

Flexible, Ultra-Low Power On-Device AI

Hoon Choi, Hussein Osman, Lattice Semiconductor

With the ever increasing amount of data being generated by scattered sensors, local on device and ultra-low power processing is becoming a requirement especially for battery powered systems. Lattice's FPGAs enable processing on device AI workloads at ultra-low power with a mix of distributed logic, memory and DSP blocks built on top of the parallel FPGA architecture. When married with sensAI solution stack, designers can develop applications faster, with flexible inferencing engines, easy to use compiler tool and end to end example applications. 3 of these applications are showcase this year:

- Run on iCE40 UltraPlus a 5.5 mm² device capable of 1 GOP peak performance
- Use VGG like 8bit compact CNN Network

Human Presence Detection

- On device, accurate human presence detection
- Wake the system up when users are detected
- Input from low power image sensor
- Runs @ 7.7 mW / 5 frames per second
- 128x128 input resolution into the NN
- The neural network can be retrained to detect other objects

Hand Gesture Detection

- On device hand gesture detection (Rock/Paper/Scissors)
- Human-Machine Interfacing using gestures
- Input from IR sensor
- Runs @ 3.3 mW / 5 frames per second
- 64x64 input resolution into the NN
- The neural network can be retrained to detect other gestures

Key Phrase Detection

- On device wake word detection, trained to detect 4 key phrases (Seven, Marvin, On, Happy)
- Wake the system up when the key words are detected
- Input from 2 I2S microphones
- 700 uW when no sound present
- 7.7 mW when sound is present
- 40 evaluations per second
- The neural network can be retrained to detect other wake words