Deployment Document

Contents

1. Setting up the wav2lip model with proposed scripts and injection on Jupyter Notebook
2. Setting up GCP.
3. Setting up Confluent Cloud for Inference

Wav2Lip Setup.

The first part of the project is to setup the Wav2Lip model. We've already identified some of the blockers that we faced and altered the methods that would be needed to work with our pipeline setup ,and the wav2lip model. This includes the usage of the model with Python 3.90. The current inference pipeline is tied to be used in the Jupyter notebook. Since we wanted to use it in a pipeline setup which was initially. We had to modify certain methods in the original model so that it works with our script.

Please follow below steps to setup :

Prerequisites :

1. Enable Text To Speech, Speech to Text and Text to Text Translations using below links

Sr. No.	Description	Setup Guidelines	References
1	Enable GPU on Jupyter Notebook	Navigate to your Colaboratory Account and Enable GPU based Notebook	https://colab.research.google.com/
2	Use script : Folder_Setup.ipnyb	Use this script to setup the Wav2Lip model with the code to interact with GCP in Saas based setup	Everything has been provided here

Parameters to be set in audio_video_handler.py

Variable name	Value
CHECKPOINT_PATH	"Wav2Lip/checkpoints/wav2lip.pth"
FINAL_OUTPUT_DIRECTOR	"converted_videos/output.mp4"
OUTPUT_TRANSLATED_AUDIO_LOCATION	"translated_audio/translated.wav"
OUTPUT_AUDIO_FILE_LOCATION	"extracted_audio"
DEFAULT_VIDEO_FILE	"data/uploaded/uploaded.mp4"
DEFAULT_IMAGE_FILE	"data/uploaded/uploaded.jpg"
DEFAULT_VIDEO_FILE_NAME	"talk_voice_English"
DEFAULT_LANGAUGE	"Spanish"