One-Shot\_Free-View\_Neural\_Talking\_Head\_Synthesis linux环境下的训练环境搭建

项目介绍

One-Shot\_Free-View\_Neural\_Talking\_Head\_Synthesis是一个用于视频会议中头部视频合成的神经网络模型。该模型能够通过一个包含目标人物外观的源图像和一个指导输出视频中动作的驱动视频来合成头部视频。其创新之处在于运动编码基于一种新颖的关键点表示，能够无监督地分解身份特定和运动相关信息。实验结果表明，该模型在基准数据集上的表现优于竞争方法，并且紧凑的关键点表示使得视频会议系统能够在保持与商业H.264标准相同的视觉质量的同时，仅使用十分之一的带宽。

VoxCeleb2 数据集介绍

VoxCeleb2是一个大规模的开源说话人识别语料库，它在VoxCeleb1的基础上增加了种族多样性，并扩大了数据规模。该数据集包含音频采样率为16kHz，16bit，单声道，PCM-WAV格式的语音数据，这些语音带有一定真实噪声，包括环境突发噪声、背景人声、笑声、语音混叠、回声、室内噪音、录音设备噪音等。视频场景包括明星红地毯、名人讲台演讲、真人节目访谈、大型体育场解说等。VoxCeleb2数据集的构建流程是在VoxCeleb1数据集的基础上经过改进取得的，涵盖了更多种族、口音和地区。

Hopenet\_robust\_alpha1.pkl 数据集介绍

Hopenet是一个深度网络，用于从图像强度中准确且鲁棒地预测头部姿态。Hopenet\_robust\_alpha1.pkl是Hopenet的一个模型文件，用于人脸姿态的欧拉角估计，输出三个角度：偏航（yaw）、俯仰（pitch）和翻滚（roll）。该方法优于使用最先进的路标检测方法的landmark-to-pose方法，并且在数据集上是通用的。

Linux训练环境搭建

为了搭建One-Shot\_Free-View\_Neural\_Talking\_Head\_Synthesis项目的训练环境，需要以下配置：

1. 操作系统：Ubuntu 20.04.4，这是一个流行的Linux发行版，以其稳定性和安全性而闻名。
2. Anaconda：一个流行的Python数据科学和机器学习平台，用于创建独立的环境和包管理。
3. Python 3.7：一个广泛使用的高级编程语言，适用于数据科学和机器学习任务。
4. PyTorch 1.7.1：一个开源的机器学习库，广泛用于计算机视觉和自然语言处理任务。
5. CUDA Toolkit 11.0：一个由NVIDIA提供的并行计算平台和编程模型，用于利用NVIDIA显卡进行高性能计算。

通过这样的环境配置，可以有效地进行One-Shot\_Free-View\_Neural\_Talking\_Head\_Synthesis项目的训练和测试。项目创新点在于采用预训练模型，并使用不同的数据集（如VoxCeleb2和Hopenet\_robust\_alpha1.pkl）进行再次训练，通过对比训练后的模型与预训练模型来判断训练效果。

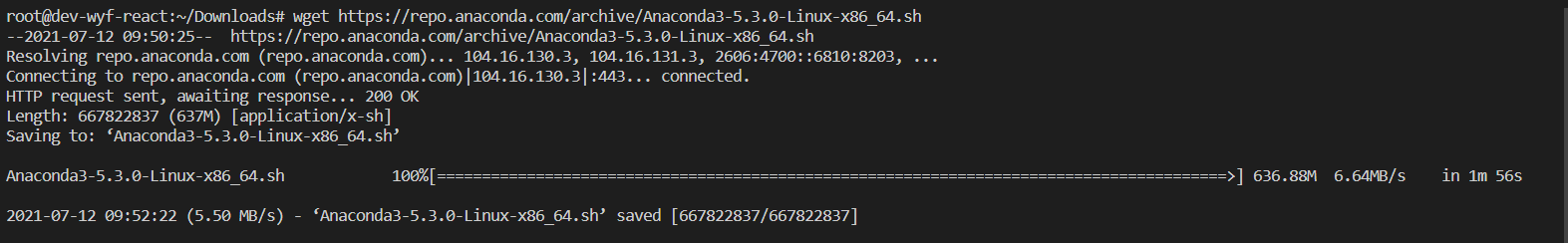
搭建流程  
1.智算云平台租赁linux环境下的云服务器



2.使用mobaxterm进行远程访问

3.在环境中搭建anaconda平台

3.1下载安装包

wget <https://repo.anaconda.com/archive/Anaconda3-5.3.0-Linux-x86_64.sh>  


3.2执行安装程序

./Anaconda3-5.3.0-Linux-x86\_64.sh

然后点击enter以及输入yes等待安装完成

3.3检验是否安装成功

conda -V

4.配置conda环境

conda create -n video python==3.7

#采用python3.7

Conda activate video#激活环境

# 配置pytorch

conda install pytorch==1.7.1 torchvision==0.8.2 torchaudio==0.7.2 cudatoolkit=11.0 -c pytorch

环境就先配到这，去把项目拉进来

5.拉取项目

使用winscp链接云服务器，从本地将压缩包上传到服务器并且解压

Unzip One-Shot\_Free-View\_Neural\_Talking\_Head\_Synthesis-main.zip

6.运行代码并补充依赖s

运行readme中的python run.py --config config/vox-256.yaml --device\_ids 0,1,2,3,4,5,6,7命令，下载依赖（这个--device\_ids是一个设备编号，根据自己设备的数目多少而设置，设成0最保险指就一个）

其实在windows下是有一个已经配好的包的，但是由于linux下有些需要的版本不同，且有些包是windows下的特定版本，所以又重新进行了一遍。（具体有啥我忘记了，吐槽这个项目没有requirements）

7.下载数据集

采用的是voxceleb2和hopenet\_robust\_alpha1.pkl

后者是项目原本采用的数据集，前者为比较有名的视频数据集（根据模型first-order-model-master找到的）

笔者最终也就进行到这步，并未真正的完成训练（voxceleb2太大了下载了两天一夜也没下完压缩包255g还是分卷压缩包，下载部分无法解压），下载完成后运行readme中的python run.py --config config/vox-256.yaml --device\_ids 0,1,2,3,4,5,6,7就等待训练即可。