ArtAdapter: Text-to-Image Style Transfer using Multi-Level Style Encoder and Explicit Adaptation

Supplementary Material

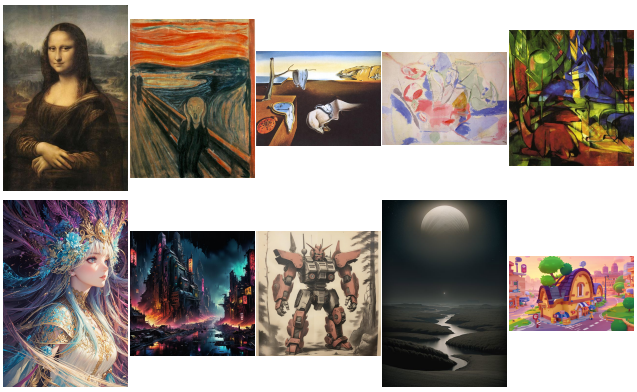

Figure 8. Style reference used in the test dataset for single-reference T2I style transfer.

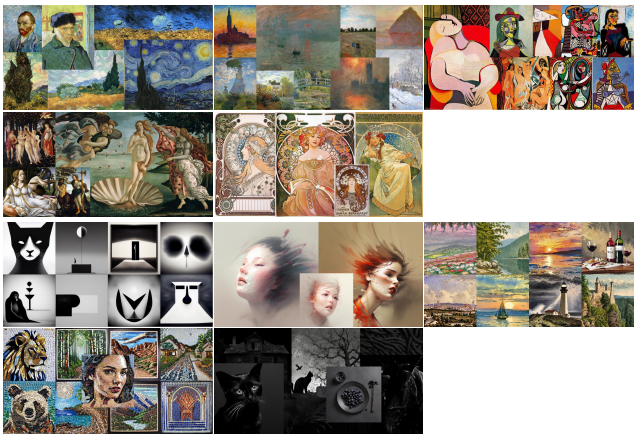

Figure 9. Collections of styles used in the test dataset for multi-reference T2I style transfer.

A. Test Dataset

We showcase the style references and prompts utilized in our quantitative evaluation and user study in Figures 8 and 9, and Table 3. Renowned artworks have been sourced from https://www.wikiart.org/, while others are synthesized images collected from https://civitai.com/.

"A robot"	"A curly-haired girl wearing a red dress, she is dancing"
"A little boy playing football"	"A monkey playing with a banana"
"A dog in the desert"	"A pine tree with a snowman hugging it"
"A bird in a wood"	"A stone with a face carved on it, standing on a pedestal in a museum"
"A kitten sleeping on a pillow"	"A modern house with a pool"
"A daisy with a ladybug on it"	"A house covered with ice and snow"
"Cherry blossom in full bloom"	"People in gondolas on the river"
"solar system and galaxy"	"A woman sitting on grass near a tree"
"A river with rapids and rocks"	"A windmill on a hillside"
"Polar bears on the ice"	"A woman sitting on grass near a tree"
"A snowy mountain peak"	"A castle with steps leading up to it"
"A motorcycle"	"A couple sitting on a swing seeing aurora"
"Sheep in autumn field"	"A woman wearing a baseball cap"
"Volcano eruption"	"Eiffel Tower in heavy lightning storm"
"A white rose"	"A house made of cardboard boxes"
"A stone with a hole in it"	"A bronze sculpture of an old man"
"A portrait of a rugby star"	"A colosseum near a waterfall"
"A girl playing electric guitar"	

Table 3. List of prompts used in the test dataset.

