

有了这个开源项目，说你是老板都有人信？

人工智能与大数据技术 昨天

来自：开源最前线 (ID: OpenSourceTop)

连接: <https://towardsdatascience.com/you-can-now-speak-using-someone-elses-voice-with-deep-learning-8be24368fa2b>

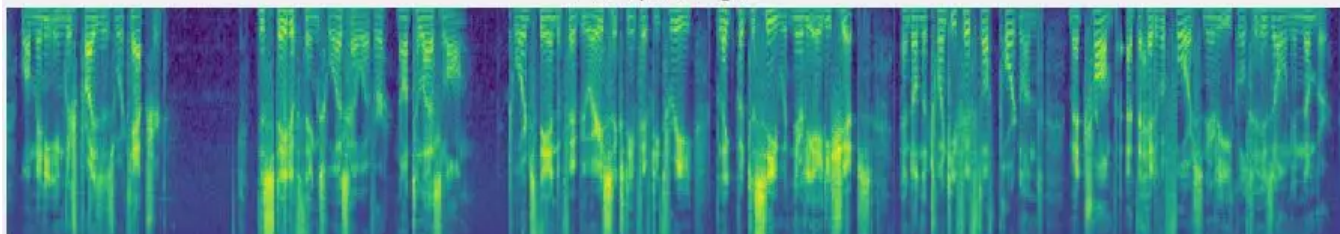
只要你采集了你Boss的语音，这个项目就可在短短几秒之内，克隆生成与之类似的语言，你想让它说什么，它就跟你说什么。

要实现这样的功能，首先不得不提文本转语音（TTS），它是指将文本转换为音频，人类只需阅读即可完成此任务，良好的TTS系统的目标是让计算机自动完成。

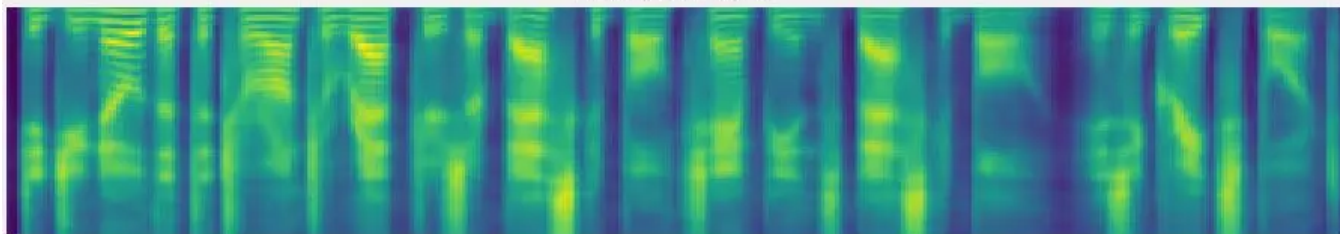
有趣的是有人构建了一个系统，可以选择你想要的声音生成音频，它可以是男声也可以是女声，你还可以选择他的声音是洪亮的还是低沉的。

在进行深度学习的TTS时，你必须收集文本-语音的数据集，并且录制这个音频的演讲者必须是固定的。如果你想要创建一个能够模仿自己/他人声音的音频，要做的第一件事就是收集音频

LibriSpeech/train-clean-100/1081/125237/1081-125237-0053.flac
mel spectrogram



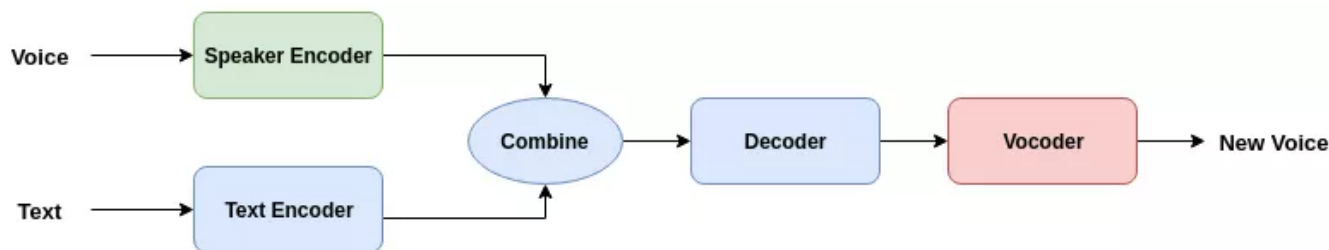
LibriSpeech/train-clean-100_1081_gen_05097
mel spectrogram



