Schedule (7.15 – 8.30)

	Target	Requirement	Milestone	Plan B
Week 1	1. Implement the simplest brutal-force search + discriminator model. Manually label a small part of the data. Walk through the code, and test with this prepared subset to make it executable. 2. Implement a simple renderer that can visualize the image based on features. 3. Assist others in labeling the text descriptions of the rest data. Describe the images independently to generate at least two descriptions for each image.	A labeled subset of the data	1. Viable code that can produce an image given the text. No concern for the reasonability of generated images. 2. Fully labeled data	1. Put aside the rendering part. Consider the pipeline from text to image embedding first. 2. Fulfill an even simplified model, e.g. with randomly generated images
Week 2	1. Train the discriminator on the entire labeled dataset. 2. Fine tune the discriminator to achieve good precision/recall on the text-image pairs, refine the network architecture if necessary. 3. Run the predictor on the test set using the fine-tuned discriminator. Perform quantitative and qualitative evaluation based on multiple metrics. Do error analyses on the preliminary results. 4. Try sophisticated text processing tools, and	1. Fully labeled data. 2. A viable model.	1. An implemented model that can produces reasonable images. 2. A knowledge of the importance of each component, including text embedding and image feature construction.	The quality of generation relies on the performance of the discriminator. If no reasonable results can be made, focus on the feature engineering part and at least gain a knowledge of promising embedding settings.

	different feature settings of the image. Find out the effect of each part on the quality of generation.			
Week 3 - 4	1. Determine the text processing tool. Finalize image feature settings. 2. Try and validate other feasible models, including statistical model, sequence model, variational autoencoder and conditional GAN. The implemented discriminator can be inserted into the conditional GAN. 3. Do a fair comparison between these models. Report on the metrics, and compare to the baseline.	An executable baseline that can produce reasonable results.	1. Implemented sophisticated model(s) that can work on the entire dataset. 2. Comparison between models.	1. Skip a model later found not feasible or too difficult to implement. 2. Implement one model at least, conditional GAN and sequence model of the priority 3. Deep learning model may require more data. Work with others to augment the dataset.
Week 5	1. Choose the most promising model, probably sequence model or conditional GAN. Train and fine-tune the model on the entire dataset. 2. Evaluate the preliminary results. Compare to the baseline model. 3. Do error analyses based on the precision of the generated image at the category level. Do user study if necessary. Optimize the model accordingly.	1. Implemented sophisticated model(s) 2. A knowledge of the pros and cons of available models.	An implemented sophisticated model that produces reasonable images.	1. If stabilized training takes too much time, stop when reasonable results are observed. Then do detailed error analyses and localize the bottleneck. 2. User study can be omitted if high generation precision can be made.

Week 6	1. Consider more image features, including position, size, etc. First test on a small subset. If feasible, work with others to refine the dataset labels. 2. Modify the model based on the new features. 3. Improve the model's architecture. Try novel tricks, and fine tune.	A trained sophisticated model	1. A dataset that incorporates more features. 2. A sophisticated model that produces acceptable results.	If no further progress can be made on the model, focus on improving the dataset and modify the model accordingly.
Week 7	Implement user interface. Fulfill data interchange details to incorporate the model into a system.	A sophisticated model	An integrated system that can produce and show an image given the text input	If an integrated generation system cannot be completed for the moment, decompose it into several parts and make sure the entire pipeline can be fulfilled by manually transmitting signals between parts

Chengyu 7.12