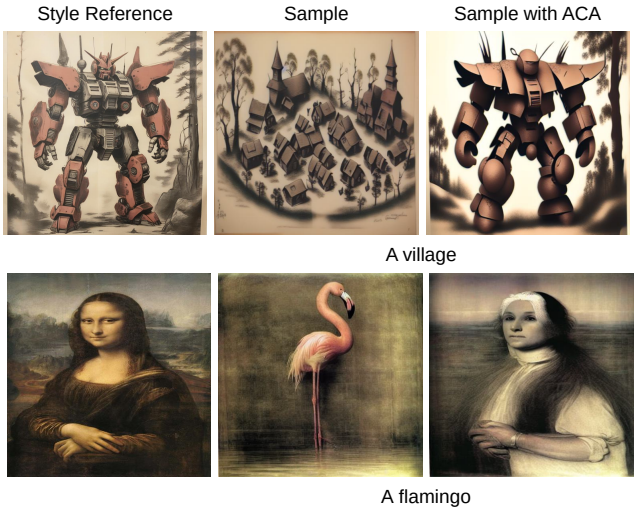

Figure 10. Ablative results of the ACA.

B. Details on User Study

We randomly selected 40 generated images, encompassing the entire spectrum of multi-reference style collections and prompts. Results from LoRA [20], TI [14], and our ArtAdapter were presented in a randomized order for users to compare and rate on a seven-point scale. Responses were solicited from 212 individuals, comprising both GenAI industry professionals and general users. Each participant evaluated five image tuples randomly drawn from our selection, yielding an aggregate of 1060 assessments.

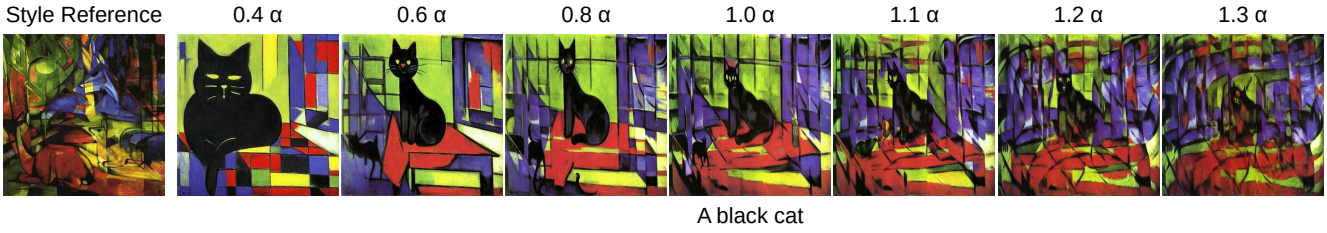

Figure 11. **Adapting α.** ArtAdapter enables users to scaling the α, effectively balance textual and style fidelity.

