100T Flagship AI Box

- AIBOX-OrinNX (16GB)
- AIBOX-OrinNano (8GB)

V1.0 2024-8-19

T-CHIP INTELLIGENCE TECHNOLOGY

Product Features

NVIDIA High-performance edge computing module

NVIDIA Jetson OrinNX edge computing module (16GB version) features an octa-core ARM CPU and a 1024-core NVIDIA Ampere architecture GPU with 32 Tensor Cores. It is capable of running multiple concurrent AI application pipelines, providing powerful inference performance.

100TOPS computing power empowers AI applications

Mainstream modern AI models can be run. For example, the ROS robot model realizes larger and more complex deep neural networks, and realizes functions such as object recognition, object detection and tracking, speech recognition, and other visual development.

The private deployment of large models

Large language models: LLaMa2, ChatGLM, Qwen, and more.
Vision models: ViT, Grounding DINO, and SAM.
AI painting: Stable Diffusion V1.5 image generation model in the AIGC field.

Multiple deep learning frameworks

Support traditional network architectures such as CNN, RNN, and LSTM;
a variety of deep learning frameworks, including TensorFlow, PyTorch, MXNet, PaddlePaddle, and ONNX, as well as custom operator development.

Product Features

AI software stack and ecosystem

NVIDIA JetPack™, Isaac ROS, and reference AI workflows enable seamless integration of cutting-edge technologies into your products, eliminating the need for costly internal AI resources. Experience end-to-end acceleration for AI applications and speed your time to market using the same powerful technologies that drive data centers and cloud deployments.

Abundant expansion interfaces

1 × Gigabit Ethernet (RJ45), 1 × HDMI 2.0, 2 × USB 3.0, 1 × Console, 1 × Type-C and other interfaces facilitate the connection of various peripherals and enable multi domain product applications
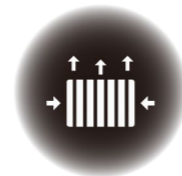
Aluminum alloy enclosure with efficient heat dissipation

