https://www.bilibili.com/video/BV1mQ4y1E7qD/?spm\_id\_from=333.337.search-card.all.click&vd\_source=aae6703b3293054ccb56e6f2fef39e3b

Fish Speech TTS 技术介绍：

Free speech,总的来说有两个模块组成,一个是VQGAN,也就是将梅尔普压缩为语义token,再重新解码为音频的模块啊,这一块更接近于一个HIFIGAN的结构,另一个模块是基于LLAMA做的自回归模型,自回归语言模型模块,在训练中我们会分开对两个模块进行训练,其实我们可以看到呃,听到我们在例子中是有一定的变音啊,音质不好的情况,这主要是因为,目前的模型更像是一个测试版的模型,我们在VQGAN的训练中,并没有花太多时间在下一个版本中,这些电音的问题将会得到修复,在推理中,我们对我们把文本或者因素推入大圆模型,然后我们让模型来生成一个个量化的语义token,再通过Stage1中训练的decoder把它们进行解码,总的来说结构非常简单,但是也是正是因为它比较简单,我们的优化以及啊性能提升都更加的好做啊,我们这个模型对系统的资源,并没有特别多的要求,对于推理来说,我们只需要2GB的显存,对于微调只需要16G或者24G,所以不需要因为是它是用语言模型和自回归,就担心它,然后我们实际上测试也可以达到300~500,token每秒的生成速度,也就是说,单卡一秒钟可以生成10~20秒的音频,这个基本上是可以满足大部分的需求了,我们嗯在最新版本中加入了梯度检查,Gradient pointing,还有flusher tension的支持,这样训练速度变得更快,以及节约了一定的显存,在之前我们也完成了web ui和HDPAPI的开发.

Github 仓库: fishaudio/fish-speech

Demo: <https://speech.fish.audio/samples/>

大象给的：

模型测试

https://www.bilibili.com/video/BV1tN4y1i7xU/?spm\_id\_from=333.337.search-card.all.click&vd\_source=3372fff6e7d234146cb42ebb89dd8445

介绍页，按照这个部署到我们的4090

<https://speech.fish.audio/>

[fish-speech]爆杀爆杀？基于fish speech的孙笑川模型测试，比更好还更好的tts

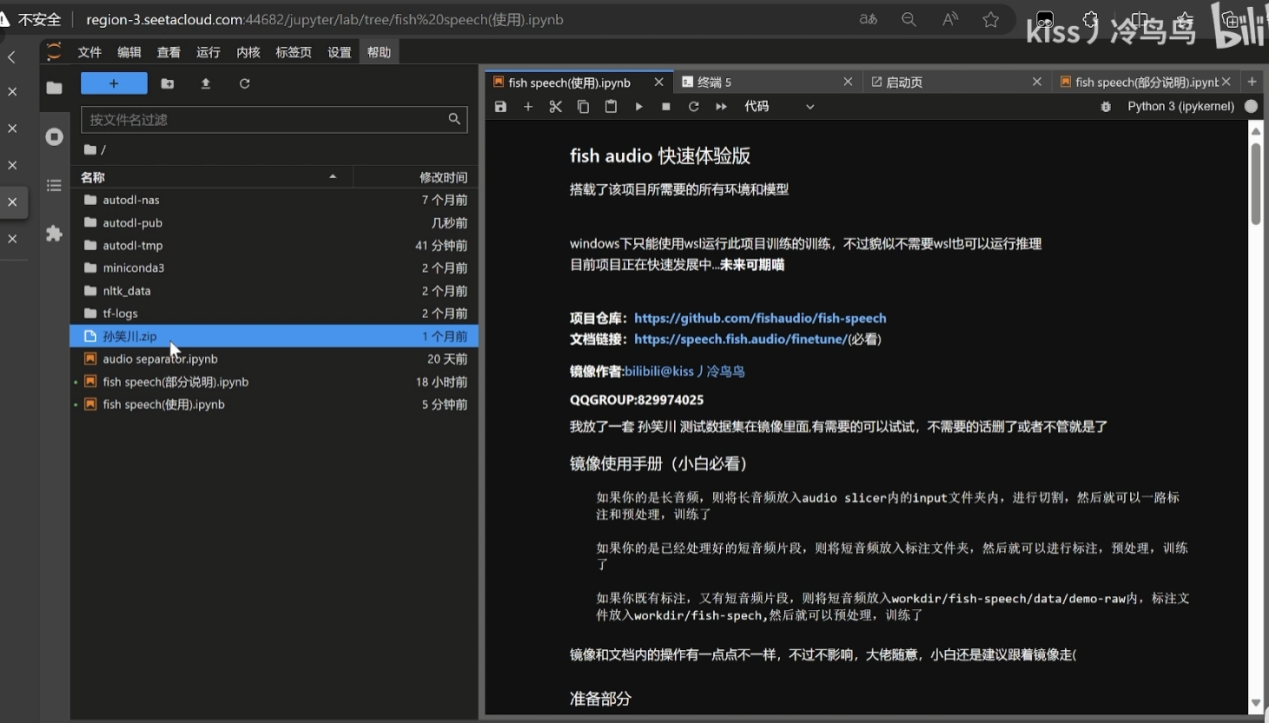
https://www.bilibili.com/video/BV1tN4y1i7xU/?spm\_id\_from=333.337.search-card.all.click&vd\_source=3372fff6e7d234146cb42ebb89dd8445

