ComfyUI-llama-cpp_vllm 安装步骤

目录

----------感谢夜风大佬制作

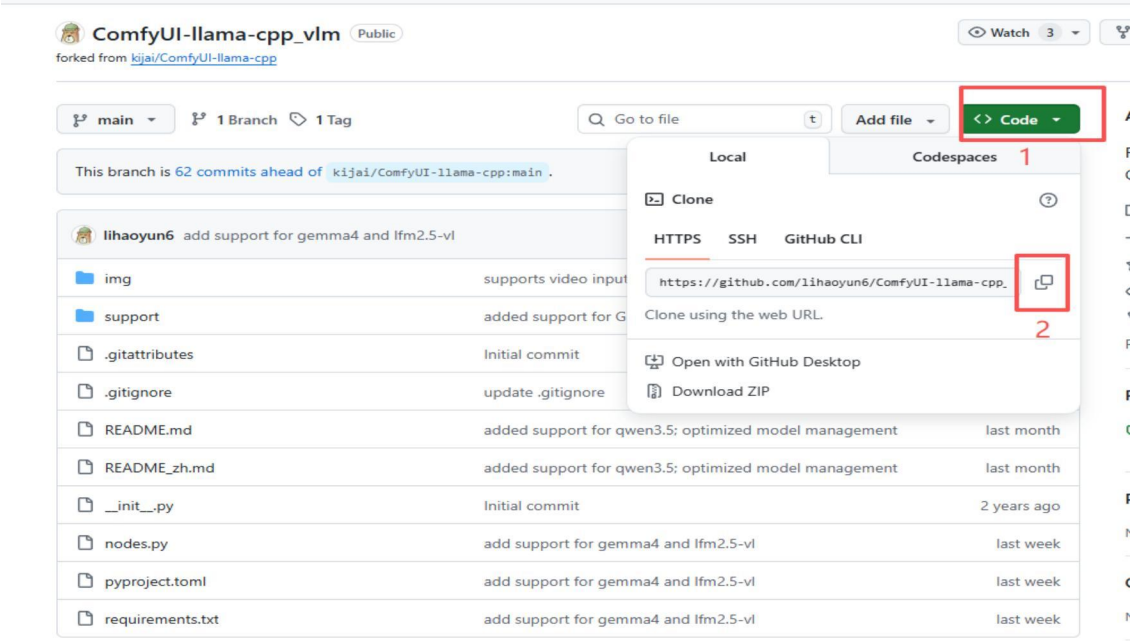
1 NB 管理器安装

安装前打开MF

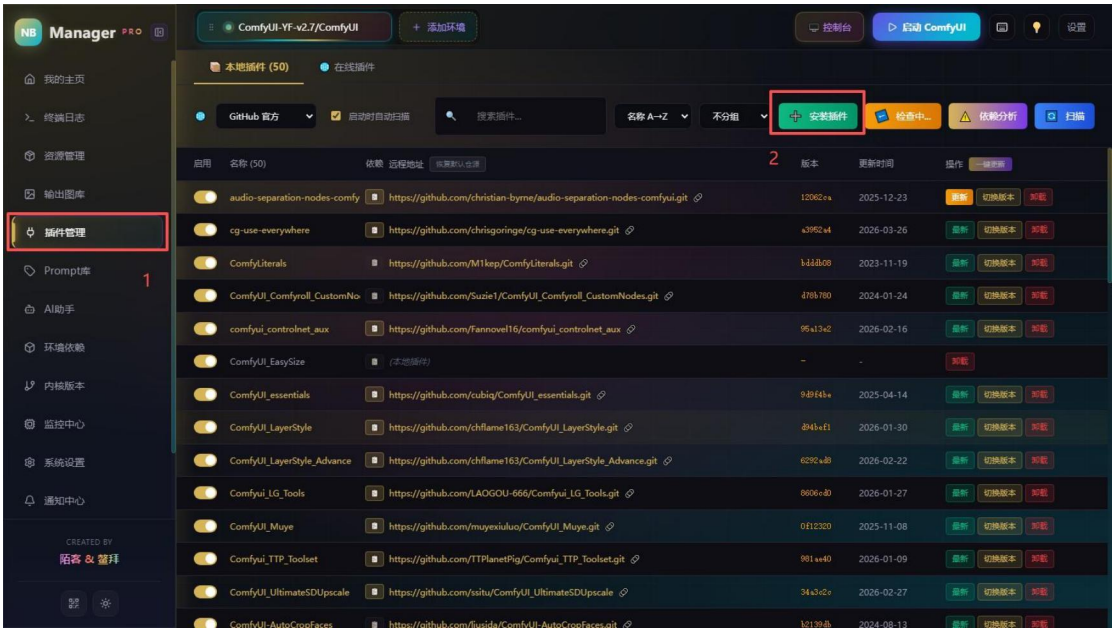
1.1 安装插件

1. 打开网址，复制 git 地址：

Github 网址： https://github.com/lihaoyun6/ComfyUI-llama-cpp_vlm



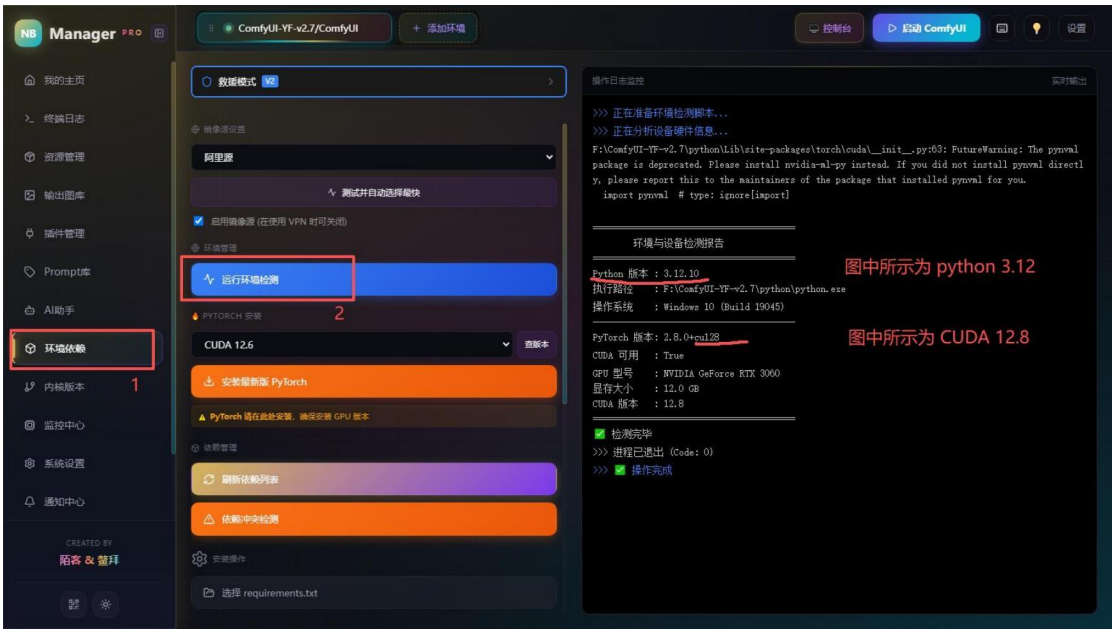
2. 打开 NB 启动器


Manager
ComfyUI-YF-v2.7/ComfyUI
添加环境
控制台
启动 ComfyUI
设置
本地插件 (50)
在线插件
GitHub 官方
启动时自动扫描
搜索插件...
名称 A→Z
不分组
安装插件
检查中...
依赖分析
扫描
我的主页
终端日志
资源管理
输出图库
插件管理
Prompt库
AI助手
环境依赖
内核版本
监控中心
系统设置
通知中心
1
2
启用
名称 (50)
依赖
远程地址
版本
更新时间
操作
audio-separation-nodes-comfy
https://github.com/christian-byrne/audio-separation-nodes-comfyui.git
cg-use-everywhere
https://github.com/chrisgoringe/cg-use-everywhere.git
ComfyLiterals
https://github.com/M1kep/ComfyLiterals.git
ComfyUI_Comfyroll_CustomNo
https://github.com/Suzie1/ComfyUI_Comfyroll_CustomNodes.git
comfyui_controlnet_aux
https://github.com/Fannovel16/comfyui_controlnet_aux
ComfyUI_EasySize
ComfyUI_essentials
https://github.com/cubiq/ComfyUI_essentials.git
ComfyUI_LayerStyle
https://github.com/chflame163/ComfyUI_LayerStyle.git
ComfyUI_LayerStyle_Advance
https://github.com/chflame163/ComfyUI_LayerStyle_Advance.git
Comfyui_LG_Tools
https://github.com/LAOGOU-666/Comfyui_LG_Tools.git
ComfyUI_Muye
https://github.com/muyexiuluo/ComfyUI_Muye.git
Comfyui_TTP_Toolset
https://github.com/TTPlanetPig/Comfyui_TTP_Toolset.git
ComfyUI_UltimateSDUpscale
https://github.com/ssitu/ComfyUI_UltimateSDUpscale.git
CREATED BY
陌客 & 蔷拜


安装新插件
输入 GitHub 或其他 Git 仓库地址来安装插件
https://github.com/lihaoyun6/ComfyUI-llama-cpp_vlm.git
GitHub 加速镜像
将刚刚复制的git地址粘贴进来
GitHub 官方
GitHubFast
GitCode
GHProxy
自定义镜像
取消
开始安装

