Automotive Vision Sensor

Deep Learning – Object Detection

(Convolutional Neural Network Engine on ZYNQ)

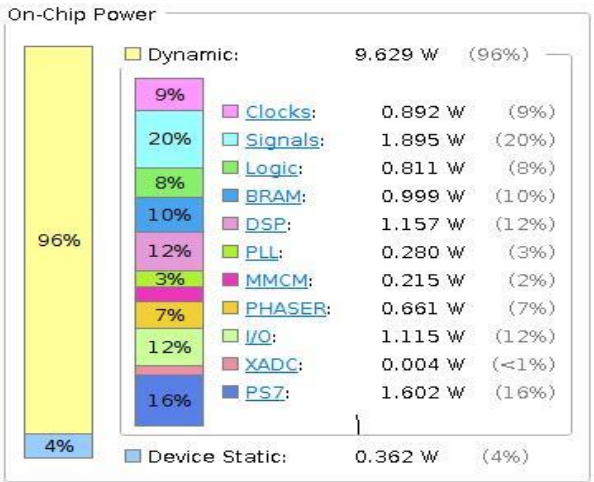
Across The Universe

❑ Convolutional Neural Network on ZYNQ Programmable SoC

- Convolutional Neural Network (CNN) achieves the state-of-art performance in object detection for the automotive camera system.
- High computation complexity in both inference and training, which needs specific hardware to accelerate.
- GPUs are an excellent alternative to performance for the deep-running algorithm, but because of the enormous power requirements, they are limited to embedded systems.
- There is a need for a processing platform capable of accelerating algorithms without excessively increasing power consumption.
- The 'Programmable SoC is an ideal alternative because it has a unique ability to easily handle multiple parallel with low power attributes and rapid replacement of advanced algorithms can be easily upgraded in-filed.

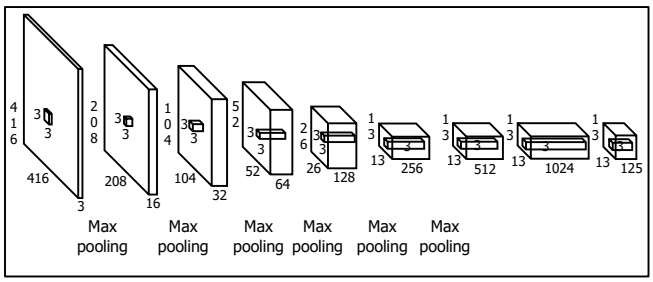
❑ ZYNQ Power Consumption

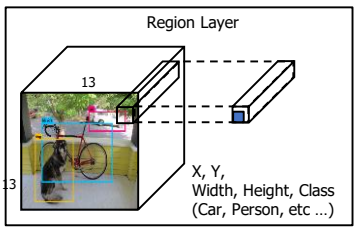
- Total On-Chip Power : 10.432 W
- Junction Temperature : 42.7
- Clock : 200 MHz (PL), 666 MHz(PS)



❑ Object Detection and Classification Model

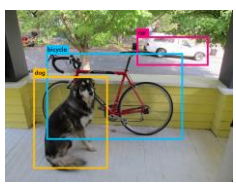
(Tiny YOLO : You only look once)

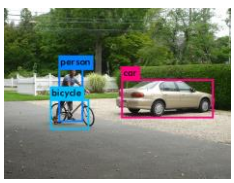




- Model : YOLO (Tiny)
- Framework : Darknet
- Precision : Fixed Point
- Train Data : VOC 2007 + 2012
- Test Data : VOC 2007

Model	mAP	FLOPS
SSD300	74.3	-
SSD500	76.8	-
YOLOv2	76.8	34.90 Bn
YOLOv2 544x544	78.6	59.68 Bn
Tiny YOLO	57.1	6.97 Bn

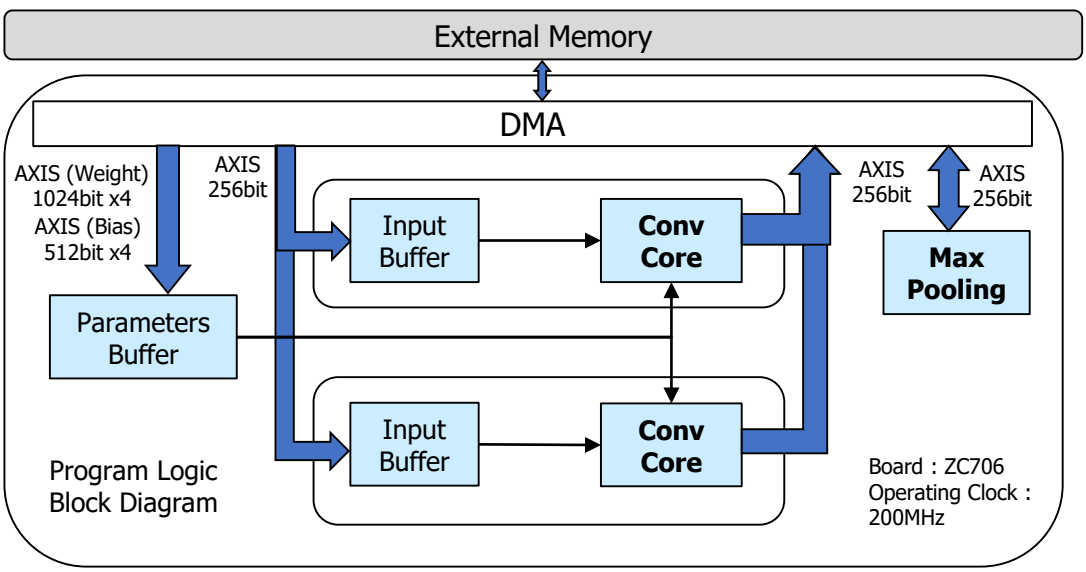




•VOC 2007 + 2012 Class

- ✓ Aero plane
- ✓ Bicycle
- ✓ Bird
- ✓ Boat
- ✓ Bottle
- ✓ Bus
- ✓ Car
- ✓ Cat
- ✓ Chair
- ✓ Cow
- ✓ Dining table
- ✓ Dog
- ✓ Horse
- ✓ Motorbike
- ✓ Person
- ✓ Potted plant
- ✓ Sheep
- ✓ Sofa
- ✓ Train
- ✓ TV monitor

