

MACx

MAC01 | Multimedia AI Accelerator Card (MAC)

Product Capabilities

Multi-modal AI Processing

Hardware-level acceleration provides synchronous support for computer vision (CV), natural language processing (NLP), and speech models.

Compute and Interconnect

The MAC01 card provides FP32 Tensor Core performance of 40 TFLOPS, FP16 Tensor Core performance of 120 TFLOPS, and INT8 Tensor Core performance of 240 TFLOPS. It features 64 GB of fast memory and a PCIe Gen5 interconnect rated at 128 GB/s.

MACA Software Development Platform

MACA (Multimedia AI Computing Architecture) comprises a hardware abstraction layer, programming model, acceleration libraries, runtime libraries, and a development toolchain. It supports a self-developed inference engine and video codec acceleration frameworks.

Framework Compatibility and Deployment

Compatible with PyTorch, TensorFlow and MXNet, the platform provides model compression toolchains and OpenCL programming support. The specification presents smart security, medical imaging and edge computing as deployment scenarios.

Product Specifications

Product Model MACx / MAC01		
Compute	FP32 Tensor Core	40 TFLOPS
	FP16 Tensor Core	120 TFLOPS
	INT8 Tensor Core	240 TFLOPS
Memory & I/O	Fast Memory	64 GB
	Interconnect	PCIe Gen5, 128 GB/s
Video Codec	Encoder	2 encoders
	Encoding formats	H.264, H.265, JPEG
	Decoder	14 units
	Decoding formats	H.264, H.265, VP9, JPEG
Hardware	Max Thermal Design Power	< 300 W
	Form Factor	Full-height, full-length, dual-slot