

SlideReader: Automated Slide Analysis from Lecture Footage (Master Thesis)

ACCESS@KIT - Zentrum für digitale Barrierefreiheit und Assistive Technologien

Slides are crucial documents for conveying lecture material. Well-designed slides efficiently summarize complex information, making them valuable study materials. Additionally, when integrated with lecture footage, static elements offer deeper insights, such as summarizing data tables or interpreting charts. Well-produced lecture footage offers an excellent resource for retrieving the element-wise reading order with the corresponding auditory explanation. When slides are combined with the reading order and audio, it aids both sighted and visually impaired students in better understanding static slide materials. Numerous other applications can be derived, including but not limited to skimming, pitch preparation, and automated slide animation (slide authoring). In this research, we aim to develop the SlideReader, a system that can analyse lecture recordings to retrieve the element-wise slide reading order and audio explanation employing state-of-the-art models.

Dataset Preparation:

- collect high-quality lecture videos along with their corresponding digital slides.
- Utilize heuristics for data curation and annotation
- Model Training and Experimentation:
 - Train diverse neural networks, with a focus on transformers and LLMs.
 - Performance Evaluation:
 - Conduct a quantitative & qualitative evaluation of the trained models.

Throughout your research, you'll benefit from experts' guidance and support. You will have access to a powerful computing cluster to facilitate your experiments. Significant findings will be submitted as a research paper at a prestigious conference.

Requirements:

- Demonstrated interest in the topic. Related work [1-5].
- Experience with deep learning models, such as transformers and LLMs.
- Familiarity with the Linux operating system and comfort with terminal commands.

If interested, please send your application, which should include your CV and transcript of records (optional: motivation letter) to any of the responsible people:

Omar Moured, omar.moured@kit.edu, Kunyu Peng, kunyu.peng@kit.edu, Dr. Saquib Sarfraz, saquib.sarfraz@kit.edu

References:

<https://ieeexplore.ieee.org/document/7333955>

<https://dl.acm.org/doi/pdf/10.1145/3544548.3580921>

<https://github.com/BradyFU/Awesome-Multimodal-Large-Language-Models>

https://cvhci.anthropomatik.kit.edu/~mhourile/papers/2019_ICDAR_WiSe.pdf