



# 6T Large Model AI Box

| AIBOX-3576

| AIBOX-3588

V1.0 2024-11-26

T-CHIP INTELLIGENCE TECHNOLOGY



# Product features



## High-performance octa-core 64-bit AIOT processor, RK3576/RK3588

RK3576/RK3588, the new octa-core 64-bit AIOT processor, features a big.LITTLE architecture, an advanced lithography process, and a frequency of up to 2.2 GHz/2.4GHz.



## 8K video decoding/video encoding

AIBOX-3576 support 8K@30fps/4K@120fps Decoding (H.265/HEVC, VP9, AVS2, AV1) and 4K@60fps Encoding (H.265/HEVC, H.264/AVC). AIBOX-3588 support 8K@60fps H.265/VP9 video decoding, supports same encoding and same decoding, can achieve up to 32 channels 1080P@30fps Decoding and 16 channels 1080P@30fps code.



## The private deployment of large language models

Support the private deployment of ultra-large-scale parameter models under the Transformer architecture, including large language models such as Gemma-2B, ChatGLM3-6B, Qwen-1.8B, Phi-3-3.8B.



## Multiple deep learning frameworks

RKNN models can be imported and exported; Support a variety of deep learning frameworks, including TensorFlow, TensorFlow Lite, PyTorch, Caffe, ONNX and Darknet. It also supports the development of custom operators.

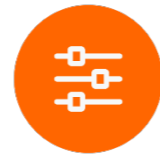


# Product features



## Powerful network communication capability/operating system

Support dual Gigabit Ethernet, high-speed and stable network communication; It supports Linux system, and the system is safe and stable to meet the needs of different application scenarios.



## Abundant expansion interfaces

Supports 2 × Gigabit Ethernet, 2 × USB 3.0, 1 × TF Card, 1 × Type-C, 1 × HDMI 2.0, 1 × Console and other expansion interfaces, making it easy to connect various external devices.



## All-aluminum alloy enclosure for heat dissipation

The industrial-grade all-metal enclosure with aluminum alloy structure for thermal conduction. The side of the top cover features a grille design for efficient heat dissipation. Its top cover is a porous hexagonal design, combining elegance with high efficiency. The compact, exquisite device operates stably and meets the needs of various industrial-grade applications.