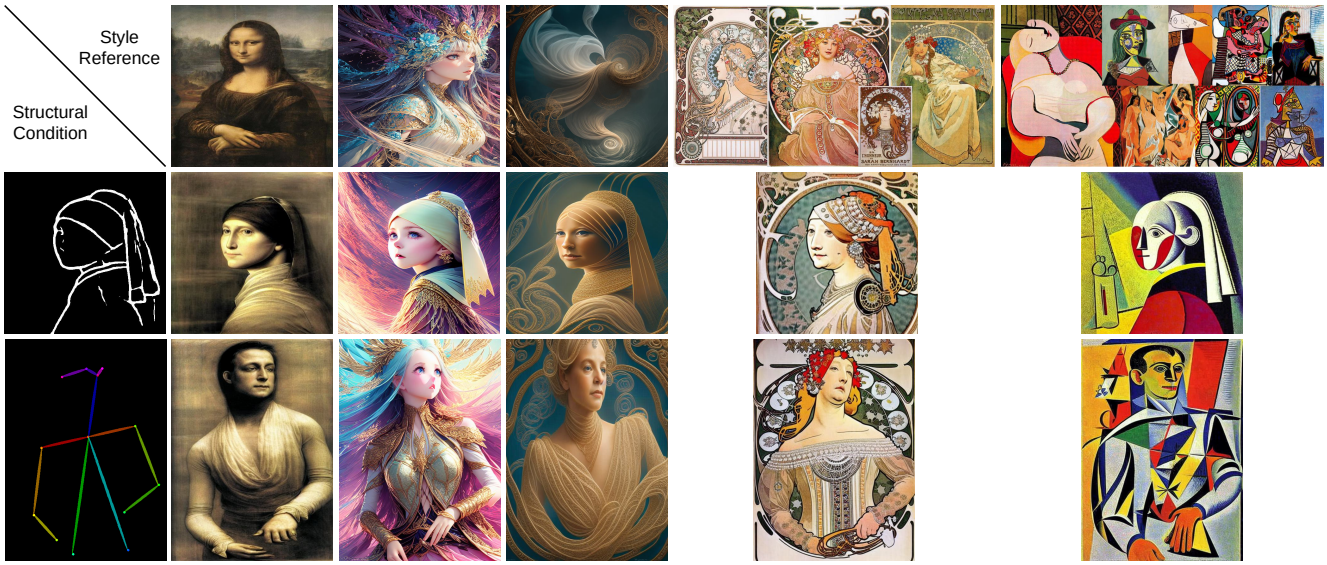

Figure 12. Results with additional structural controls.

C. Extended Ablation Study

C.1. Auxiliary Content Adapter

Figure 10 demonstrates how the Auxiliary Content Adapter (ACA) captures rough content structures, influencing the final results. When employing ACA during the denoising process, the content structures from style references will be presented. Consequently, by omitting ACA during inference, we can effectively remove these content structures while retaining the style representation.

C.2. Adaptive α

Figure 11 illustrates the impact of the adaptive α, as defined in Equation 2, during sampling. By adjusting the scaling factor of α, users can tailor the balance between the style elements from style references and the textual semantics. A higher α scaling emphasizes style traits in the results, albeit at the expense of textual semantics alignment, while a lower scaling does the opposite. This flexibility empowers users to generate results that align with their own preferences.

D. More Qualitative Results

In Figure 12, we demonstrate how effectively ArtAdapter integrates with the existing T2I-Adapter [31], showcasing its adaptability in exerting additional structural controls. Further evaluations of ArtAdapter across a broader range of styles are presented in Figures 13, 14, and 15. These results underscore ArtAdapter's impressive performance across diverse styles.

