

Expand LLaVA-Phi3

Expand LLaVA-Phi3

[Model scale](#)

[Performance](#)

[Got LLaVA-Phi3](#)

[Use LLaVA-Phi3](#)

[Run LLaVA-Phi3](#)

[Have a conversation](#)

[End the conversation](#)

[References](#)

Demo Environment

Development Boards: Jetson nano

SD(TF) card: 64G

It is recommended to run the 4B and below parameter models

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.

Model scale

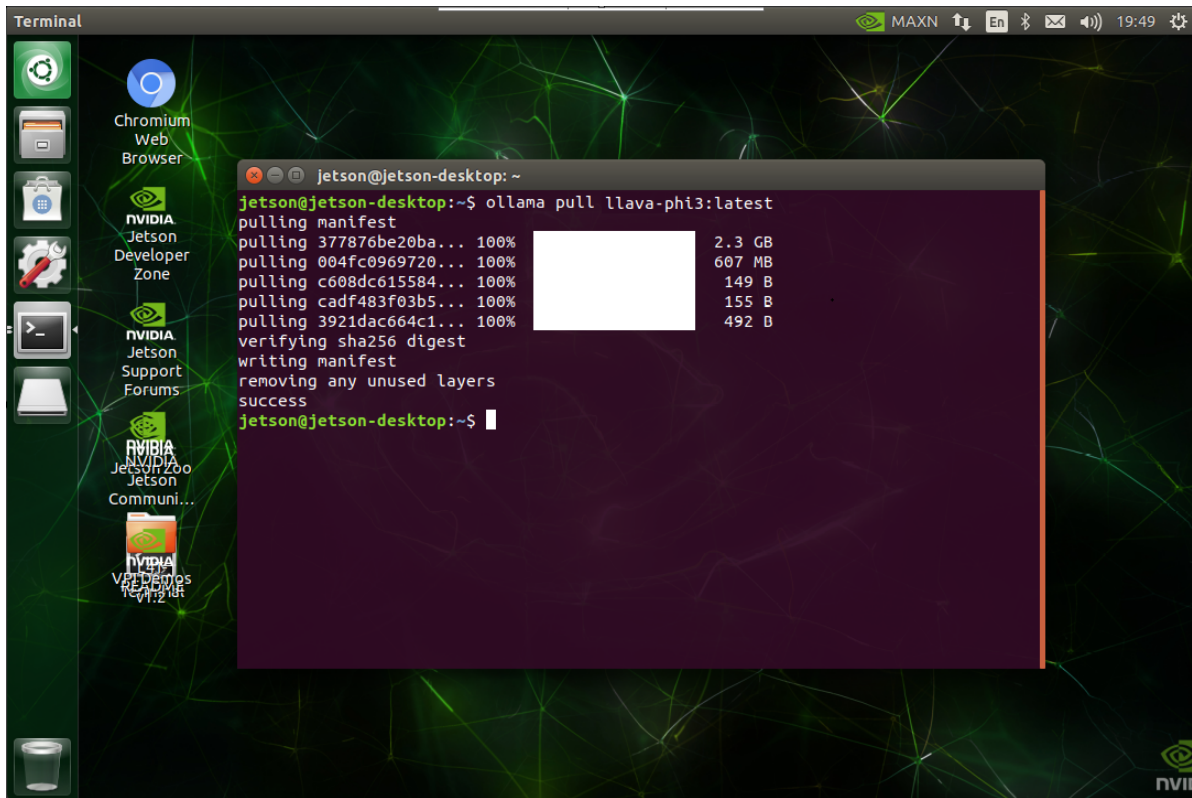
Model	Parameter
LLaVA-Phi3	3.8B

Performance

Got LLaVA-Phi3

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

```
ollama pull llava-phi3:latest
```



Use LLaVA-Phi3

Use LLaVA-Phi3 to identify local image content.

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.

