Yuvraj Singh
AI/ML Intern Assesment
https://yuvraj153989.github.io/Portfolio/
+91 8591476799

Lip Sync Model using wav2lip

TASK: The objective of this assignment is to demonstrate your skills in creating an AI model that is proficient in lip-syncing i.e. synchronizing an audio file with a video file. Your task is to ensure the model is accurately matching the lip movements of the characters in the given video file with the corresponding audio file.

Solution:
In provided video, face was not detected in every frame, as per pre-trained wav2lip model's rule, we have to use a video having valid face in every frame.

For Bypassing this, I have made some minor changes in inference.py code
Add a condition in "for" loop for face detection

```
for p, f, c in zip(pred, frames, coords):
  # y1, y2, x1,x2=c <-Move this line…

  if len(c) == 0:  # Check if no face detected
      out.write(f)  # Write the original frame without lip-syncing
  else:
      y1, y2, x1, x2 = c        ..moving here
      p = cv2.resize(p.astype(np.uint8), (x2 - x1, y2 - y1))
      f[y1:y2, x1:x2] = p
      out.write(f)  # Write the lip-synced frame to the output video
```

In this way, I have achieved the result.

Steps to view:
1) Open colab file, and check for GPU, dependecies, update them, and switch runtime to T4 GPU.
2) Mount your g-drive
3) Get the code and model by cloning from Rudrabha/Wav2Lip repo
4) Create a new folder in g-drive and rename it Wav2Lip.
5) Get pre-requisites like installing requirements.txt, face-detection model etc
6) Upload your input video in mp4 and audio in wav format in g-drive/Wav2Lip
7) Update librosa library (Shown in the code)
8) You'll find the imported Wav2Lip folder, locate inference.py and paste the above code (Find line no. 267)
9) Run the next command in code, after 3-5 minutes, check for results in Wav2Lip/results/result_voice.mp4