

全爱科技 AI 推理服务器 HOUYI-300U1 技术白皮书

文档版本

02

发布日期

2025-10-25



全爱科技（上海）有限公司

版权所有 全爱科技（上海）有限公司 2025. 保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



和其他全爱商标均为全爱科技（上海）有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受全爱科技商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，全爱公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

全爱科技（上海）有限公司

地址：上海市闵行区剑川路 920 号 2 栋 3 层 邮编：200240

网址：www.quanaichina.com

修改记录

文档版本	发布日期	修改说明
01	2025-10-10	初版发布
02	2025-10-25	规格细节修改

目 录

- 1 产品简介.....1
 - 1.1 概述.....1
 - 1.2 产品特点.....1
 - 1.3 外观结构.....2
- 2 产品规格.....4
 - 2.1 基本规格.....4
 - 2.2 环境条件.....6
- 3 快速使用.....7
 - 3.1 默认软件使用.....7
 - 3.2 默认测试 demo.....8
 - 3.3 软件资料.....10
- 4 接口说明.....12
 - 4.1 千兆以太网口.....12
 - 4.2 USB 3.0 Type-A 接口.....12
 - 4.3 HDMI 接口.....12
 - 4.4 VGA 接口.....12
 - 4.5 音频接口.....12
 - 4.6 电源接口.....12
- A 缩略语.....14

1 产品简介

1.1 概述

全爱科技推出的 HOUYI-300U1 AI 推理服务器是一款面向人工智能边缘计算的高性能、多功能设备，专为开发者、系统集成商及工商业用户打造，致力于全面支持多形态 AI 应用的开发、算法验证与系统评估。

该服务器采用高性能国产 x86 CPU，搭配华为 Atlas 300V/300I 系列 AI 加速卡，面向高密度 AI 推理、大规模音视频智能解析以及边缘侧大模型部署等前沿应用场景，提供业界领先的边缘算力。系统深度融合工业级可靠性与强劲 AI 性能，可从容应对智能交通、智慧城市、大型园区安防等复杂严苛的边缘计算环境，为用户提供高效、稳定、可扩展的 AI 基础设施支撑。

1.2 产品特点

- 提供高性能 x86CPU，具备良好的软件兼容性，支持多种操作系统；
- 可提供 100 TOPS INT8 算力
- 支持 H.264 / H.265 硬件解码，100 路 1080P 25FPS / 80 路 1080P 30FPS / 10 路 4K 60FPS
- 支持 H.264 / H.265 硬件编码，30 路 1080P 25FPS / 24 路 1080P 30FPS / 4 路 4K 60FPS
- JPEG 解码能力 4K 384FPS，编码能力 4K 192FPS，最大分辨率：8192 x 8192
- 提供丰富的外设接口，满足多种产品形态开发需求。

📖 HOUYI-300U1 默认使用 Atlas 300V 加速推理卡，并支持选配 Atlas 300I Pro、Atlas 300IDuo。性能更高，更多详情可参考 [2.1 基本规格](#)

1.3 外观结构

外观图

HOUYI-300U1 AI 推理服务器采用紧凑的结构设计，外观如图 1-1 所示。



图 1-1 外观图

尺寸图



图 1-2 尺寸图（单位：mm）

接口与按键说明

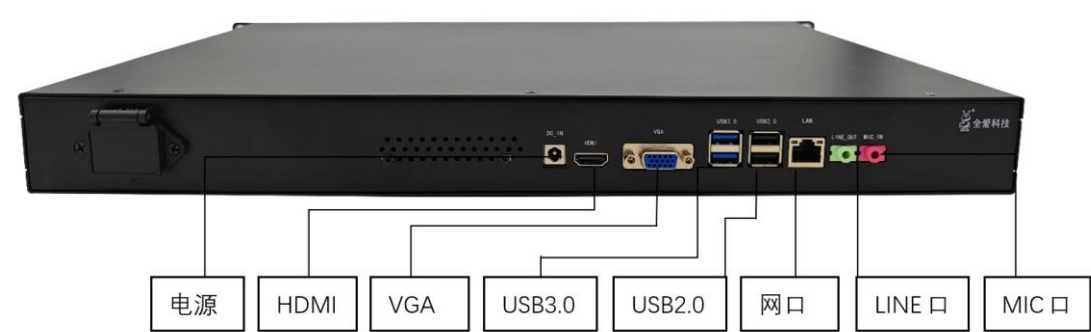


图 1-3 按键与接口说明

1	电源（默认不使用）	2	HDMI 显示接口
3	VGA 显示接口	4	USB 3.0 接口
5	USB 2.0 接口	6	千兆网口
7	LINE 口（音频输出）	8	MIC 接口（音频输入）

