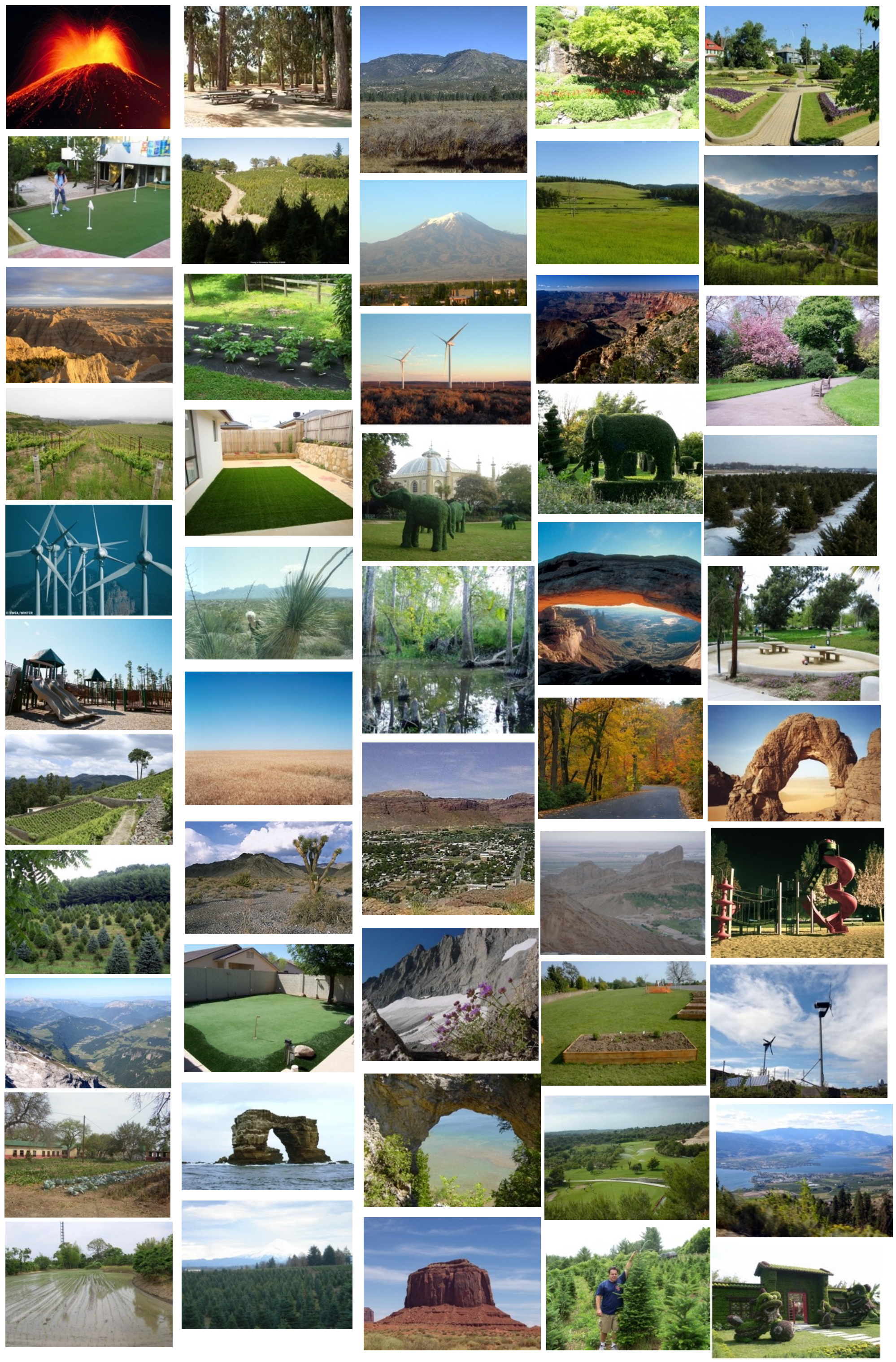
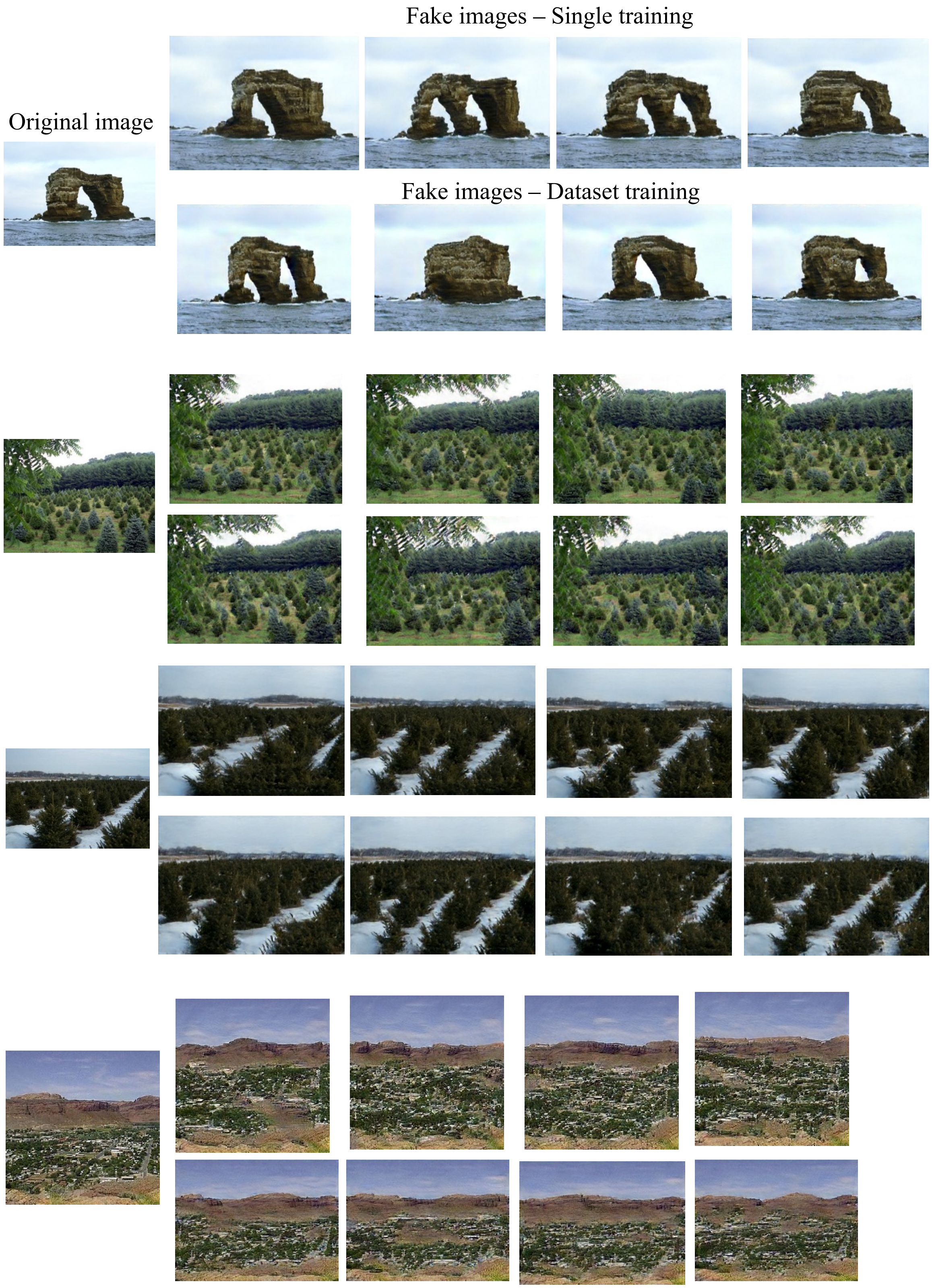
