2020 embedded VISION summit®

Machine Learning Based Perception on a Tiny Low-Power FPGA

Hoon Choi
Fellow, R&D
Sept. 2020

LATTICE SEMICONDUCTOR®

iCE40 UltraPlus

- 5.4mm² package
- Sensor fusion
- LUT 5280
- Static current 75uA
- Low power, size, and cost
- ML Inferencing at the edge
- SPRAM 1Mbits
- EBR 120kbits
- MUL 8 of 16x16 / 16 of 8x8

Power and Performance Chart

24Mhz clock	Layers	Resolution	MOPS	FPS	Avg. pwr	FPGA	SPI	Max FPS
Character classification	4,4,1	32	3	10.0	1	1	0	100
Static gesture	6,4,1	32	6	10.0	1	1	0	48
Dynamic gesture	7,4,1	32	8	10.0	2	2	0	35
Human presence	8,4	64	10	1.0	1	1	0	N/A
	8,4	64	32	5.0	7	7	0	10
	8,4	64	32	10.0	13	13	0	10
	8,4	128	191	2.0	19	13	6	2
Front+Shoulder	7,3	64	34	9.5	19	13	6	9.5
	9,4	160	282	1.1	19	13	6	1.1
Place	9,5,1	160	411	0.8	19	13	6	0.8

Layers = CONV & BN, POOL, FC; PWR=mW

For VGG type network. Mobilenet delivers 3~4x FPS

Network Selection & Optimization

Header: YOLO, SqueezeDet, SSD

Base: VGG, Mobilenet

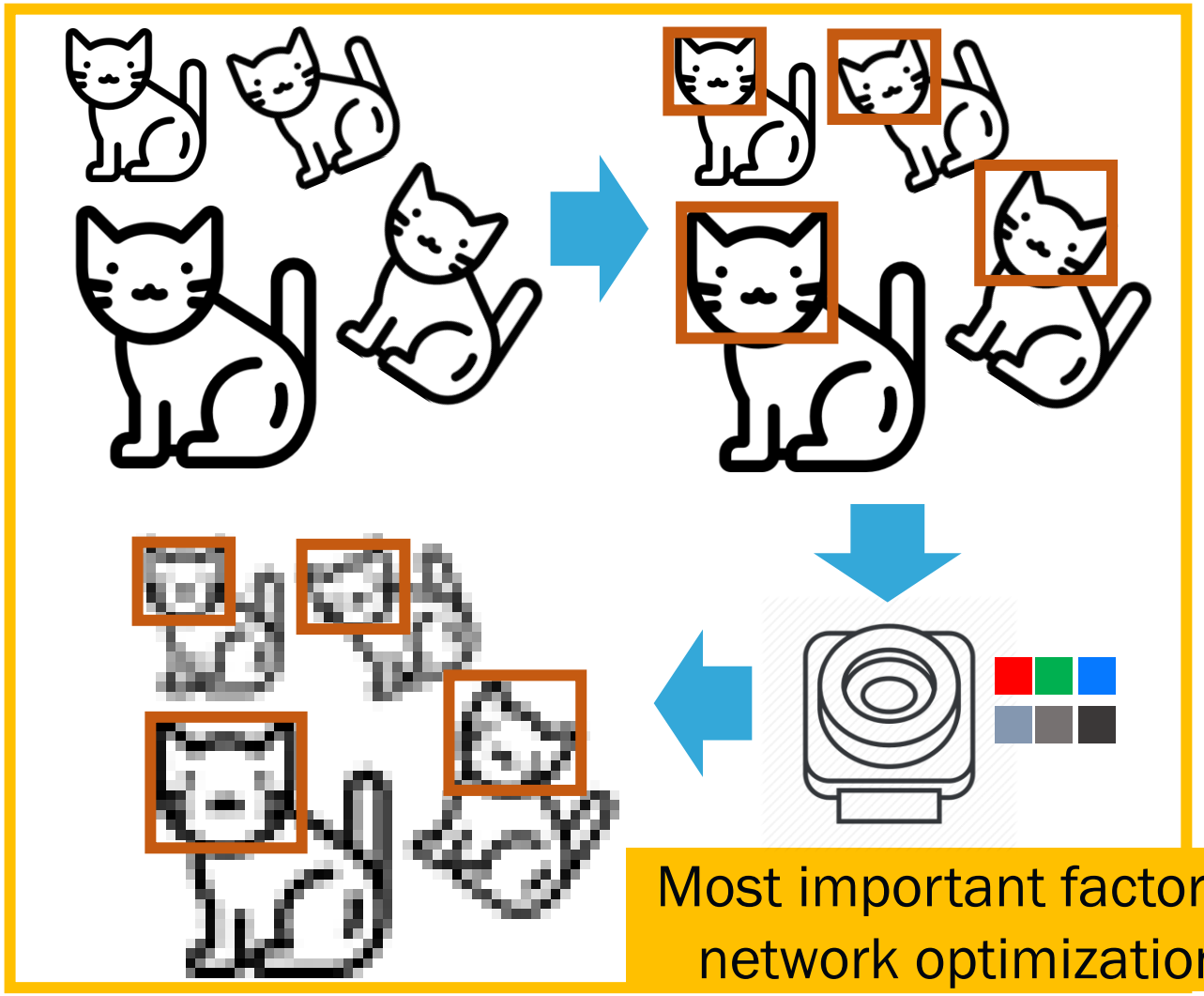
Good for limited resource

Quant.

