



Lightning-Fast Modulation Classification with Hardware-Efficient Neural Networks

BacalhauNet Team

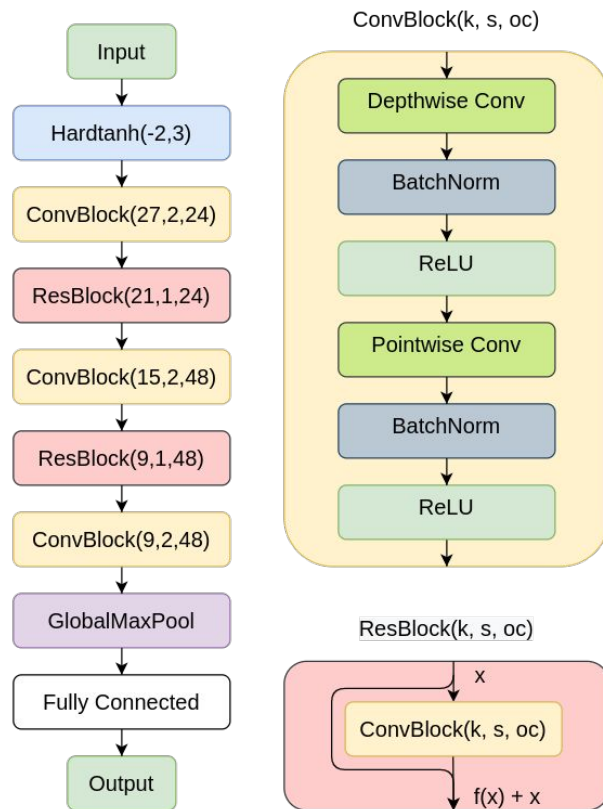
José Nuno Grácio Rosa¹, Guilherme Carvalho², Daniel Granhão², Tiago Filipe Gonçalves²

¹Escola Superior de Tecnologia e Gestão - Politécnico de Leiria

²INESC TEC and Faculty of Engineering - University of Porto

Architecture

- A new architecture based on commonly used network structures;
 - Depthwise separable Convolutions;
 - Residual connections.
- Architecture obtained through **design space exploration**;
- **Kernel length, stride length and number of output channels** adjusted for each layer.





Quantization

- BacalhauNetV1 with **floating-point (FP)** inputs, weights and activations achieved a **good initial inference cost score**;
- **Quantization** of input, weights and activations down to **5 bits** were tested;
- **Bit-width of 6** selected due to **good compromise** between accuracy and inference cost.

Bit Width	Test Accuracy Reached	Inference Cost Score
FP	59.09%	1.4155
7 bit	58.35%	0.1002
6 bit	58.67%	0.0781
5 bit	55.89%	0.0562



Pruning

- Both structured and unstructured pruning methods were tested;
- **Unstructured pruning** was selected as that is the method **most favored** by the challenge evaluation method;
- **Several iterations of sparsity inducing training + prune** were performed;
- Exploration of 2 variables:
 - **Weight Decay**;
 - **Minimum Weight Absolute Value.**

Step	Weight Decay	Min. Weight Abs. Value	Accuracy	Inference Cost Score
Original	0.0001	-	57.62 %	0.0735
1st prune	0.00005	0.15	58.26 %	0.0348
2nd prune	0.0001	0.25	57.55 %	0.0235
3rd prune	0.00005	0.25	56.24 %	0.0162



Thank You

BacalhauNet Team

