Delay Removal on Television Interviews

CSE495

3^{nd} Presentation

Şeyda Özer

Project Advisor: Prof. Dr. Yusuf Sinan AKGÜL

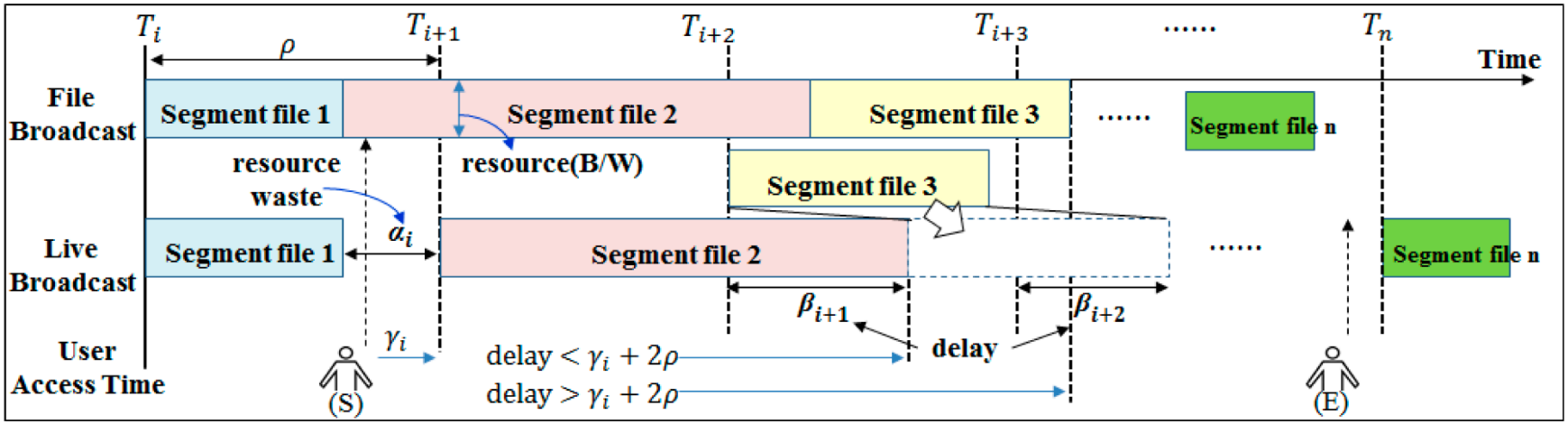
January 2023

Content

- Project Scheme and Description
- Project Design Plan
- Voice Activity Detection
- Delay Detection
- Project Timeline
- Environment
- Success Criterias
- Resources

