Weekly Report

5/21

1. Summary of Weekly Progress

- Create better solution
- Summary and outlook
- Image alignment finished
- Adjusted the prompt of the subtitle generation
- Semantic segmentation transition algorithm

2. Task Assignments and Contributions

Team Member	Assigned Tasks	Completion Status
徐浩哲(01157030)	1. Improve human figure in stable diffusion 2. Paper work of summary and outlook 3. Image alignment	1. In progress 2. In progress 3. Complete
翁子翔(01157048)	1. Semantic segmentation transition algorithm 2. Paper work of summary and outlook	1. Complete 2. In progress
蔡豐蔚(01157010)	1. Adjusted the prompt of the subtitle generation 2. Write program report	1. In progress 2. In progress

2.1. Comments

Comments
Find out that if change the input and output resolution can strongly improve figure, so i think it is no need to use inpainting
This week, I continued researching algorithms for transitioning between semantic segmentation maps. I found that panoptic segmentation isn’t quite suitable for this task when there are multiple objects, it requires an additional custom-defined color map, which makes it difficult to generalize across different images. As for the "connected components" method that the professor mentioned last time, it seems less necessary in this context, since the class IDs in a semantic segmentation map essentially serve the same purpose. I think the current version of the algorithm is generally workable. I also did some preliminary testing on a model that generates videos from a single image. I think the results are impressive, but it can only generate up to 8 seconds.
Adjusted the prompt of the subtitle generation so that it analyzes the description of 10 keyframes to generate subtitles of appropriate length. The current problem is that the subtitles don't match the screen.

3. Challenges and Issues Faced

- still need to improve transition
- The subtitles are out of sync with the screen

4. Next Steps and Goals for Next Week

- adjust prompts
- Improve the algorithm
- The video timeline needs to be adjusted
- Add functionality to the frontend
- Connecting algorithms
- Subtitle generation may require another run of the Florence model

5. Additional Notes/Comments