1.3.3 Vision Language Model: Object Detection

Note:

1) **This section requires the configuration of the API key in "1.3.2 Vision Language Model Accessing" before proceeding. Additionally, ensure that the images to be used in this section are imported.**

2) **This experiment requires either an Ethernet cable or Wi-Fi connection to ensure the main control device can access the network properly.**

3) **In this course, we will use a program to transmit an image to the large model for recognition, which will then identify and locate the objects within the image by drawing bounding boxes around them.**

1. Experiment Steps

1) Execute the following command to navigate to the directory of Large Model.

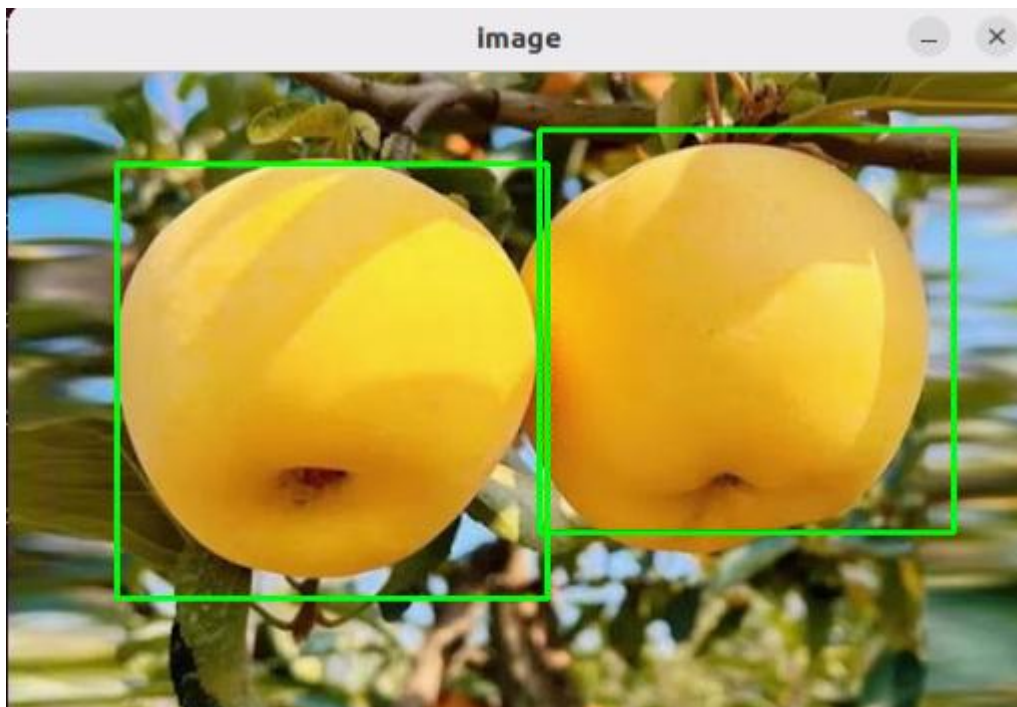**cd large_models/**

```
cd large_models/
```

2) Run the program:

python3 qwen_vllm_detect_demo.py

```
python3 qwen_vllm_detect_demo.py
```

2. Function Realization

After running the program, the positions of the fruits in the image will be circled.

```
{
    "persimmon1": ["persimmon", 0.107, 0.143, 0.526, 0.817],
    "persimmon2": ["persimmon", 0.518, 0.088, 0.923, 0.713],
}
{'persimmon1': ['persimmon', 0.107, 0.143, 0.526, 0.817], 'persimmon2': ['persimmon', 0.518, 0.088, 0.923, 0.713]}
341 512
```
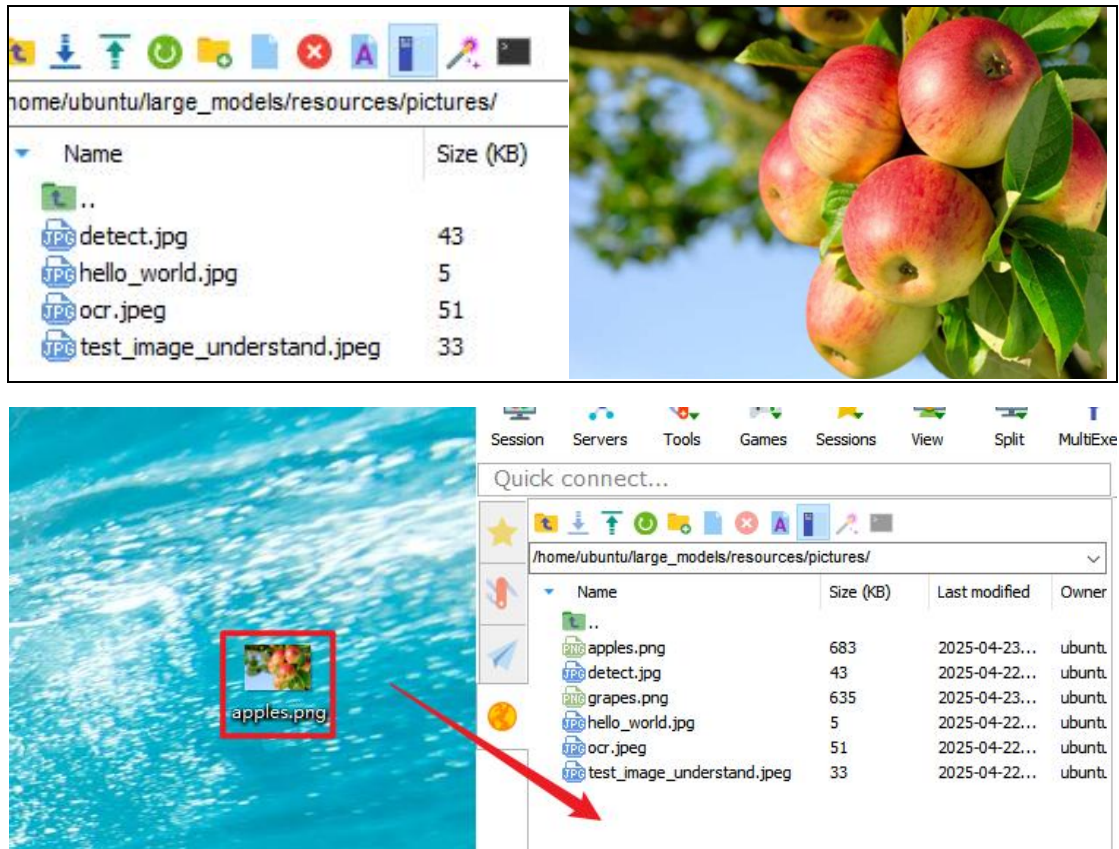
3. Function Expansion

We can switch the image and change the large model to experience different functionalities of various models.

4. Change Pictures

1) Click on the path box to navigate to the following directory:
   /home/ubuntu/large_models/resources/pictures/

Here, you can drag in other images, for example, in the apples.png format.
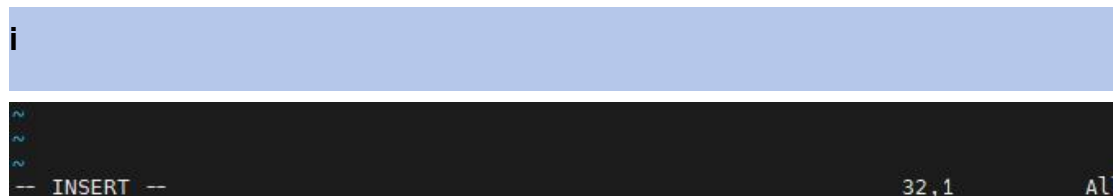

2) Then, input the command:

```
vim large_models/qwen_vllm_detect_demo.py
```

```
> vim large_models/qwen_vllm_detect_demo.py
```

3) Press the "i" key on your keyboard, which will display "INSERT" at the bottom.

```
i
```

```
~
~
~
-- INSERT --                                              32,1        Al
```

4) Change the image recognition path from:

./resources/pictures/test_image_understand.jpeg

To: image = cv2.imread('./resources/pictures/apples.png')

```
30    image = cv2.imread('./resources/pictures/apples.png')
```

5) Next, input the following command and execute the program again to see
the results

python3 qwen_vllm_detect_demo.py

```
python3 qwen_vllm_detect_demo.py
```