❑ BLOCK DIAGRAM



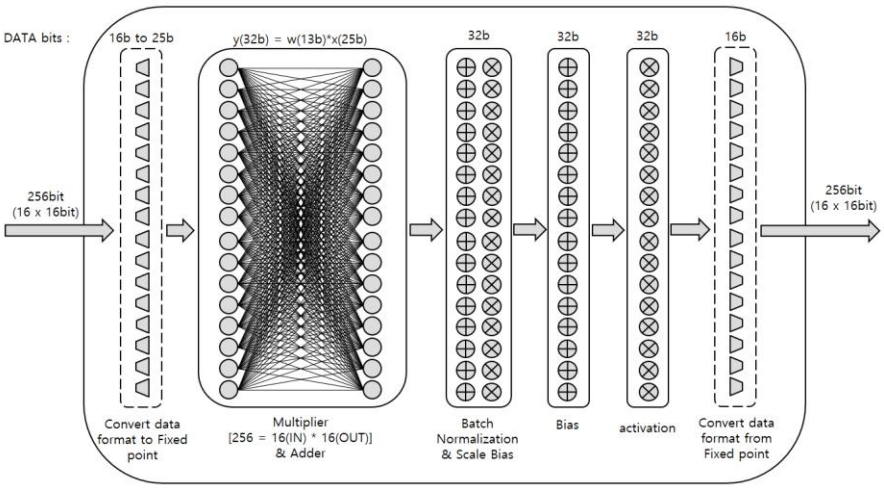
- Input / Output Bus Width
  Number of input : 16
  Number of output : 16
  Data bit : 16bit
  Bus Width = 256bit
  (16bit x 16)
- Parameters
  Weight
  Bias
  Batch Normalization
  (Mean, Variance, Scale Bias)

https://youtu.be/Ahhf1D5x9bM

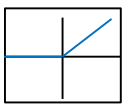
Copyright © 2017 ATUS Inc. All rights reserved.

Automotive Vision Sensor
Deep Learning – Object Detection
(Convolutional Neural Network Engine on ZYNQ)

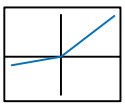
❑ Conv Core (Convolution Unit)



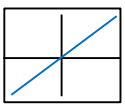
- Support Activation
  - ReLU
  - Leaky
  - Leaner



- ReLU



- Leaky



- Leaner

• Input Buffer
✓ DATA INPUT : 256bit (16 * 16bit)
✓ DATA OUTPUT : 256bit (16 * 16bit)
✓ Reduce read data from external memory

• Parameters Buffer
✓ Buffer Size
- Weight : 1024 * 3328bit (16[I] * 16[O] * 13bit)
- Mean : 64 * 512bit (16 * 32bit)
- Variance & Scale Bias : 64 * 512bit (16 * 32bit)
- Bias : 64 * 512bit (16 * 32bit)

• Conv Core
✓ DATA INPUT : 256bit (16 * 16bit)
✓ DATA OUTPUT : 256bit (16 * 16bit)
✓ Convolution Layer : 256 multiplier per clock (16[IN] * 16[OUT])
✓ Support Batch Normalization & Scale Bias
✓ Support Bias operation
✓ Support Activation List (ReLU / Leaky / Leaner)

• Max Pooling
✓ DATA INPUT : 256bit (16 * 16bit)
✓ DATA OUTPUT : 256bit (16 * 16bit)

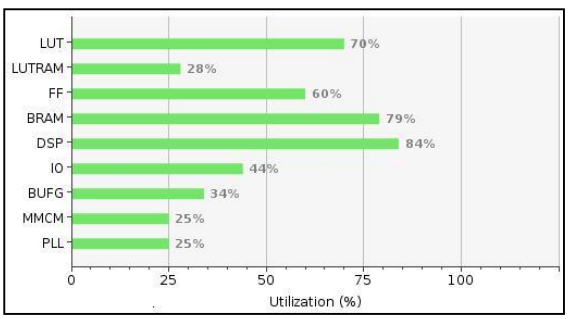
❑ Run Time and Performance

• Operating Clock : 200MHz
• Performance : 17.53 fps (57.028 msec/image)

Index		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Layer		Conv	Max	Conv	Max	Conv	Max	Conv	Max	Conv	Max	Conv	Max	Conv	Conv	Conv	Total
Performance Estimates	Max Latency (Clock Cycles)	173086	86547	389406	43283	389406	21651	389406	10835	389406	5427	389406	6291	1677342	3354654	46622	-
	Time (msec)	0.865	0.433	1.947	0.216	1.947	0.108	1.947	0.054	1.947	0.027	1.947	0.031	8.387	16.773	0.233	36.862
Real	Time (msec)	1.933	0.856	3.698	0.494	2.974	0.253	2.051	0.111	2.244	0.069	3.427	0.087	12.818	25.626	0.387	57.028

❑ Logic Utilization

Target Device : XILINX ZYNQ XC7Z045-2FFG900 [Vivado HLS 2016.4]



❑ CNN Evaluation Kit (Xilinx ZC706 Base)

https://youtu.be/Ahhf1D5x9bM

Copyright © 2017 ATUS Inc. All rights reserved.