D.1 Places-50 real images



D.2 Places-50 fake images (single training)

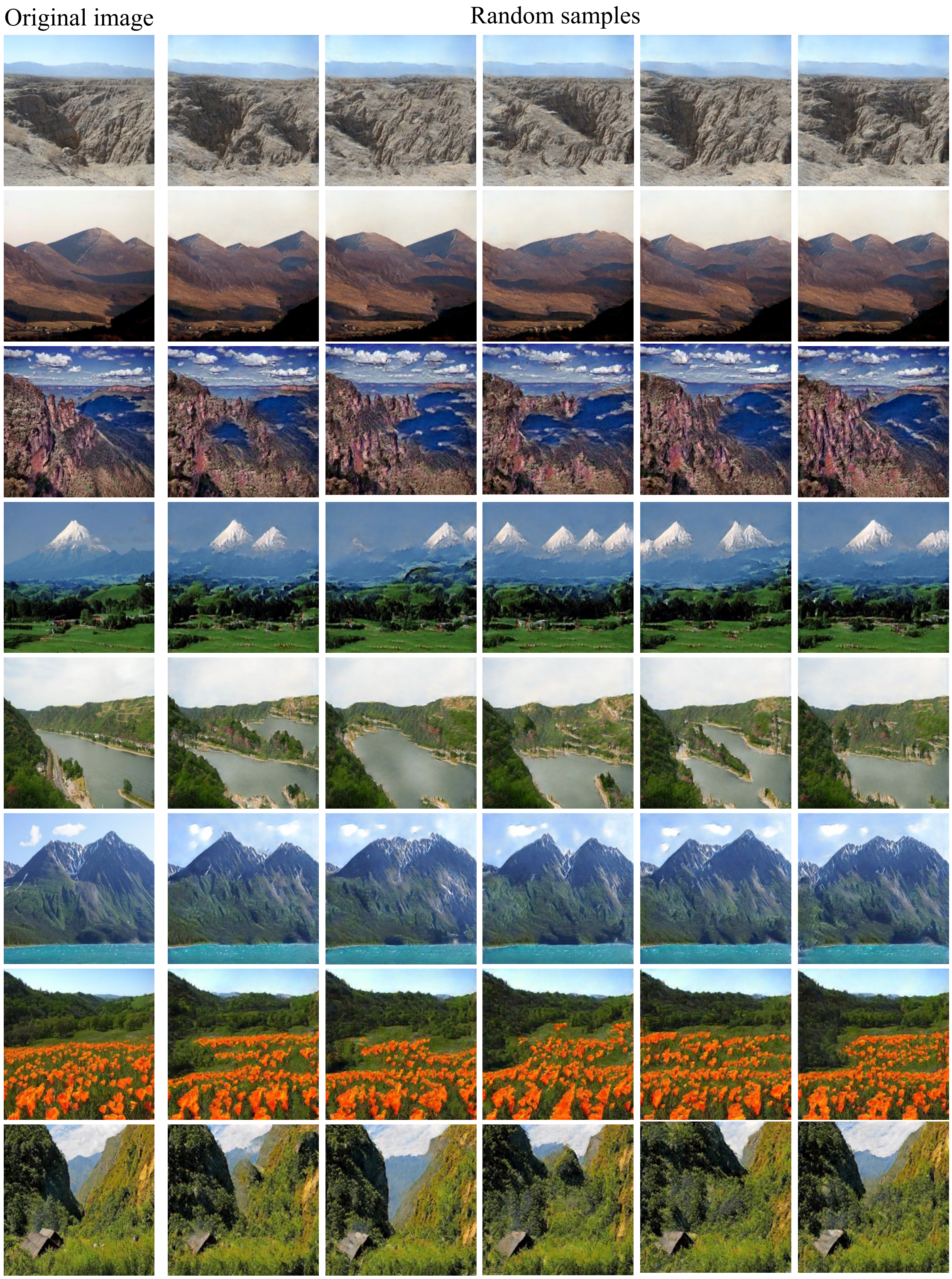