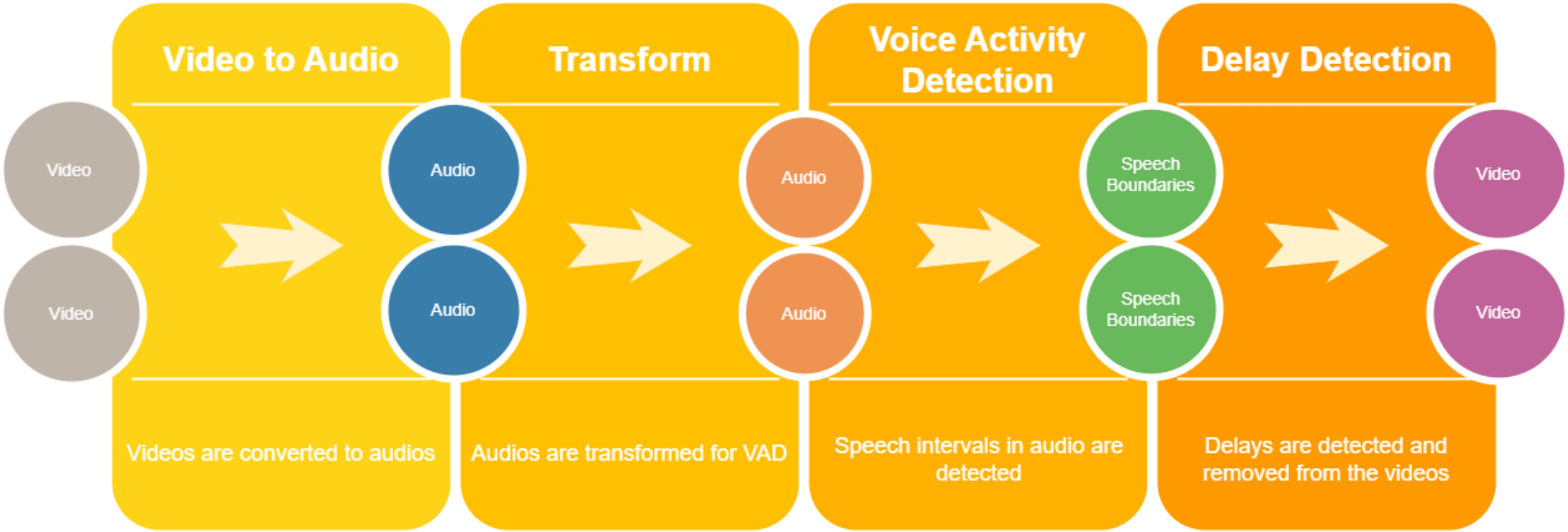
Project Scheme and Description

Delays can occur in TV interviews. When studio presenter questions a remote interviewee, they cannot hear each other in real time.



The aim of this project is to remove these delays. The audience should watch the video as if there were no delays and should not be aware of the delays.

Project Design Plan

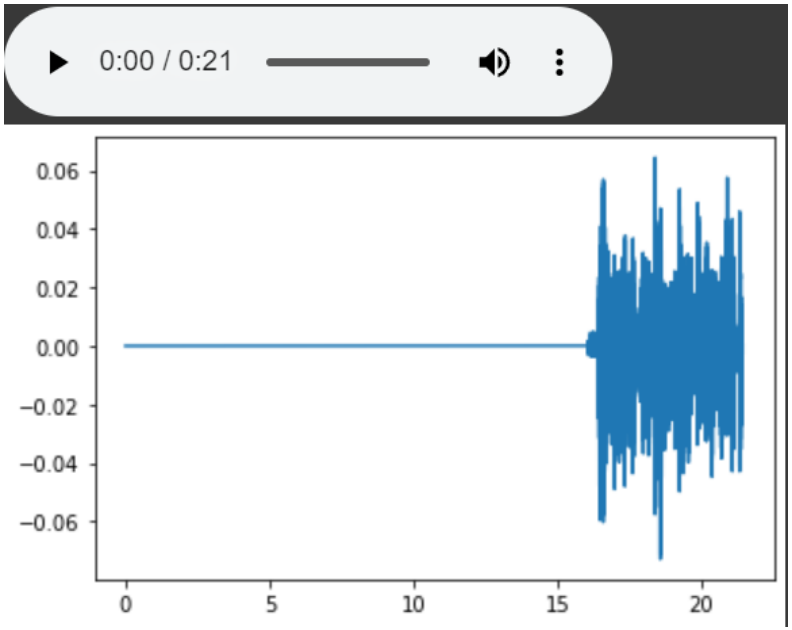


Voice Activity Detection

- SpeechBrain is an open-source all-in-one speech toolkit based on PyTorch.
- Speechbrain's pre-trained VAD model can be used for voice activity detection.

