Make a MLLM with just training on single RTX 3090 24G

Train demo:

First step: you need to download the weight of the vit+q-former structure, and the weight of the llm model Vit weights：

**1.Weights on blip2**

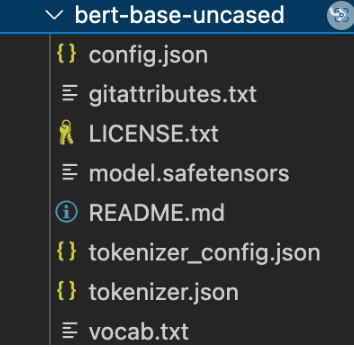
(a) eva vit-g (image encoder)

<https://storage.googleapis.com/sfr-vision-language-research/LAVIS/models/BLIP2/eva_vit_g.pth>

q-former weights：

(b) bert-base-uncased

<https://huggingface.co/bert-base-uncased/tree/main>



(c)blip2\_pretrained\_flant5xxl

<https://storage.googleapis.com/sfr-vision-language-research/LAVIS/models/BLIP2/blip2_pretrained_flant5xxl.pth>

**2.weight on LLM**

(a) qwen-7B-chat

<https://huggingface.co/Qwen/Qwen-7B-Chat>

**Directory structure**

├── cache

│ ├── ckpt

│ │ ├── bert-base-uncased

│ │ ├── blip2

│ │ │ ├── blip2\_pretrained\_flant5xxl.pth

│ │ ├── eva

│ │ │ ├── eva\_vit\_g.pth

│ │ ├── Qwen7B-chat

**Test the model:**

python test\_model\_chat.py

**Run the demo:**

**Without rag : python cli\_demo.py --checkpoint-path xxxxxx**

**With rag:** **python cli\_demo2.py --checkpoint-path xxxxxx**

After running, you need to enter the image path, enter the dialog

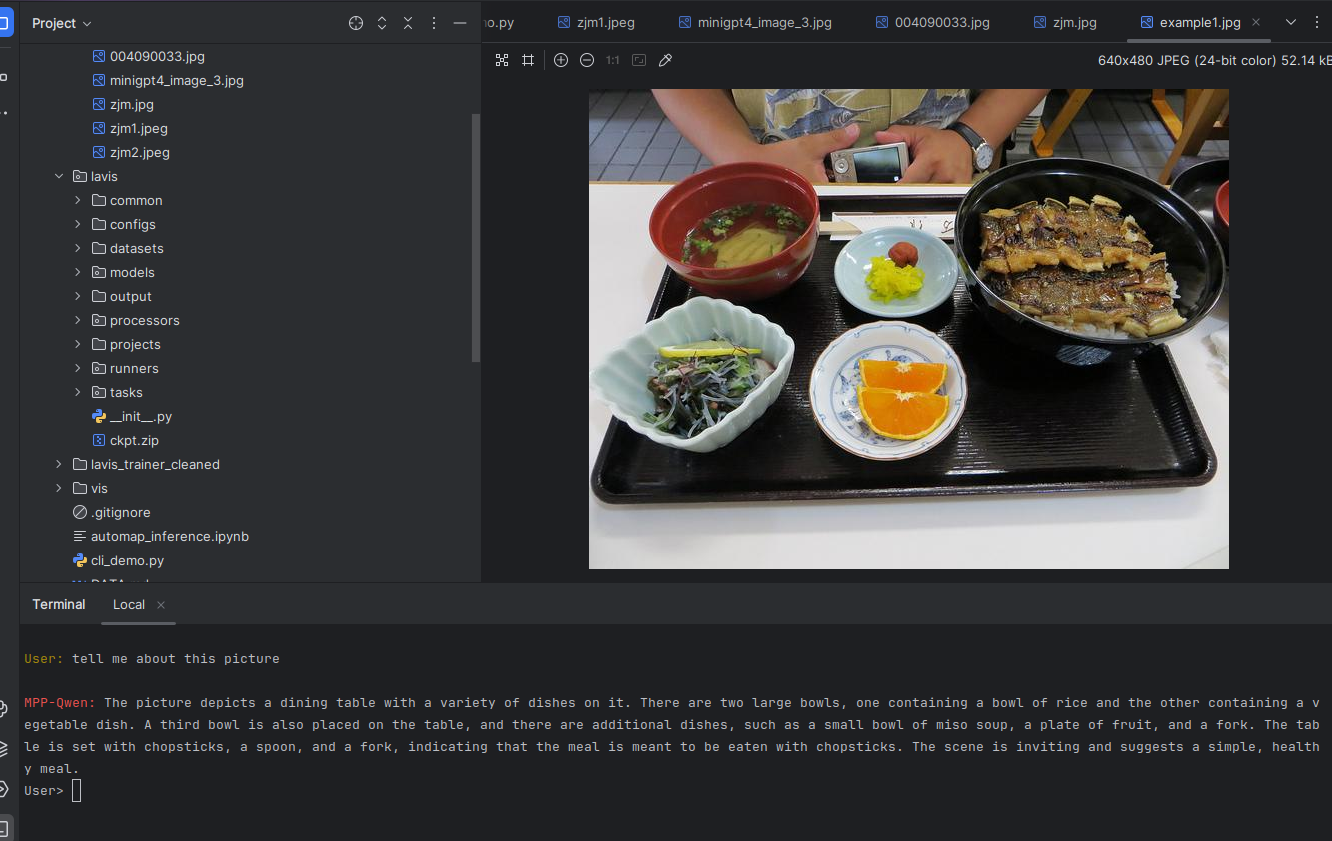
Common operations:

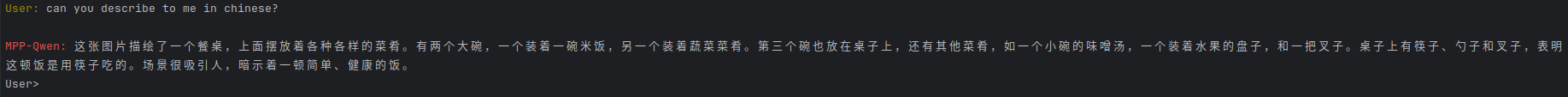
:help View the help

:clear Clears the current command line

:clh clears the conversation history (but the image input does not change)

:his View the conversation history









If you want to see the structure of the entire model, here it is:

NLP\_termporject/lavis/models/minigpt4qwen\_models/minigpt4qwen.py

#############################################################################################

If you need to download data for training:

<https://huggingface.co/datasets/deepHug/minigpt4_training_for_MMPretrain>

**Directory structure**

├── cache

│ └── dataset

│ ├── llava

│ │ │ ├── llava\_minigpt4qwen\_format.json

│ │ │ ├── image

│ ├── minigpt4

│ │ │ ├── image

│ │ │ ├── minigpt4\_minigpt4qwen\_format.json

Train:

CUDA\_VISIBLE\_DEVICES=xxx python train.py --cfg-path lavis/projects/instruction\_tuning/train.yaml

################################################################################################