1.2 安装轮子

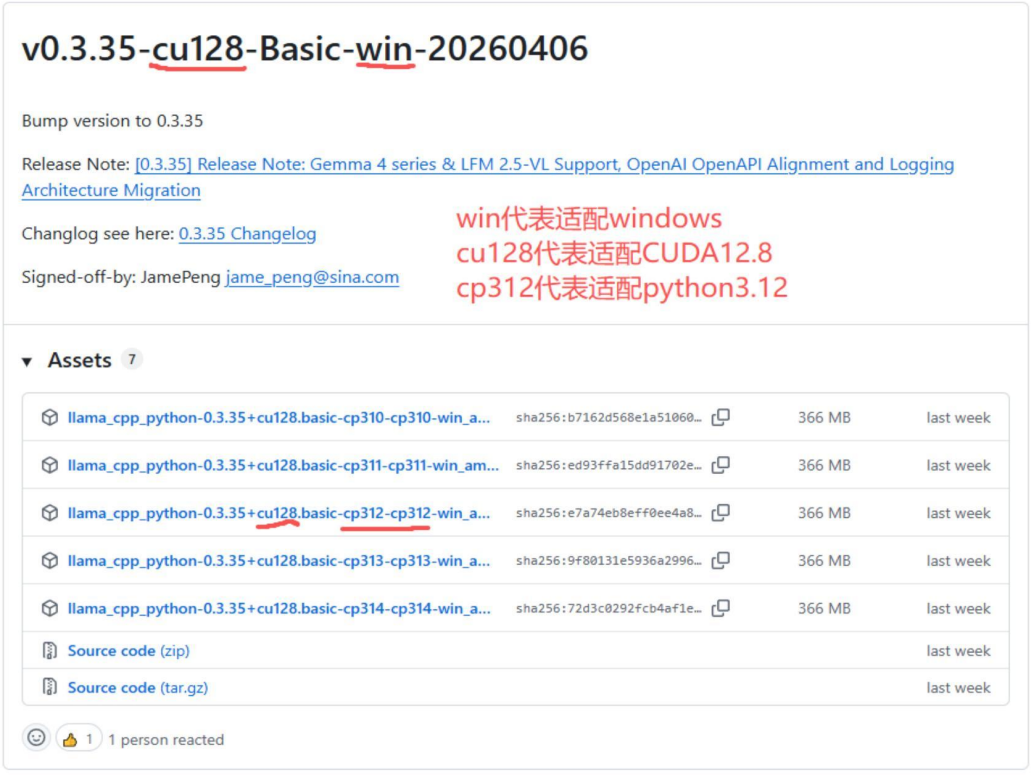
1. 查看 python 版本、 CUDA 版本



2. 下载对应版本的轮子.whl 文件， 此处以上图版本为例

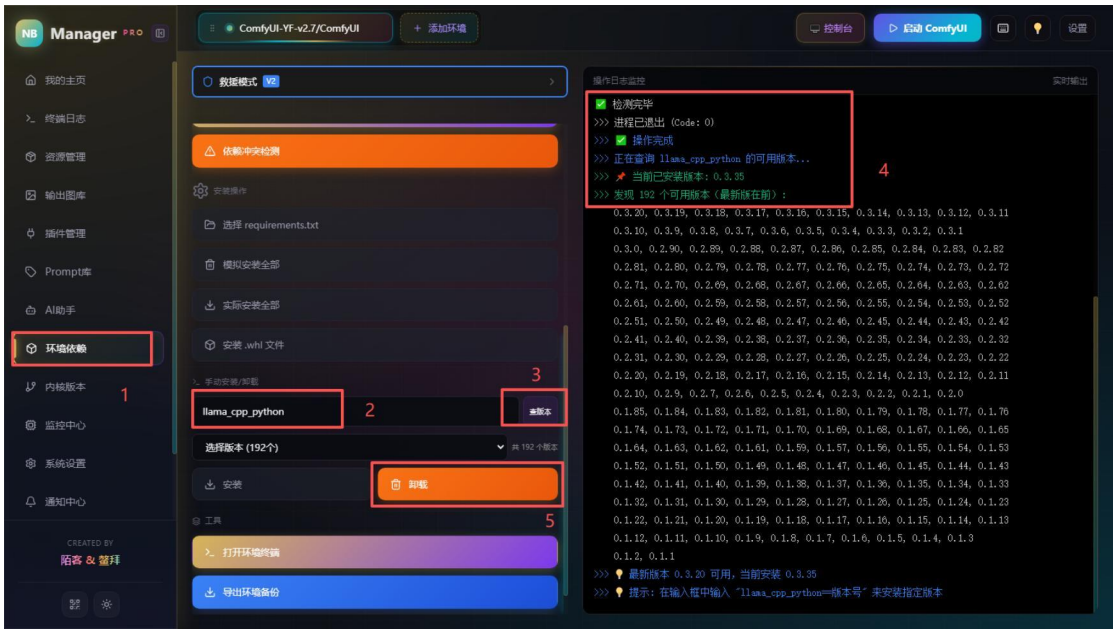
打开网址： https://github.com/JamePeng/llama-cpp-python/releases