2 产品规格

2.1 基本规格

默认配置为 Atlas 300V。支持选配 Atlas 300I Pro、Atlas 300IDuo，性能更高，规格如下。

表 2-1 加速卡基本规格

特征	Atlas 300V（默认配置）	Atlas 300I Pro	Atlas 300IDuo
AI 处理器	1x 310P 系列处理器， 8 个 DaVinci AI Core 8 个自研 CPU 核 100T（INT8）	1 x 310P 系列处理器 8 个 DaVinci AI Core 8 个自研 CPU 核 140T（INT8）	2x 310P 系列处理器： 16 个 DaVinci AI Core 16 个自研 CPU 核 280T（INT8）
CPU 算力	8 core * 1.0 GHz 或 8 core * 1.9 GHz	8 core * 1.9 GHz	16 core * 1.9 GHz
内存	- LPDDR4X - 容量：24GB - 位宽：384bit - 速率：4266Mbps - 总带宽：204.8GByte/s - 支持 ECC	- LPDDR4X - 容量：24GB - 位宽：384bit - 速率：4266Mbps - 总带宽：204.8GByte/s - 支持 ECC	- LPDDR4X - 容量：96GB - 总带宽（整卡）：408GByte/s - 支持 ECC
编解码能力	- 支持 H.264 / H.265 硬件解码，100 路 1080P 25FPS / 80 路 1080P 30FPS / 10 路 4K 60FPS - 支持 H.264 / H.265 硬件编码,30 路 1080P 25FPS / 24 路 1080P 30FPS / 4路 4K 60FPS - JPEG 解码能力 4K 384FPS，编码能力 4K 192FPS，最大分辨率：8192 x 8192	- H.264、H.265 视频编解码 - JPEG 图片编解码	- 支持 H.264 硬件解码，256 路 1080P 30FPS（32 路 3840 * 2160 60 FPS） - 支持 H.265 硬件解码，256 路 1080P 30FPS（32 路 3840 * 2160 60 FPS） - 支持 H.264 硬件编码，48 路 1080P 30FPS（6 路 4K 60 FPS）

特征	Atlas 300V（默认配置）	Atlas 300I Pro	Atlas 300IDuo
			● 支持 H.265 硬件编码，48 路 1080P 30FPS（6 路 4K 60 FPS） ● JPEG 解码能力 4K 768 FPS，编码能力 4K 384 FPS ● 最大分辨率：8192 * 8192
功耗	功耗为 72W	功耗为 72W	功耗为 150W

默认配置为八核 CPU，可选四核 CPU。

表 2-2 CPU 基本规格

特征	处理器（默认配置）	处理器
CPU	八核 2.7G CPU	四核 2.6G CPU
内存	NB-DDR4 内存，支持 2133/2400/2666MHZ 内存，默认 16G	
存储接口	MSATA\默认 NVME256G/SATA	
USB 功能	2*USB3.0 接口、 2*USB2.0 接口	
外设接口	● HDMI 接口：1 个 ● 千兆网口：1 个 ● 电源接口：1 个	
网络功能	● 1/2*LAN 接口	
声音功能	● 1*LINE_OUT 接口，支持音频输出（绿色） ● 1*MIC_IN 接口，支持麦克风输入（红色）	
运行环境	● 工作温度：-10℃~60℃； ● 工作湿度：5%~95%相对湿度，无冷凝	
显示功能	● 显卡：集成 ZXC960DX11 超核心显卡，支持单显、双显/三显复制、双显/三显扩展 ● 1*VGA DB15 接口：支持分辨率 1920*1080@60HZ	
功耗	● 70W	35W

表 2-3 结构基本规格

结构尺寸	300mm x 440mm x44mm（长 x 宽 x 高）
净重	5.15kg
a：稳定提供的峰值算力。	

3.5 寸硬盘重量：550g

表 2-4 软件基本规格

特征	规格
操作系统	openEuler22.03
固件与驱动版本	23.0.RC3

特征	规格
CANN 版本	7.0.0

2.2 环境条件

表 2-5 环境要求

环境指标	规格
温度	工作温度：0℃～+40℃ 存储温度：0℃～+85℃
湿度（RH，无冷凝）	工作湿度：5%～90% 存储湿度：5%～95%
海拔高度	小于 3000m。1800m～3000m，海拔每升高 220m 最高温度规格降低 1℃。