Depth

Width

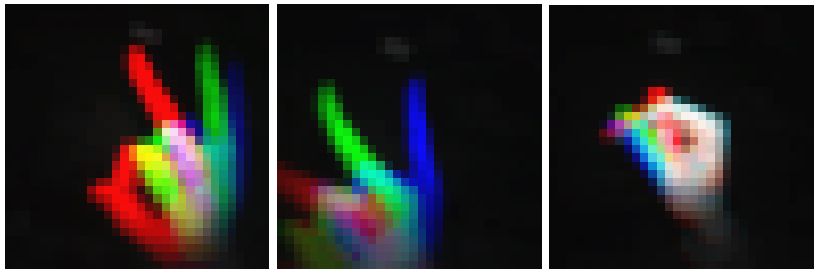
Channel

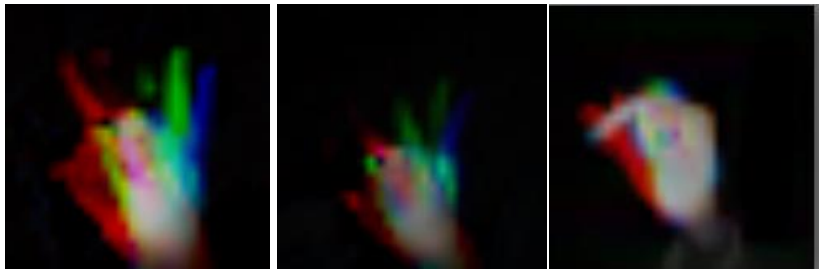


Most important factor in network optimization

LATTICE SEMICONDUCTOR

Dynamic Gesture





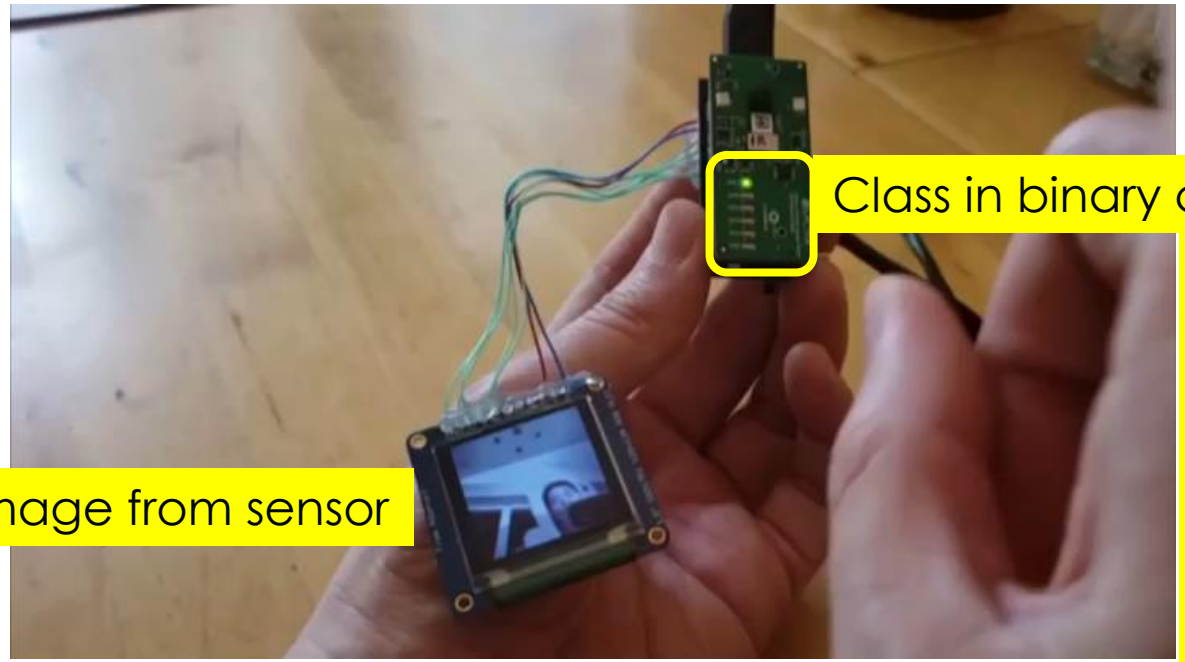


Image from sensor

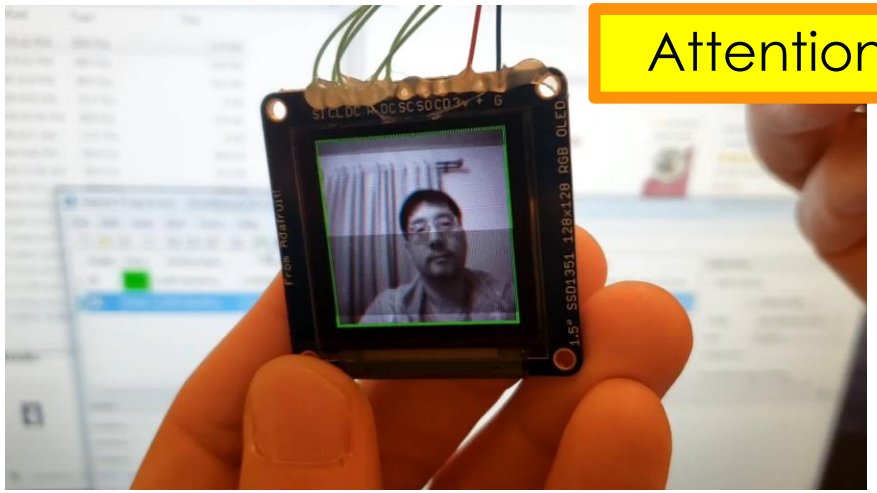
Class in binary code

1: one finger tap
2: two finger tap
3: thumb from knuckle to tip
4: thumb from tip to knuckle
5: index forward
6: index backward
7: up
8: down
9: none of above