The AI box features an industrial-grade all-metal enclosure with an aluminum alloy structure for thermal conduction. The side of the top cover features a grille design for external airflow and efficient heat dissipation, ensuring computing performance and stability even under high-temperature operating conditions.
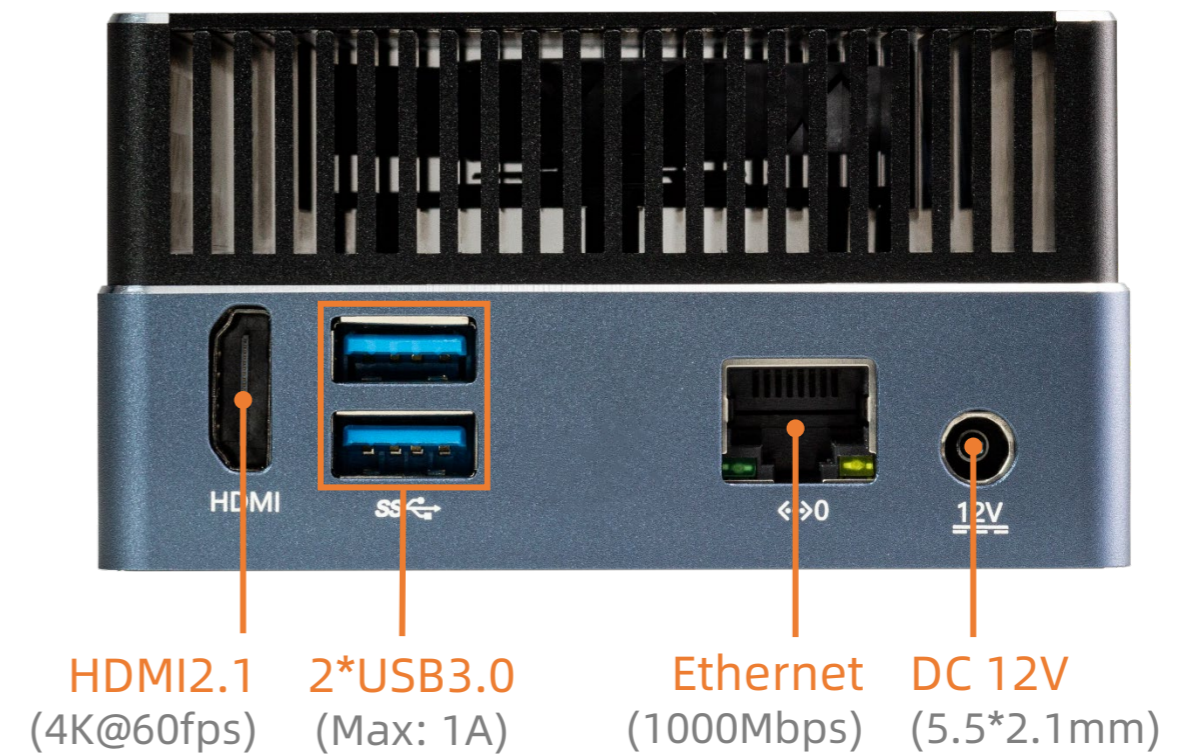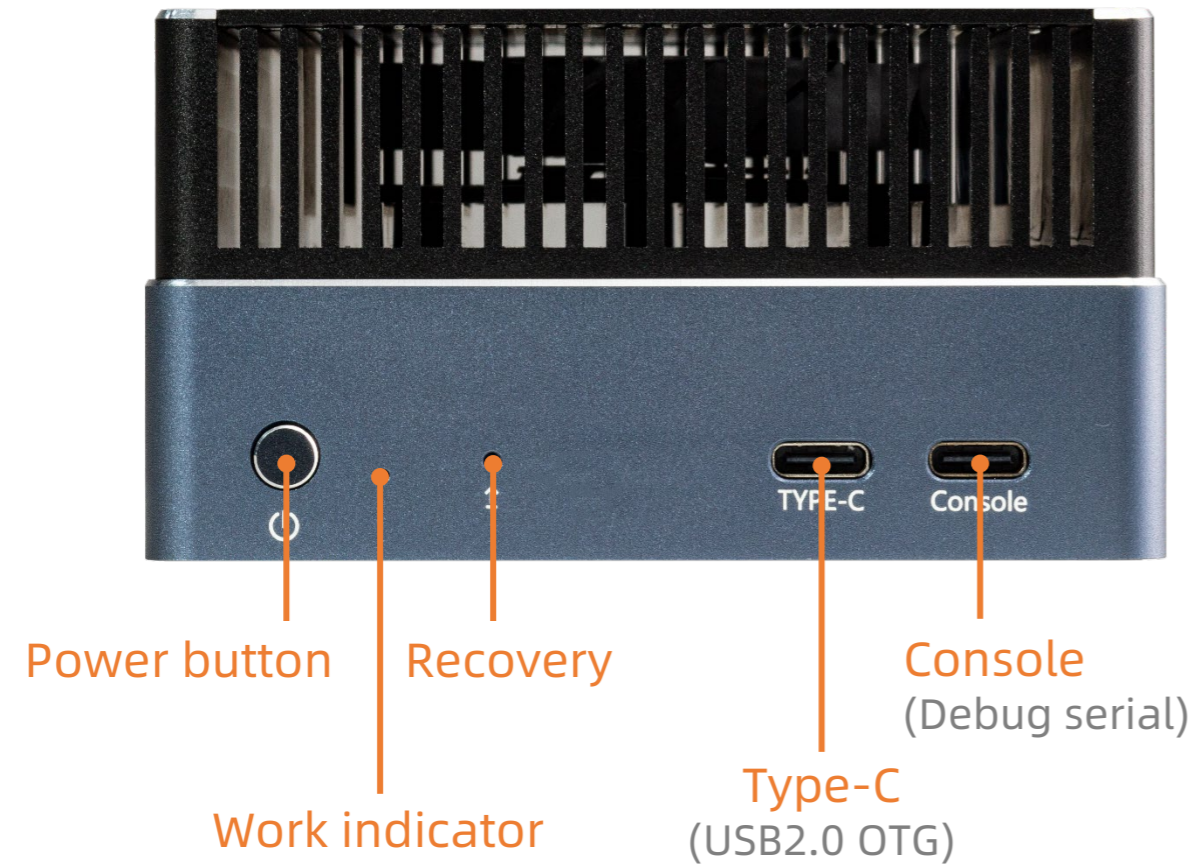
A wide range of applications