D.3 Places-50 fake images (dataset training)

D.4 Places-50 random samples (Single vs dataset training)

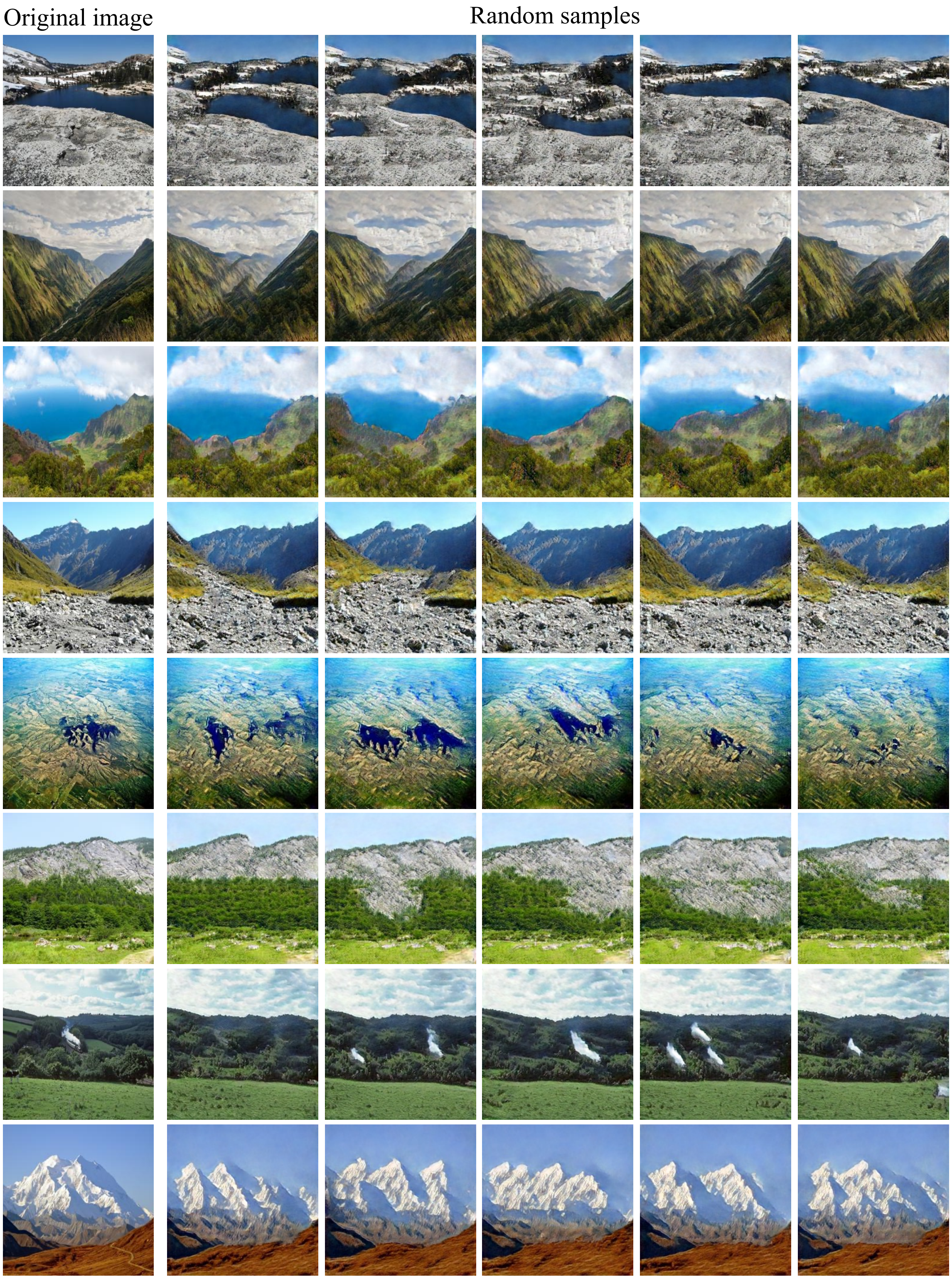


D.5 Single mini-batch training experiment

