Person-Tracking Security Camera

Team A5:
Jerry Ding, Nathan Levin, Karthik Natarajan

Application Area

- The primary distinguishing feature of our security camera system is the ability to use optical zoom and tracking to more clearly show a person’s face.

- The product, in one sentence:

  A compact and self-contained security camera that automatically tracks and zooms into any suspicious person, and that an average store or homeowner can easily install and use.

Solution Approach

- “Automatically tracks and zooms ”
  - A user interface for moving the camera is insufficient. We use computer vision algorithms.
- “Compact and self-contained”
  - Central server is out of the question. We use a small FPGA known to be a good fit for running computer vision algorithms.
- “Any suspicious person”
  - Multiple targets are possible if they are all suspicious. Add a scoring system to pick the best targets and the amount of time focused on them.

Solution Approach

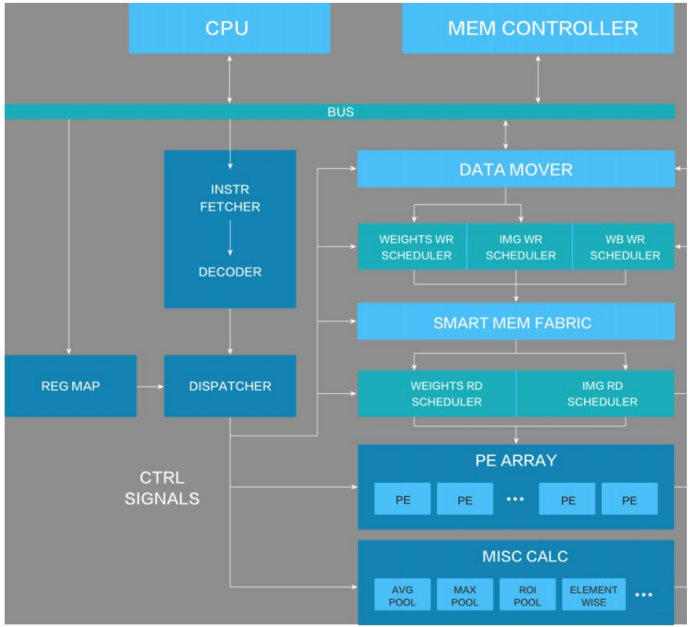
- “Can easily install”
  - Some people but not all people have convenient ways to plug in the camera to wall power. Need to support battery operation.

  - For battery users, minimize the inconvenience of needing to recharge.
    - The competition: ~500 minutes of active operation, ~30 days idle state
    - Easy to recharge without disassembling the whole system.

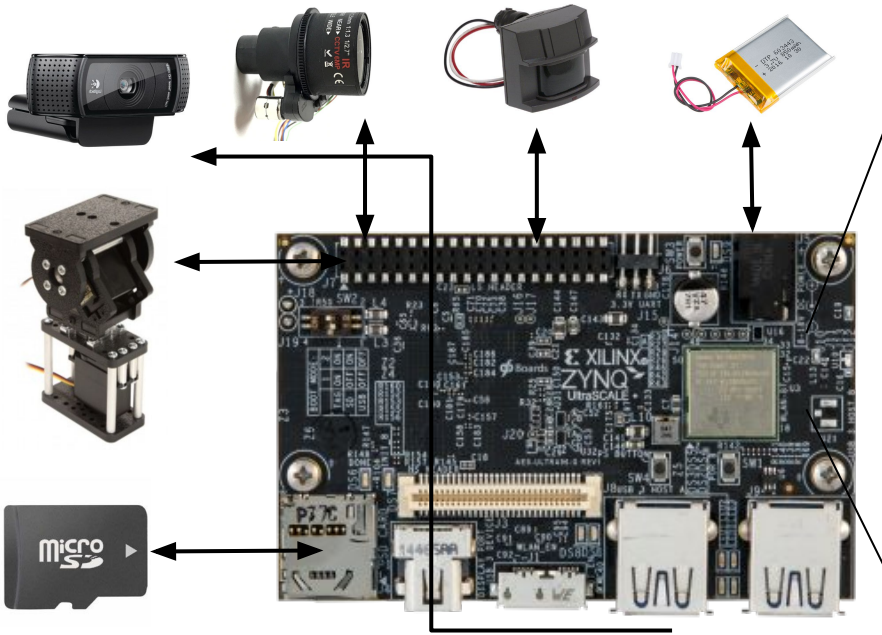
    - Use a removable battery module containing a pair of 5V, 13Ah battery packs.
    - Generally run in low power mode, wake up when activity is detected
    - High sensitivity / spurious wakeups are OK to a certain extent

