XiNet: Efficient Generative Networks for Edge Devices

FONDAZIONE BRUNO KESSLER

ICT DAYS

E3DA ENERGY EFFICIENT EMBEDDED DIGITAL ARCHITECTURES

XiNet

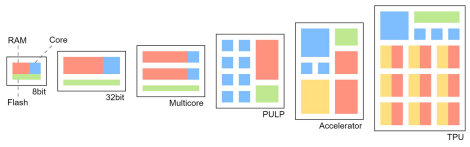
Challenging the **efficiency** of depthwise convolutions for **edge and tinyML**:

- Novel convolutional block optimizing **latency** and **energy** usage
- **Benchmarking** on multiple embedded platforms
- **Hardware Aware Scaling**: from hardware constraints to neural architecture

Hardware-Aware Scaling

Three main computational constraints in different embedded devices:

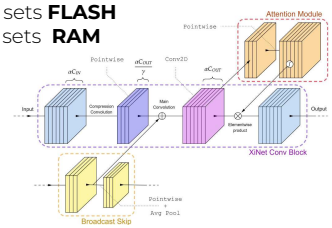
- **FLASH**: stores network parameters
- **RAM**: stores intermediate tensors
- **MAC/s**: determines latency & energy



Convolutional Block

Designed from **real world** efficiency measurements on various platforms.
Three hyperparameters:

- **α**: sets **MAC**
- **β**: sets **FLASH**
- **γ**: sets **RAM**



Object/Pose Detection

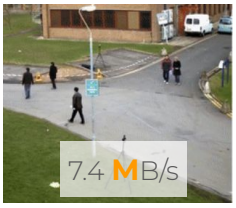
Typical pipeline:

Video transmission → Analysis (Large neural networks)

Privacy problems
High bandwidth
Not scalable

Expensive, power hungry

Reducing bandwidth by 3 orders of magnitude!



On-device:

Move processing to the edge device

XiNet Pose + Analysis (Simple neural networks)

Privacy, Low bandwidth, Scalable

Low power, simpler processing

Results

Networks scaled using **Hardware Aware Scaling**:

MCU:	STM32 -	100MMAC/s,	2MB Flash,	1MB Ram
TPU:	K210 -	1GMAC/s,	16MB Flash,	5MB Ram
MPU:	rPi 4B -	16GMAC/s,	16GB SD,	4GB Ram

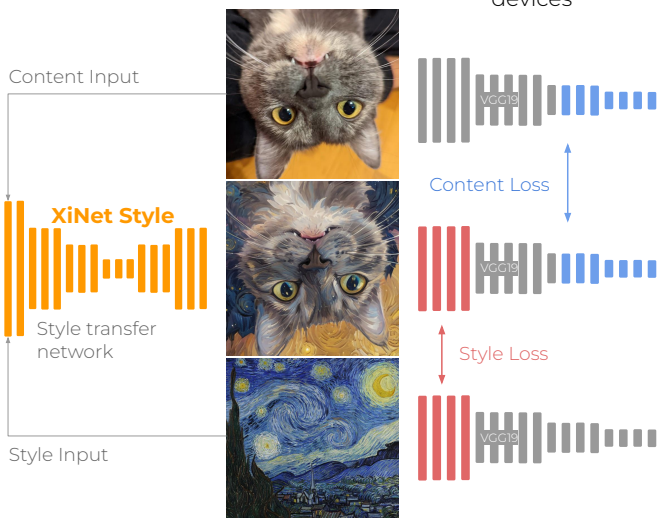
Raspberry Pi 4B:
α= 1.0 β= 1.0 γ=4.0
-Speed: **63.6 fps**
-Power: **14.89 W**

Kendryte K210:
α= 0.75 β= 1.0 γ=4.0
-Speed: **31.7 fps**
-Power: **410 mW**

ST STM32H743:
α= 0.33 β= 0.8 γ=5.0
-Speed: **5.5 fps**
-Energy: **72.4 mW**

Style transfer

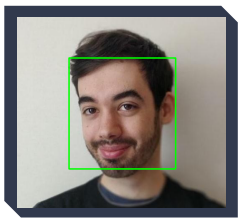
Efficient **image generation** on edge devices



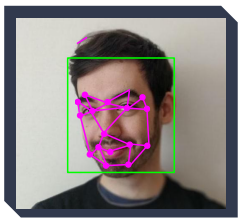
One shot image generation can be used for **anonymization** while preserving semantic content - removing personal information for **downstream tasks**

XimSwap

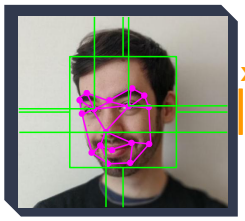
Single target face swapping in 4 steps using 3 networks:



1. Face Detection — XiNet + Yolo



2. Landmark Detection — XiNet + PFLD



3. Face Alignment — XiNet + PFLD

XiNet FS

4. Face Generation — **XiNet GAN**

Ancilotto, A., Paissan, F & Farella, E.
XiNet: Efficient Neural Networks for tinyML
International Conference on Computer Vision - ICCV 2023

Micromind
TinyML made simple
github.com/micromind-toolkit/micromind