
EE/CprE/SE 491 BIWEEKLY REPORT 3

September 28 - October 11

Group number: sddec21-06

Project title: DigiClips Media Design

Client: DigiClips

Advisor: Ashfaq Khokhar

Team Members/Role:

Sam Massey - Assignment planning, Research, Work on speech-to-text

Tyler Johnson - Planning and implementing test cases

Maxwell Wilson - Primary point of contact with client, Research, Work on speech-to-text, Team Leader

Max Van De Wille - Documenting architecture changes, Research, Work on video-to-text

o **Weekly Summary**

This past week, our team continued development of speech-to-text and video-to-text elements. While using DeepSpeech open API for speech-to-text, we have made progress to allow recordings to be broken up into multiple segments, transcribed, then concatenated back to give a final result of a recording converted to a transcript. On the video-to-text side, the refactoring process is mostly complete but in need of some final adjustments. This week we will work on Docker-izing the video-to-text API then begin performance testing on our client's system.

o **Past week accomplishments**

Max Wilson:

- Researched Python multiprocessing and issues with Python multiprocessing pools
- Implemented multiprocessing for speech recognition
- Researched Docker wsl2 issues
- Optimized multiprocessing pool so that deadlocks don't happen
- Optimized Docker container so it is 30% faster

Sam Massey:

- Experiment with PyDub to try and increase speed of transcripts created.
- Research more ways to increase speed in the program.
- Research specific ways to increase the speed of the program when running within Docker.
- Try to optimize audio files for the speech-to-text program.

Max Van De Wille:

- Continued refactoring video processing script into API
- Refactored some utility methods into separate utils file/folder to clean up main running process
- Began working on duplicate filtering for multiple instances of text within a segment
- Standardized pre-processing methods after testing across different samples

Tyler Johnson:

- Continued collecting data from video clips
- Created basic testbench for testing accuracy between speech to text and transcribed text
- Began creating testbench for testing accuracy between video to text and transcribe text

o **Pending issues**

- No unified/standardized testing set to compare performance of one iteration to the next makes it hard to benchmark progress/performance improvements.
- One of the main pending issues our team as well as the team at DigiClips is facing will be ensuring accuracy in our speech-to-text and video-to-text.
- Speech-to-text app is currently splitting the audio file without overlap. This could split in the middle of a word and could cause data to be lost. We need to implement overlap in the chunking logic so that we don't cut any words in half.
- Certain fonts displayed in sample videos are not detected as well by tesseract

o Individual contributions

Team Member	Contribution	Weekly Hours	Total Hours
Sam Massey	Deepspeech work, PyDub research and experimentation	7	40
Tyler Johnson	Testbench creation and proofing	7	40
Maxwell Wilson	Docker and multiprocessing experimentation	7	42
Max Van De Wille	Video-to-text development, generating benchmark samples for client	7	40

o Plans for the upcoming week

Max Wilson:

- Implement chunking overlap to catch any words that may have been cut in half
- Improve multiprocessing system that was implemented last week
- Design method of tagging audio chunks so that Digiclips can see where certain words occurred in the audio file
- Identify max audio clip length DeepSpeech can process so future audio clips can be chunked dynamically

Sam Massey:

- Look into PyDub to see about the possibility of reducing background noise.
- Try implementing more ways to increase speed in the program.
- Research specific ways to increase the speed of the program when running within Docker.
- Continue to optimize audio files for the speech-to-text program.

Tyler Johnson:

- Add features to text-based testbench, such as potentially time measurements
- Determine best solution for video-based testbench options
- Continue working on video-based testbench

Max Van de Wille

- Finalize refactoring video processing script into API
- Continued work on duplicate filtering, trying out a couple of new approaches
- Alter output method from one long string to time stamped indices of written text for better search potential