Attention Tracking & Shoulder Surfer Detection

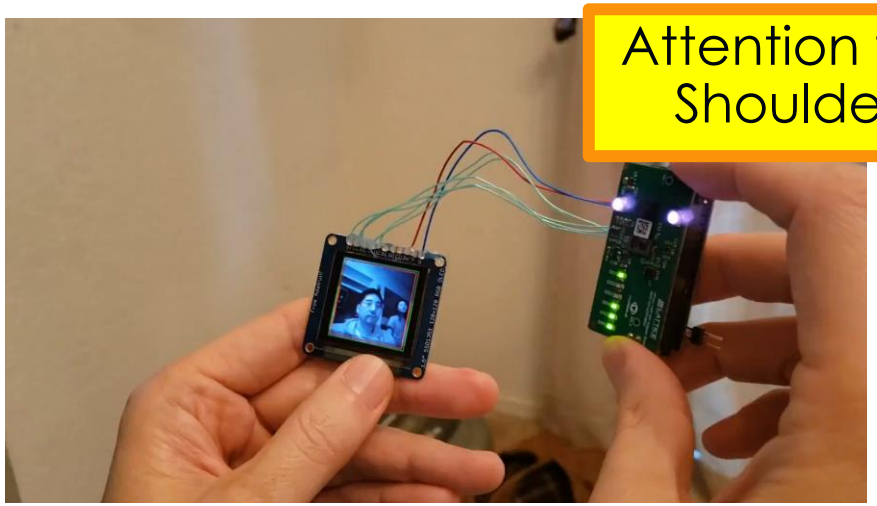
Attention tracking @ECP5

Attention tracking



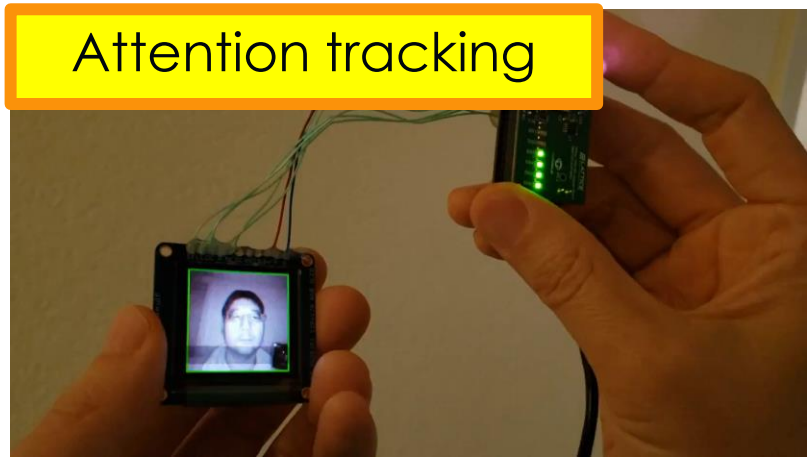
Shoulder surfing detection

Attention tracking + Shoulder surfing

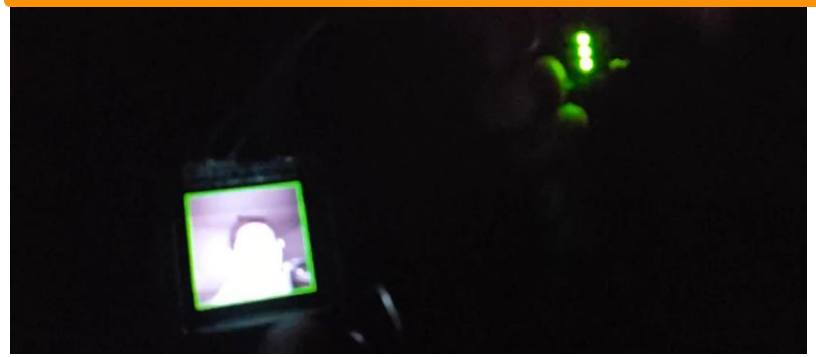


IR-based Attention Tracking

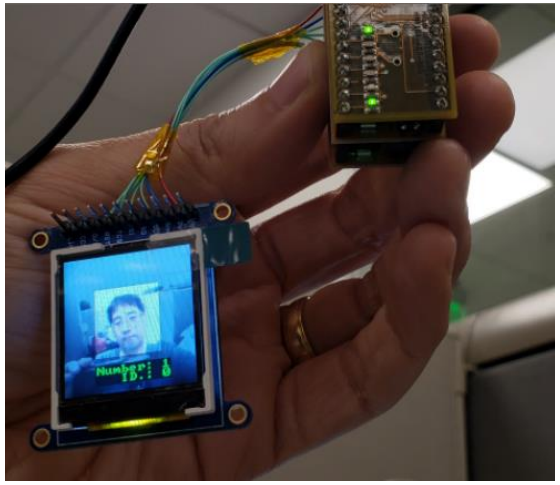