仓库地址:https://github.com/fishaudio/fish-speech

autodl官网:https://www.autodl.com/

region-3.seetacloud.com:31129/? theme=light（打不开）

这期视频是对fish speech,也就是雨雨的效果做一个简单的展示,这是fish speech的文档,然后呢,这是fish speech的项目仓库,我在摇头,迪奥上发布了一个fish speech快速体验版的镜像,直接搜就能搜到,记得仔细阅读官方给出的文档,还有就是window下不支持训练,只能推理,不过可以使用WSIO进行训练



fish-speech&bert-vits2

<https://www.bilibili.com/video/BV1Jp4y1Z7mS/?spm_id_from=333.337.search-card.all.click&vd_source=aae6703b3293054ccb56e6f2fef39e3b>

fish-speech项目地址:

https://github.com/fishaudio/fish-speech/

<https://speech.fish.audio/zh/>

今天给大家介绍一个新项目,然后是这个项目也是原项目组,还是那个原项目团队,还是这个B2T这个项目团队是另外一个项目,然后他这个是最新版本,B2T是2.3了,然后你对B2T的话,,然后就趁着这个机会,然后就把这个群六层全部写一遍吧,嗯到时候我会在B站做一个相关的合集,预测是这个星期弄完嗯,这个是B2对,我对B2T的,然后我们继续回到主题上来啊,因为2.3的话是嗯近期嗯近期最终版吧,应该就是近期已经不会做较大的更新了,所以说刚好那个教程可以完美的契合到,目前的2.0以上的版本,然后我的所有教程的话,嗯云上云上的话,所有的那个嗯云上的所有那个教程全部是无,主要是2.0以上的嗯,主要是200以上的版本的话,我这里所有的都是支持的,就是所有流程全都是一样的,没有任何区别,主要是2.0以上的,就是全部通用,基本上,OK就是回到我们的主题上来嗯,这个新项目还是那个原著的原创团队,就可能重新用哪个项目,然后我觉得这个项目的前景比这个,其实你这个B2T的前景更大,因为他这个的话是更符合起就是商用类型的,不就是你只要上传一段比较少,因为BC的话你需要上传一个嗯,就是你要训练好,效果好的话,最起码要一个小时左右的数据集吧,就半个小时到一个小时或者朝上的数据集,也是对数据集的要求可能更大,但是这个这个项目的话,他对那个数据集的要求,就如果你不微调的话,就是训练集要求不是很大,你上传需要生成一个十秒到30秒左右的,那个音频文件,就可以复刻那个声音的,那个就可以完美的复刻那个声音,当时当然这个项目还是在这个初期阶段嘛,就是我个人觉得嗯不微调的话也还行啊,当然这个还是初期,所以后面还会进行更新,就是然后我再来看一下相关项目的相关介绍,然后的话这个项目在利率就是推理的话,至少两届减肥了,然后微调的话16G显存朝上,LINUXCC现在可以完美的嗯推理还有微调什么的,都是支持的,但是问的话请支持推理,但是推理的话他也不支持加速嗯,到时候到时候会解释相关原因啊,就是大概的原因就是因为windows不支持这两个东西,然后你的话这个主要是不持这个,它就不支持那个全部的加速,然后的话windows建议是用WSL或者是GD,也可以用虚拟机里面安装一个其他系统,也可以是进行完成的啊,但是我个人建议是使用DOTA2的,因为它是非常方便,就是你如果使用这个WCL2的话,它会跟某些环境进行相冲突,但是这个其实也很好,就是看你自己,比如你这个什么安卓的模拟器,雷电模拟器,什么各种实现了模拟器之类的吧,你如果使用了这个的话,它就就是相互冲突,他就无法相互兼容的运行,包括你这个虚拟机里面的那些某些虚拟内存的,一些应用的话,它都产生冲突,就说你使