## A wide range of applications

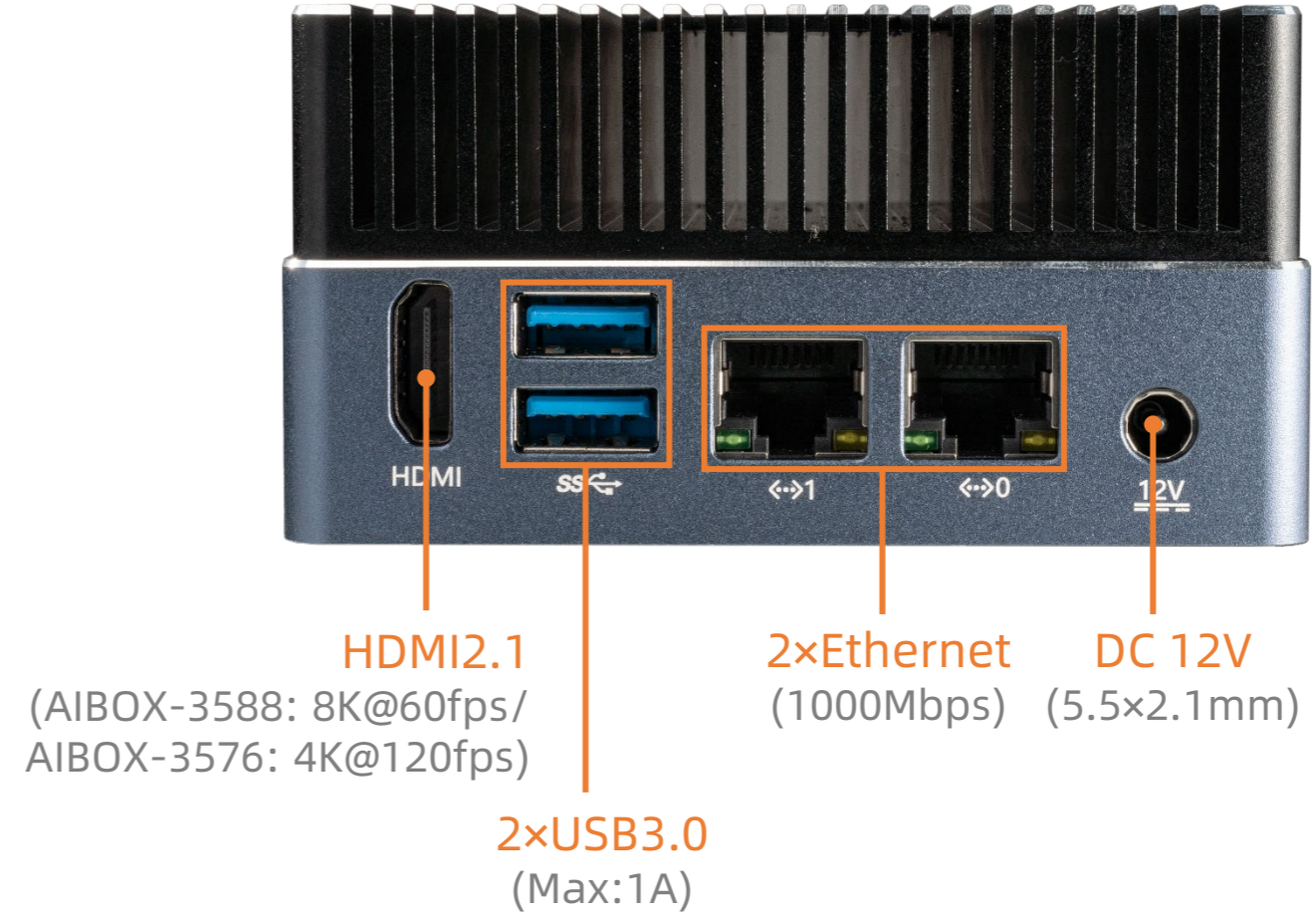
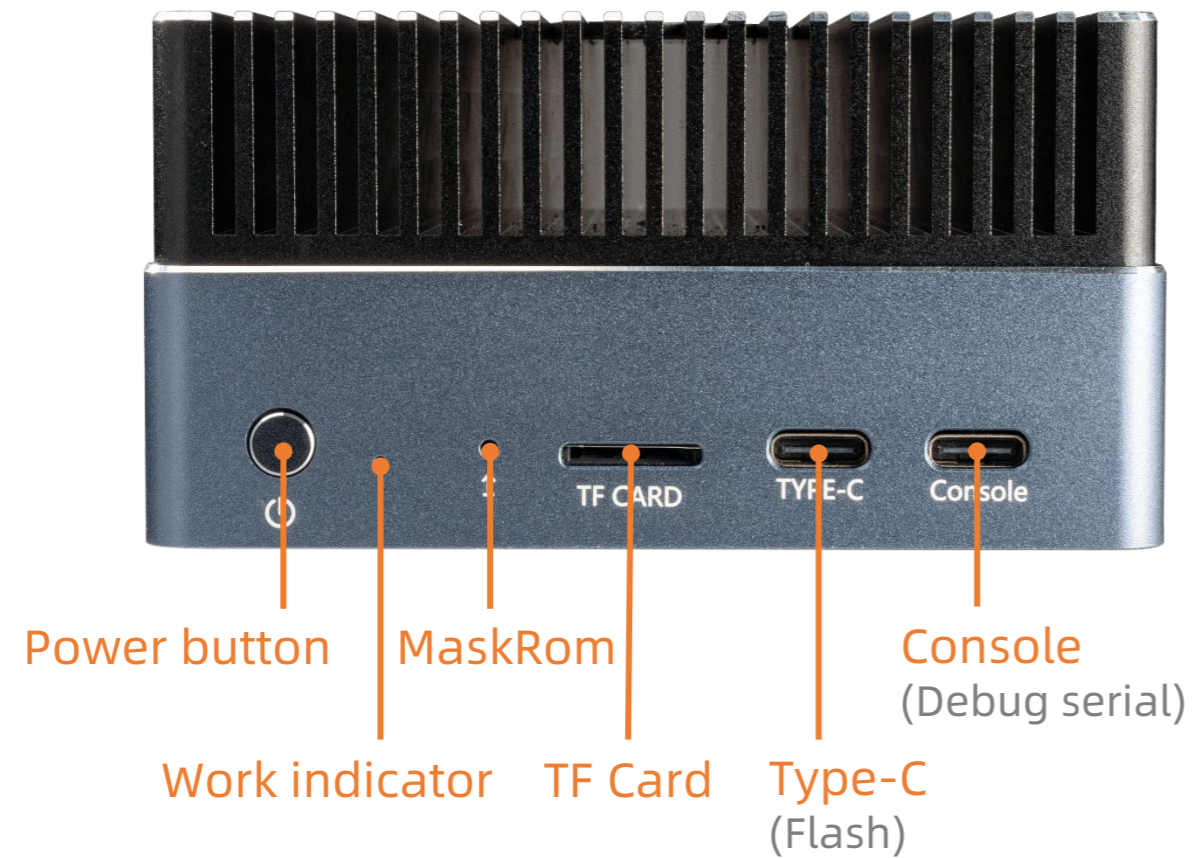
The device is widely used in intelligent surveillance, AI education, services based on computing power, edge computing, private deployment of large models, and data security and privacy protection.

