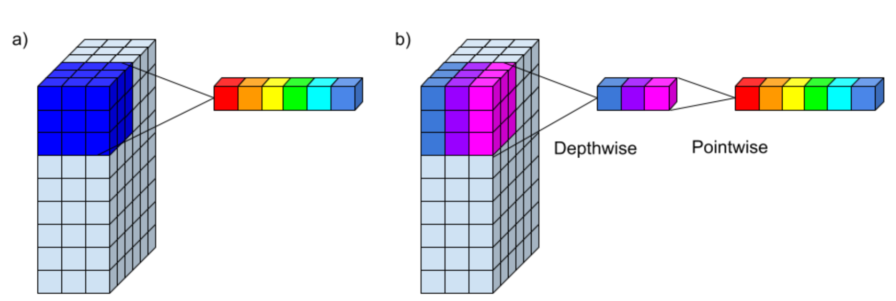

Figure 2: Standard convolution (a) and Depthwise separable convolution (b)

- GiordyNet [3] is a traffic sign recognition model with **domain-specific preprocessing**.

Motivation

No existing architecture combines domain-specific optimization with computational efficiency

Methodology

Data and Preprocessing

- AykoNet is trained on the **GTSRB** dataset.
- **43 classes** with class imbalance ratio of **1:11**
- All images converted to **grayscale** and resized to **32×32 pixels**, as shown below.

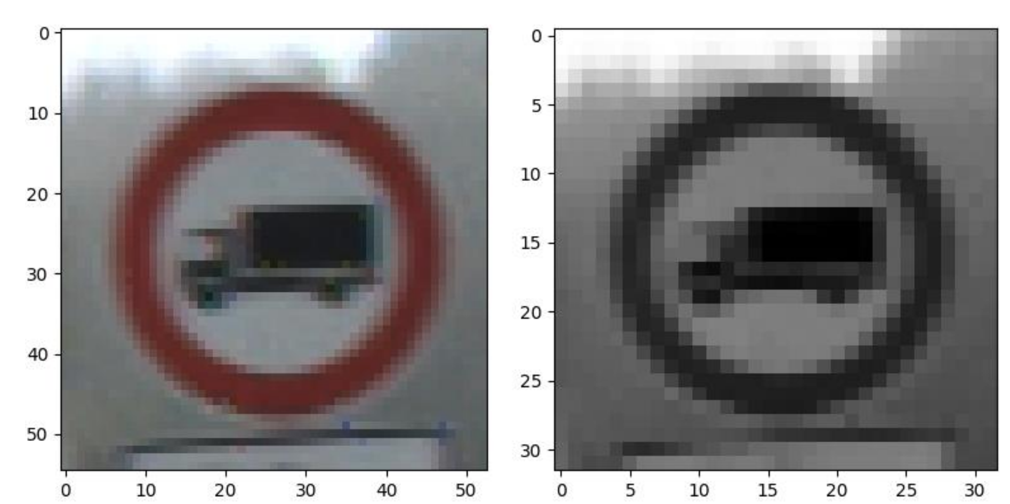

Figure 3: Preprocessing of an image

Class-Aware Data Augmentation

- **Strategy:** Apply proportionally to class size
- **Techniques:** Rotation, translation, shearing, gamma correction