用WSL的话,你的信息里面的某些环境他就用不了了,这个的话你自己嗯自己看着办吧,到时候嗯有时间的话会出这个项目叫什么,其实大部分情况下windows进行推理就够了,然后LINUX的话你可以在这个嗯平台上,云平台上进行微调,而微调模型下载下来,使用windows进行推理,其实也是差不多够了,然后我们这些环境的话,其实也没什么可以说的,就是这一部分的话,这个安装的话在windows上是无法安装的,就是说这一步你如果是windows进行仅进行推理的话,你这一步可以去掉,这一步可以去掉,不用管它,然后直接进行下一步,然后我们来看一下这个,例子吧,首先来看一下例子,然后这里的话是有一个五秒钟的,那个纳西达的这个文件,然后他这里就是生成这个文本,你可以看一下效果,这是原音频,然后因为是项目初期的话,这个暂时还是能接受的,因为后期作者可能会对这些进行优化,嗯就你只要上传一个六秒的音频,能合成这个效果,因为后面因为毕竟项目只是初期阶段嘛,啊还有很大的进步空间,然后我们对比一下这个嗯商用的这个吧,就是这个火山的这个火山引擎（https://www.volcengine.com/product/voicecloning）,它这个是你上传一个十秒到30秒的音频,然后可以生成就复刻这个声音,然后他的官方API文档的话,它是支持上传20分钟的嗯音音频文件,然后进行训练成模型,然后价格是150块钱,一年使用那个模型啊,不包括他那个语语句生成,只是150块钱,包括那个模型的保管以及合成,但是你每每使用一次音频,还是每合成一次音频还是要收费的,还是我们来试一下这个效果,我上传的是神里绫华的声音,听一听,听一下,像这样悠闲安稳的时光,乌鸦的叫声一样的难听,我们身上有着的黑色羽,差不多就是这个效果,然后他的话是你可以你可以去搜火山以前嗯,就是搜火山那个声音复刻,然后可以进入这个体验页面,使用次数是免费使用十次嗯,就是大概这样你可以休息一下,OK这就回到我们主题上来,可以自己试一下这个合成,我直接进入这个推理的界面,这里的话就是嗯昨天我看的话,还是不支持这个外部UI来处理的,但是今天已经支持了呃,就是作者跟我的很就是项目更新很频繁啊,这是一个很好的趋势,OK这里来稍微看一下,他这个九分给那个5~10秒的音频,然后的话我们来直接试一下效果吧,然后我们需要把这个下载下来,就是直接点这个吧,然后下载JP,然后你下载下来之后停在这里,就是来解压,然后到这个目录来,然后我们直接打开这个嗯,先D后直接进去,我们直接回到这个主持来,这里的话,直接复制这个,然后的话收音频吊牌,我们这里的话就是神秘礼包,然后后面就是那个模型的位置,然后直接回车,回车之后,它会在目录下生成一个NPNPY的文件,稍等一下,这里的话嗯原项目也是说了,就是这个在合成的速度上,要是其实是比那个B2T微慢一点的当然这里的话其实是作者是支持外部UI,但是我这里的话还是用命令行来测试一下,然后的话他是音色,就是说画饼的音色是否不同或者质量不高,如果你的音色有问题的话,你可以微调这个模型,我们来听一下这个啊,是否音色跟那个原石一样,然后的话我们直接来进行这一步,然后这里的话把这条命令复制过来,说明这个GTT我们就不用管了,然后这个我们改一下,要稍微改一下,然后改成跟这个刚刚的那个一样的吧,嗯就改成这个跟那个火山形成的那个音频里面,然后的话这里是默认的音频文件输入,就是你合成的音频特色TT,然后这里的话我就免得在命令行输入了,然后直接在这里改一下,所以说我们这一步就可以直接去嗯,在这里这一步我们可以直接去掉了,我们直接使用默认的,就是我们刚刚说的这一部分的话,特TT的话,就是我们原始原始的这个输入的那个,神里绫华WV的,它的转成中文的那个格式,就是这个它转转写成中文的那个音频嘛,就是这个,然后直接把它复制到这里来嗯,就是这一部分,然后后面的话我们按部就是按照他的命令,他的后面那个格式来进行直接复制粘贴,这个也是一样的,还有就是为什么是生成这个这个FAKE啊,因为这里的话他这里就开始我们生成的,就是嗯生成这个就是如果你需要改的话,你可以自己转来这里指定,因为它默认的话是到时候再说吧,后面会出现个教材啊,这些都是默认的,就是这个参数吧,默认的,然后我们两个发文件的是一个后后,这样的话可以继续复制,这里就是到这个的话,我们在windows上是无法使用这个命令的,因为我们可以看到下面说这个,如果我们使用这个的话,我们来试一下啊,其实我们在这里的话,他是已经说了,在介绍里面,就是你在嗯windows