3. 找到适配自己 python 版本+CUDA 版本的轮子， 点击蓝字下载

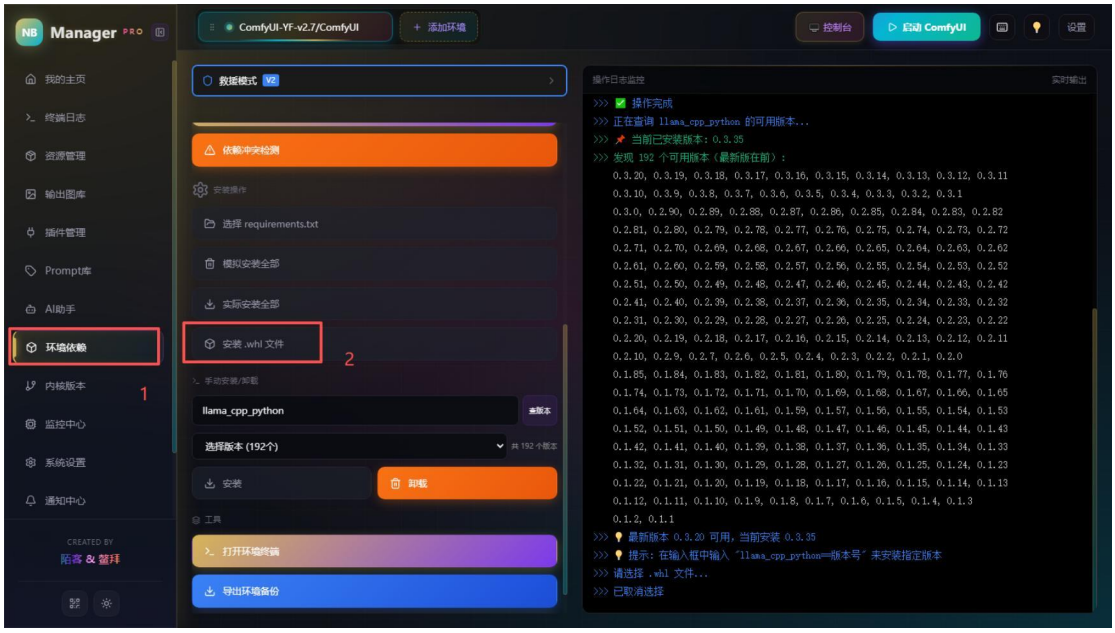


4. 卸载当前安装的轮子（由于不知道你之前是否安装对了轮子，所以先卸载）：

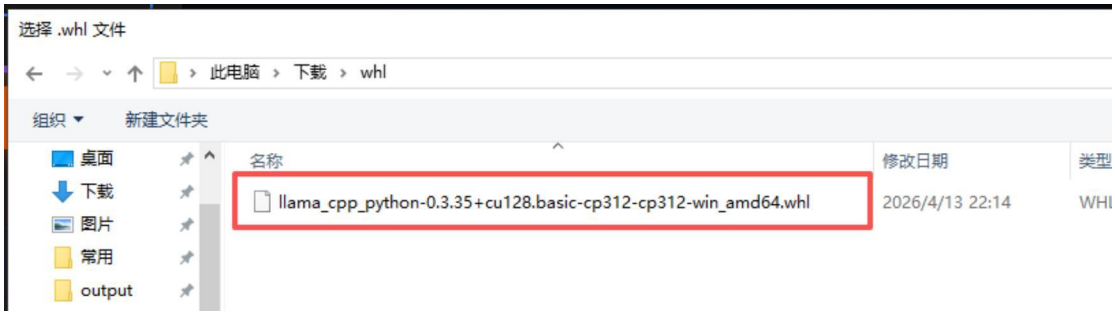
输入查找： llama_cpp_python



5. 安装正确版本的轮子：



6. 选择刚刚下载的轮子文件



等待安装结束即可。

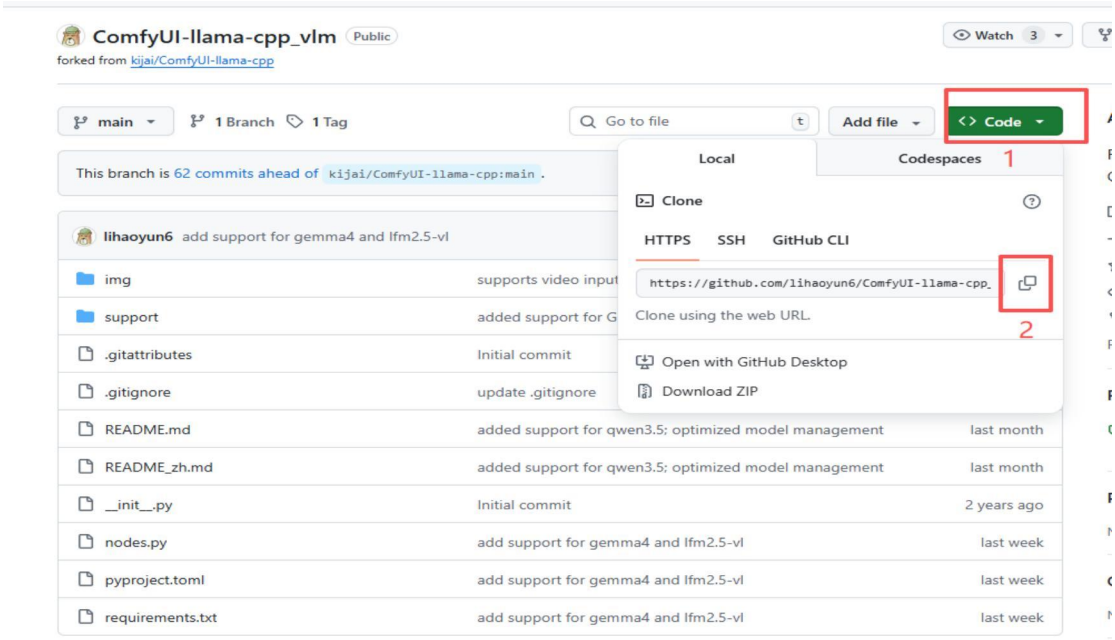
2 TE 启动器安装

安装前打开MF

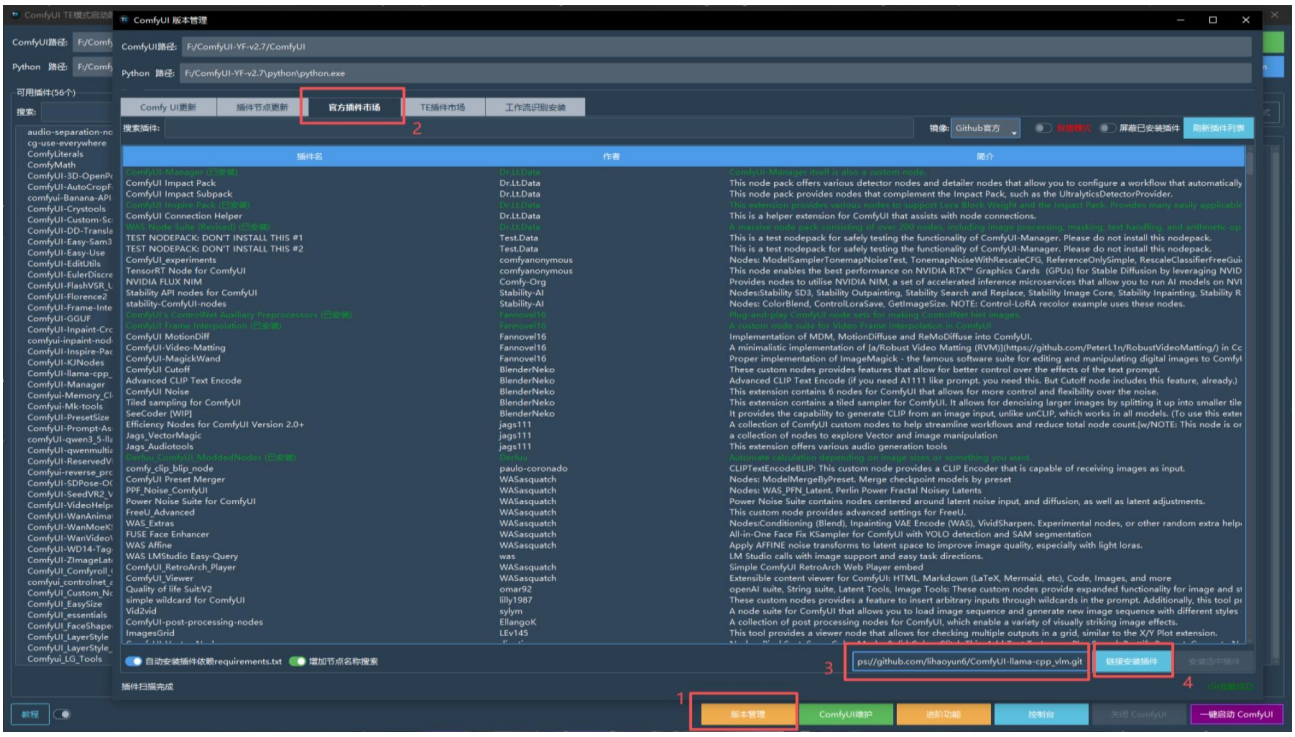
2.1 安装插件