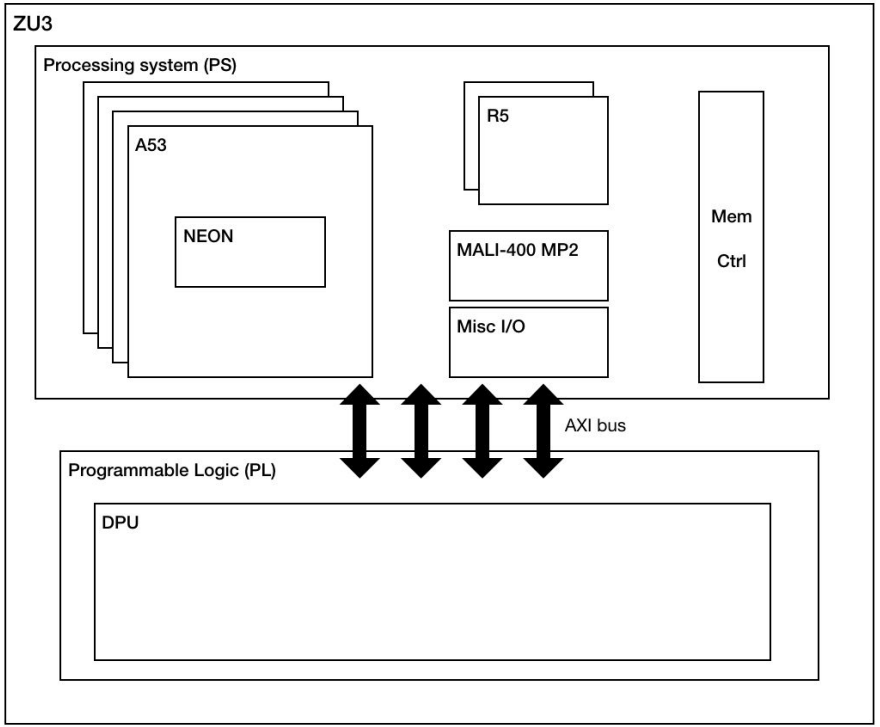
System Architecture

- Implementation of the Deephi Inference Accelerator (B1152F)
- Motivating factors
  - Extremely new ecosystem
    - Room to try unexplored possibilities
  - Robust Xilinx documentation
  - Optimized for low power (edge) inference
  - Highly configurable/customizable

