A comparison of Edge AI Processors

This list is not exhaustive and is provided as a summary of some popular available options (last updated – March 2024)

//consult.red

	Processor											
Feature	**Ambarella CV25S**	**GreenWaves GAP8**	**Hailo Hailo-15**	**Kneron KL720**	**NXP i.MX 8M Plus**	**Nvidia TM660M-A2 (Tegra X1)**	**Syntiant NDP200**	**ST STM32MP25**	**Syntiant NDP101**	**TI AM68A**	**XMOS XCORE.AI**	**Xilinx Versal Edge VE2002**
Description	SoC	SoC - Applications processor	VPU (Vision Processing Unit) SoC	SoC - Applications processor with dedicated NPU	SoC - Applications processor with dedicated NPU and GPU	SoC Processor	Compact neural processor (Neural Decision Processor)	SoC	Compact neural processor (Neural Decision Processor)	SoC - Vision Processor	SoC - Applications processor with multiple cores	SoC
Typical Application(s)	Vision including: Auto parking assist, Electronic mirror, Advanced driver-assistance systems, Depth perception cameras.	Vision including: People counting, Road monitoring, Consumer robotics, Gesture recognition, Face detection, Autonomous drones.	Vision including: Intelligent Cameras tagetting security, retail, transportation, industrial Automation.	Vision icluding: High-end IP Cameras, Smart TVs, AI glasses/headsets.	Smart Applications, Industrial IoT.	Application include: Image classification, Object detection, Segmentation and Speech processing.	Audio/Vision including: Object detection & classification, Wake word detection, Motion tracking, Scene classification.	Industry 4.0, Machine learning, Smart homes, Smart metering, Payment terminals, IoT, Edge computing	Audio/Vision including: Object detection & classification, Wake word detection, Motion tracking, Scene classification.	Vision including: Cameras, Smart agriculture, Video surveilance, Traffic monitoring, Drones, Human Machine Interfaces (HMI).	Audio/Vision including: AI IoT, Presence detection, Voice interfaces, Communications and control	Vision AI applications including: Advanced driver-assist systems, Robotics, Unmanned Aeriel Vehicles, Ultrasound Imaging.
System Role (Primary)	Host	Host	Host combining VPU with CPU and DSP	Host	Host	Host	Host	Host	Coprocessor	Host	Host	Host
Processing Resource	Quad-core Arm Cortex-A76 up to 1.0 GHz with NEON™ SIMD and FPU acceleration Ambarella Image Signal Processor (ISP) and CVflow Computer Vision Processor	8 x RISC-V cores - Convolution Neural Network Accelerator	Quad-core ARM A53 1.3 GHz (12k DMIPS)	NPU (700MHz), Cadence DSP (500 MHz), Arm Cortex M4 (400MHz	NPU (Vivante VIP8000), 4 x Arm Cortex-A53 (1.8 GHz), Arm Cortex-M7 (800MHz), DSP (Cadence Tensilica HiFi 4), GPU (GC7000UL)	GPU (28-core Maxwell @920MHz), 4 x Arm A57 (1.43 GHz	ARM Cortex-M0 HiFi-3 DSP	Arm Cortex-A35 @ 1.5GHz, Arm Cortex-M33 @ 400MHz, NPU 3D GPU (1080p)	ARM Cortex-M0	Upto 2 x 64-bit Arm Cortex-A72, Dual-core Arm Cortex, R5F, Deep Learning Accelerator, 3D GPU (IMG BSX-64-4)	16 xCore logical processing blocks based on STM32M7 (2 tiles, 8 logical cores per tile)	2 x Arm Cortex-A72, 2 x Cortex-R5F, 8 x AI Engine-ML, 90 x DSP Engines
RAM	LPDDR4 / LPDDR4x / DDR4 (32-bit)	Integrated 512 kB SRAM	LPDDR4 (32-bit)	Integrated 128 MB LPDDR3	LPDDR4/DDR4/DDR 3L (16/32-bit)	Up to 4 GB LPDDR3/LPDDR4 (64-bit	48 kB SRAM (M0) 96 kB I SRAM (DSP) 192 kB D SRAM (DSP)	DDR4/LPDDR4	Integrated 112 kB SRAM	Upto 4 MB Integrated RAM LPDDR4	integrated 1 MB SRAM (2 x 512 kB Modules), LPDDR1	256 KB internal (Scalar Engine)
Performance		22.65 GOPS (235 GOPS/W)	Three Variants: Hailo-15H - 20 TOPS Hailo-15M - 11 TOPS Hailo-15L - 7 TOPS	1.5 TOPS (0.9 TOPS/W)	2.3 TOPS	2.3 TOPS	6.4 GOPS	1.35 TOPS	0.24 GOPS	8 TOPS	Peak: 51.2GMACC/s at 8-bit Sustained: 40.96GMACC/s	11 TOPS (INT4) 5 TOPS (INT8) 0.6 TOPS DSP (INT8) 2 TOPS P/Logic (INT4) 1 TOPS P/Logic (INT8)