1. 打开网址复制 git 地址：

   Github 网址： https://github.com/lihaoyun6/ComfyUI-llama-cpp_vlm

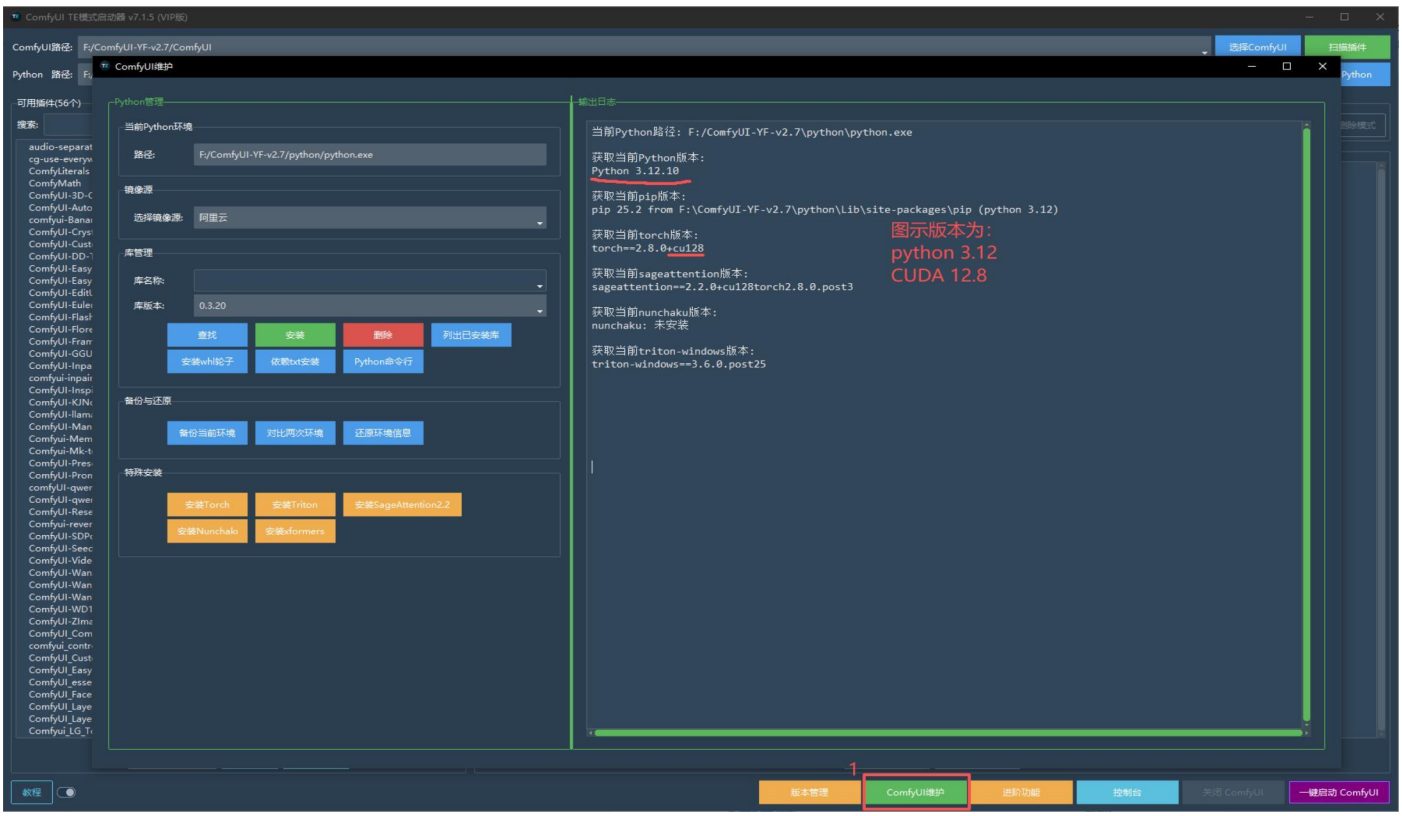


2. 打开 TE 启动器



2.2 安装轮子

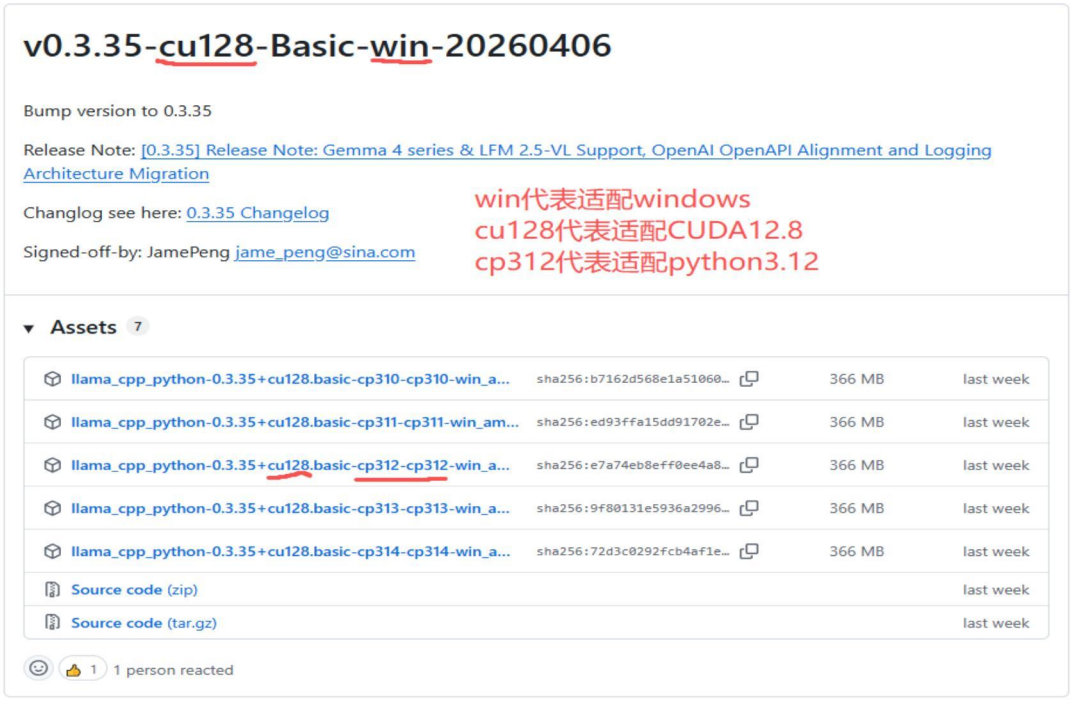
1. 查看 python 版本、 CUDA 版本



2. 下载对应版本的轮子.whl 文件， 此处以上图版本为例

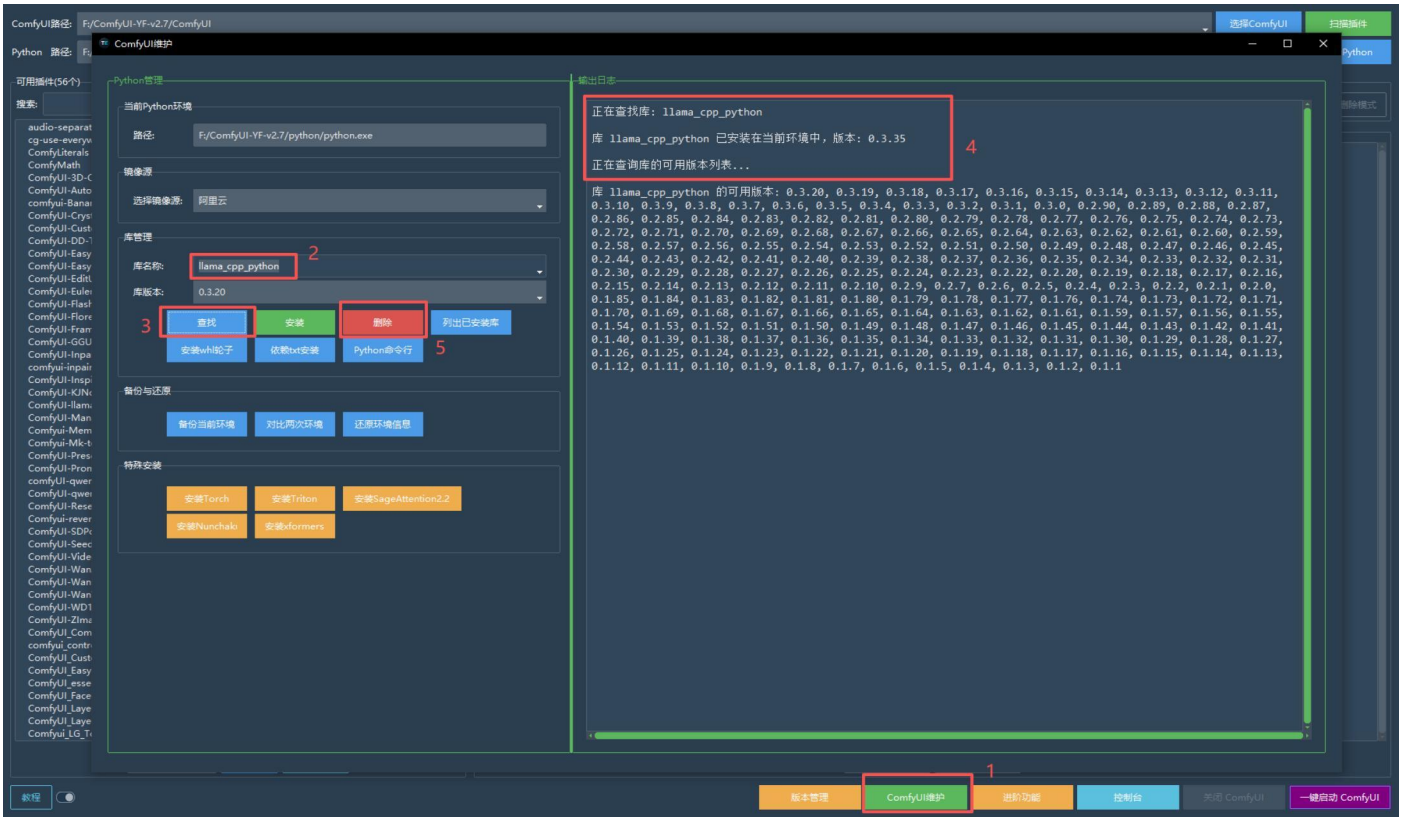
打开网址： https://github.com/JamePeng/llama-cpp-python/releases