A widely used in intelligent surveillance, AI education, services based on computing power, edge computing, private deployment of large models, data security, and privacy protection.
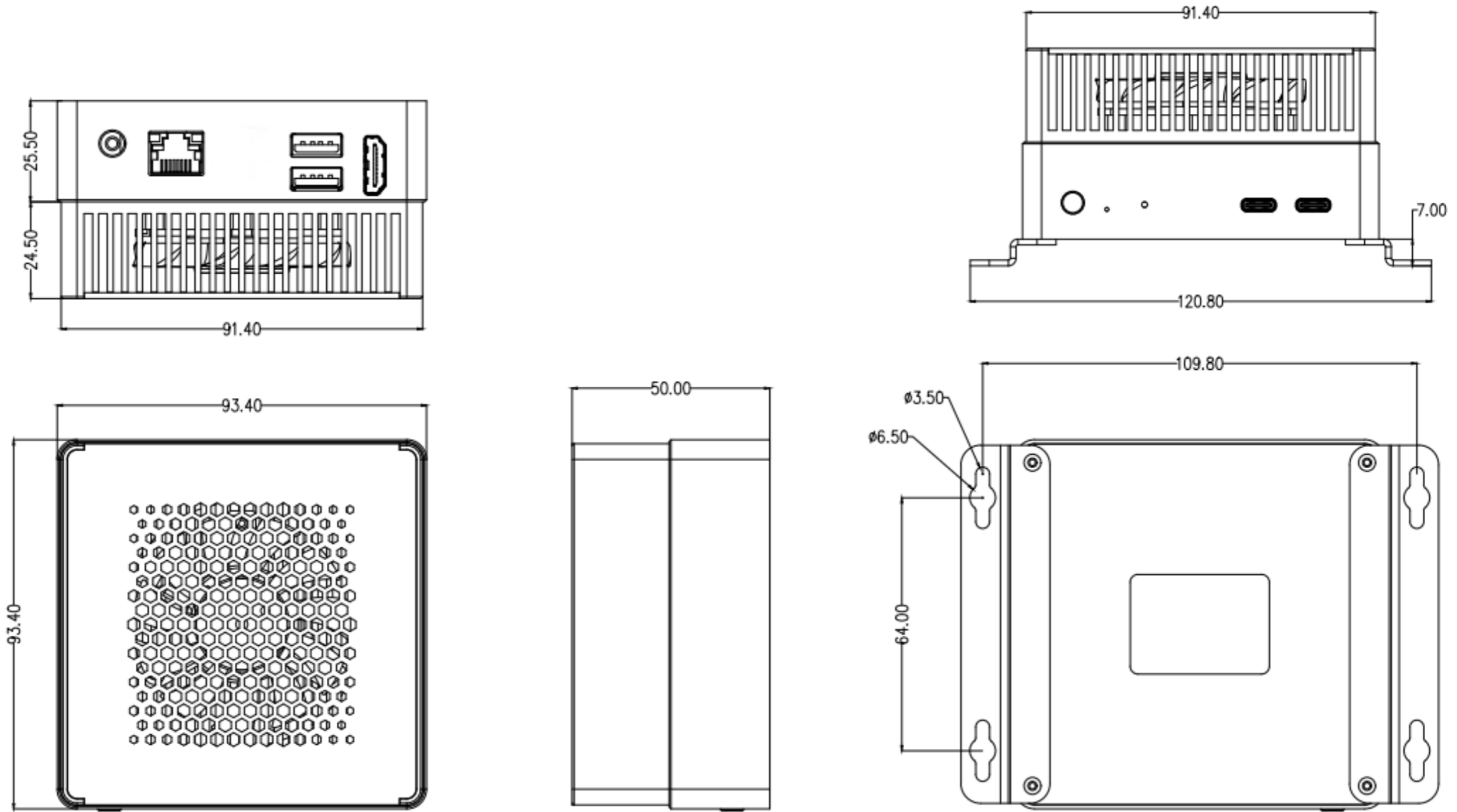
Specifications