Attention tracking



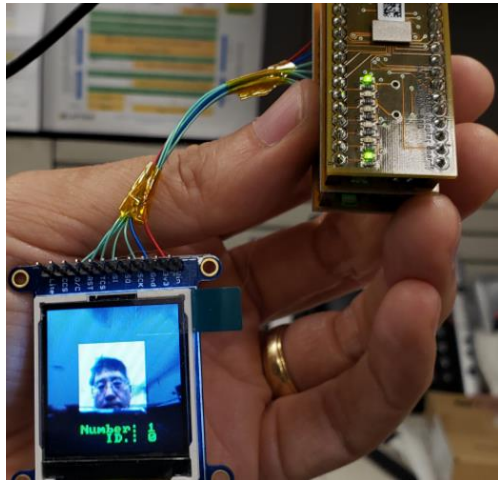
Attention tracking in darkness

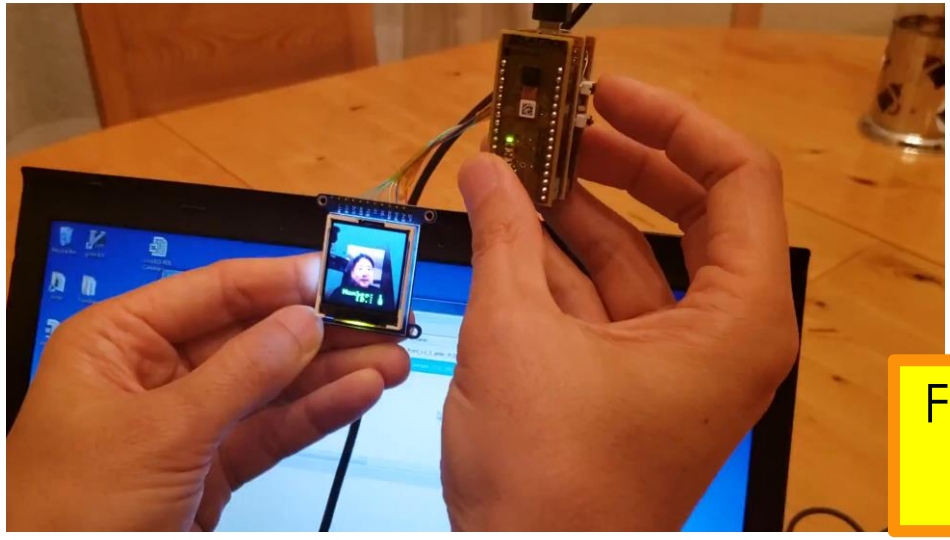


Local Face Recognition + Smile Detection



Recognize a person among registered persons





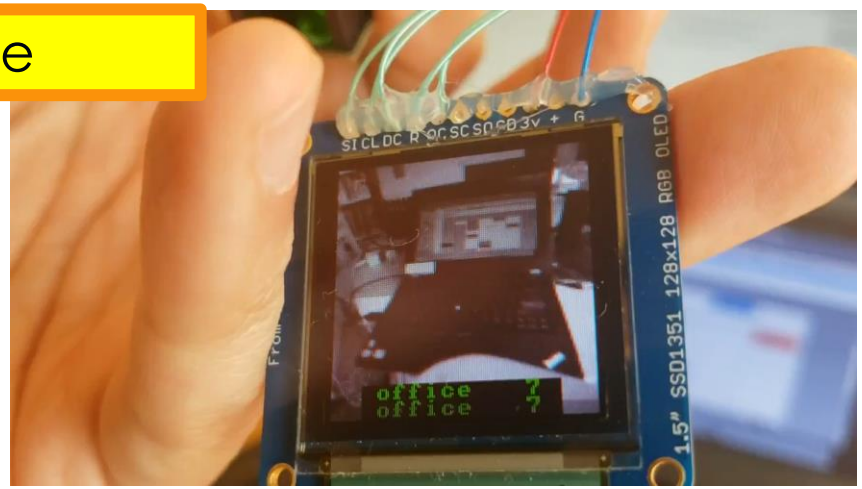
Face recognition and then smile detection

Place Classification (Vision only)