3. 找到适配自己 python 版本+CUDA 版本的轮子， 点击蓝字下载

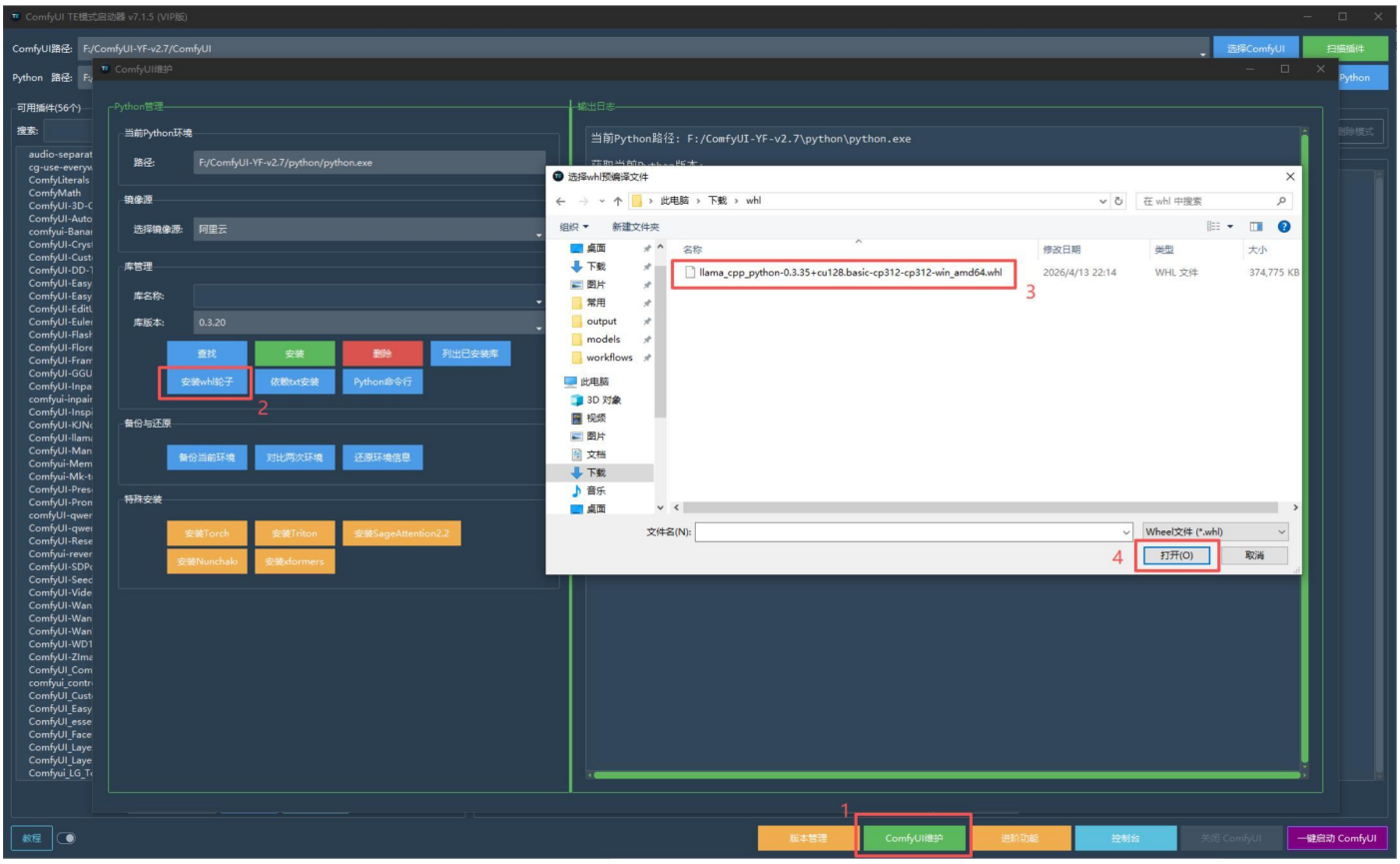


4. 卸载当前安装的轮子（由于不知道你之前是否安装对了轮子，所以先卸载）：

输入查找：llama_cpp_python



5. 安装正确版本的轮子：



等待安装结束即可。

3 CMD 命令安装

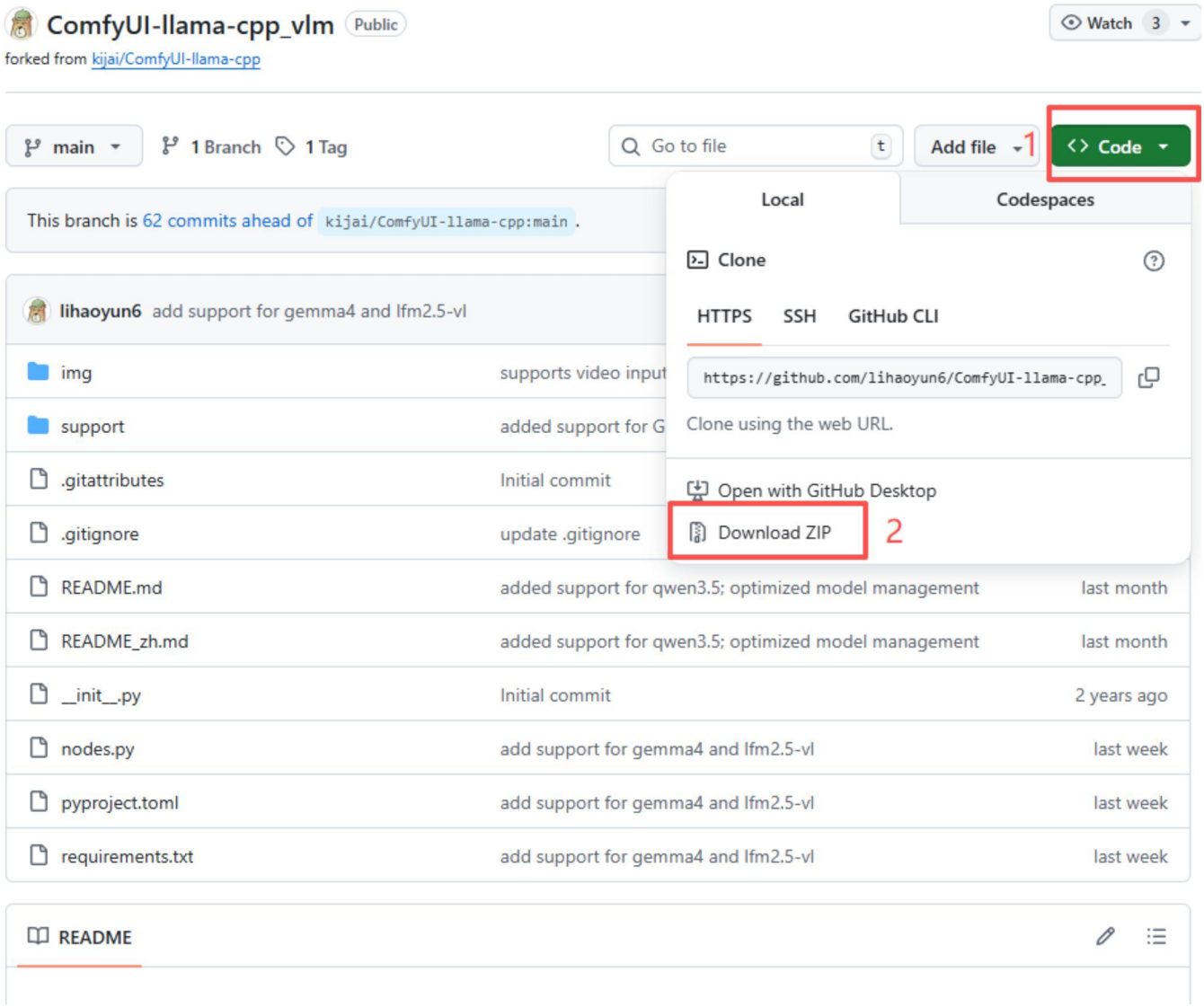
安装前打开魔法

3.1 安装插件

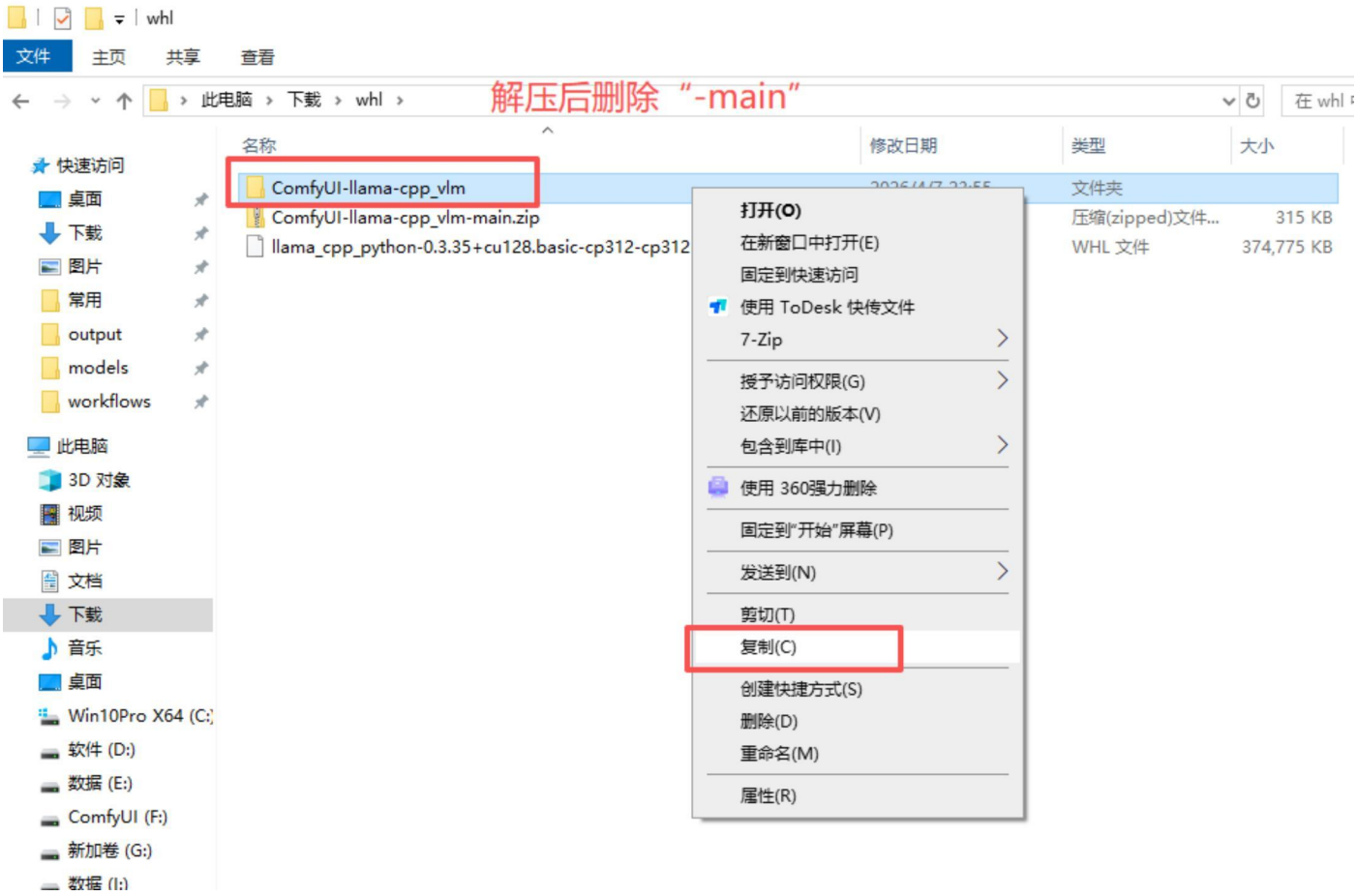
1. 打开网址复制 git 地址：

   Github 网址： https://github.com/lihaoyun6/ComfyUI-llama-cpp_vlm

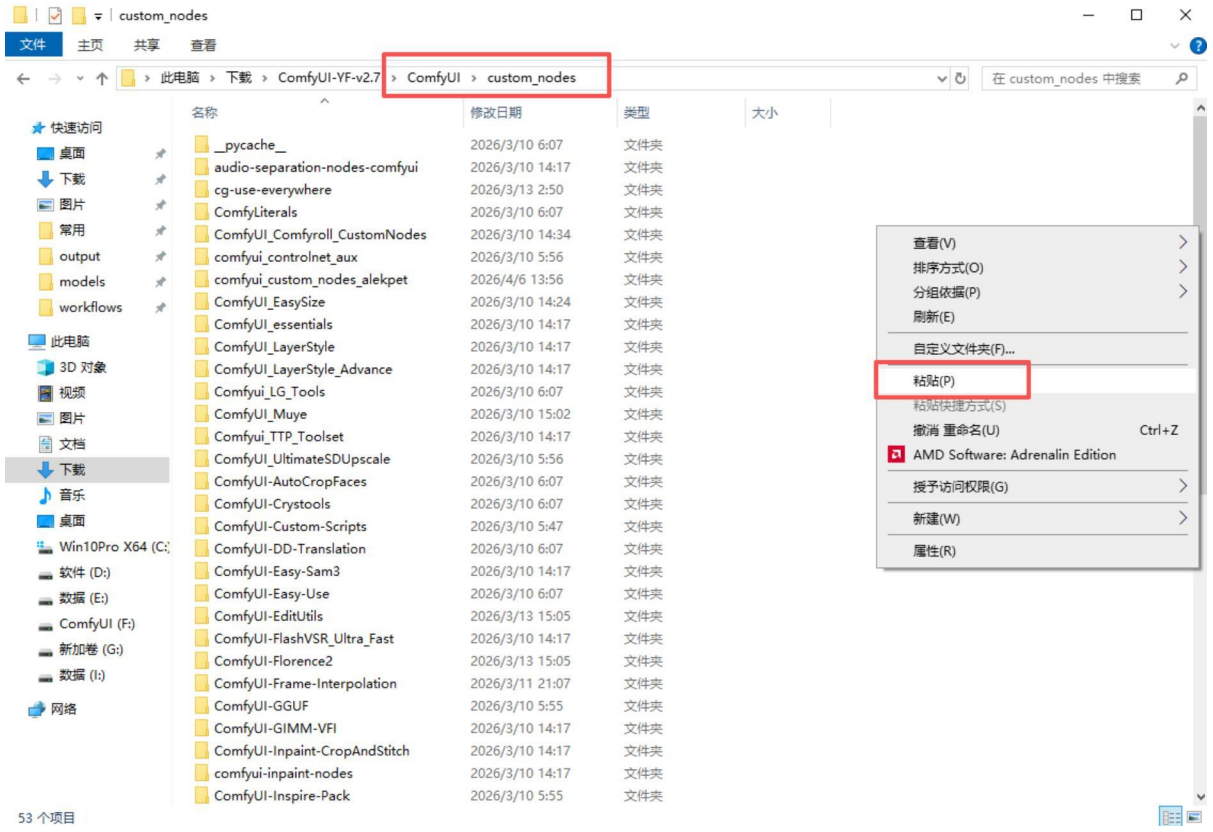
