

LLaVA

LLaVA

- Model scale
- Got LLaVA
- Use LLaVA
 - Run LLaVA
 - Dialogue
 - End conversation
- Reference material

Demonstration environment

Development board : Raspberry Pi 5B (8G RAM)

SD(TF)card: 64G (Above 16G, the larger the capacity, the more models can be experienced)

Raspberry Pi 5B (8G RAM): Run 8B and below parameter models
Raspberry Pi 5B (4G RAM): Run 3B and below parameter models, can't run LLaVA
Raspberry Pi 5B (2G RAM): Run 0.5B and below parameter models, can't run LLaVA

LLaVA (Large scale Language and Vision Assistant) is a multimodal model aimed at achieving universal visual and language understanding by combining visual encoders and large-scale language models.

Model scale

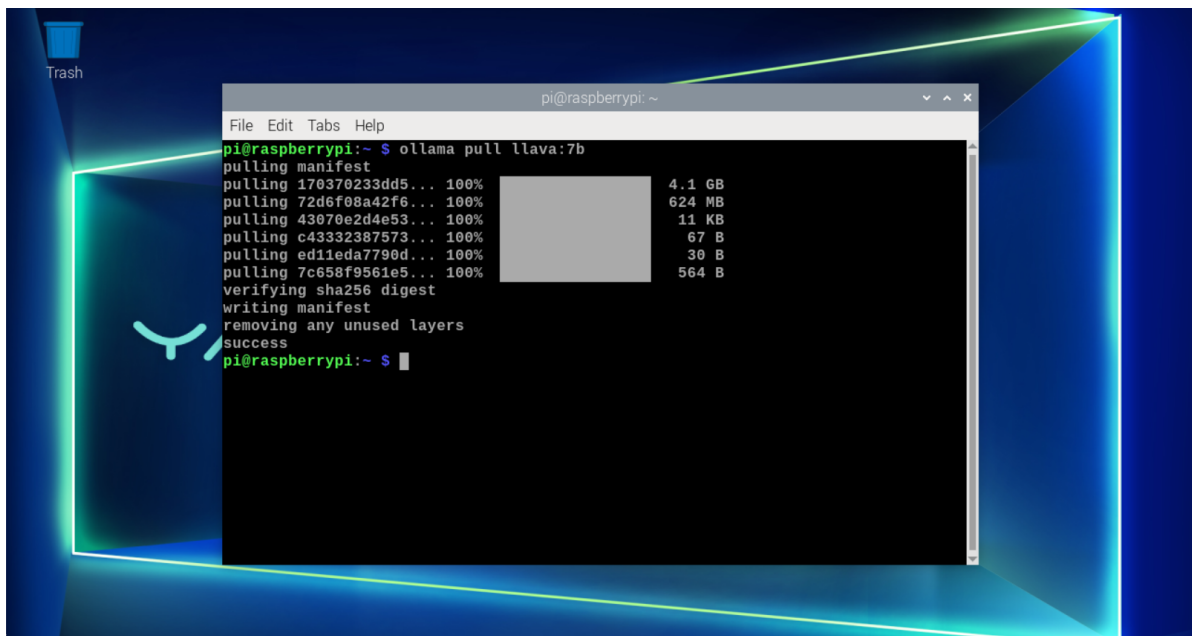
Model	Parameter
LLaVA	7B
LLaVA	13B
LLaVA	34B

Raspberry Pi 5B (8G RAM): LLaVA model testing with 7B parameters.

Got LLaVA

Using the pull command will automatically pull the models from the Ollama model library.

```
ollama pull llava:7b
```



Use LLaVA

Use LLaVA to recognize local image content.