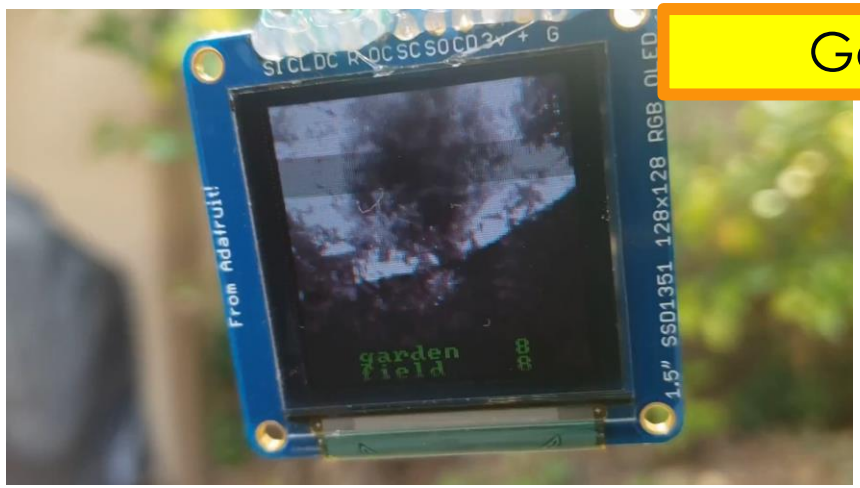
Interior of vehicle

Restaurant

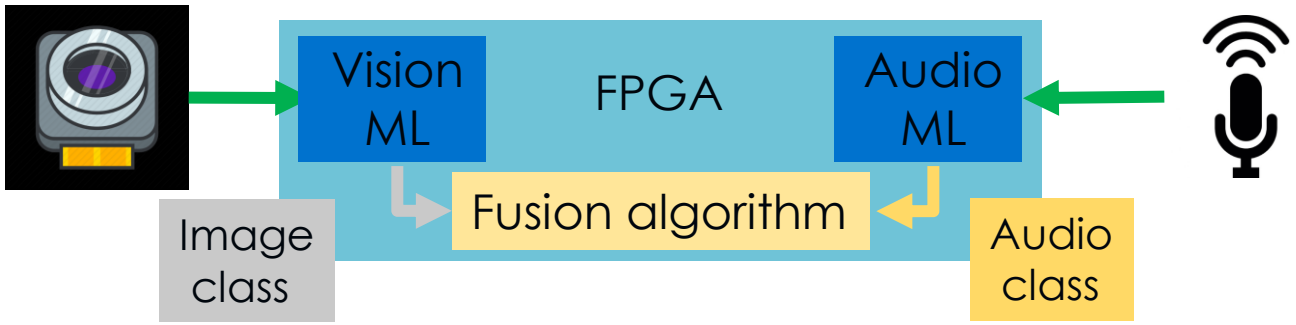
Office

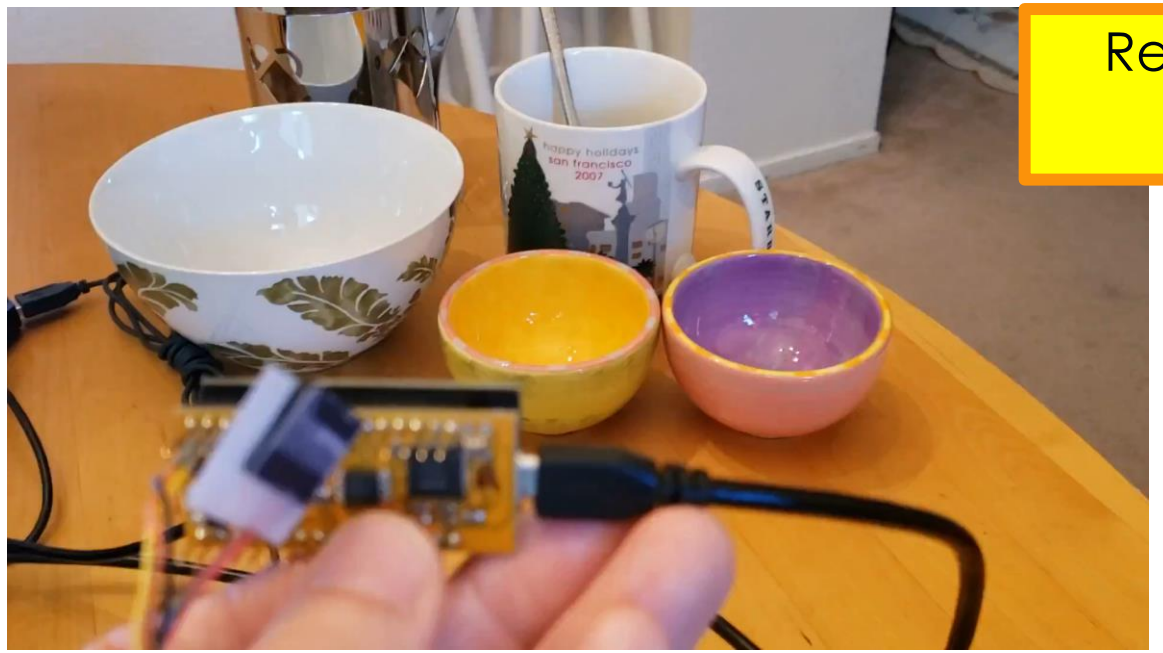


Garden



Place Classification (Vision + Audio)

