LLM (MiniCPM-V-2_6) for Image and Video Understanding: How to Install and Run MiniCPM-V

MiniCPM is series of end-side multimodel LLM designed for visual language understanding. It will answer questions related images and videos you uploaded.

sudo apt update
sudo apt upgrade

sudo apt install git-all

Then, make sure that Git support for large-scale files is installed by running
curl -s https://packagecloud.io/install/repositories/github/git-lfs/script.deb.sh | sudo bash

sudo apt-get install git-lfs
git lfs install

git clone https://github.com/OpenBMB/MiniCPM-V.git

cd MiniCPM-V
ls -la

create virtual environment;
python3 -m venv minicpm
source minicpm/bin/activate

Lets install the required libraries:

pip install setuptools

to see the list >> pip list

pip install -r requirements.txt

Then, we need to install three additional packages to be able to run the model
pip install huggingface_hub
pip install wheel
pip install flash_attn

Lets download the model:

mkdir model_files

Download model using hugging face hub. Run this script:
from huggingface_hub import snapshot_download
snapshot_download(repo_id="openbmb/MiniCPM-V-2_6", local_dir = "/home/modelscope/shan/MiniCPM-V-2_6")

Run the main script:
python Chat-with-multiple-images.py

Create conda environment
conda create -n MiniCPM-V-2_6 python=3.10 -y && conda activate MiniCPM-V-2_6

Install required packages
pip install -r requirements.txt -i https://pypi.tuna.tsinghua.edu.cn/simple --default-timeout=1000

Install jupyter to run code on jupyter (Optional)
conda install jupyter -y
uninstall charset_normalizer -y
pip install charset_normalizer -y
jupyter notebook

Download model:

modelscope download --model AI-ModelScope/MiniCPM-V-2_6 --local_dir /home/modelscope/shan/MiniCPM-V-2_6/model