Run LLaVA

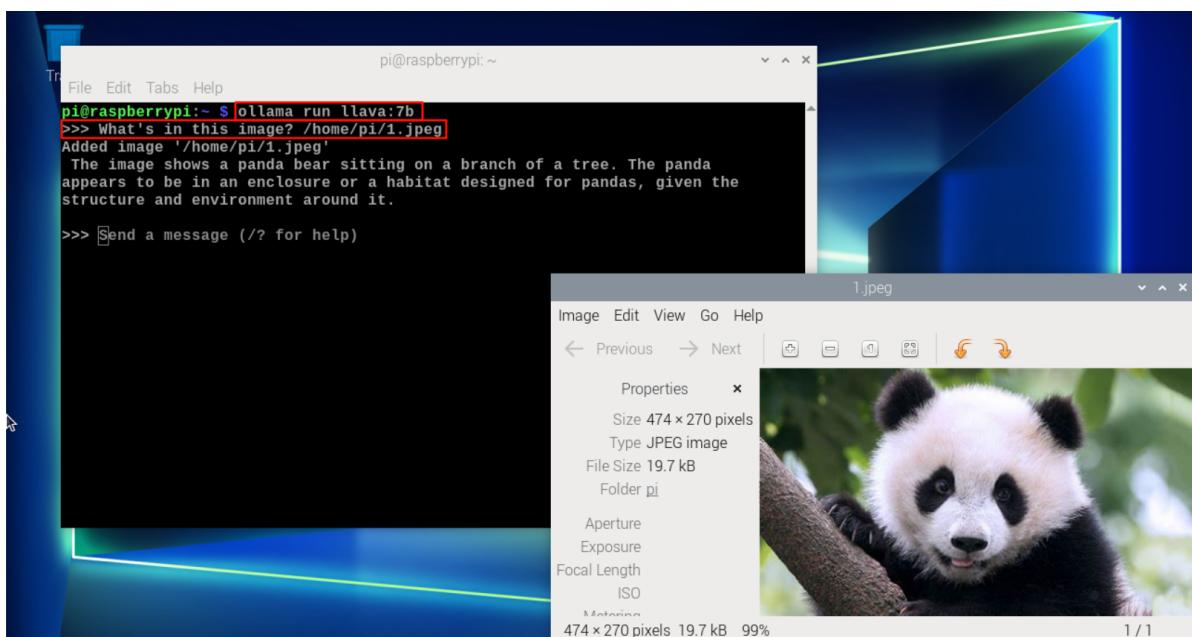
If the system does not have a running model, the system will automatically pull the LLaVA 7B model and run it.

```
ollama run llava:7b
```

Dialogue

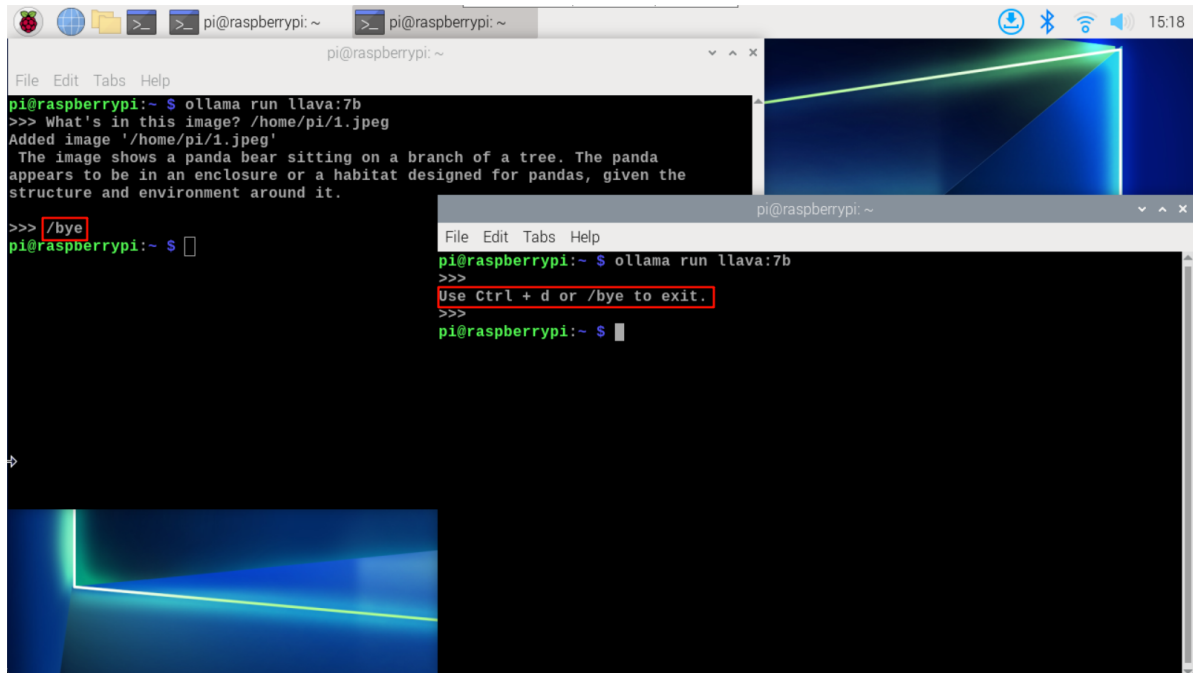
```
what's in this image? /home/pi/1.jpeg
```

The time to reply to the question is related to the hardware configuration, please be patient.



End conversation

You can end the conversation by using the shortcut key 'Ctrl+d' or '/bye'.



Reference material

Ollama

Website: <https://ollama.com/>

GitHub: <https://github.com/ollama/ollama>

LLaVA

GitHub: <https://github.com/haotian-liu/LLaVA>

Ollama model: <https://ollama.com/library/llava>