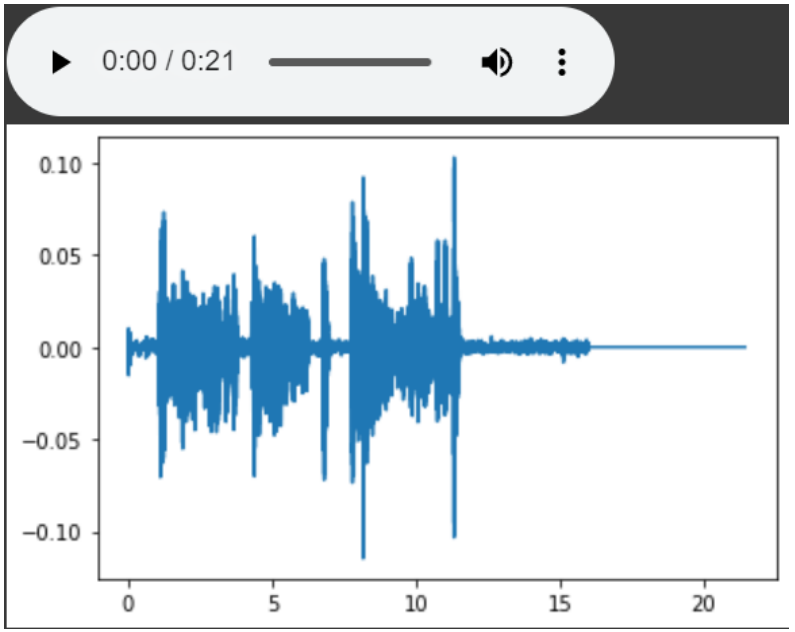
Voice Activity Detection

First video :



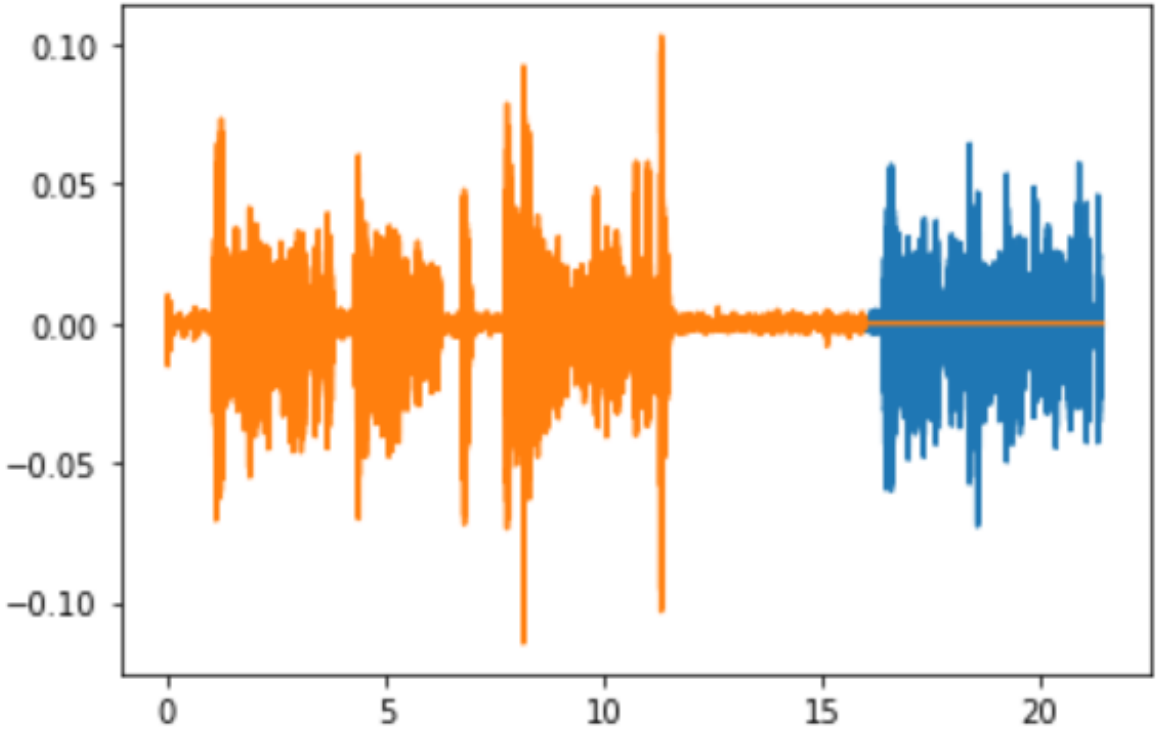

```
segment_001  0.00  16.32 NON_SPEECH
segment_002  16.32  19.99 SPEECH
```

Second video :