2. 下载 Zip 压缩包



3. 解压刚刚下载的 Zip 压缩包，然后删除后缀“-main”



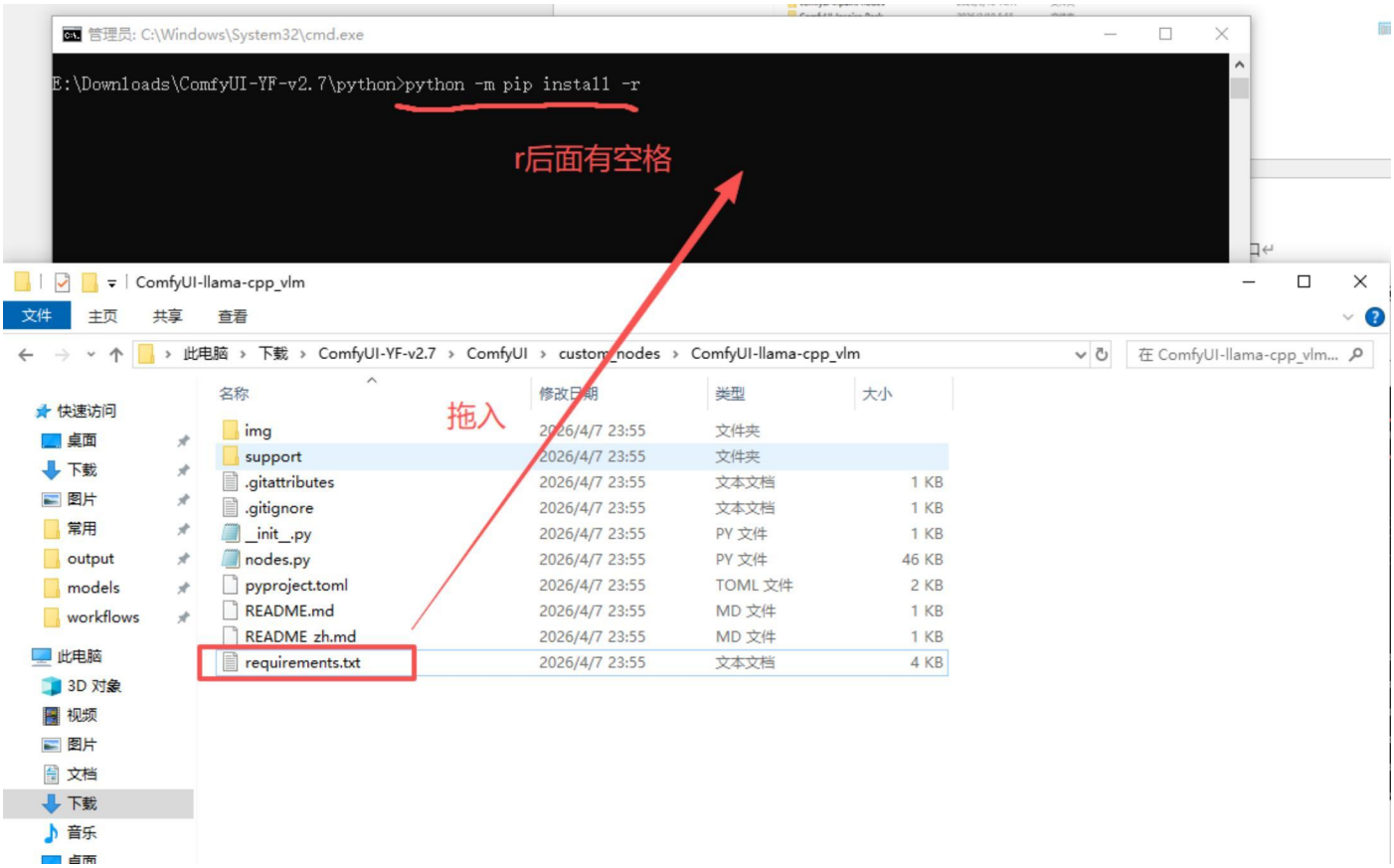
4. 打开 ComfyUI 插件文件夹 custom_nodes 所在目录 X:\ ComfyUI-YF-v2.7\ComfyUI\custom_nodes，将刚刚解压改名的文件夹复制进去

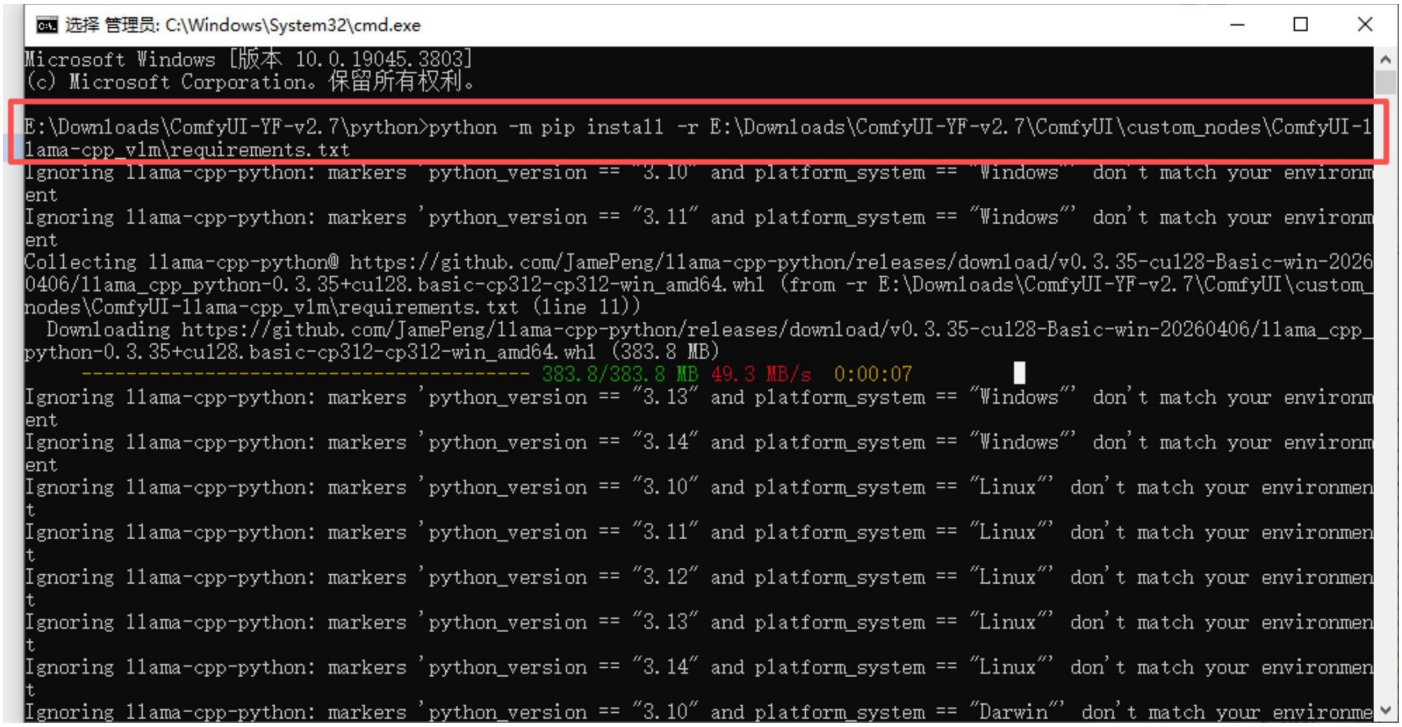


3.2 安装插件依赖：

1. 进入 ComfyUI\python 文件夹，在地址栏输入 CMD 回车，打开命令提示符窗口

2. 输入 python -m pip install -r [然后将 omfyUI-llama-cpp_vlm 文件夹中的 requirements.txt 文件拖动进 CMD 黑框中， 回车]