订货可选配置

订货型号	CPU	内存	NPU/内存	加速卡	功耗	mSATA	备注
HOUYI-300U1-8-100	八核 2.7G	16G	100T/24GB	Atlas 300V	142W	1TB	标准配置
HOUYI-300U1-8-140	八核 2.7G	16G	140T/24GB	Atlas 300I Pro	142W	1TB	需订货
HOUYI-300U1-8-280	八核 2.7G	16G	280T/96GB	Atlas 300IDuo	220W	1TB	需订货
发货清单	设备 1 台、电源线 1 根						

- 可选配外设提供 2.5 寸、3.5 寸硬盘
- 可选配四核 2.6G CPU

3 快速使用

3.1 默认软件使用

开机后进入系统

输入账户：root

密码：ok20250101*

```
Authorized users only. All activities may be monitored and reported.
Activate the web console with: systemctl enable --now cockpit.socket

localhost login: root
Password:
Last login: Wed Nov 12 08:47:16 from 192.168.1.87

Authorized users only. All activities may be monitored and reported.

Welcome to 5.10.0-182.0.8.95.el2203sp3.x86_64

System information as of time: Wed Nov 12 10:54:03 AM CST 2025

System load: 1.01
Processes: 211
Memory used: 5.1%
Swap used: 0%
Usage On: 18%
IP address: 192.168.1.171
IP address: 192.168.122.1
Users online: 2

[root@localhost ~]# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: ens5s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 8e:c5:8e:0d:21:60 brd ff:ff:ff:ff:ff:ff
    inet 192.168.1.171/24 brd 192.168.1.255 scope global dynamic noprefixroute ens5s0
        valid_lft 251531sec preferred_lft 251531sec
    inet6 2409:8a3e:2a20:3ed0:8e:c5:8e:0d:21:60/64 scope global dynamic noprefixroute
        valid_lft 259190sec preferred_lft 172790sec
    inet6 fe80:8e:c5:8e:0d:21:60:64 scope link noprefixroute
        valid_lft forever preferred_lft forever
3: enp5s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 10:1b:54:3a:01:d3 brd ff:ff:ff:ff:ff:ff
4: virbr0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc noqueue state DOWN group default qlen 1000
    link/ether 52:54:00:8e:6e:a4 brd ff:ff:ff:ff:ff:ff
    inet 192.168.122.1/24 brd 192.168.122.255 scope global virbr0
        valid_lft forever preferred_lft forever
5: virbr0-nic: <BROADCAST,MULTICAST> mtu 1500 qdisc fq_codel master virbr0 state DOWN group default qlen 1000
    link/ether 52:54:00:8e:6e:a4 brd ff:ff:ff:ff:ff:ff
[root@localhost ~]#
```