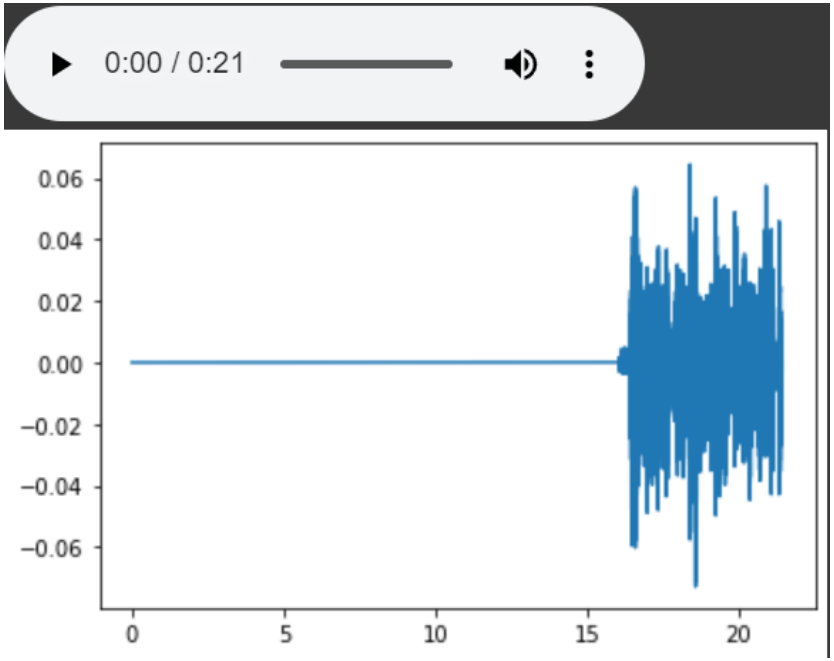

```
segment_001  0.00  11.82 SPEECH
```

