

### Appendix A. Top Volunteers by Contributed Compute Time

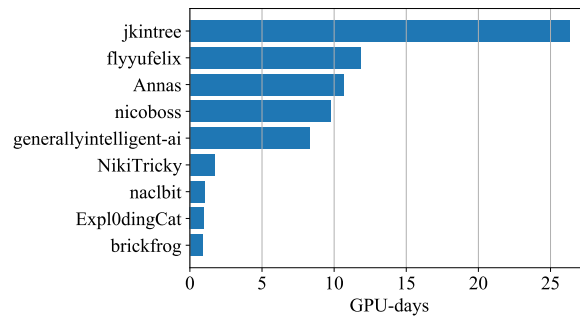
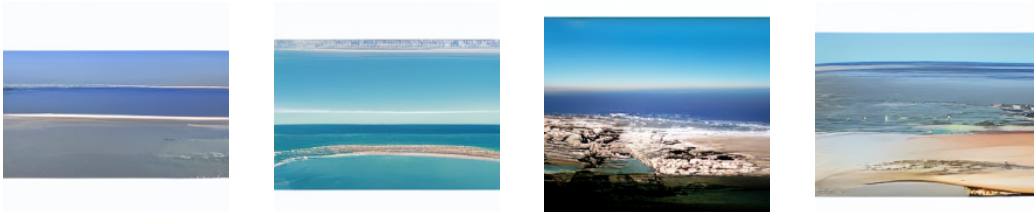


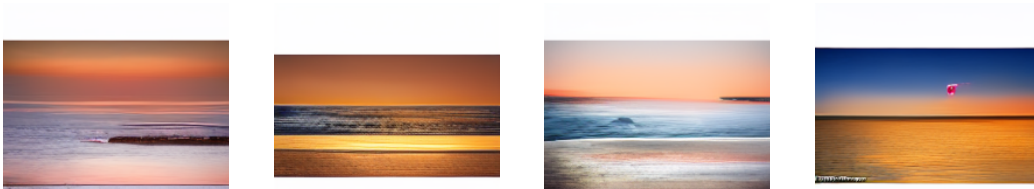
Figure 1: Hugging Face usernames of volunteers who contributed the most compute time.

### Appendix B. Model Inference Results

(a) aerial view of the beach during daytime



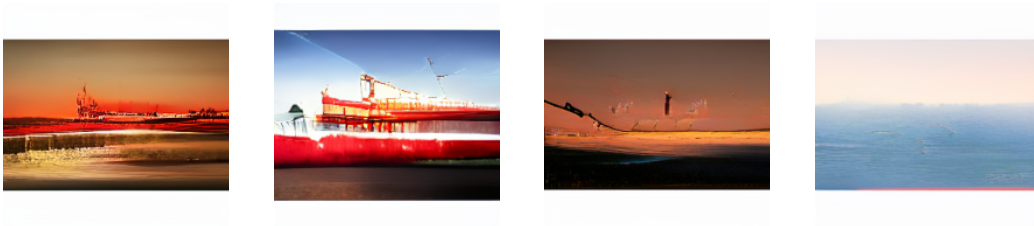
(b) a beautiful sunset at a beach with a shell on the shore



(c) flower dress, size M



(d) a photo of san francisco golden gate bridge



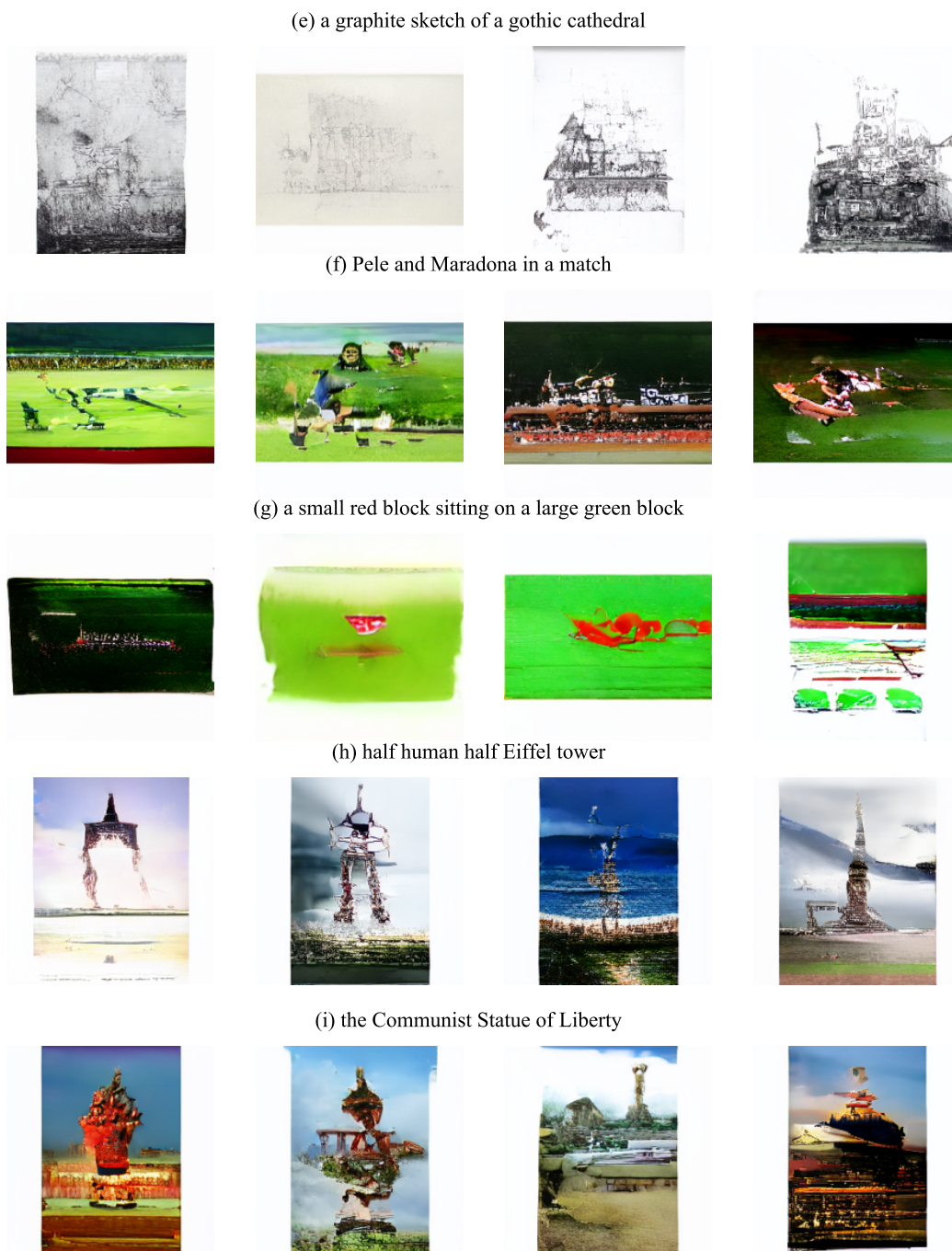


Figure 2: Examples of the inference results of the final model:  
 (a)–(c) Prompts leading to realistic outputs.  
 (d)–(f) Prompts where the model fails to draw the correct object shapes, but uses the appropriate image style, textures, and colors.  
 (g)–(i) Prompts where the model is able to generalize and draw the concepts not present in the training dataset. This is checked by inspecting images whose CLIP embeddings are the most similar to the embeddings of the prompt.