可通过 ip a 查询 ip 地址

通过 SSH 登录推理服务器

步骤 1 双击“MobaXterm_Personal_22.2.exe”启动 SSH 登录工具。

步骤 2 点击左上方的"Session"进入界面。

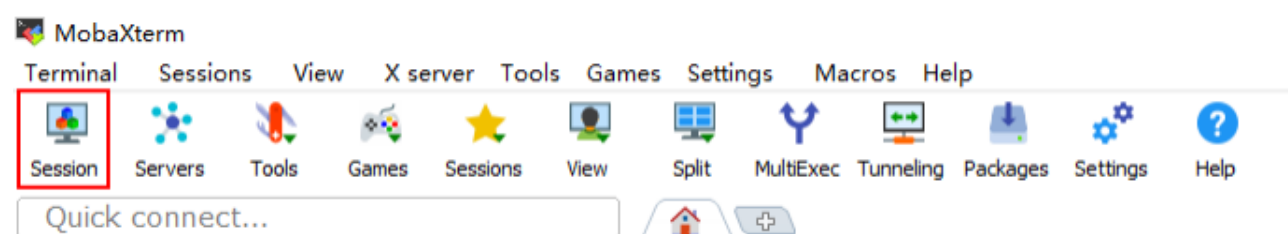


图 3-1 通过 SSH 登录推理服务器界面图

步骤 3 点击左上方的“SSH”进入 SSH 连接配置界面。

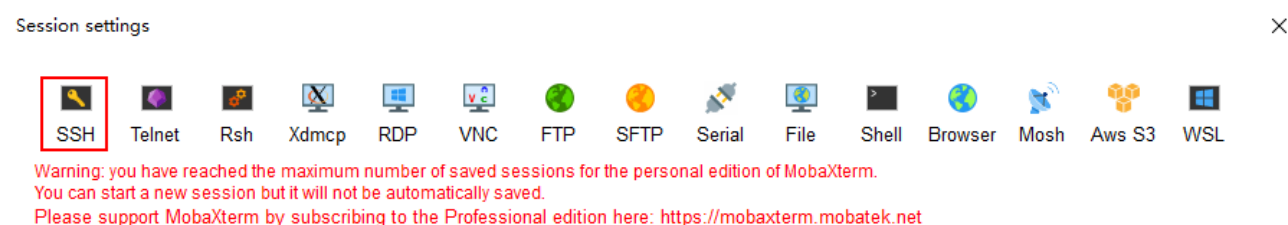


图 3-2 通过 SSH 登录推理服务器界面图

步骤 4 填写 AI 推理服务器网口 IP 地址和 root 用户。

📖 IP 地址可通过桌面登录后查询（以 192.168.1.171 为例）

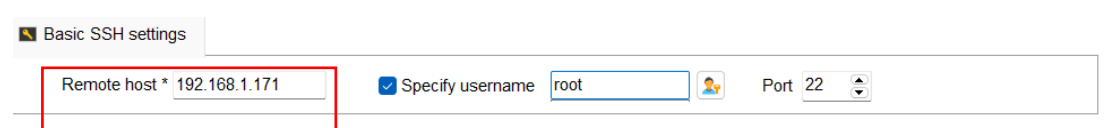


图 3-3 通过 SSH 登录 AI 推理服务器界面图

步骤 5 点击 “OK” 按钮，进入远程登录界面，

输入用户：root

密码：ok20240101*

3.2 默认测试 demo

1. 默认测试目录

```
cd ~/samples/inference/modelInference/sampleYOLOV7/scripts
```

2. 默认测试 demo 的启动指令

```
bash ./sample_run.sh
```

3. 默认测试 demo 的效果

测试模型: yolov7x.om, 138M, FP16 量化

8核设备静态功率: 43.7W

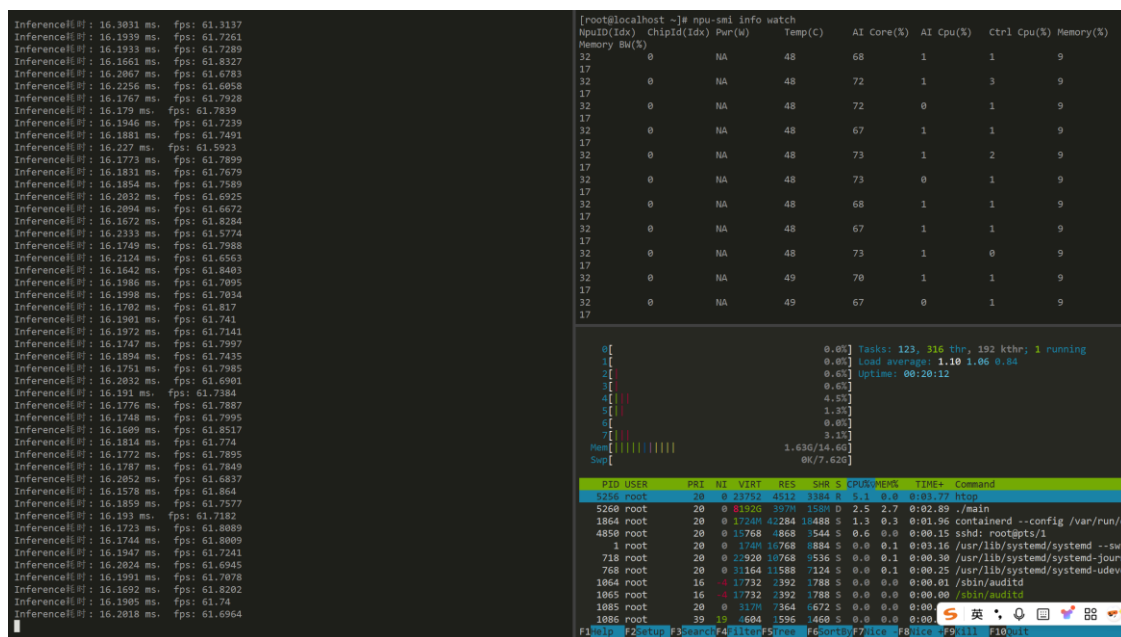
测试程序: 基于

<https://gitee.com/ascend/samples/tree/master/inference/modelInference/sampleYOLOV7>, 修改为: 图像预处理 -> 循环推理, 最大限度的用满npu。