windows上是不支持这里的东西,然后你其实也是无法使用的,嗯我们首先加一个试一下哈,就是如果如果我们加了这个,我们来看一下它会是什么结果,嗯稍微等待一会儿,这里思路其实也不算慢,稍微等一会儿就好了,你看这里他就是windows啊,就是什么不支持这个东西,所以说这个命令我们是无法添加的啊,当然你不加这个面的话,直接把这个面去掉啊,就是嗯这样也是可以的,你这样回车它也是可以生成的,当然它下面也是提供另外一个参数,对于不知手这个可以使用这个参数嗯,先把这个复制过来,然后粘贴这里也可以进行完成的,然后这里稍微等待一会儿,哦这里是不需要全部跑完的,如果你是没有指定这个,如果你没有指定这个参数,就是这个嗯嗯TXT嗯这个参数的话,它是会本时代所有的跑完,而且会出现各种各样的问题,先把这一条就是最后一步这个命令复制过来啊,我们我就是分流程分步完成的,嗯其实如果你要使用外部UI的话,它会一次性生成,就是你不需要这么麻烦的操作,我这里是稍微整理一下各个参数的使用什么,这东西,就是等你以后进行,去外部UI方面的那些操作的话,可以更清楚的知道报错的原因,以及各项问题是什么的,然后这里的话他是最后一步的话,就是指定这个就是我们刚刚生成的,这里执行这个命令的话,它会生成这个嗯营业PY文件指定之后,我们这里直接最后使用这个模型来进行,生成这个文件,生成这个音频文件,这一步还是很快的,稍微等待,然后这里的话会形成一个F开头的,那个是需要哎FKFAKE那个WA文件嗯,看这里已经这里生存的这个FK就是这里,我们来稍微听一下,我们几乎长得没有分别,可是人类却把你们分为吉祥的象征,把祸端全归就在我们身上,你们叫声就是报喜,而我们的叫声就是不响,嗯怎么说呢,嗯嗯呃这不好说,但是我前面生成的那些音频还行,可能是也有可能是这里太长了,原因吧,啊当然我们还是这个是没有使用微调的模型哈,嗯就是没有使用微调啊,所以暂时的这个效果还是还是这样,嗯也可能是那个嗯这里嗯删除文件太太长了,因为我前面一次的话它是很少,就是这个就是来试下这个吧,我们默认不改的话,我们来试一下效果,就是前面两个,这是第一步吗,嘴巴还是把这个生成,OK第一步就完成了,那这里的话就是专业音色的,然后我们执行第二步,第二步就是这里,然后的话就是那个嗯,嗯这个文件的话对应的是这个转写的文字,然后生成这个前面一个OK直接回车,它就会稍微等待一会就好了,就说你这个嗯合成的,这个就是你说的要需要合成这个文字,数量越少的话,这里是合成的越快,这里就很快生成了,OK我们进行最后一步了,是嗯就是这一步,最后一步我直接进行生成那个最终的W软件,可以看一下效果,嗯就是生成的文件在这,你,你说的对,但是原神是一款由米哈这个效果我觉得嗯,因为毕竟你输入的音频只有呃六七秒吧,这个效果其实还是可以接受的,因为而且项目也是在初期嘛,刚开始初期能达到这个效果,其实已经嗯很好了,嗯包括我们这里还可以,如果对效果不满意的话,还是可以进行微调,什么方面的还没有尝试,所以说当时我觉得还行啊,当然如果等项目成熟的时候,会再做更详细的教程了,看一下,包括他这里是支持英语和日语,也都是支持。

天空配音软件

window版下载：

https://wwp.lanzouq.com/b02upfdxe

密码8888

备用 链接：https://pan.baidu.com/s/1O0Sk-AA1fmtlvBNUz9gk8A

提取码：tk99

关注UP主更新