# Weekly Report

## 3/26

### 1. Summary of Weekly Progress

- Connect frontend and backend by internet
- Test back-front connect by Interpolated result
- Manipulate the latent space through text operations
- Create a basic demo
- Try to use  stable diffusion with out pipeline

### 2. Task Assignments and Contributions

| Team Member | Assigned Tasks | Completion Status |
|---|---|---|
| 徐浩哲(01157030) | 1. Learn about latent space<br>2. Deploy ngrok to server<br>3. Try to use stable diffusion with out pipeline | Complete |
| 翁子翔(01157048) | 1. Using Florence for image-to-text and Phi4 to generate a step-to-step transition between two images<br>2. Attempt to manipulate the latent space through text operations<br>3. Generate transitional images between pictures by manipulating the latent space and interpolation | 1. Complete<br>2. Complete<br>3. Complete |
| 蔡豐蔚(01157010) | 1. Connect frontend and backend<br>2. Use the result image of the interpolated method when test backend is connected<br>3. read *Latent Diffusion Model (Pt.4.2~Pt.6)* | 1. Complete<br>2. Complete<br>3. Complete |

### 2.1. Comments

| Comments |
| --- |
| Finished to put ngork to let server on internet.<br>Learn latent space by youtube video, paper reading , chatGPT. Find out that it is difficult to find the vector by two photo. trying to use stable diffusion without pipeline, able to better to Full control over the latent space, Ability to add specific vectors to modify style or features, Allows gradual adjustment of influence (e.g., alpha interpolation) |
| Use Florence to process two images and generate a detailed textual description of the scenes. Then, use phi4 to let the model imagine what would realistically happen between the two images. Format the output to apply to photos with different interpolation levels. Last, use interpolation to connect the images in between. |
| I finished reading the paper *Latent Diffusion Model* and understood each step of the operation in general, so I tried to summarize its. I also completed the connection between the frontend and backend, as well as tested it by the result image of interpolated method. |

## 3. Challenges and Issues Faced

- The old picture and the new picture may need to be aligned
- use to photo compare in latent space trying to find the key vector, but failed.
- struggle to find a way to convert prompt to key vector.
- how can Florence generate better image descriptions
- Text manipulation of the latent space may not always be ideal parameters will need to be adjusted
- The video is not smooth enough
- The generation takes a lot of time
- The steps for generating images
- hard to understand how to use latent space to convert photo.

## 4. Next Steps and Goals for Next Week

- Digging deeper into the operation of another latent space
- Maybe we need to confirm the size and shooting angle of the input pictures

## 5. Additional Notes/Comments

Take easy on spring break and focus on mid exam.