
EE/CprE/SE 491 BIWEEKLY REPORT 5

October 25 - November 8

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. On the speech-to-text side of our app, we have begun adding chunk overlap and dynamic chunking features to split the input file into manageable sizes for output timestamps. For video-to-text we have started working on a duplicate filtering method to try and limit the repetition of constant phrases across many different timestamps. This week we will continue work on chunking and duplicate filtering processes and work on expanding our documentation to be clearer and more encompassing. In the near future we will be packaging both portions of our application using the driver microservice which is yet to be developed.

o **Past week accomplishments**

Max Wilson:

- Initial implementation of dynamic chunking and chunk overlap
- Researched docker volumes as a possible method of linking client audio files to docker container
- Wrote documentation for speech-to-text service describing setup and use of the API

Sam Massey:

- Get more results with the DigiClips system using our speech-to-text code.
- Try our speech-to-text code on DigiClips radio recordings for accuracy and possible errors.

Max Van De Wille:

- Finalized duplicate filtering approach after testing
- Began implementing duplicate filtering process
- Swapped output to timestamped indices w/ text output

Tyler Johnson:

- Continuing work on video-to-text issues,
- Continuing research into adding testbench parameters

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.
- Speech-to-text needs chunking overlap
- Need to setup docker so the docker image can access the clients filesystem when running on their host machine
- Docker optimizations on the client's machine
- Need to write Documentation for the speech-to-text application
- Certain fonts displayed in sample videos are not detected as well by tesseract
- Need to begin packaging apps and developing driver microservice

o Individual contributions

Team Member	Contribution	Weekly Hours	Total Hours
Sam Massey	Deepspeech work, PyDub research and experimentation	7	68
Tyler Johnson	Video to text, adding features to speech-to-text	8	66
Maxwell Wilson	Docker and multiprocessing experimentation	7	68
Max Van De Wille	Video-to-text development, generating benchmark samples for client	7	68

o Plans for the upcoming week

Max Wilson:

- Optimizations for dynamic chunking and chunk overlap
- Docker volume experimentation
- Run our container on the clients machine and do some optimization to improve performance

Sam Massey:

- Run container on client machine and look for ways to increase speed.
- Test our program on 10 minute segments to compare speeds for clients' new recording process.

Tyler Johnson:

- Continue to complete video-to-text
- Continue to implement more speech-to-text options

Max Van de Wille

- Continue to implement duplicate filtering process
- Test dockerized API on client machine w/o duplicate filtering
- Performance test dockerized api on client's machine
- Alter output method to client-approved JSON formatting