

AI Chip BM1684X



SOPHON cloud and edge AI chip BM1684X is oriented at AI inference in cloud and edge scenarios. It can be integrated into AI micro servers, AI servers, IPC and other types of products and effectively adapts to all AI algorithms in the market. It can enable AI for various industries through the application of image classification, target detection, instance segmentation, semantic segmentation, behavior analysis, character recognition, language processing, search recommendation, etc.

Outstanding Performance

- 32 TOPS INT8, 16 FLOPS FP16, 2T FLOPS FP32
- 32 channels 25fps HD video decoding
- 12 channels 25fps HD video encoding

Mature SDK

- Adaptable to various mainstream deep learning frameworks
- Mature and easy-to-use SDK and toolchain; rapid algorithm migration

Strong Environment Adaptability

- Working temperature range: -40°C to 105°C
- Typical power consumption: 16W

Integration into Various Products



AI Module



AI Accelerator Card



AI Micro Server



AI Server



Edge Gateway



Intelligent NVR



Robot



UAV

Product Specification

Product	BM1684X
AI Computing Power	32 TOPS INT8 16 TFLOPS FP16/BF16 2 TFLOPS FP32
Multimedia	Video Decoding
	Resolution
	Video Encoding
	Resolution
	Image Decoding
	Image Encoding
	Max. Resolution
Memory	Interface
	Max. Capacity
Main CPU	ARM A53 8-core 2.3GHz
Interface	1* PCIE EP X16 Gen3 or 1* PCIE EP X8 Gen3 / 1* PCIE RC X8 Gen3 10/100/1000 RGMII X2 I2C/UART/PWM/SPI/GPIO
Typical Power Consumption	16W
Working Temperature	-40°C ~ +105°C
Package Size	21 mm * 21 mm
Ball Pitch	0.4 mm
Deep Learning Framework	TensorFlow / Pytorch / Paddle / Caffe / MxNet / DarkNet