Towards Real-time Object Detection on Mobile Devices

Byeongkyo Cheong*, Donghee Ha, Jinse Kwon, Hyungshin Kim
Department of Computer Science and Engineering
Chungnam National University
Daejeon, South Korea

RTSS 2018
Nashville, TN, USA
We challenge to run CNN on mobile devices in **REAL** real-time.
What We Did

Current State-of-the-art
- NVIDIA CUDA
- Darknet / TensorFlow / Caffe Framework
- Matrix Multiplication Library for CNN (CUDA)

Our Work
- Without Framework
- For all kinds of Mobile GPUs
- Matrix Multiplication Library for CNN (OpenCL)

Our Library vs CLBlast
- 1.3 times faster than CLBlast
- Non-Square Matrix Optimization
At Our Demo

Input: Camera

Input: VOC Images

Mobile Device

ARM Big.LITTLE

Mali-T628 MP6 GPU

CNN

Real Detection Challenge!

Result

Real Detection

Exynos 5422

Samsung Exynos 5 Processor