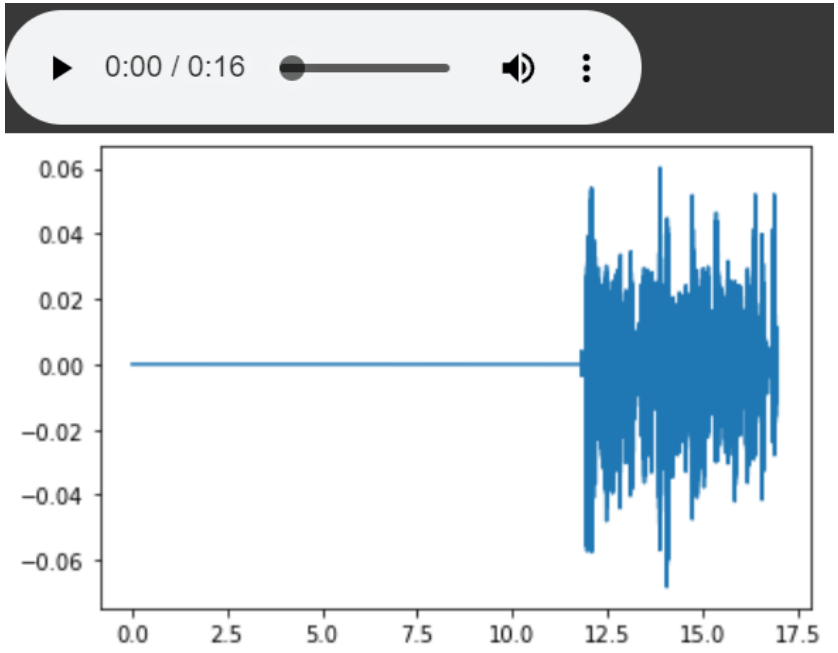
Delay Detection



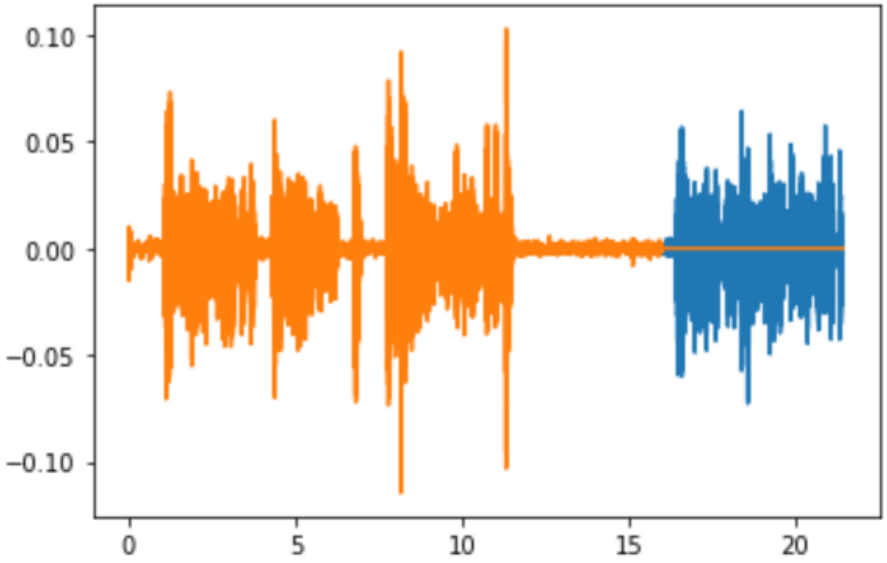
Delay Detection

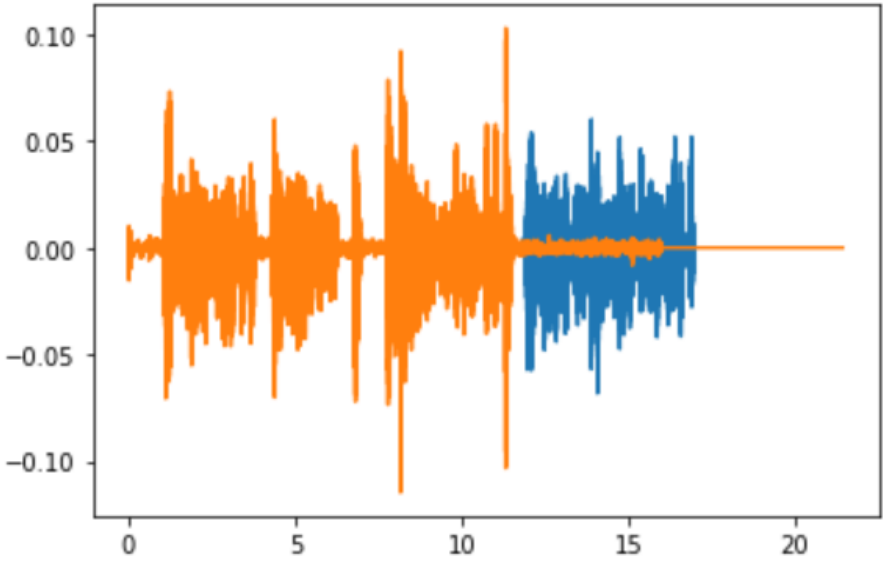
One delay is detected in the first video.





Delay Detection





Project Timeline

Weeks	17 October	24 October	31 October	7 November	14 November	21 November	28 November	5 December	12 December	19 December	26 December	2 January	9 January	16 January
Preparing the development environment, learning the necessary libraries, literature review														
Collecting datasets														
Algorithm implementation														
Test														
Preparing the report														

Environment

Success Criterias

- TV interviews with no delay (75% accuracy rate)
- 1 minute video will be handled in 1 minute.
- At least 100 videos will be used as data

Resources

1. Seo, Hyungyoon. Kim, Goo. “DASH Live Broadcast Traffic Model: A Time-Bound Delay Model for IP-Based Digital Terrestrial Broadcasting Systems” Applied Sciences 2021, 11(1), 247, https://doi.org/10.3390/app11010247
2. “Achieving Broadcast-Grade Low Latency in Live Streaming” in Streaming Media. Nov/Dec 2018, Vol. 15 Issue 8, p16-25. 10p.
3. Zhang, C. Liu, J. “On Crowdsourced Interactive Live Streaming: A Twitch. TV-Based Measurement Study” in NOSSDAV ’15: Proceedings of the 25th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video, March 2015, Pages 55-60, https://doi.org/10.1145/2736084.2736091.
4. https://speechbrain.github.io