● 8 核设备-单线程

8 核设备下, 单线程跑满 (仅循环处理推理), 推理耗时约为: 16.2104 ms/帧, fps: 61.6642 帧/秒。npu 利用率约为 67~73%。

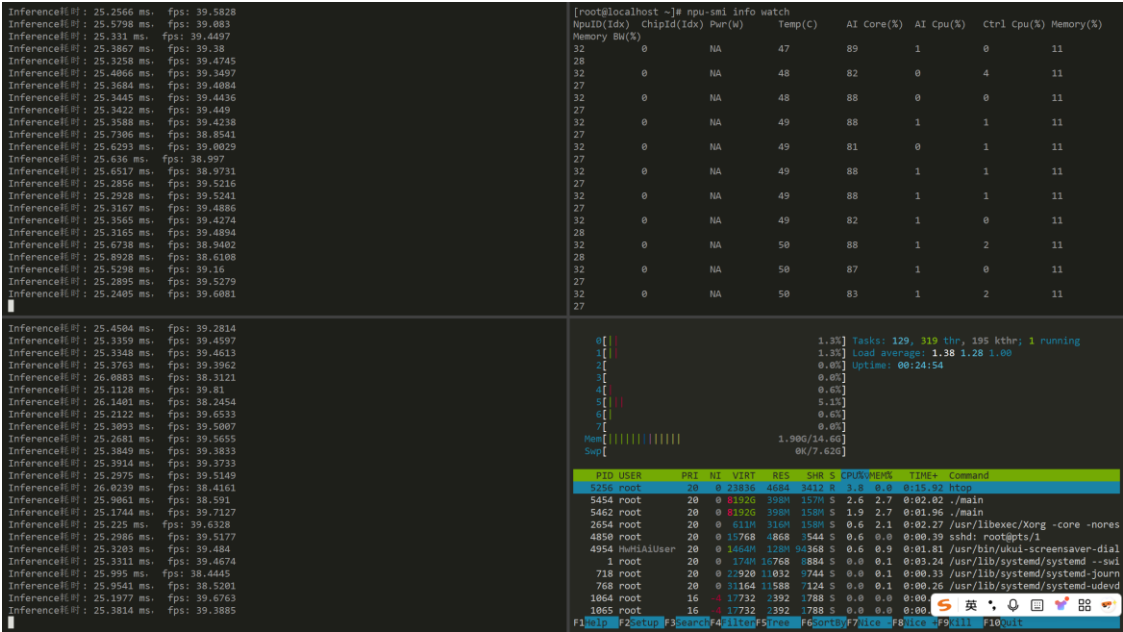
功率约为: 80.5~81.1W



● 8 核设备-2 个线程

8 核设备下, 2 个线程跑满, 每个线程推理耗时约为: 26.0416 ms/帧, fps: 38.3882 帧/秒。npu 利用率约为 81~89%。

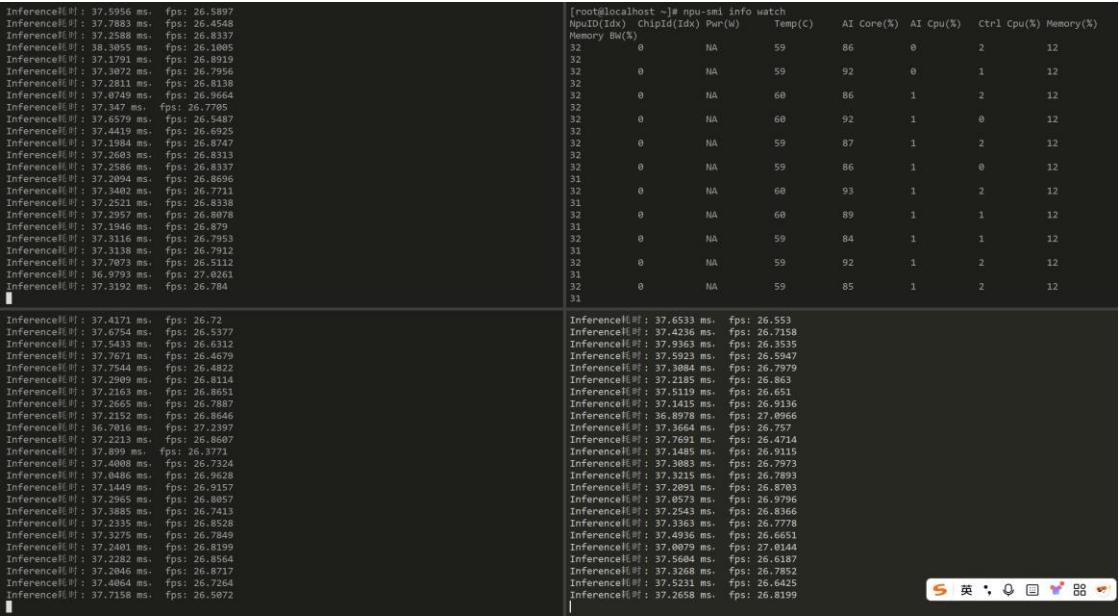
功率约为: 88.1~88.7W



● 8 核设备-3 个线程

8 核设备下，3 个线程跑满，每个线程推理耗时约为：37.4313 ms/帧，fps：26.7101 帧/秒。npu 利用率约为 85~93%。

功率约为：90.5~91.0W



3.3 软件资料

- 昇腾驱动与固件：
链接：<https://pan.baidu.com/s/10qa9y4Y0SrYS-f9vk9zf-A?pwd=prqv> 提

取码：prqv

- 测试例程的 gittee 地址：[samples: CANN Samples](https://gitee.com/Ascend/samples)



- 例程目录

/home/HwHiAiUser/samples/inference/modelInference/sampleY0
LOV7MultiInput

- 例程源码下载地址 <https://gitee.com/Ascend/samples>

- 系统镜像 通过网盘分享的文件：HOUYI-300U1 系统镜像

链接：https://pan.baidu.com/s/1n17E5h2pJss_M-2Wilh_jA?pwd=fy44

提取码：fy44

4 接口说明

4.1 千兆以太网口

HOUYI-300U1 AI 推理服务器对外提供 1 个 10/100/1000M Base-T 接口，接口类型为 RJ45，使用普通网线接入网络。

4.2 USB 3.0 Type-A 接口

HOUYI-300U1 AI 推理服务器对外提供 4 个 Type-A 接口类型 USB 接口，两路 USB 3.0 (SuperSpeed)，两路 USB 2.0 (HighSpeed)。

4.3 HDMI 接口

HOUYI-300U1 AI 推理服务器对外提供一个 HDMI 接口，最大支持连接 4K@60Hz 分辨率设备。

4.4 VGA 接口

HOUYI-300U1 AI 推理服务器对外提供一个 VGA 接口。

4.5 音频接口

HOUYI-300U1 AI 推理服务器对外提供一个 LINE 输出接口，一个 MIC 输入接口。