# Specifications

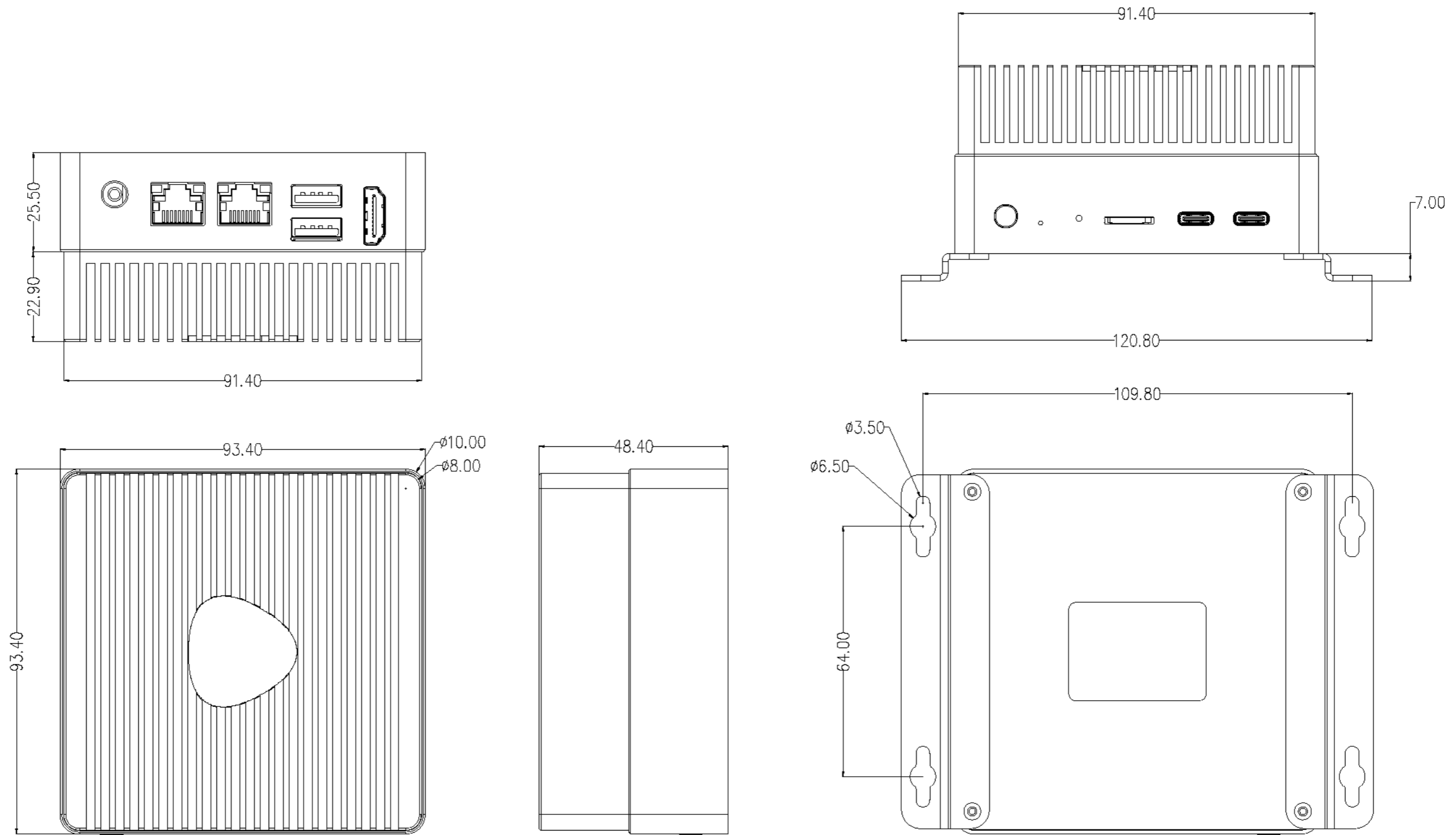


		AIBOX-3576	AIBOX-3588
Basic Specifications	SOC	Rockchip RK3576	Rockchip RK3588
	CPU	Octa-core 64-bit processor(4xA72+4xA53), main frequency up to 2.2GHz	Octa-core 64-bit processor(4xCortex-A76+4xCortex-A55), main frequency up to 2.4GHz
	GPU	G52 MC3@1GHz, supports OpenGL ES 1.1/2.0/3.2, OpenCL 2.0, Vulkan 1.1, embedded high-performance 2D acceleration hardware	ARM Mali-G610 MP4 quad-core GPU, supports OpenGL ES3.2/ OpenCL 2.2/Vulkan1.1, 450 GFLOPS
	NPU	6 TOPS NPU, supports INT4/8/16/FP16/BF16/TF32 mixed operations	6 TOPS NPU, supports INT4/INT8/INT16 mixed operations
	ISP	Built-in 16 million pixel ISP, support low-light noise reduction, support RGB-IR sensor, support up to 120dB HDR, AI-ISP to improve low-noise image effect	Integrated 48MP ISP with HDR&3DNR
	Encoding/Decoding	Decoding: 4K@120fps: H.265/HEVC, VP9, AVS2, AV1, 4K@60fps: H.264/AVC Encoding: 4K@60fps: H.265/HEVC, H.264/AVC	Decoding: 8K@60fps/4K@120fps H.265/VP9/AVS2, 8K@30fps H.264 AVC/MVC, 4K@60fps AV1, 1080P@60fps MPEG-2/-1/VC-1/VP8 Encoding: 8K@30fps H.265/H.264
	RAM	LPDDR4 (4GB/8GB/16GB optional)	LPDDR4 (4GB/8GB/16GB optional, up to 32GB)
	Storage	eMMC (16GB/32GB/64GB/128GB/256GB optional), UFS2.0 (Only AIBOX-3576 optional)	
	Storage Expansion	1 × M.2 (Expandable SATA 3.0/PCIe NVMe SSD, supports 2242/2260/2280; Inside the computer), 1 × TF Card	
	Power	DC 12V/2A (5.5 × 2.1mm)	DC 12V/3A (5.5 × 2.1mm)
