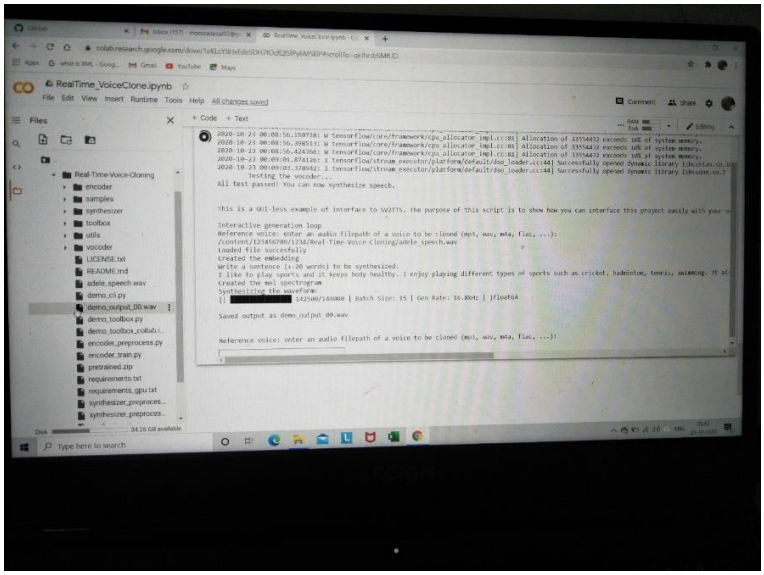
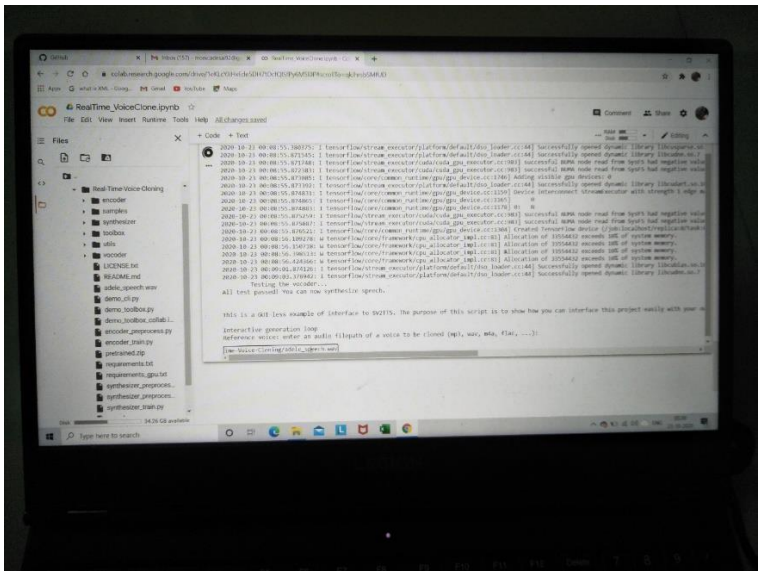
Project Name: Real Time Voice-Cloning