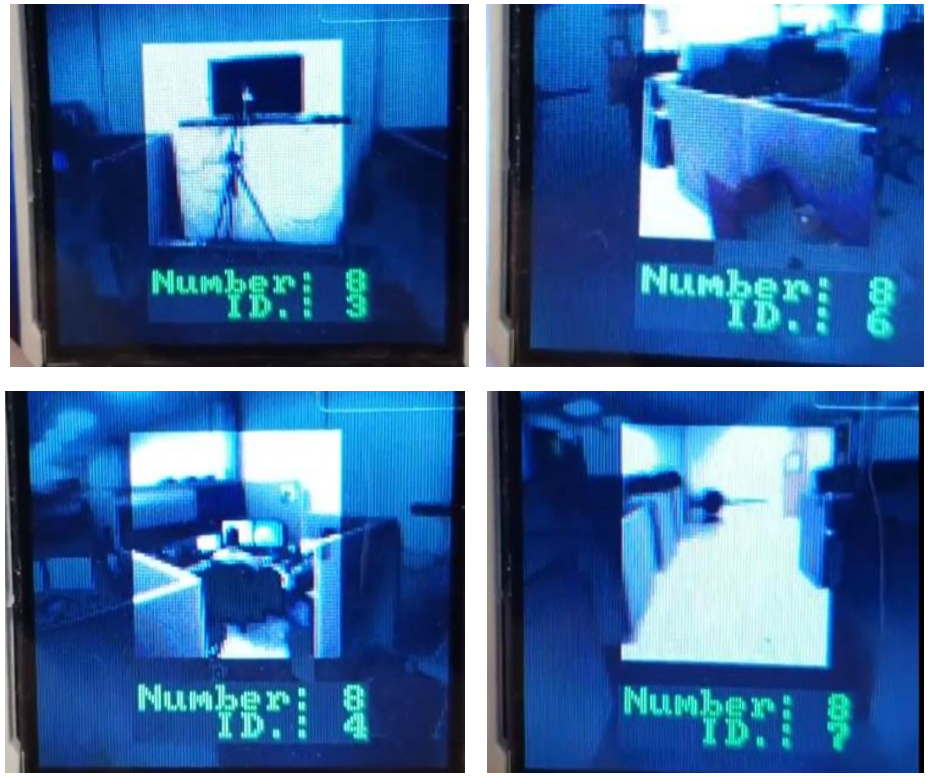




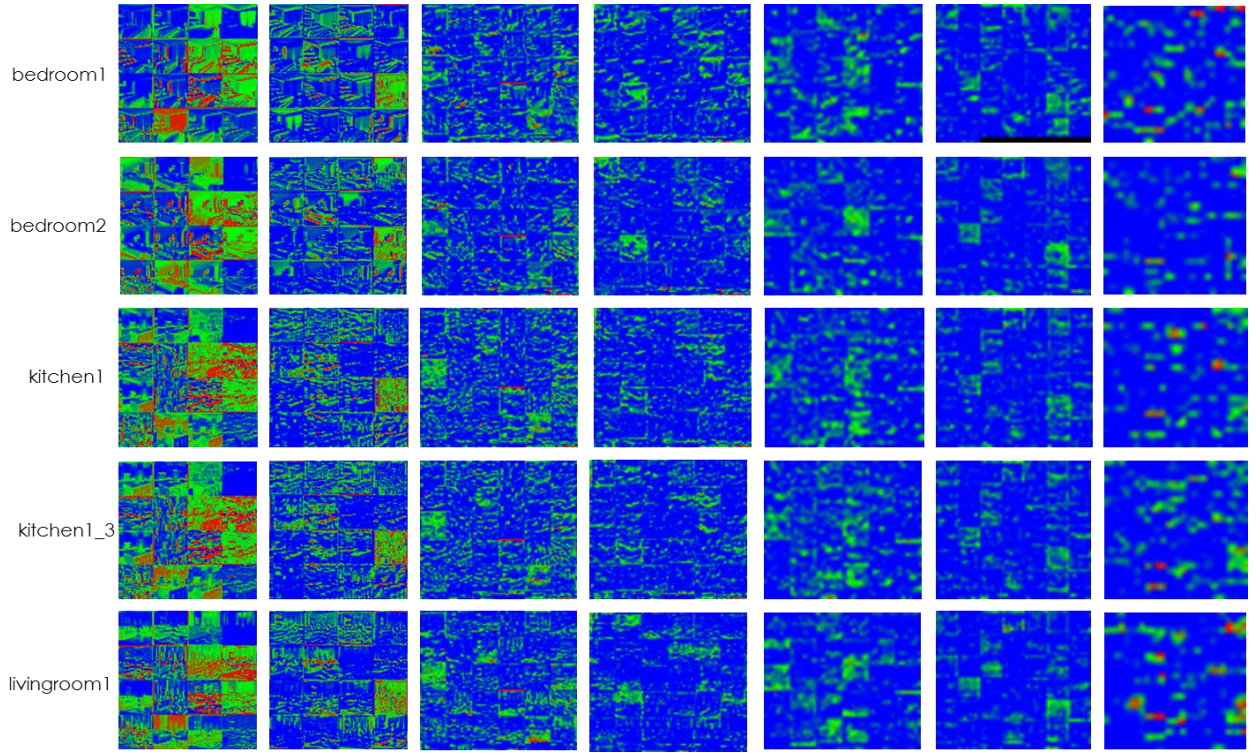
Restaurant with human chatting noise

Place Recognition

Registered places
(one shot learning)



