Deep Learning Visual Sensor for Industrial Applications

Jerry Byungik Ahn
[]> Neurocoms Inc.
2018.7.21
Contents

• About the presenter
• Comparison of mobile deep learning hardware
• Our hardware architecture
• Deep Runner Visual Sensor
• Training deep learning models
• Industrial applications
• Conclusion
About the Presenter

• Developer
• Computer architect
  • we value efficiency (speed/resource) rather than just speed
• Founder / CEO of Neurocoms Inc.

CPU  
GPU  
New Architecture
A Comparison of Mobile DL Hardware

- NVIDIA GTX980M GPU: 25.4 FPS
- NVIDIA Jetson TX1 GPU: 3.3 FPS
- Qualcomm Snapdragon 820 with GPU: 5 FPS
- Neurocoms Deep Runner: 30 FPS

SSD300/MobileNet Object Detection

> 122 Watts

< 10 Watts
Hardware Architecture: Neuron Machine


"Computation of Deep Belief Networks Using Special-Purpose Hardware Architecture", IEEE IJCNN, 2014


"Izhikevich Spiking Model, An FPGA achieved 1200x over best PC

"Hodgkin-Huxley Neuromorphic Model, FPGA 1241x over Xeon PC"
Hardware Architecture: Neuron Machine

• Key ideas
  • Computational circuit same as the computation model (the shape of neuron)
  • Special memory circuit: (1) No inter-stage data movement required, (2) large number of slow memories

• Other properties
  • All self-contained in hardware: no processor involved
  • Fully pipelined and no idle clock cycle for arithmetic operators
  • High utilization of multipliers (see next page)
  • Sort of CISC computer - Each instruction for one CNN layer
## Multiplier Utilization Comparison

<table>
<thead>
<tr>
<th>Hardware</th>
<th># of cores (multipliers)</th>
<th>Clock freq. (GHz)</th>
<th>Peak speed (AxB, Gops)</th>
<th>Optimal FPS (C/1.27)</th>
<th>Actual FPS</th>
<th>Multiplier utilization (E/D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTX980M</td>
<td>1536</td>
<td>1.038</td>
<td>1594</td>
<td>1255</td>
<td>25.5</td>
<td><strong>2.03 %</strong></td>
</tr>
<tr>
<td>Jetson TX1</td>
<td>256</td>
<td>1.68</td>
<td>430</td>
<td>338.7</td>
<td>3.3</td>
<td><strong>0.96 %</strong></td>
</tr>
<tr>
<td>Deep Runner</td>
<td>256</td>
<td>0.2</td>
<td>51.2</td>
<td>40.3</td>
<td>29.5</td>
<td><strong>73.20 %</strong></td>
</tr>
</tbody>
</table>

1) 1.27 Giga operations are required for a single SSD300/MobileNet inference

---

### Xilinx 7-Series FPGA

- **Maximum Capability**
  - Logic Cells
  - Block RAM
- **Lowest Power and Cost**
  - 12 Mb
- **Industry's Best Price/Performance**
  - 34 Mb
- **Industry's Highest System Performance**
  - 65 Mb

### Resource Utilization

- LUT: 87%
- DSP: 75%
Deep Runner Visual Sensor device

• Industrial device with built-in deep learning algorithms
  • To be used as a component for building intelligent systems
  • No prior knowledge of deep learning is required for user

• Supports multiple DL algorithms
  • GoogLeNet (classification)
  • YOLO, Tiny YOLO (detection)
  • SSD/MobileNet (detection)
  • MobileNet, Xception (classification)
Deep Runner Products

Deep Runner Module
- 4 x 5cm size
- Mounted as a part on user’s PCB
- Input video signal: YCbCr4:2:2
- Recognition result: Ethernet, Serial port
- Power consumption: 5 watts

Deep Runner Device
- Video Input: HDMI (1920x1080@30, 1600x900, 1280x720)
- Recognition Result: Ethernet, Serial port, GPIO pinout
- Simultaneous recognition of up to 16 cameras from split screen
- Power consumption: 8 watts

Deep Runner CCTV
- For CCTV Surveillance
- Input: IP cameras
- Recognition Result: Ethernet
- Simultaneous recognition of up to 8 IP cameras
- Video recording function
Training Procedure

- Deep Trainer
  - Windows software
  - Train classification algorithms
- Darknet
  - Train YOLO and Tiny YOLO object detection algorithms
- TensorFlow
  - Train SSD object detection algorithm
Deep Trainer
## Reference Users

<table>
<thead>
<tr>
<th>Customers</th>
<th>Area</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>CCTV Surveillance</td>
<td>Show Room</td>
</tr>
<tr>
<td>N</td>
<td>CCTV Surveillance</td>
<td>Highway</td>
</tr>
<tr>
<td>S</td>
<td>CCTV Surveillance</td>
<td>CCTV Monitoring Center</td>
</tr>
<tr>
<td></td>
<td>CCTV Surveillance</td>
<td></td>
</tr>
<tr>
<td>V, Turkey</td>
<td>CCTV Surveillance</td>
<td></td>
</tr>
<tr>
<td>O, Japan</td>
<td>CCTV Surveillance</td>
<td>Security Camera</td>
</tr>
<tr>
<td>U</td>
<td>CCTV Surveillance</td>
<td></td>
</tr>
<tr>
<td>I, Spain</td>
<td>CCTV Surveillance</td>
<td>Cloud system</td>
</tr>
<tr>
<td>H, Poland</td>
<td>Quality Inspection</td>
<td></td>
</tr>
<tr>
<td>W, China</td>
<td>Quality Inspection</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Quality Inspection</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Automotive</td>
<td>Digital Room Mirror</td>
</tr>
<tr>
<td>H</td>
<td>Automotive</td>
<td>Around view of Excavators</td>
</tr>
<tr>
<td>H</td>
<td>Automotive</td>
<td>Around view of Excavators</td>
</tr>
<tr>
<td>E, A Univ., S Univ.</td>
<td>Research, Education</td>
<td></td>
</tr>
</tbody>
</table>
Limitation of the use of classification

Because of the softmax function, the output of the classifier does not indicates exact score.
Quality Inspection

• The use of classification
• Special features
  • Find small defects in high resolution product images
• PLC pinout communication

"defect product"

CLASS0
CLASS1
CCTV Surveillance

- The use of object detection
- Special feature
  - Recognize multiple cameras simultaneously
Conclusion

• Embedded deep learning will become mainstream

• As a leading company, we shared
  • Our hardware architecture
  • Device specification
  • Applications

• We are seeking
  • Funding for ASIC
    • Typically 50 times more power efficiency could be achieved with ASIC - 200mW with the same speed as Deep Runner
    • Applicable to millions of CCTV equipment
  • Collaboration projects
  • Recruits
  • Distributors