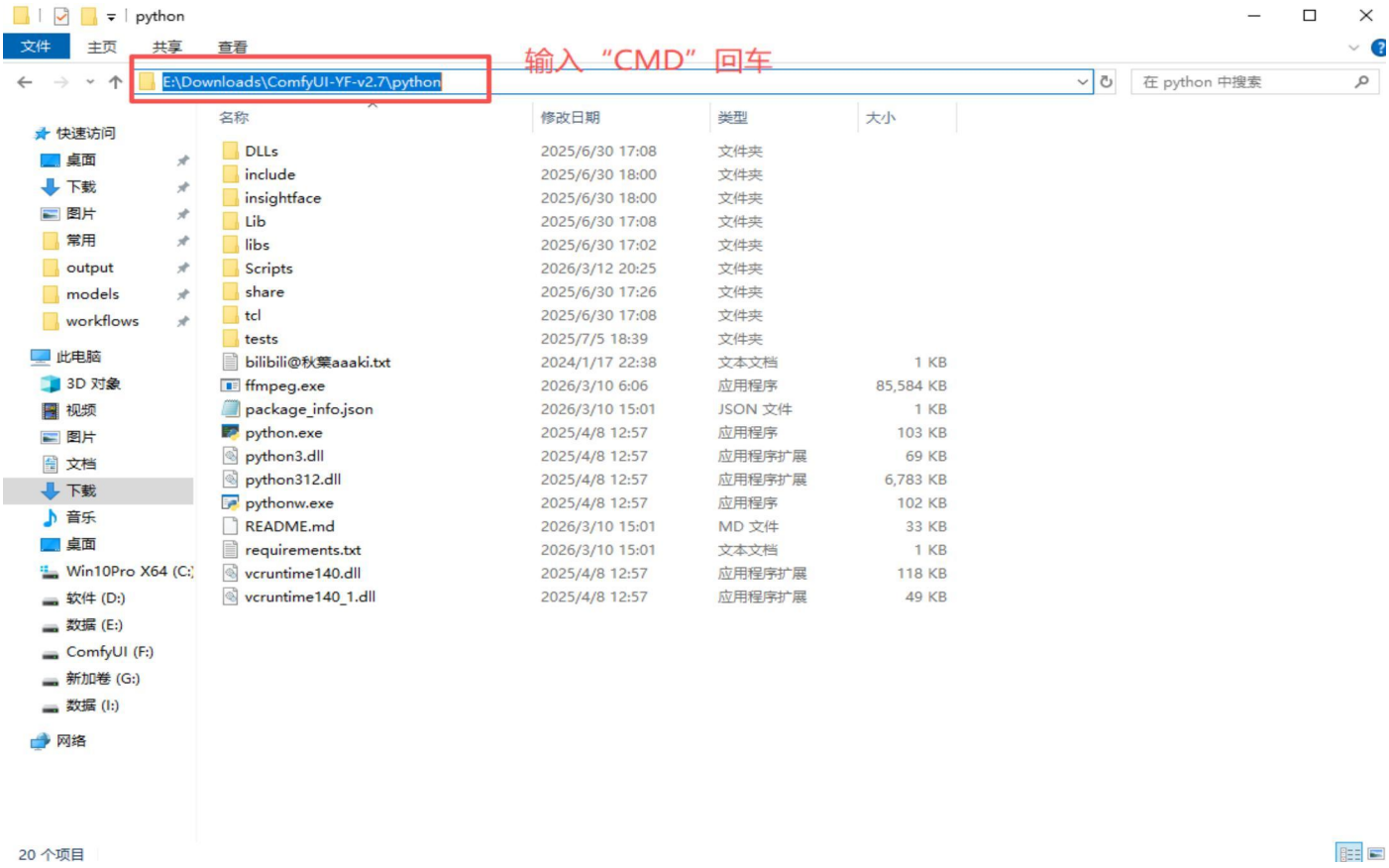




3.3 安装轮子

1. 查看 python 版本、 CUDA 版本

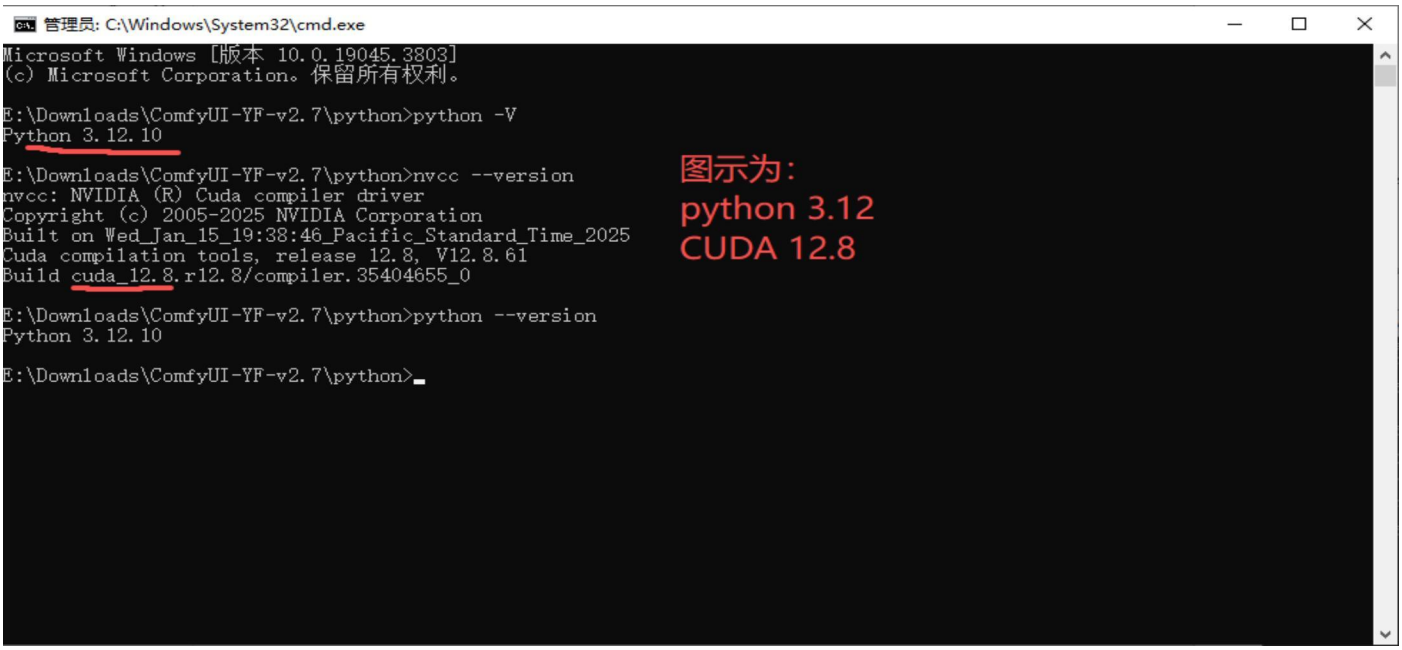
   打开 ComfyUI/python 文件夹所在目录，在地址栏输入“CMD” 回车