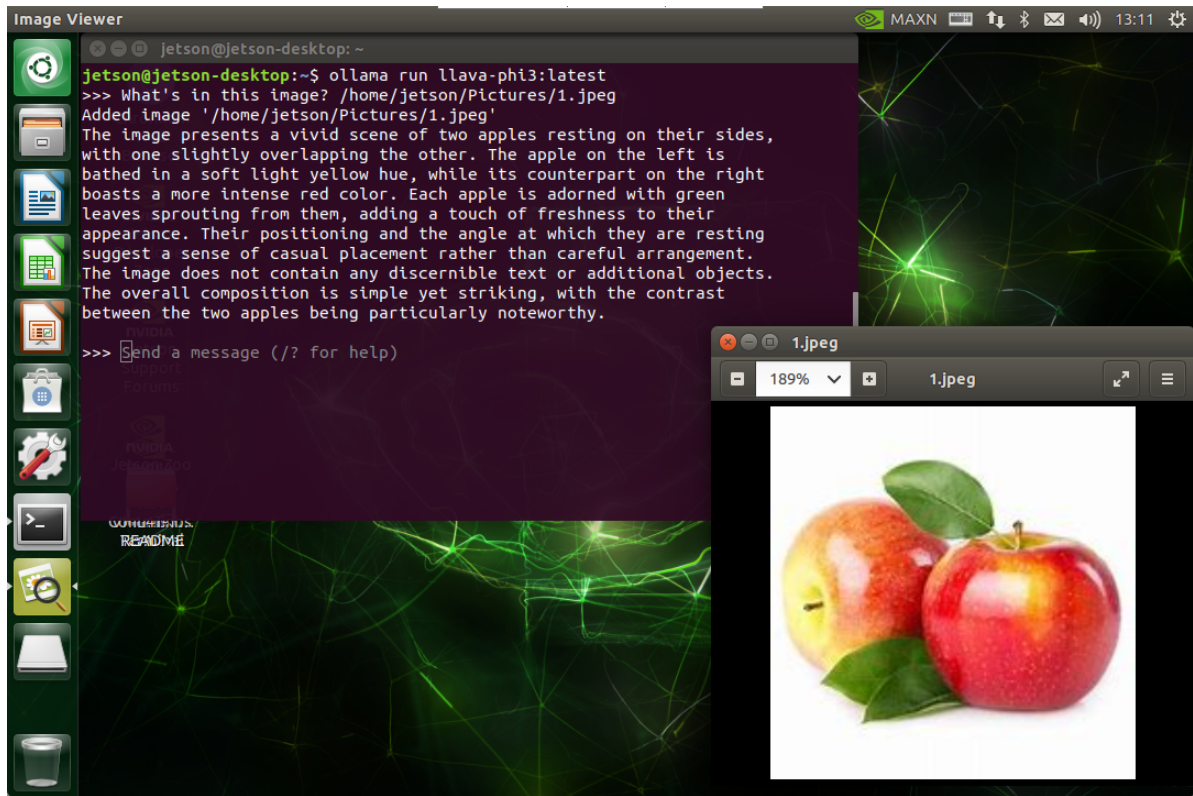
```
ollama pull llava-phi3:latest
```

If an error occurs during operation, you can restart the system and try again.

Have a conversation

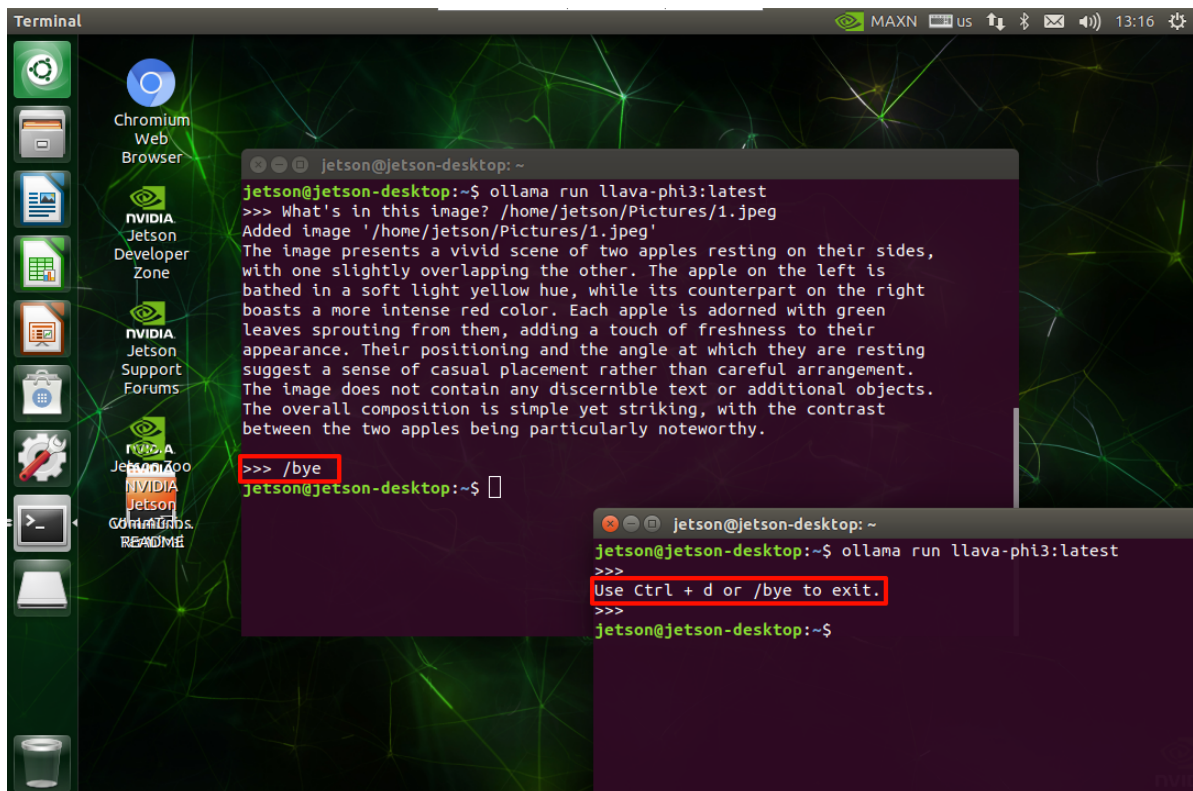
```
what's in this image? /home/jetson/Pictures/1.jpeg
```

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



End the conversation

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



References

Ollama

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>