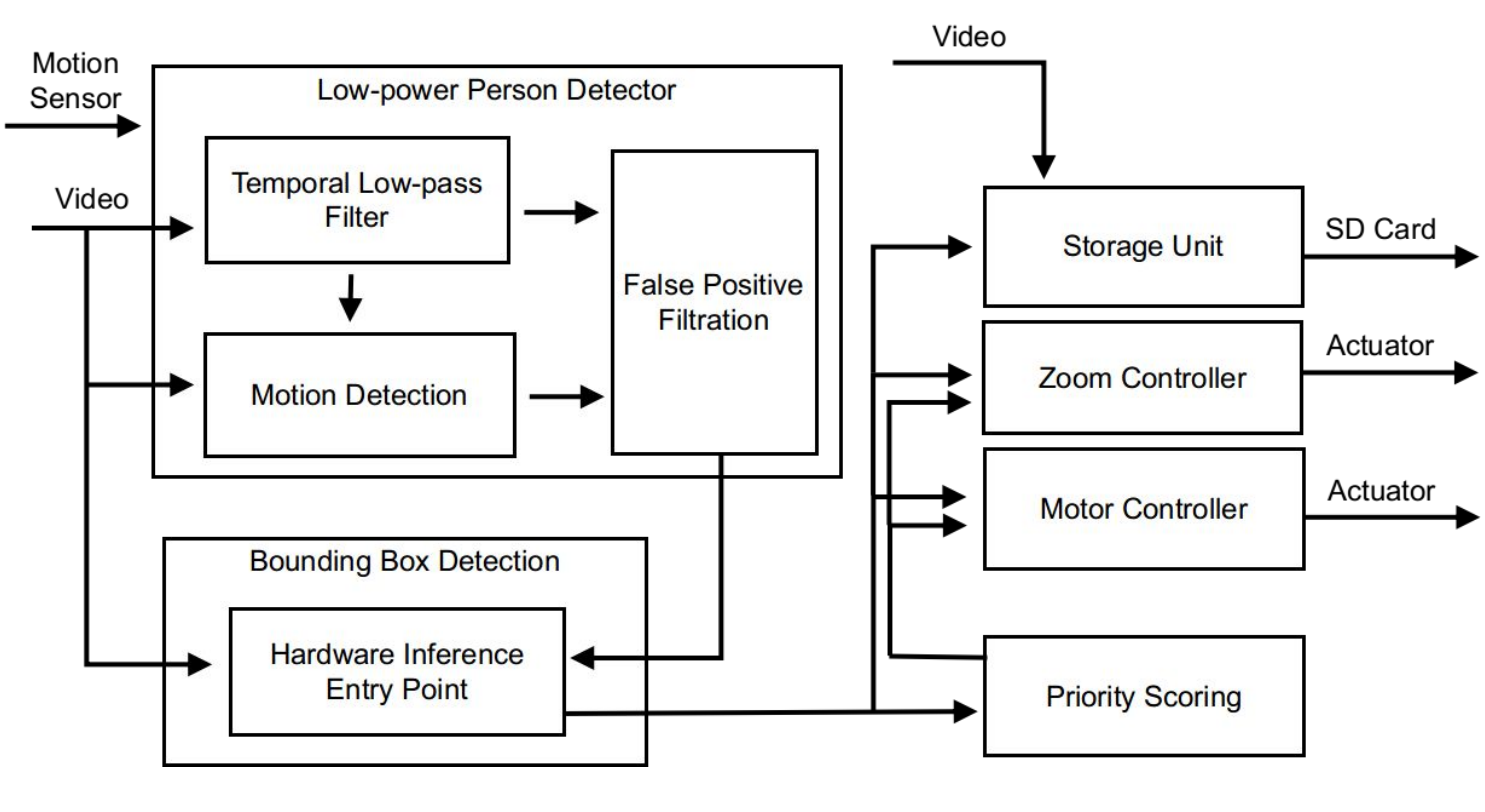


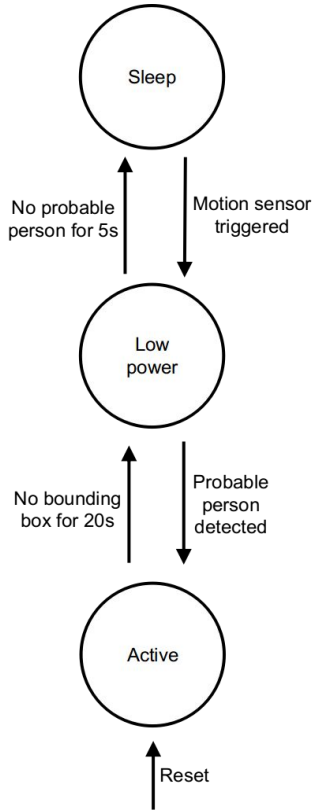
Hardware Block Diagram





Software Block Diagram / State Diagram





Contributions

Hardware

- Off the shelf
  - Ultra96
  - Motion sensor
  - Battery
- Customized
  - Deephi DPU core
  - C920 Pro camera
  - Optics
- New
  - Power control (systems level)
  - Mechatronics
  - Enclosure

Software

- Off the shelf
  - Linux operating system
  - Gstreamer (video streaming)
  - OpenCV
  - Yolo-v3 Tiny
  - Xilinx (Vivado, SDSoC)
  - Deephi DNNDK (inference engine)
- New
  - Low power object detection algorithm
  - Motor control
  - Zoom control
  - Priority scoring
  - Firmware level (sensor interrupts, etc.)

Hardware Utilization - Reference Implementation

	LUT	Slice_reg	Block Ram	DSPs
All logic	70560	141120	216	360
DPU_B1152	36560	68729	115.5	288
Utilization ratio	51.81%	48.70%	53.47%	66.67%

Power consumption of programmable logic ~= **3.5W**
(based on ZU2 implementation)

Performance Baseline

Neural Network	MAC (GOPS)	fps
ResNet-50	7.7	25
Inception-v1	3.2	58
MobileNet	0.56	116
Face detection	1.1	133
Video analysis	5.5	35
Pose detection	5.0	48
ADAS detection	5.5	30
Semantic segmentation	8.8	24

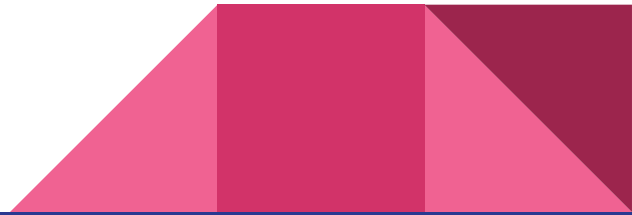
Performance with DPU at 500MHz

Goals:

- Meet performance requirements with greater power efficiency than the reference design.
- Derive performance through methodology, not brute force hardware. Don’t have power to spare.

Metrics

- Success rate for detecting at least one person on time, starting in sleep mode.
  - Unlikely to buy 50 packages in a year, let alone be targeted 50 times in a year
  - Goal: At least 50 trials between failures ⇒ more than 98% success rate
- Percentage of people correctly framed within the bounding box.
  - Goal: Given successful detection of at least 1 person, at least 50 trials between failures, where each failure only omits at most one person when three are in view
- 30 days idle time, 500 minutes active time
  - Idle time includes losses caused by false positives



Schedule & Division of Labor