Real-Time-Voice-Cloning就可以帮你实现，这是一个基于深度学习的语音合成项目，它通过采集分析一段具体的声音样本，可在 5 秒内生成与之类似的克隆语音。

那么它的工作原理是怎样的呢？

比如，假如我们想让蝙蝠侠读“我爱披萨”这句话，那么我们需要提供两样东西给系统：一个“我爱披萨”的文本和一段蝙蝠侠的语言样本，这样它才知道蝙蝠侠的声音是怎样的，以及他要读的文本是怎样的，具体流程如下：



在过去几年中，文本转语音系统在深度学习社区中得到了很多研究关注，如果你也想尝试，首先你需要克隆存储库：

```
git clone https://github.com/CorentinJ/Real-Time-Voice-Cloning.git
```

安装所需的库。一定要使用Python 3：

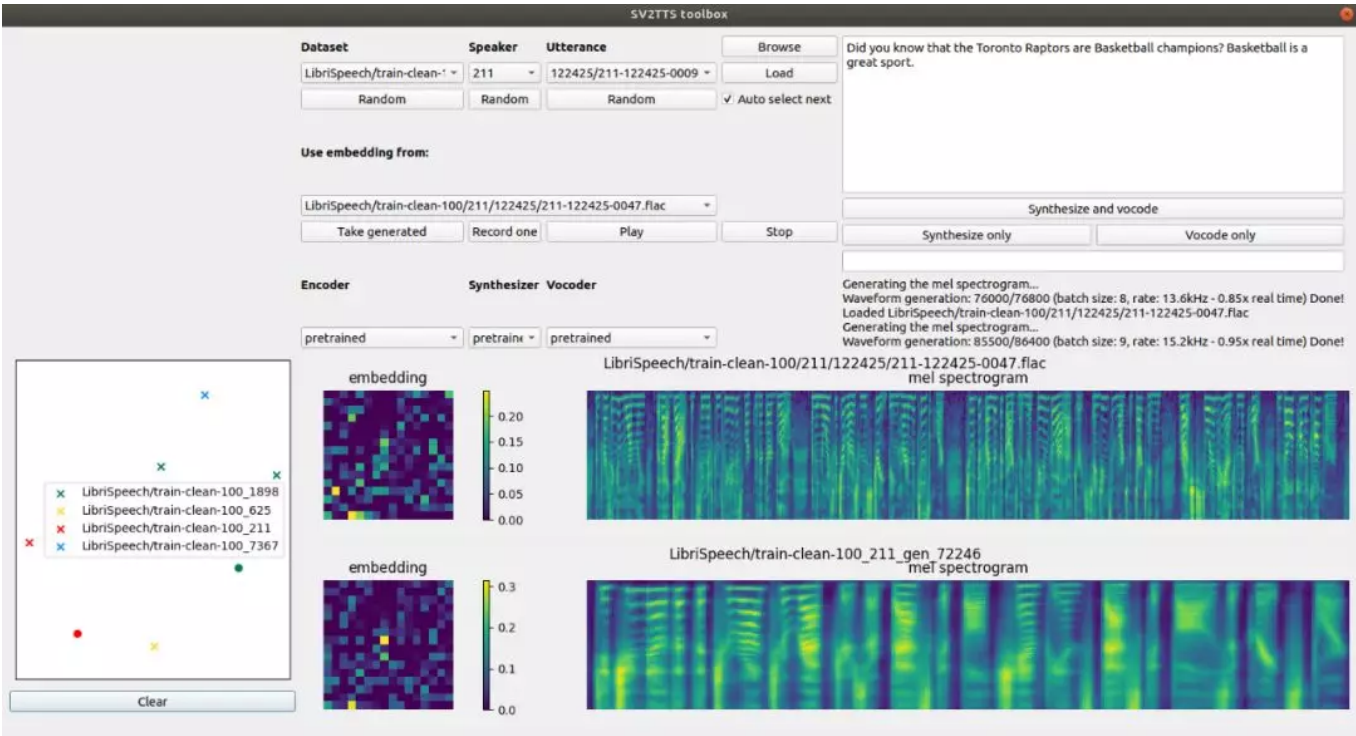
```
pip3 install -r requirements.txt
```

