

Speaking Test Video Subtitles Project

At the English center where I work, the staff spend a lot of time adding Vietnamese subtitles to students' speaking test videos. Currently, this is done using CapCut, which requires them to translate and type the text in over the video, caption by caption. I wish to create a more efficient process for adding Vietnamese subtitles to speaking test videos.

I propose utilizing **ffmpeg** and **openai-Whisper** to do this.

Here is the process.

1. Extract audio from mp4 as mp3
2. Transcribe the audio with Whisper
3. Save subtitles as .srt
4. Allow a human to check the English transcription, then to automatically translate it to Vietnamese.
5. burn the Vietnamese .srt to mp4

I would like this to be usable from a simple front-end interface.

Here are some resources I found while researching:

- [How to generate and add subtitles to videos using Python, OpenAI Whisper, and FFmpeg | DigitalOcean](#)
- <https://github.com/openai/whisper>

Here is a script created by ChatGPT:

<https://chatgpt.com/canvas/shared/67bd5ed0ead88191bfaa0cb86efe5db8>