	Power consumption	Normal: 1.2W(12V/100mA) Max: 7.2W(12V/600mA) Min: 0.72W (12V/6mA)	Normal: 3.6W(12V/300mA) Max: 13.2W(12V/1100mA) Min: 1.38W(12V/115mA)
	OS	Linux OS (Ubuntu)	
	Software support	<ul style="list-style-type: none"> <li>Support the privatization deployment of ultra-large-scale parametric models under the Transformer architecture, such as Gemma-2B, ChatGLM3-6B, Qwen-1.8B, Phi-3-3.8B and other large language models</li> <li>It supports traditional network architectures such as CNN, RNN, and LSTM, and supports the import and export of RKNN models; Support a variety of deep learning frameworks, including TensorFlow, TensorFlow Lite, PyTorch, Caffe, ONNX and Darknet. It also supports the development of custom operators</li> <li>Support Docker container management technology</li> </ul>	
	Size	93.4mm × 93.4mm × 50mm	
	Weight	≈ 500g	
Environment	Operating Temperature: -20°C ~ 60°C, Storage Temperature: -20°C ~ 70°C, Storage Humidity: 10% ~ 90%RH (non-condensing)		
Interface Specifications	Ethernet	2 × Gigabit Ethernet (1000Mbps/RJ45)	
	Video output	1 × HDMI2.1 (4K@120fps)	1 × HDMI2.1 (8K@60fps)
	USB	2 × USB3.0 (Max: 1A)	
	Other interfaces	1 × Type-C (Flash), 1 × Console (Debug serial), 1 × Power button, 1 × MaskRom	

# Interface description



# Dimension





## T-CHIP INTELLIGENCE TECHNOLOGY

---



Contact Us  
(+86)18688117175



E-mail  
global@t-firefly.com



Website  
<https://en.t-firefly.com/>



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