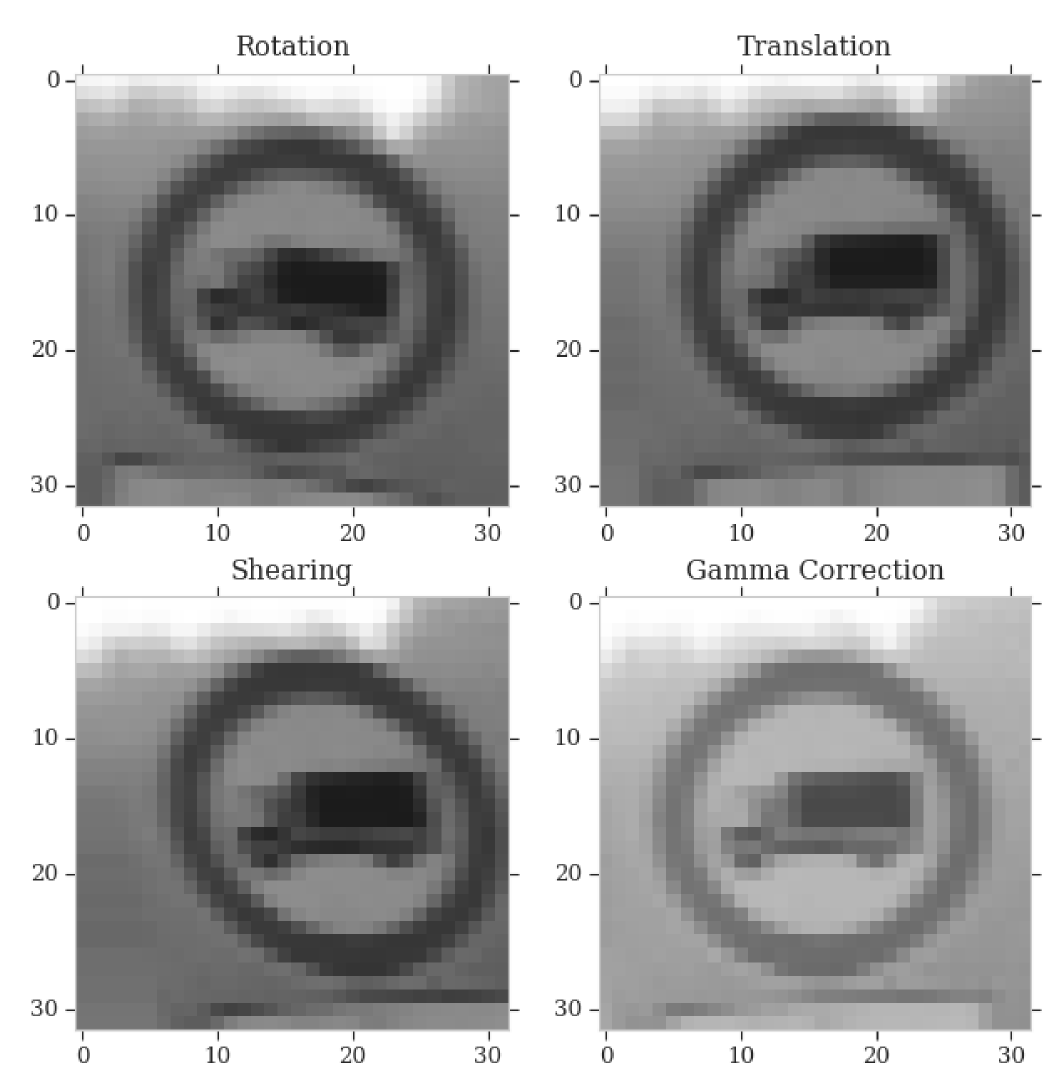

Figure 4: Effects of different techniques

Architecture

- Depthwise separable convolution blocks
- Power-of-two channel progression (8→16→32→64→128)
- AykoNet-Lite prioritizes minimal **model size** and fast **inference**
- AykoNet-Pro prioritizes **classification accuracy**

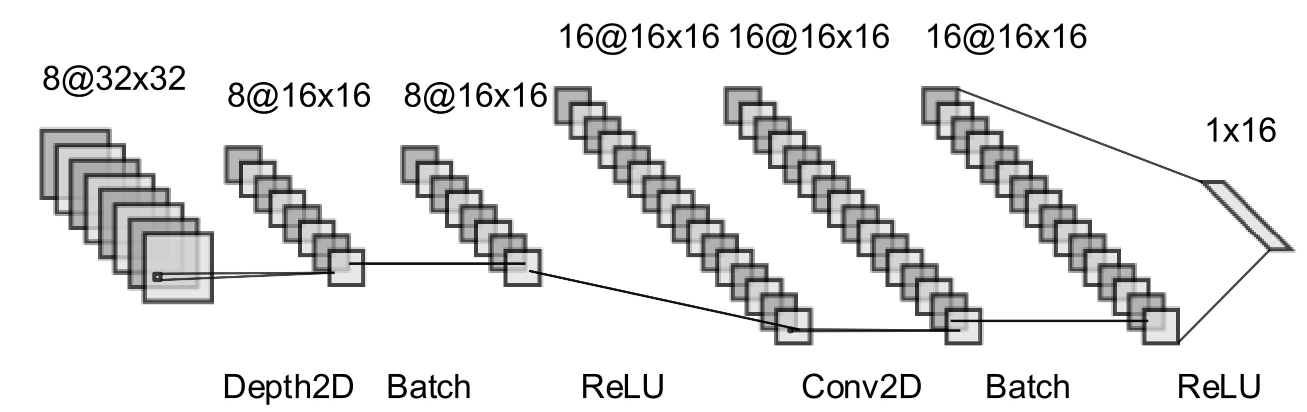

Figure 4: First depthwise block in AykoNet-Lite (8→16)