在README文件中，你还可以下载预先训练的模型和数据集，尝试一些样本。

最后，你可以通过运行以下命令来打开GUI：

```
python demo_toolbox.py -d <datasets_root>
```

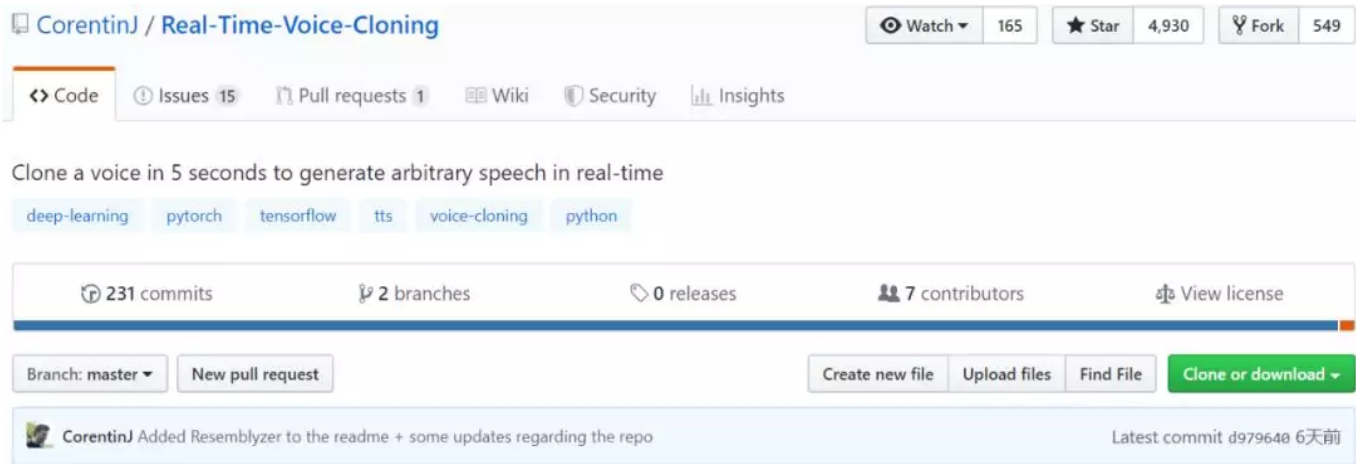
如下图所示，系统将输出的语音文本为：“你知道多伦多猛龙队是篮球冠军吗？篮球是一项伟大的运动。”



你可以单击“Random”按钮随机化语音输入，然后单击“加载”将语音输入加载到系统中。

你还可以通过以下视频查看视频演示：https://www.youtube.com/watch?v=-O_hYhToKoA

如果你感兴趣也可以用自己的声音试一下，或者把Boss的声音录下来，然后你想让他说什么就说什么，比如：这个月的Bug都不用找了。所以，躲在音频背后，说你Boss，都有人信。



目前，这个项目已经在GitHub上获得 **30.5K** 个Star， **549** 个Fork（GitHub地址：<https://github.com/CorentinJ/Real-Time-Voice-Cloning>）

不过，对于这项技术，有网友评论道：虽然很酷，但也是不道德的，容易造成谣言传播，使用需谨慎，对此，你怎么看呢？

--- EOF ---

推荐↓↓↓



程序员头条

10:24 一起看程序员热点新闻。把握行业热点，紧跟行业脚步



公众号

喜欢此内容的人还喜欢

AI大厂迎最冷冬天？网传阿里拟裁员2万人，净利下滑高管离职
人工智能与大数据技术



12星座婚前婚后差异，变化巨大！
唐绮阳看星星



生了小孩，我很抱歉
和毛利午餐

