

Many Face QA

Functionality Analysis:

Workflow:

The workflow for this project is divided into two distinct steps. When users have provided all the required inputs including the selection of an audio voice, or uploading a custom audio and the input script., in the first step of **voice cloning**, they are sent to the relevant model, which processes the data to create a cloned voice audio with the user script using text-to-speech. This cloned voice audio is then saved on the server for further use.

Moving on to the second step of our application, we use the cloned audio file along with the user's selected avatar to generate a final lip-synced video output. This process involves advanced algorithms that synchronize the audio with the avatar's movements, creating a realistic and engaging video experience. The resulting video is saved in high-quality 30 frames per second (FPS) and 1080p resolution, ensuring a professional-grade output.

When users initiate the video processing command, they can easily monitor the progress and status of their project directly from the Homepage. The status can be "pending" while waiting in the processing queue, "running" during the actual processing, or "succeeded" once the video generation is completed successfully. At this point, users have the option to download or delete the generated video from the project tab, providing them with full control over their creations.

It's worth noting that the processing time for a 200-word script is approximately 10 minutes. However, this duration may vary depending on various factors such as the complexity of the script, server load, and other operational considerations. Despite these variables, we strive to deliver efficient and timely processing to meet our users' expectations for quality and speed.