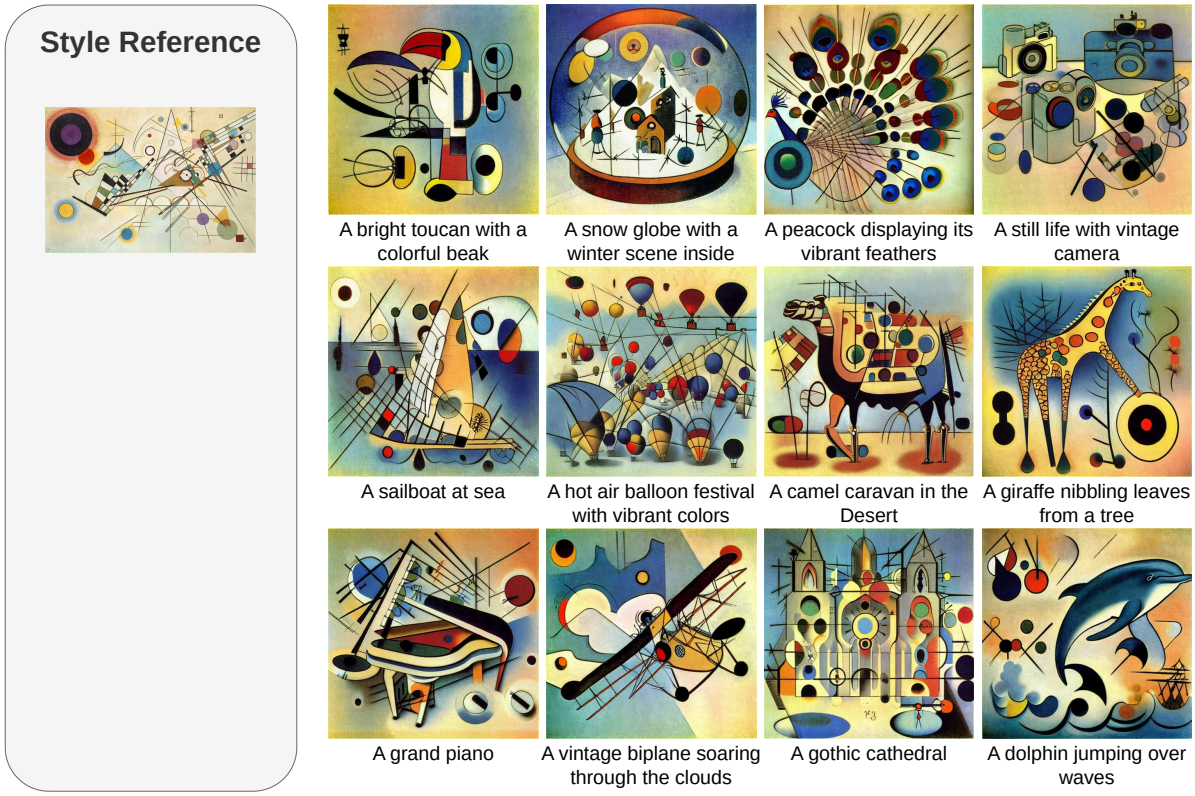

Figure 13. More results on single-reference T2I style transfer.