	Ambarella CV25S	GreenWaves GAP8	Hailo Hailo-15	Kneron KL720	NXP i.MX 8M Plus	Nvidia TM660M-A2 (Tegra X1)	Syntiant NDP200	ST STM32MP25	Syntiant NDP101	TI AM68A	XMOS XCORE.AI	Xilinx Versal Edge VE2002
AI Frameworks	Several including: Caffe, TensorFlow, and ONNX	Several including TensorFlow	Several Including: Keras, TensorFlow, TensorFlow Lite, PyTorch, ONNX	Several including: ONNX, TensorFlow, Keras, Caffe	Several including: Caffe, TensorFlow, TensorFlow Lite, and ONNX	Several including: TensorFlow, PyTorch, Caffe, Keras, Darknet, MXNet	Several including: TensorFlow		Several including: TensorFlow	TensorFlow, PyTorch, ONNX, tvm	TensorFlow, PyTorch, Caffe 2, MXNet	
Interfaces	MIPI CSI/DSI Sub-LVDS Input HiSPi Input SLVS Input HDMI 2.0 Output BT.601/656 eMMC 10/100/1000 Ethernet USB 2.0 Host/ Device	MIPI CSI (8 bit)	MIPI CSI/DSI PCIe 3.0 4-Lanes Ethernet RMII/ RGMII SDIO3.0/ eMMC5.1 (HS200) USB 3.1 Gen2 Host/Device USB 2.0 Host	MIPI CSI DVP VI/VO USB 2.0/3.0	MIPI DSI HDMI 2.0 LVDS PDM USB 2.0/3.0 Gb Ethernet CAN PCIe	USB 3.0 MIPI DSI eDP MIPI CSI-2 HDMI 2.0 eMMC 5.x SPI SDIO	SPI Master SPI Slave PDM 8x GPIO	MIPI CSI-2 MIPI DSI/LVDS Gb Ethernet PCIe Gen2 USB 3.0 CAN-FD	Image Interface (11 Wire)	MIPI CSI/DPI eDP LVDS Ethernet RMII/ RGMII McASP eQEP CAN SDIO eMMc 5.1 PCIe Gen3 USB 3.0 QSPI or HyperBus	USB 2.0 Phy MIPI 128 x G+L16:M17PIO LPDDR DRAM	Gb Ethernet CAN-FD USB 2.0
Power Consumption				1.2W		10W TDP			< 200 µW	6W (Min)		
Package	361 TFBGA (13mm x 13mm)	88 aQFN (7mm x 7mm	FCCSP (15mm x 15mm)	9mm x 9mm 11mm x 11mm	15mm x 15mm		40 QFB (5mm x 5mm)		32 QFN (5mm x 5mm)	770-pin FCBGA (23mm x 23mm)	60 QFN (7mm x 7mm) 265 BGA (14mm x 14mm)	SBVA484 (19mm x 19mm) SBVA625 (21mm x 21mm) SFVA784 (23mm x 23mm)
Differentiating Features		Small package, Low power, consumption,	Full SoC, Implementation, Good power efficiency	Extensive I/O	Extensive I/O	Large developer community, Excellent support, GPU based AI acceleration, Jetson Nano development kit	-Ultra low power (1 mW Vision Processing with Inference Power Consumption), Small package		Increased I/O when compared to NDP100, Small package	8 RGB Cameras (Max)	Proprietary & flexible, processing architecture, Flexible I/O	-Highly flexible/ configurable heterogenous architecture, Excellent power efficiency and performance

Edge AI, innovation and engineering
consult.red/edge-ai/