Results

Data Augmentation

- **Reduced class imbalance** from 1:11 to 1:3 ratio
- **Total dataset size** increased by 57.4%

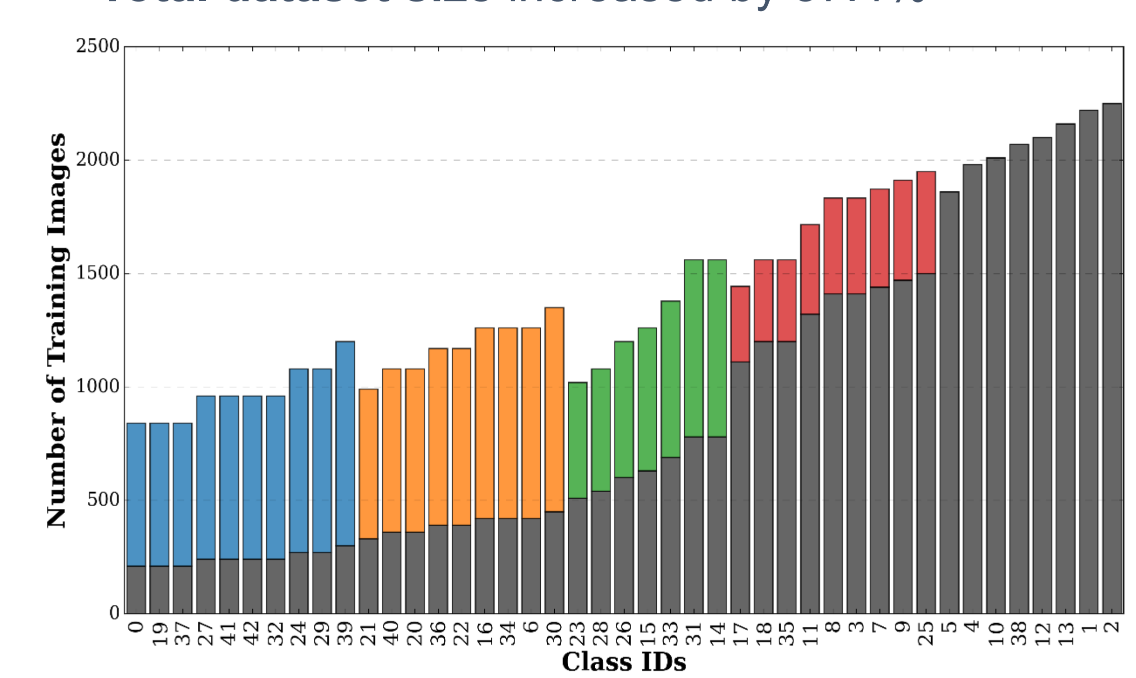

Figure 5: Class distribution before and after augmentation. Blue +300%, Orange +200%, Green +100%, Red +30%.

Performance Comparison

- AykoNet-Lite delivers the **smallest** model size (36.80KB) and the **fastest** inference time (55.34ms).
- AykoNet-Pro achieves the **highest** accuracy (95.90%).

Model	Size (KB)	Accuracy	Time (ms)
MobileNetV1_25-int8	307.59	87.50%	-
MobileNetV1_20-int8	217.79	79.80%	77.29
GiordyNet-int8	106.87	95.50%	204.08
AykoNet-Lite-int8	**36.80**	94.60%	**55.34**
AykoNet-Pro-int8	80.18	**95.90%**	87.13

Table 1: Performance comparison

Conclusion

AykoNet's performance demonstrates the effectiveness of:

- **domain-specific preprocessing**
- **class-aware data augmentation**
- **depthwise separable convolutions**
- **channel progression optimization**

for creating an optimal TinyML model for real-time sign recognition on microcontrollers.

Specifically**, AykoNet-Lite** strikes an optimal balance for practical deployment.

Our results validate the feasibility of real-time traffic sign recognition in resource-constrained embedded systems.

References

[1] Pete Warden and Daniel Situnayake, *TinyML: Machine Learning with TensorFlow Lite on Arduino and UltraLow-Power Microcontrollers*, 2019.

[2] A. G. Howard *et al.*, *MobileNets: Efficient convolutional neural networks for mobile vision applications*, 2017

[3] M. Giordano, *Traffic Sign Recognition, CNN on Microcontrollers*, 2020.