2. 分别输入以下命令查看 python 版本、 CUDA 版本

   python -V

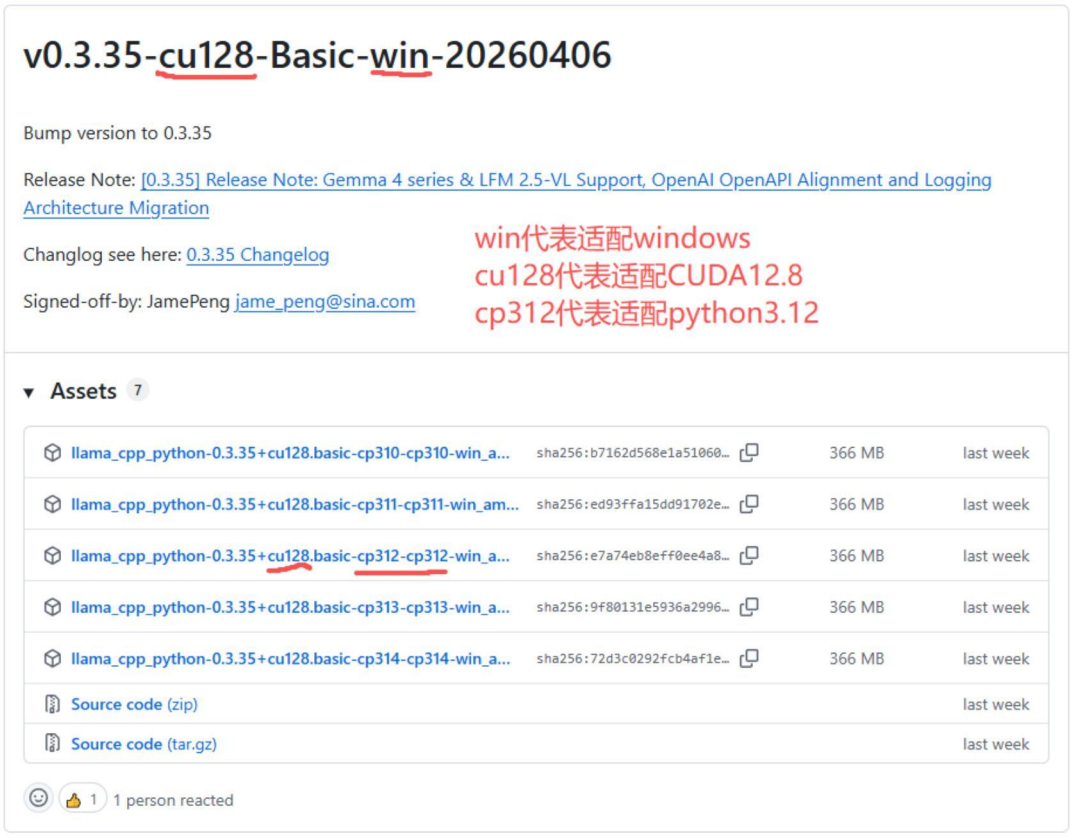
   nvcc –version



3. 下载对应版本的轮子.whl 文件，此处以上图版本为例

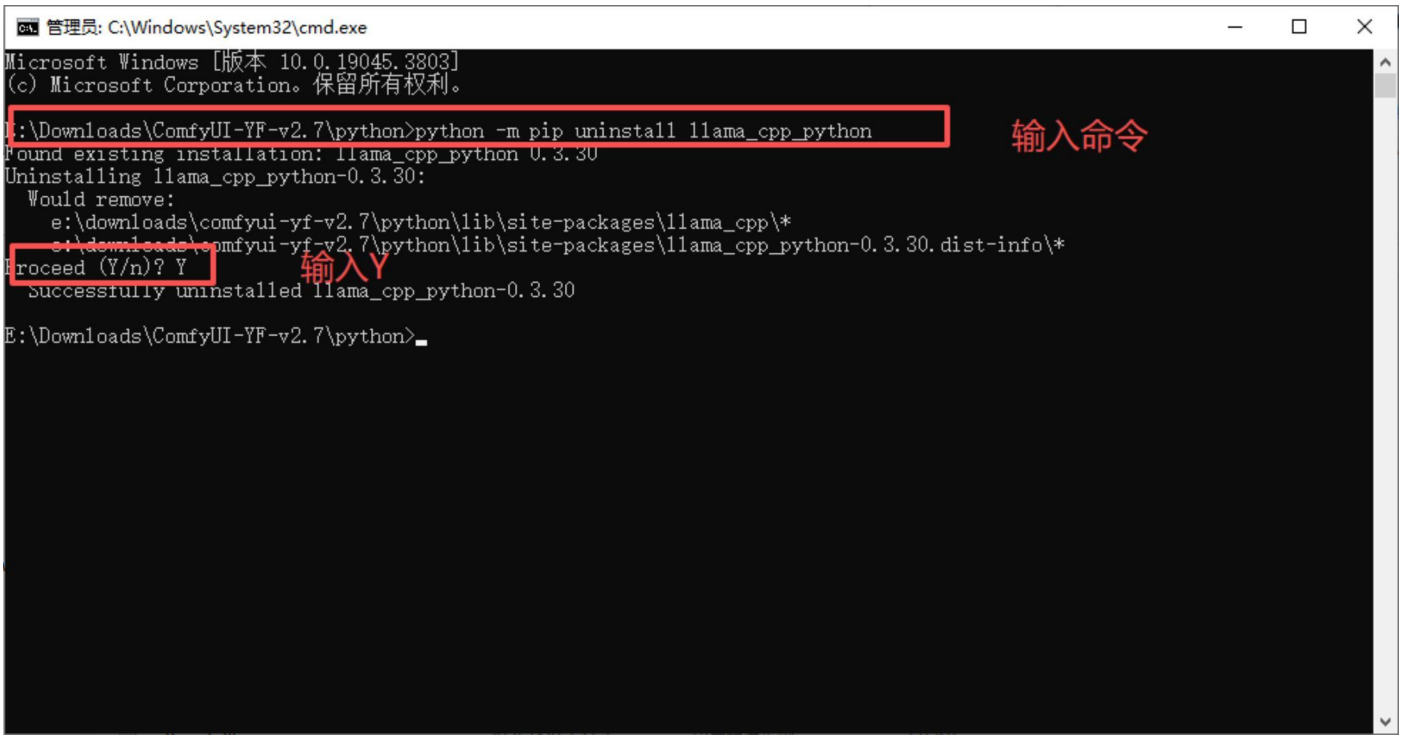
   打开网址：https://github.com/JamePeng/llama-cpp-python/releases

4. 找到适配自己 python 版本+CUDA 版本的轮子，点击蓝字下载



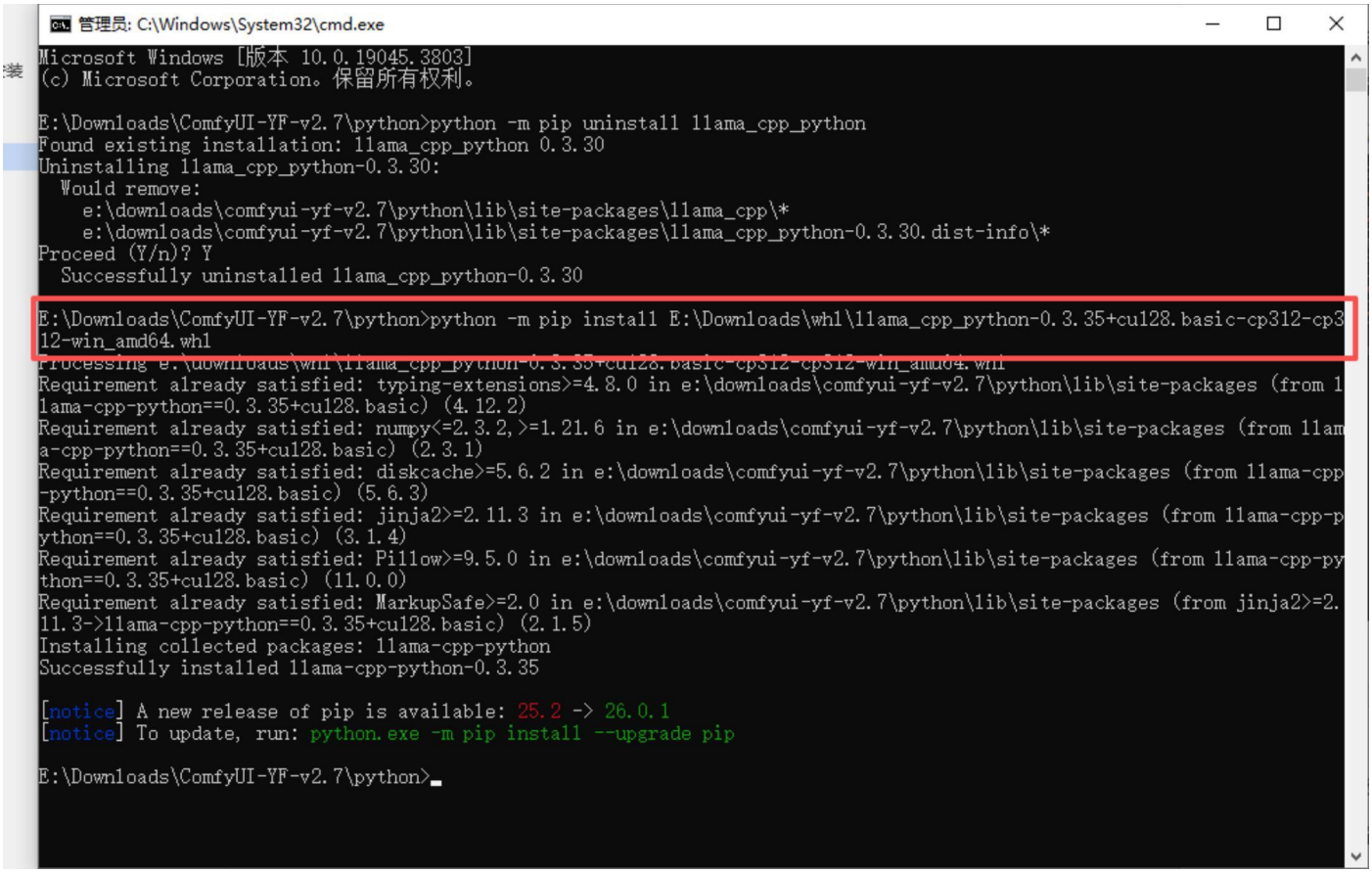
5. 卸载当前安装的轮子（由于不知道你之前是否安装对了轮子，所以先卸载）：

   输入卸载命令： python -m pip uninstall llama_cpp_python



6. 安装正确版本的轮子：

输入安装命令： python -m pip install [然后将下载好的轮子文件.whl 拖动进 CMD 黑框中， 回车]



等待安装结束即可。