How the most popular Edge AI processors on the market compare

//consult.red

Below we've compared the most popular AI frameworks presently on the market.

The list is not exhaustive and we aim to keep this updated*

Feature	Processor											
	Arm ETHOS-N57	**GreenWaves** GAP8	**Hailo** Hailo-8	**Kneron** KL520	**Kneron** KL720	**Nvidia** Jetson Nano	**NXP** i.MX 8M Plus	**ST** STM32F769xx	**Syntiant** NDP100	**Syntiant** NDP101	**TI Sitara** AM5749	**XMOS** XCORE.AI
Description	IP Block	SoC - Applications processor	NPU SoC	SoC - Applications processor with dedicated NPU	SoC - Applications processor with dedicated NPU	Single board Development Kit	SoC - Applications processor with dedicated NPU and GPU	MCU	Ultra compact neural decision processor	Compact neural decision processor	SoC - Applications processor	SoC - Applications processor with dedicated NPU
Typical Application(s)	**Mid range devices balancing performance, cost and battery life. Applications include:** Mainstream phones and digital TVs	**Application include:** People counting, Road monitoring, consumer robotics, gesture recognition, Face detection, Autonomous drone, Key word identification, Surveillance camera	**Primarily Focused toward automotive sector, including:** advanced driver assistance systems (ADAS) and autonomous driving applications	**General smart home applications including**: Smart door locks, Drones, Smart doorbells, IP cameras, Robot vacuums	**General connected devices including:** high-end IP Cams, Smart TVs, AI glasses and headsets, and AIoT Gateways	**Application include:** Image classification, Object detection, Segmentation and Speech processing	Smart Applications, Industrial IoT	Person Presence Detection, Data mining, Facial/ voice recognition, predictive maintenance, financial fraud detection	**Voice applications, with support for wake-word detection. Applications include:** Mobile phones, Ear buds and hearing aids, Bluetooth headsets, Smart watches, IoT endpoints, Remote controls, Smart speakers	**Voice applications, with support for wake-word detection. Applications include:** Mobile phones, Smart watches, IoT endpoints, Smart home, Remote Controls	Automated sorting equipment, Optical inspection, Vision computer, Code readers, Industrial robots, Logistics robots, Currency counters, ATMs, Patient monitors, Building automation, Industrial transport, Space, avionics & defence	**General IoT application requiring AI including:** Audio interface, Presence detection, Person identification, Actuation, Voice interfaces, Communications and control
System Role (Primary)	Coprocessor (IP Block)	Host	Coprocessor	Host	Host	Host (Standalone development board)	Host	Host	Coprocessor	Coprocessor	Host	Host
Processing Resource	NPU (8 Cores)	• 8 x RISC-V cores • Convolution Neural Network Accelerator	Proprietary structure-driven data flow architecture	• NPU • Dual Arm Cortex M4	• NPU (700MHz) • Cadence DSP (500 MHz) • Arm Cortex M4 (400MHz)	• GPU (28-core Maxwell @920MHz) • 4 x Arm A57 (1.43 GHz)	• NPU (Vivante VIP8000) • 4 x Arm Cortex-A53 (1.8 GHz) • Arm Cortex-M7 (800MHz) • DSP (Cadence Tensilica HiFi 4) • GPU (GC7000UL)	• ARM Cortex-M7 • FPU	ARM Cortex-M0	ARM Cortex-M0	• 2 x Cortex-A15 • C66x VLIW DSP • 2 × Dual ARM Cortex-M4 • GPU (Dual Core SGX544) • 2 × EVE Analytic Processor	• NPU • STM32M7 Architecture based on 16 configurable 'logical cores
RAM	512 kB SRAM	512 kB SRAM	Embedded	Embedded - LPDDR2	Embedded - 128 MB LPDDR3	4 GB LPDDR4 (64-bit)	LPDDR4/DDR4/ DDR3L (16/32-bit)	512 kB SRAM	Embedded - 112 kB SRAM	Embedded - 112 kB SRAM	512 kB SRAM	1 MB SRAM (2 x 512 kB Modules)