4.6 电源接口

HOUYI-300U1 AI 推理服务器使用 220V 供电。供电方式如下图。



 AC220V 供电后按此按钮即可开关机

A 缩略语

A.1 A-E

A

AI	人工智能（Artificial Intelligence）
----	-------------------------------

B

BTB	板对板连接器（Board to Board Connector）
-----	----------------------------------

E

ECC	错误检查和纠错技术（Error Checking and Correcting）
eMMC	嵌入式多媒体卡（Embedded Multimedia Card）

A.2 F-J

F

FLOPS	每秒浮点运算次数（Floating-point Operations Per Second）
FCC	美国联邦通信委员会（Federal Communications Commission）
HDMI	高清多媒体接口（High-Definition Multimedia Interface）

I

I ² C	内部整合电路（Inter-integrated Circuit）
------------------	----------------------------------

A.3 K-O

L

LPDDR	低功耗双倍速 (Low-power Double Data Rate)
-------	-------------------------------------

A.4 P-T

P

PWM	脉冲宽度调制 (Pulse-width Modulation)
PCIe	快捷外围部件互连标准 (Peripheral Component Interconnect Express)

R

RGMII	精简的千兆比媒介独立接口 (Reduced Gigabit Media Independent Interface)
RS232	串行数据传输的EIA标准 (Recommended Standard 232)
RS485	通信标准 (Recommended Standard 485)

S

SPI	串行外设接口 (Serial Peripheral Interface)
-----	--------------------------------------

T

TFLOPS	每秒万亿次的浮点运算 (teraFLOPS)
--------	------------------------

A.5 U-Z

U

UART	通用异步收发传输器 (Universal Asynchronous Receiver/transmitter)
USB	通用串行总线 (Universal Serial Bus)