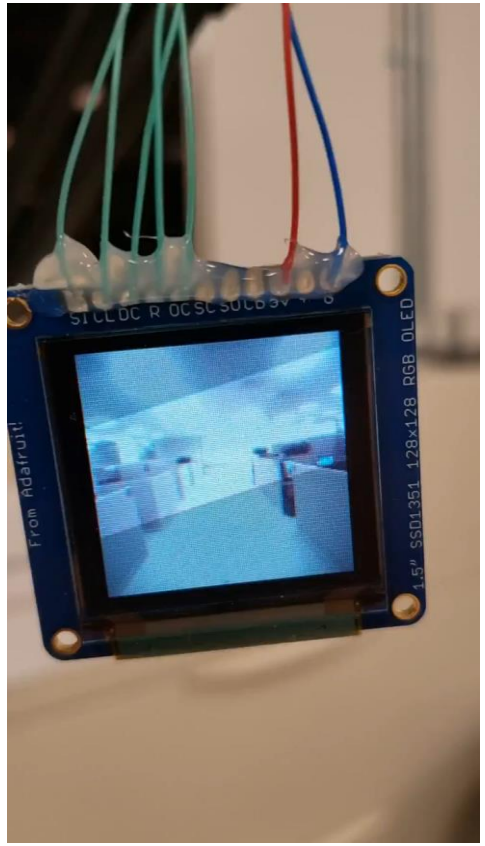
Place Recognition

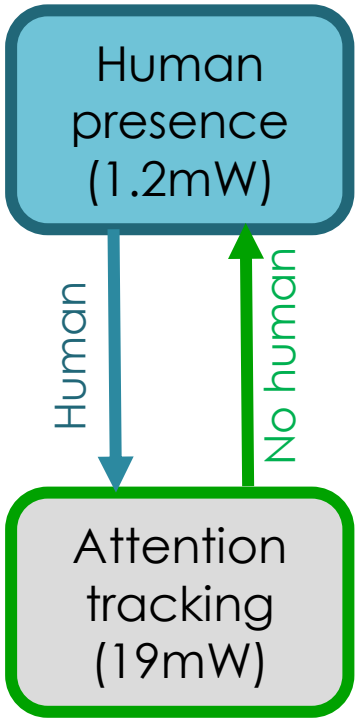


On-the-fly Reprogramming

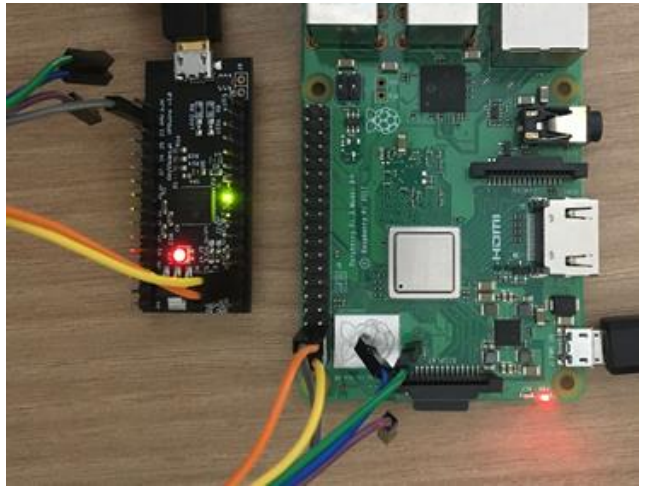
2020 embedded VISION summit

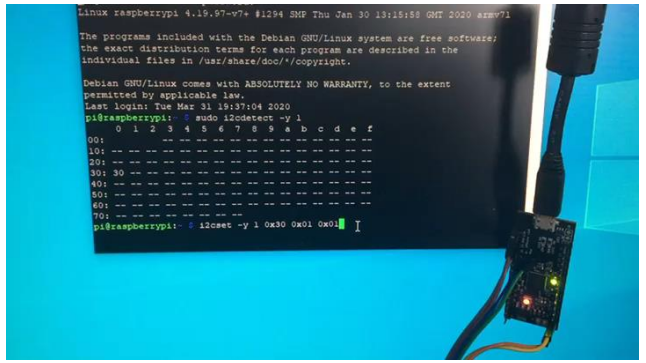
Self reprogramming





Control of reprogram by device driver





Supports multiple HWs & ML FWs for different applications

Programming in milliseconds

LATTICE SEMICONDUCTOR

LATTICE sensAI™

CUSTOM DESIGN SERVICES

Smart Home | Smart City | Smart Factory | Smart Car

REFERENCE DESIGNS / DEMOS

Face Detection | Hand Gesture Detection | Key Phrase Detection | Human Presence Detection | Face Tracking | Object Counting | Speed Sign Detection