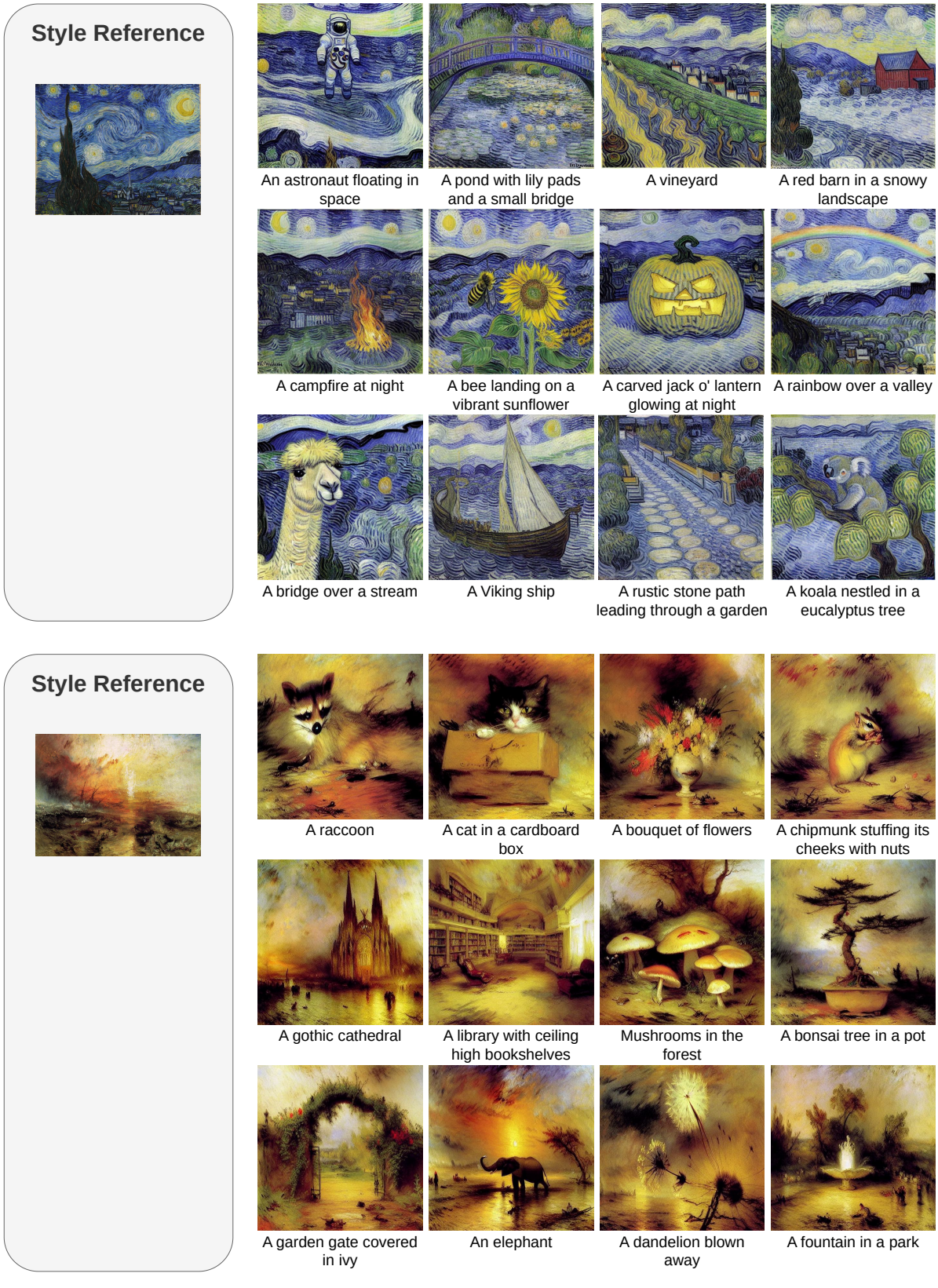

Figure 14. More results on single-reference T2I style transfer.

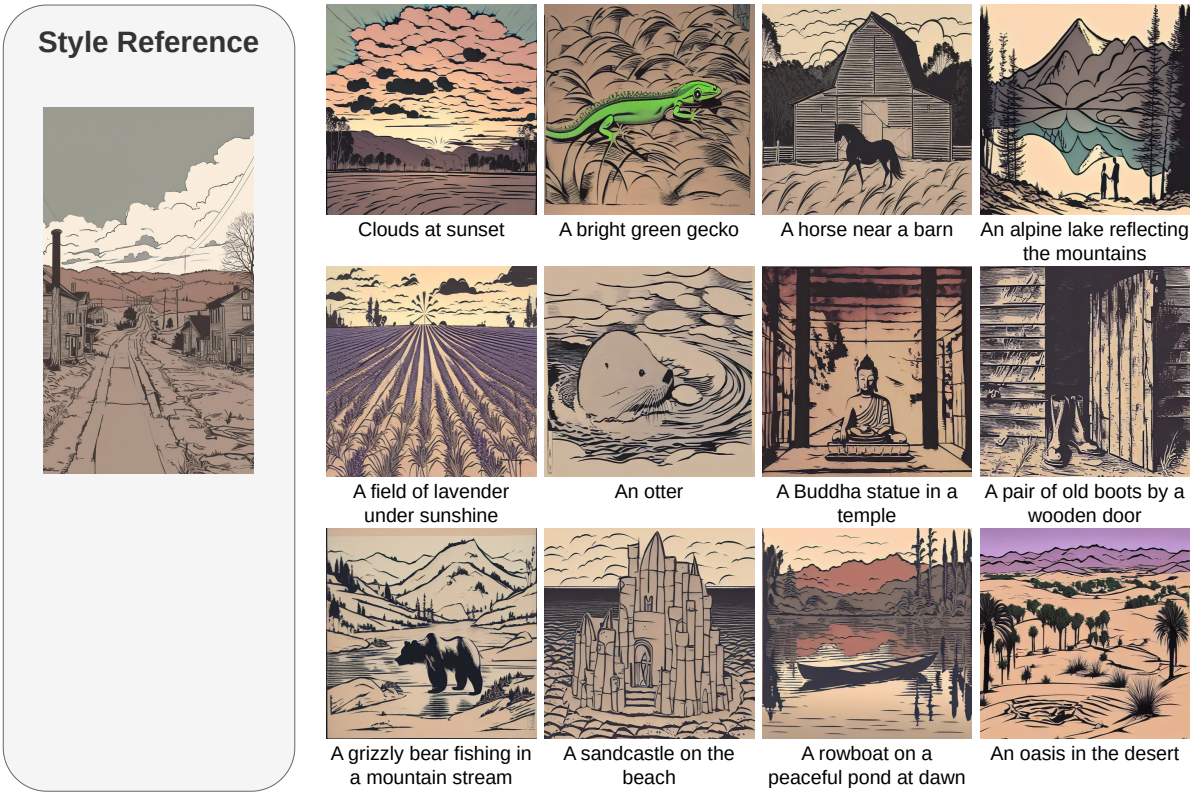

Figure 15. More results on single-reference T2I style transfer.

