

# LLaVA-Phi3

## LLaVA-Phi3

- 1. Model scale
  - 2. Performance
  - 3. Pull LLaVA-Phi3
  - 4. Use LLaVA-Phi3
    - 4.1. Run LLaVA-Phi3
    - 4.2. Have a conversation
    - 4.3. End the conversation
- [References](#)

### Demo Environment

**Development board:** Jetson Orin series motherboard

**SSD:** 128G

**Tutorial application scope:** Whether the motherboard can run is related to the available memory of the system. The user's own environment and the programs running in the background may cause the model to fail to run.

Motherboard model	Run directly with Ollama	Run with Open WebUI
Jetson Orin NX 16GB	√	√
Jetson Orin NX 8GB	√	√
Jetson Orin Nano 8GB	√	√
Jetson Orin Nano 4GB	√	√

LLaVA-Phi3 is a LLaVA model fine-tuned from the Phi 3 Mini 4k.

LLaVA (Large-scale Language and Vision Assistant) is a multimodal model that aims to achieve general vision and language understanding by combining a visual encoder and a large-scale language model.

## 1. Model scale

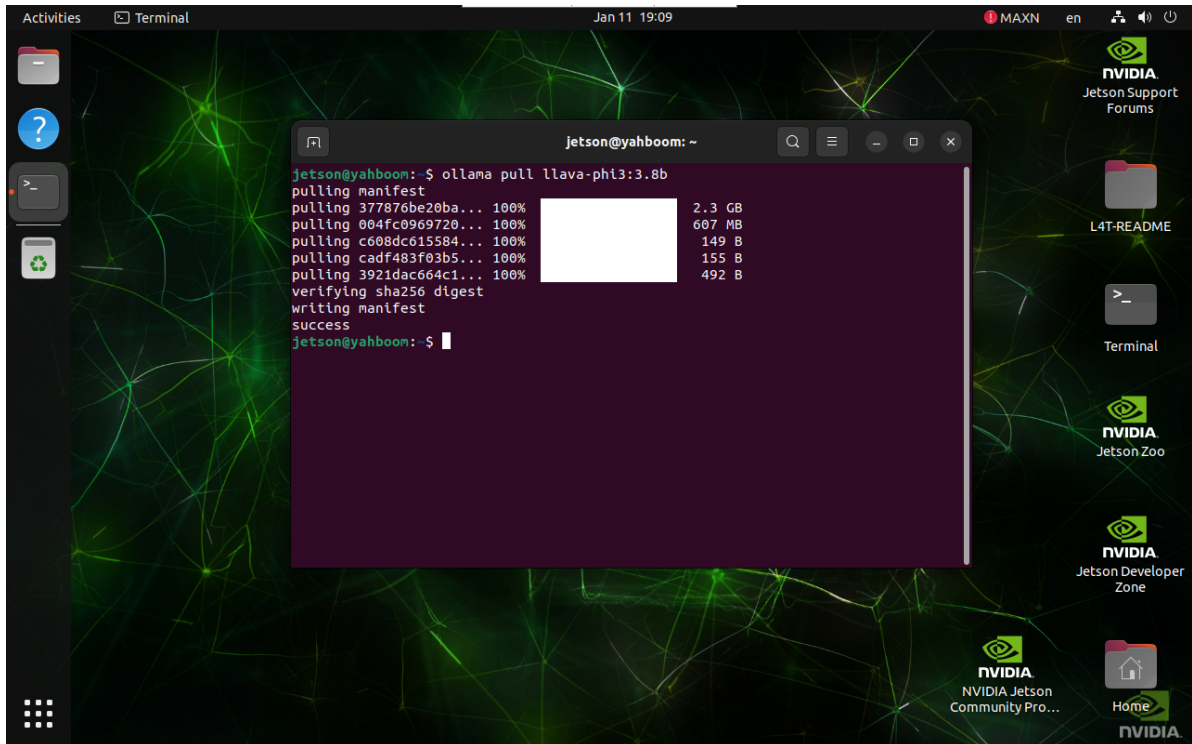
Model	Parameters
LLaVA-Phi3	3.8B

## 2. Performance

## 3. Pull LLaVA-Phi3

Using the pull command will automatically pull the model of the Ollama model library:

```
ollama pull llava-phi3:3.8b
```



## 4. Use LLaVA-Phi3

Use LLaVA-Phi3 to identify local image content.

### 4.1、 Run LLaVA-Phi3

If the system does not have a running model, the system will automatically pull the LLaVA-Phi3 3.8B model and run it:

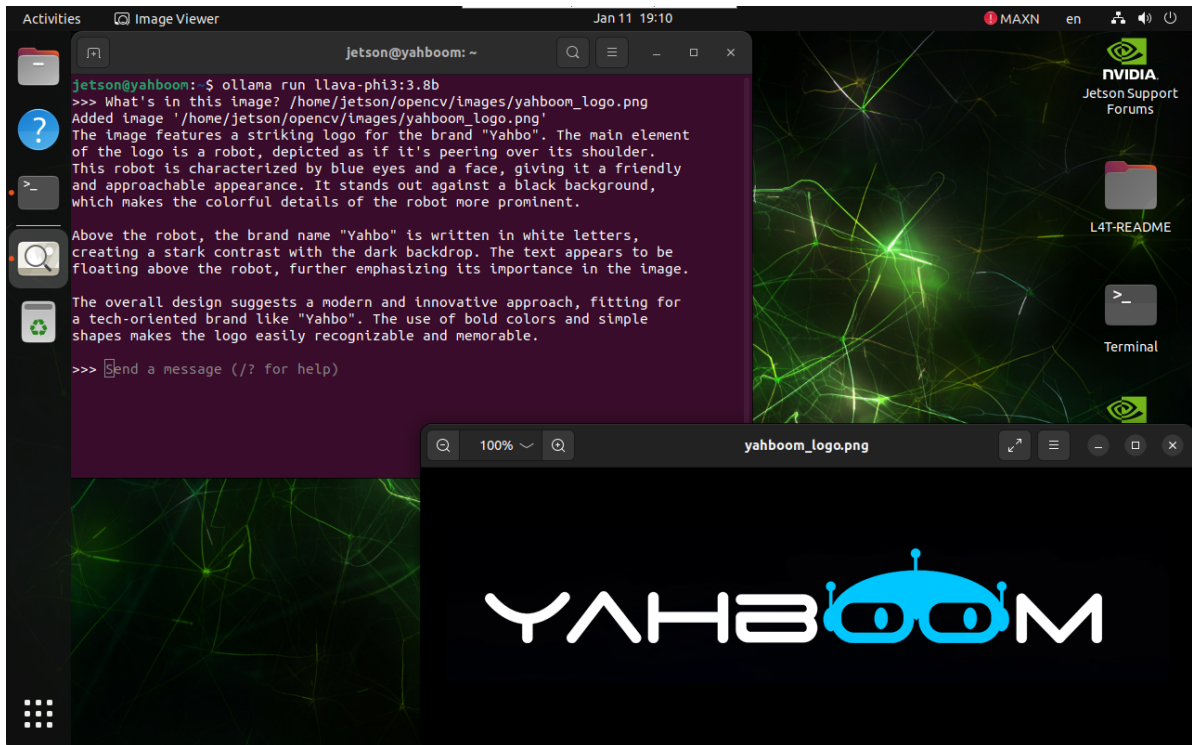
```
ollama run llava-phi3:3.8b
```

### 4.2、 Have a conversation

```
what's in this image? /home/jetson/opencv/images/yahboom_logo.png
```

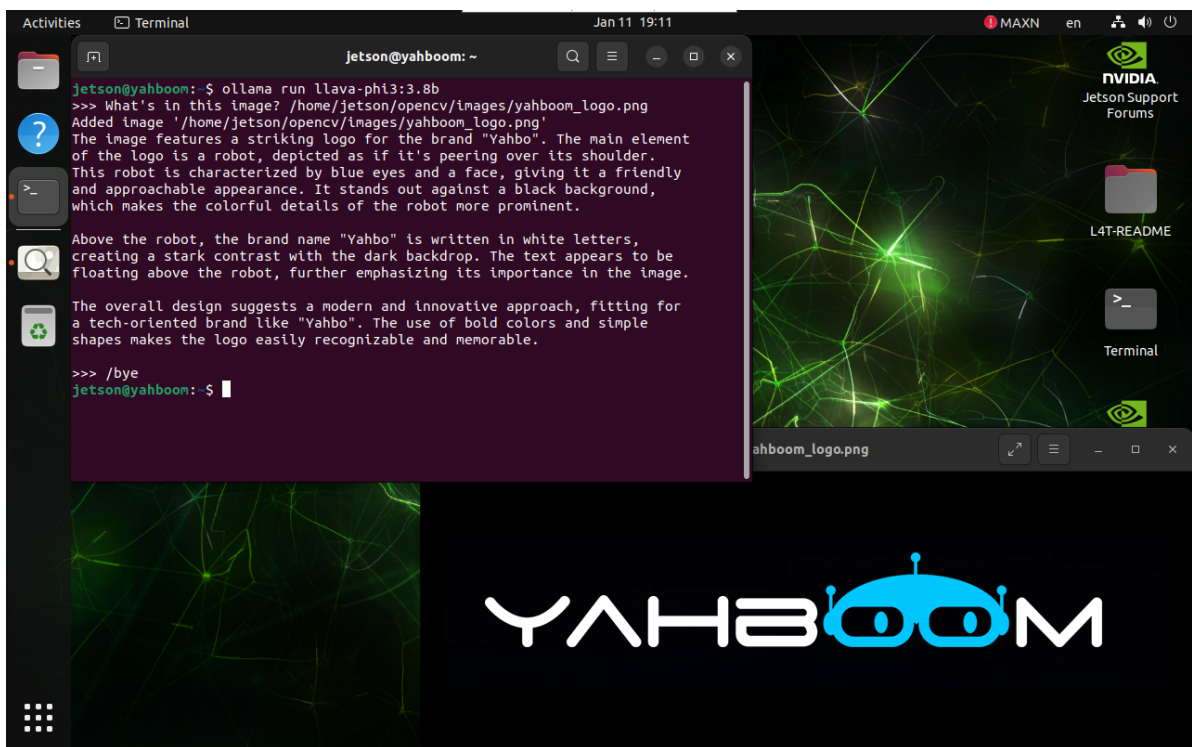
The time to reply to the question depends on the hardware configuration, so be patient!

If the image does not have a corresponding image, you can download the image yourself (the resolution should not be too large), and put the image path after the question!



### 4.3. End the conversation

Use the `Ctrl+d` shortcut key or `/bye` to end the conversation!



## References

### Ollama

Official website: <https://ollama.com/>

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

### LLaVA-Phi3

GitHub: <https://github.com/InternLM/xtuner/tree/main>

Ollama corresponding model: <https://ollama.com/library/llava-phi3>