SOFTWARE TOOLS

LATTICE RADIANT DESIGN SOFTWARE

Neural Network Compiler

Keras | TensorFlow | Caffe

LATTICE DIAMOND DESIGN SOFTWARE

IP CORES

CNN Compact Accelerator

Neural Network Accelerators

CNN Accelerator

HARDWARE PLATFORMS

UPduino Himax Shield – iCE40 UltraPlus FPGA

Embedded Vision Dev Kit – ECP5 FPGA

1 mW, 5.5 mm², 1/8/16 bits

1 W, 100 mm², 1/8/16 bits

HOT 100 EDN 2018

THE ELECTRONICS INDUSTRY AWARDS 2019

编辑选择奖

Award 2019 Assodel

AI BREAKTHROUGH AWARDS

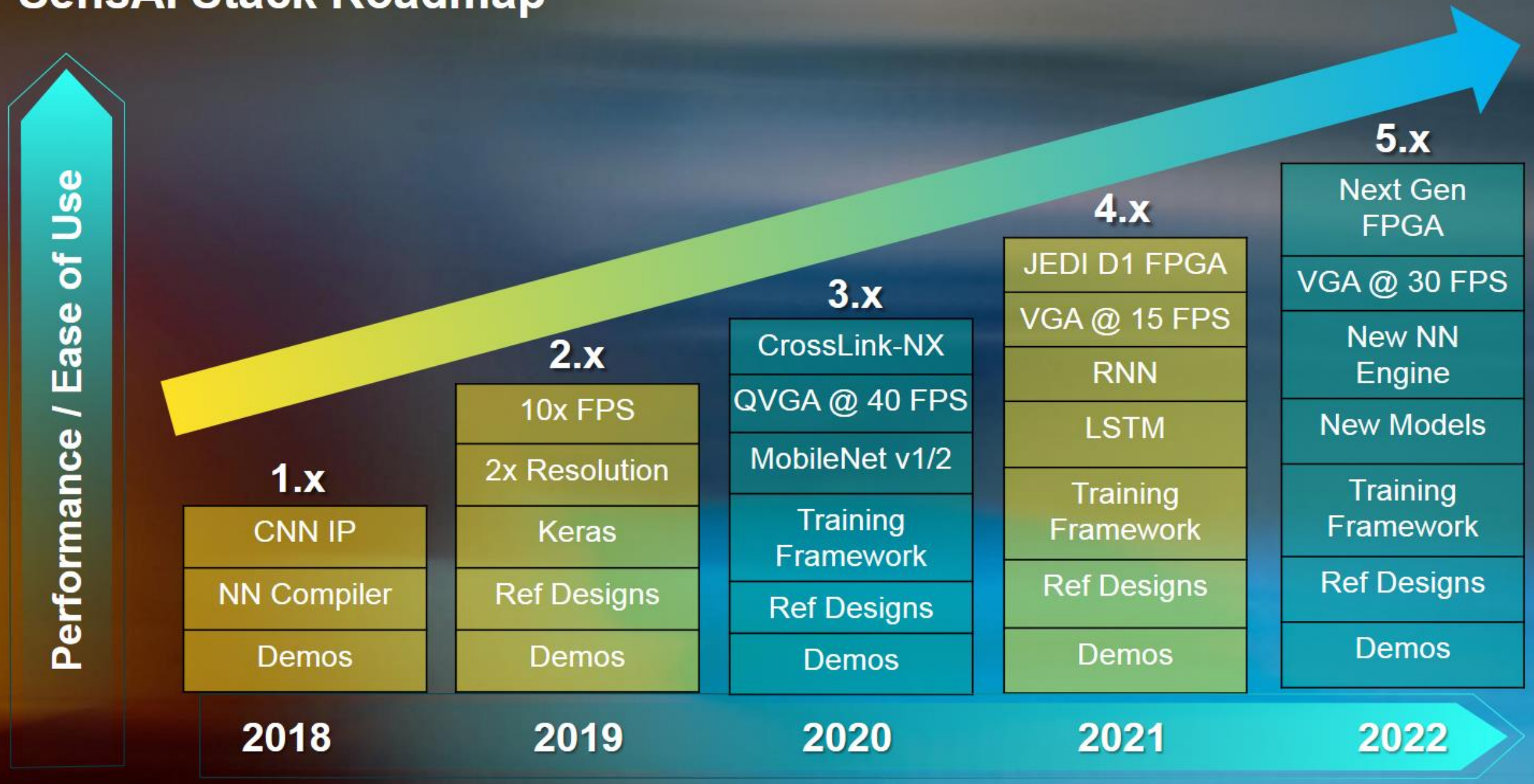
Ultra Low Power

Small Form Factor

Customizable

Design Methodology

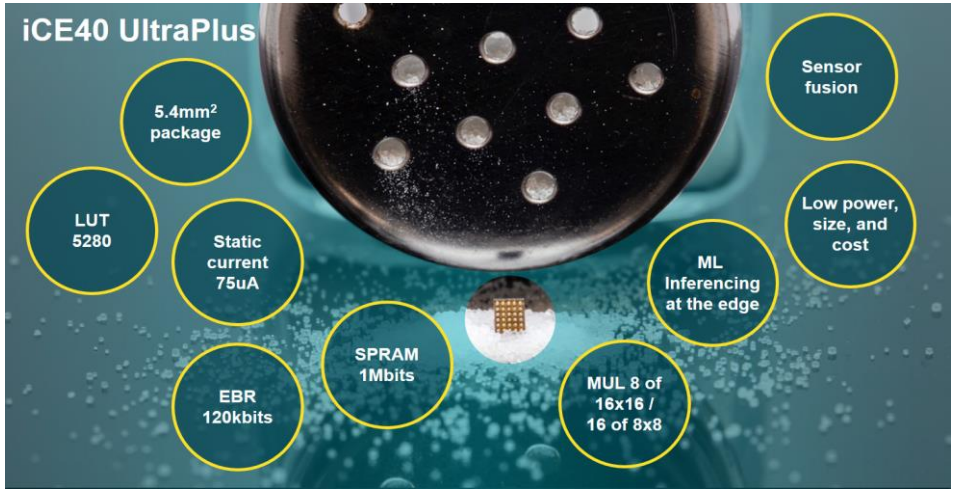
SensAI Stack Roadmap

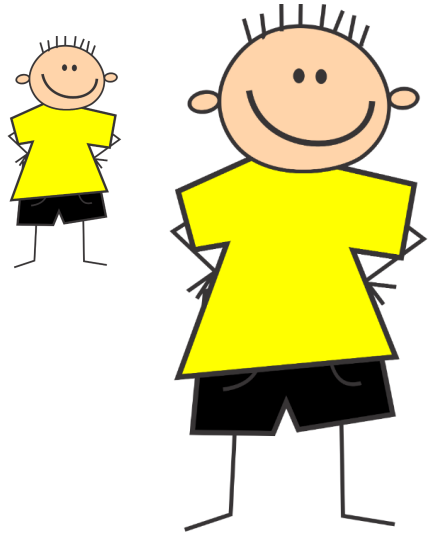


Silicon Performance

	iCE40UP	CrossLink-NX 17	CrossLink-NX 40	ECP5	Jedi 100
Footprint (mm)	2.15 x 2.55	3.7 x 4.1	6 x 6	10 x 10	9 x 9
# of DSPs	8	24	56	156	156
Distributed Memory (kbits)	120	432	1512	3744	3400
SPRAM (kbits)	1024	2560	1024	-	3072
Quantization	8b	8b	8b	8b	8b
Speed (MHz)	40	150	150	100	150
GOPS	0.72	7.2	10.8	28.8	43.2
Power (W)	0.01	0.05	0.2	0.8	0.45
GOPS/W	72	144	54	36	96
Power	★★★★★	★★★★☆	★★★☆☆	★★☆☆☆	★★☆☆☆
Performance	★★☆☆☆	★★★☆☆	★★★☆☆	★★★★☆	★★★★★
Size	★★★★★	★★★★☆	★★★☆☆	★★☆☆☆	★★☆☆☆

Summary





Resource

Lattice sensAI Stack

https://www.latticesemi.com/en/Solutions/Solutions/SolutionsDetails02/sensAI

Neural Network Compiler

https://www.latticesemi.com/en/Products/DesignSoftwareAndIP/AIML/NeuralNetworkCompiler

sensAI boards, demos, IP cores, and reference designs

https://www.latticesemi.com/solutionsearch?qtag=1d7036d97cf446db8bd20f19a07d702f&active=sensai