Table of Contents

Demo





Overview

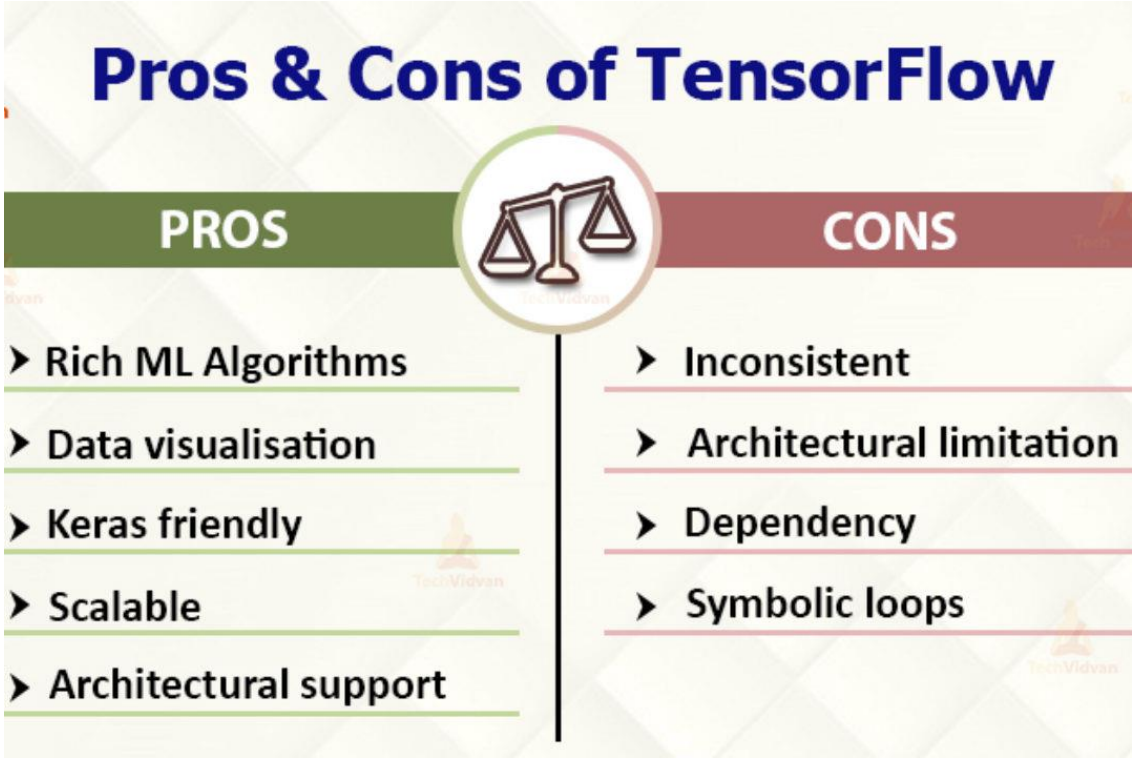
This is diving into Real Time Voice Cloning Concept.
SV2TTS is a three-stage deep learning framework that allows to create a numerical representation of a voice from a few seconds of audio, and to use it to condition a text-to-speech model trained to generalize to new voices.
This repository contains the code for Feature Engineering using python's various libraries.
It used Numpy and Tensorflow libraries.
These libraries help to perform individually one particular functionality.
Numpy is used for working with arrays. It stands for Numerical Python.
TensorFlow is a Python library for fast numerical computing created and released by Google.
The purpose of creating this repository is to gain insights into how to clone voice live.
These python libraries raised knowledge in discovering these libraries with practical use of it.
It leads to growth in my AI repository.
This above few screenshots will help you to understand flow of output.