		AIBOX-OrinNano (8 GB)	AIBOX-OrinNX （16 GB）
Basic Specifications	Module	Original NVIDIA Jetson OrinNano (8GB) module	Original NVIDIA Jetson OrinNX （16GB） module
	CPU	Hexa core 64 bit ARM Cortex-A78AE v8.2 processor Up to 1.5GHz	Octa core 64 bit Arm Cortex-A78AE v8.2 processor Up to 2.0GHz
	AI performance	40 TOPS	100 TOPS
	GPU	1024 core NVIDIA Ampere architecture GPU with 32 Tensor Cores	
	Video encoding	H.265: 1080p30	H.265: 1×4K60, 3×4K30, 6×1080p60, 12×1080p30
	Video decoding	H.265: 1×4K60, 2×4K30, 5×1080p60, 11×1080p30	H.265: 1×8K30, 2×4K60, 4×4K30, 9×1080p60, 18×1080p30
	Memory (Video Memory)	8GB LPDDR5	16GB LPDDR5
	Storage	128GB PCIe NVMe SSD(Installed inside the device)	
	Power	DC 12V(DC 5.5 * 2.1mm)	
	Power consumption	Normal: 7.2W(12V/600mA) Max:18W(12V/1500mA)	Normal: 7.2W(12V/600mA) Max: 33.6W(12V/2800mA)
	Size	93.4mm * 93.4mm * 50mm	
	Weight	≈ 500g	
	Environment	Operating Temperature: -20℃～60℃, Storage Temperature: -20℃～70℃, Storage Humidity:10%～90%RH(non-condensing)	
Software support	OS	Jetson systems based on Ubuntu 22.04 provide a complete desktop Linux environment with graphics acceleration and support for libraries such as NVIDIA CUDA, TensorRT, CuDNN, and more	
	Large model	Robot model: ROS robot model is supported Large language models: Support the private deployment of ultra-large-scale parametric models under the Transformer architecture, such as LLaMa2, ChatGLM, Qwen, and other large language models Large visual models: Support the privatization deployment of large visual models such as ViT, Grounding DINO, and SAM AI Painting: Supports the private deployment of the Stable DiffusionV1.5 image generation model in the AIGC field	
	Traditional network architecture	Supports deep learning frameworks for traditional network architectures such as CNN, RNN, LSTM, etc.: Supports a variety of deep learning frameworks, including TensorFlow, PyTorch, MXNet, PaddlePaddle, ONNX, and Darknet, and supports custom operator development Docker containerization: Docker containerization technology is supported, which can be easily used for image deployment	
	AI software stack	The NVIDIA Jetson Orin series delivers powerful AI compute power, massive unified memory, and a comprehensive software stack to power the latest generative AI applications. It enables fast inference on any generative AI model powered by the Transformer architecture, enabling superior edge performance on MLPerf	
Interface Specifications	Internet	1 × Gigabit Ethernet (1000Mbps/RJ45)	
	Display	1 × HDMI2.1(4K@60fps)	
	USB	2 × USB3.0 (Max: 1A)	
	Watchdog	Support external watchdog	
	Other interfaces	1 × Type-C (USB2.0 OTG), 1 × Console (Debug serial), 1 × Recovery, 1 × Power button	

Interface description

Power button   Recovery   Console
                                                                  (Debug serial)

          Work indicator   Type-C
                                      (USB2.0 OTG)

HDMI2.1   2*USB3.0   Ethernet   DC 12V
(4K@60fps)   (Max: 1A)   (1000Mbps)   (5.5*2.1mm)

Dimension

T-CHIP INTELLIGENCE TECHNOLOGY

Contact Us
(+86)18688117175

E-mail
global@t-firefly.com

Website
www.t-firefly.com

Address
Room 2101, Hongyu Building, #57 Zhongshan 4Rd, East District, Zhongshan, Guangdong, China.