	Arm ETHOS-N57	GreenWaves GAP8	Hailo Hailo-8	Kneron KL520	Kneron KL720	Nvidia Jetson Nano	NXP i.MX 8M Plus	ST STM32F769xx	Syntiant NDP100	Syntiant NDP101	TI Sitara AM5749	XMOS XCORE.AI
Performance	2 TOPS	22.65 GOPS (235 GOPS/W)	26 TOPS (2.8TSOP/W)	0.28 TOPS (0.56 TOPS/W)	1.5 TOPS (0.9 TOPS/W)	472 GFLOPs	2.3 TOPS	1082 CoreMark /462 DMIPS at 216 MHz (7 CoreMark/ mW)				3,200 MIPS, 51.2 GMAC, and 1,600 MFLOPS
AI Frameworks	**Several including:** TensorFlow, TensorFlow Lite, Caffe2, PyTorch, MXNet, ONNX	**Several including:** TensorFlow	**Several including:** TensorFlow and ONNX	**Several including:** ONNX, TensorFlow, Keras, Caffe	**Several including:** ONNX, TensorFlow, Keras, Caffe	**Several including:** TensorFlow, PyTorch, Caffe, Keras, Darknet, MXNet	**Several including:** Caffe, TensorFlow, TensorFlow Lite, and ONNX	**Several including:** Caffe, Keras, TensorFlow Lite and ONNX	**Several including:** TensorFlow	**Several including:** TensorFlow	Caffe, Keras, TensorFlow, TensorFlow Lite, GluonCV, MXNet, PyTorch, ONNX	TensorFlow Lite
Interfaces	N/A	SPI (Slave) 2 x I2S (Rx-Only) CSI (8 bit) 2 x I2C UART 32 x GPIO 16 x PWM	PCIe RGMII 2 x MIPI CSI SDIO USB 2.0 I2C I2S GPIO UART C2C		SPI UART I2S I2C USB 2.0 USB 3.0 PWM 2 x MIPI CSI SPI DVP VI DVP VO	microSD 2 x MIPI CSI-2 (12 lanes) Gb Ethernet M.2 HDMI 4 x USB 3.0 I2C I2S SPI UART GPIO	HDMI 2.0 MIPI DSI LVDS (4 or 8 Lane) 18 x I2S S/PDIF 8 x PDM Microphone I/P 2 x USB 2.0/3.0 2 x Gb Ethernet 2 x CAN PCIe 4 x UART	**Subject to device variant:** 2 x USB 2.0 2 x SDIO UART Quad SPI I2C I2S CAN 2.0 HDMI Ethernet MDIO MIPI DSI SPDIF	SPI Slave 2 x PDM Microphone I/P inputs	SPI Master SPI Slave 2 x PDM microphone inputs 8x GPIO	USB 3.0 USB 2.0 2 x PCIe 4 x MMC/SD SATA HDMI 1.4a 10 x UART 4 x SPI 2 x CAN 5 x I2C	USB 2.0 Phy MIPI 128 x GPIO LPDDR DRAM
Power Consumption			<5W	0.5W	1.2W	10W TDP			< 200 µW	< 200 µW		
Package	N/A	88 aQFN (7mm x 7mm)	324 BGA (15mm x 15mm)		9mm x 9mm 11mm x 11mm	260-pin edge connector (69 mm x 45 mm)	15mm x 15mm	100/144/176/208 LQFP 100/216 TFBGA 176 UFBGA 180 WLCSP	12 WLBGA (1.4mm x 1.8mm)	32 QFN (5mm x 5mm)	760-pin BGA (23mm × 23mm)	
Differentiating Features		Small package Low power consumption	Very good power efficiency		Extensive I/O	• Large developer community • Excellent support • Tegra processor • GPU based AI acceleration • Extensive I/O	Extensive I/O	• RISC architecture • Extensive I/O • ST support forum	• Ultra low power (< 140 µW recognising words) • Very small package	• Increased I/O when compared to NDP100 • Small package	• High performance • TI open source community	

Contact us:

Stuart Griffin, Founding Director & Technologist
stuart.griffin@consult.red

Adam Hoare, Business Development
adam.hoare@consult.red

consult.red/edge-ai

***Created January 2021.** Click here **to ensure you have the most up to date version.**