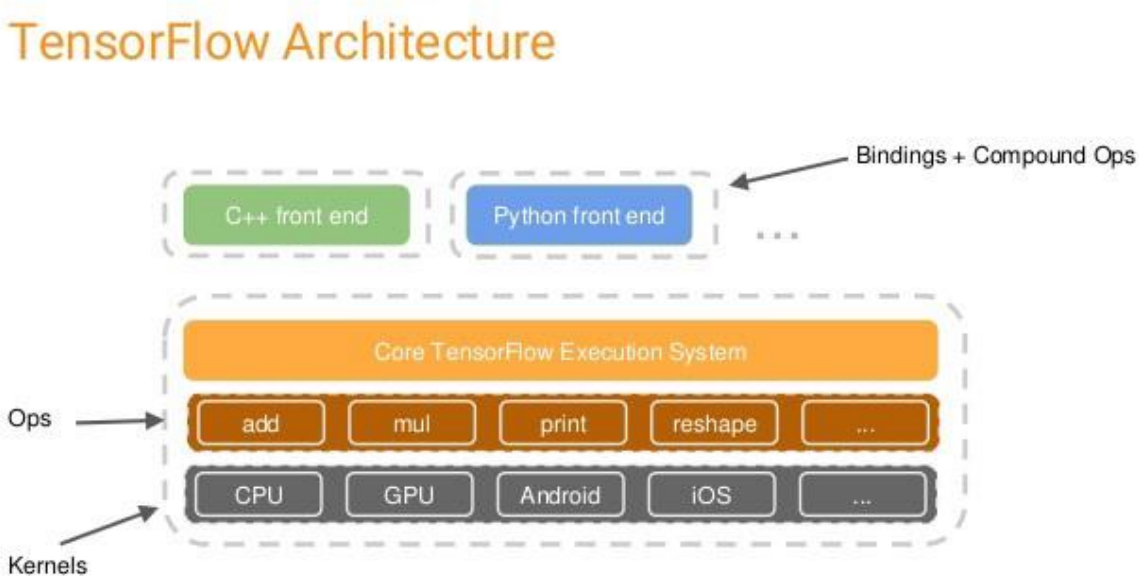
Motivation

The reason behind building this is, one of my friends stays abroad and does not have guardian. Therefore, I thought this might help him to feel presence at least in few celebration and good times. This helps to speak someone dear one's voice and feel the existence. Especially in this lockdown time, it has become more and more important to stay connected. Hence, I continue to gain knowledge while practicing the same and spread literary wings in tech-heaven.

Technical Aspect

Numpy contains a multi-dimensional array and matrix data structures. It works with the numerical data. Numpy is faster because is densely packed in memory due to its homogeneous type. It also frees the memory faster.
Tensorflow has excellent community support and provides scalability. TensorFlow provides a better way of visualizing data with its graphical approach. It also allows easy debugging of nodes with the help of TensorBoard. This reduces the effort of visiting the whole code and effectively resolves the neural network. Both TensorFlow and PyTorch provide useful abstractions to reduce amounts of boilerplate code and speed up model development. The main difference between them is that PyTorch may feel more "pythonic" and has an object-oriented approach while TensorFlow has several options from which you may choose.





Installation

Using intel core i5 9th generation with NVIDIA GFORECE GTX1650.
Windows 10 Environment Used.
Already Installed Anaconda Navigator for Python 3.x
The Code is written in Python 3.8.
If you don't have Python installed then please install Anaconda Navigator from its official site.
If you are using a lower version of Python you can upgrade using the pip package, ensuring you have the latest version of pip, *python* -m pip install --*upgrade pip and press Enter*.

Run/How to Use/Steps

Keep your internet connection on while running or accessing files and throughout too.
I recommend you to open Gmail first and then open Colab.
I have created this project in Google Colab.
Just Open my Colab File, Copy the code in your Colab. Then download adele_speech.wav from my repo and upload to your Colab.
Everytime you have to upload this .wav file to Colab before running because it stays temporarily in Colab
Reminder, uploaded files will get deleted when this runtime is recycled.
First upload adele_speech.wav in Real-Time-Voice-Cloning Folder by clicking three dots next to it and selecting upload option.

Then you will see adele_speech.wav file uploaded.

Then Right-click on it and select option 'Copy Path' and Paste it in 6th cell when asked for.

Then type few sentences that you want to listen.

Click on Refresh.

Then you will see 'demo_output_00.wav' file which is an output file, you can click on it and listen output voice of sentence you typed in.

If you want you can download it by right-clicking on it and selecting download option.

Run each cell individually and not all at once because you have to upload a .wav file after running 5th cell and before running 6th cell.

Directory Tree/Structure of Project

Folder: 21)RealTime_VoiceCloning_in6linesofcode

RealTime_VoiceClone.ipynb

To Do/Future Scope

Can try voice to be fetched from audio.

Technologies Used/System Requirements/Tech Stack

Credits

Practical Learning Channel