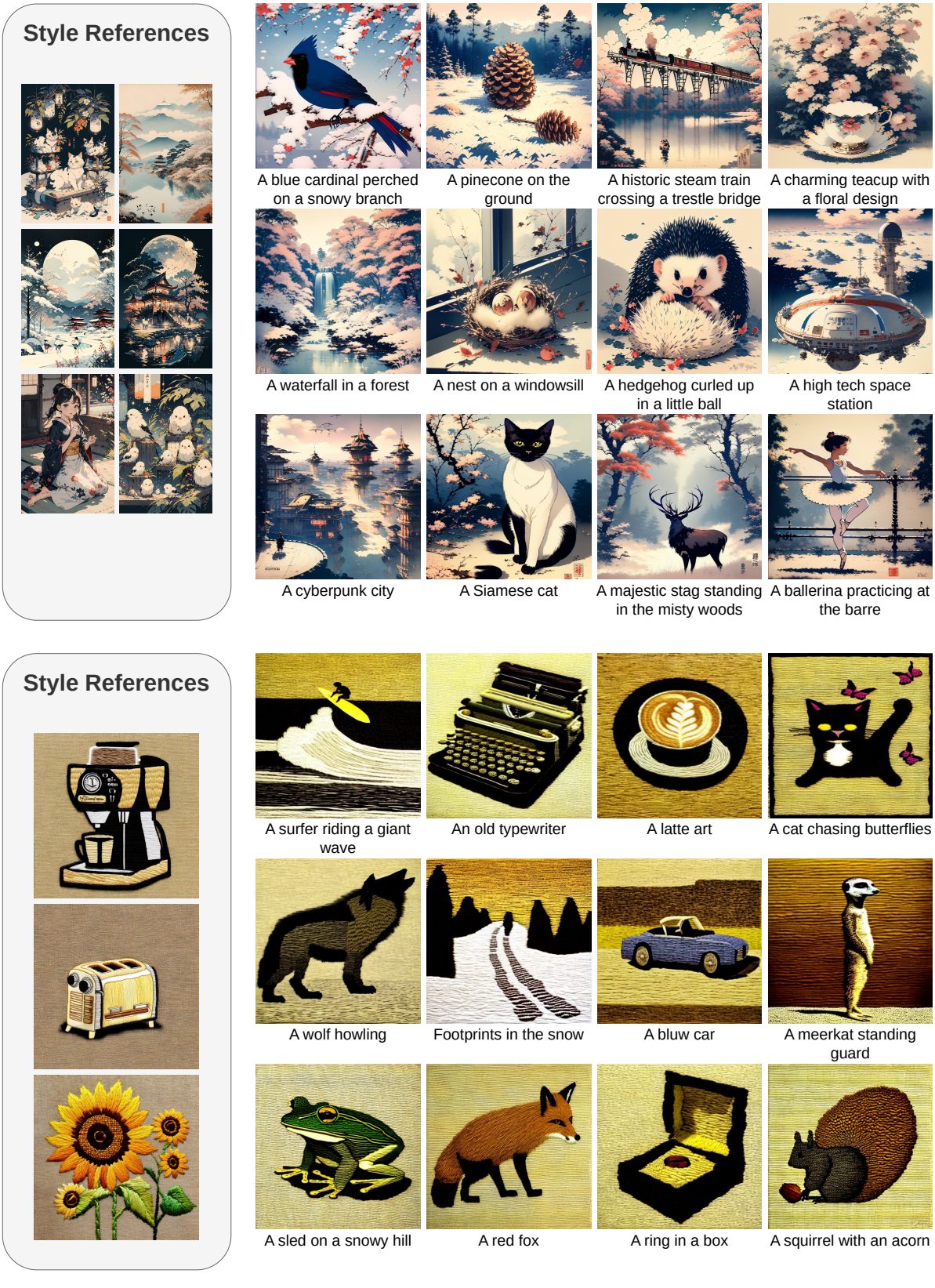

Figure 16. More results on multi-reference T2I style transfer.