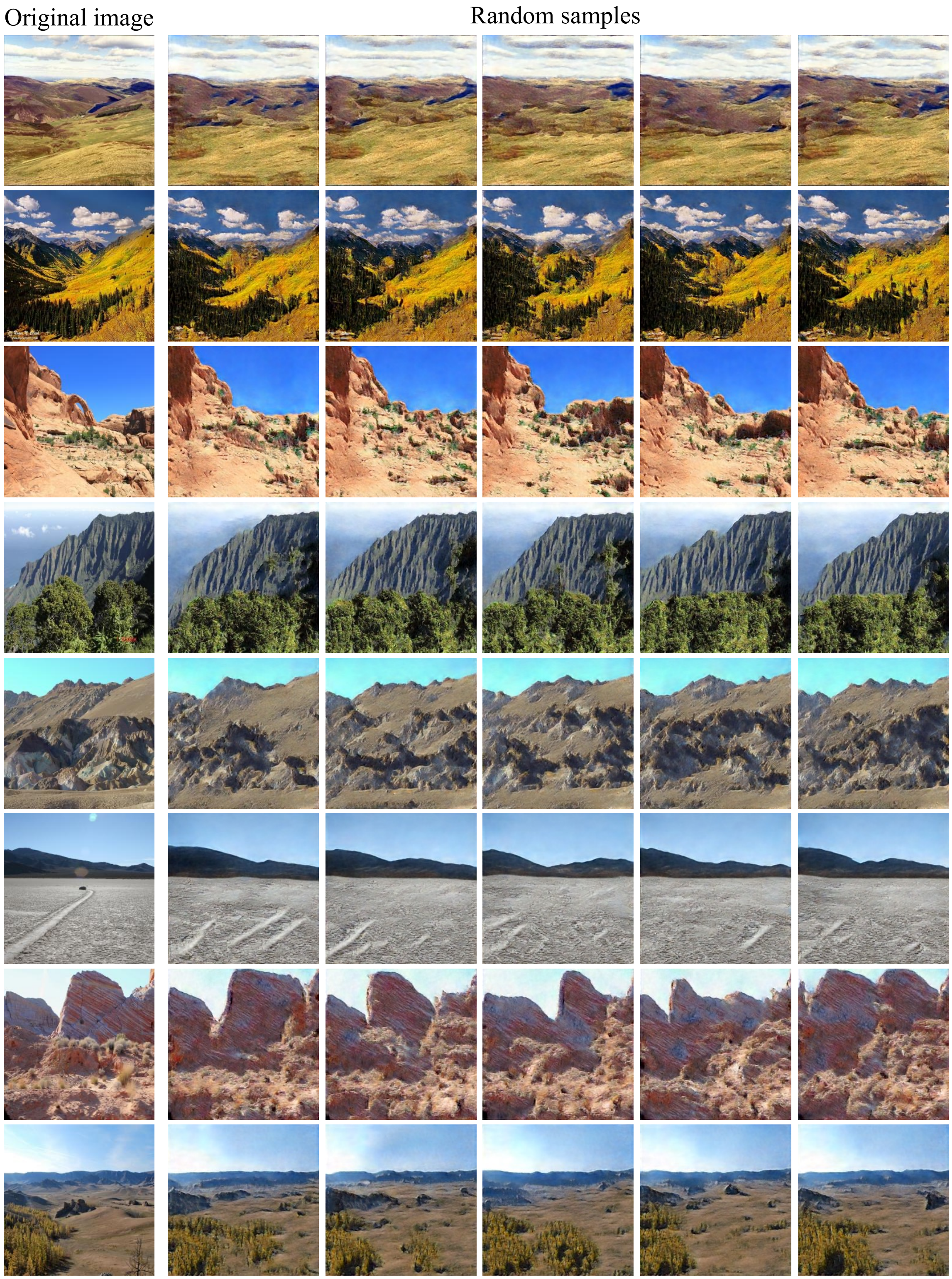
D.6 V500 - Original images and random samples



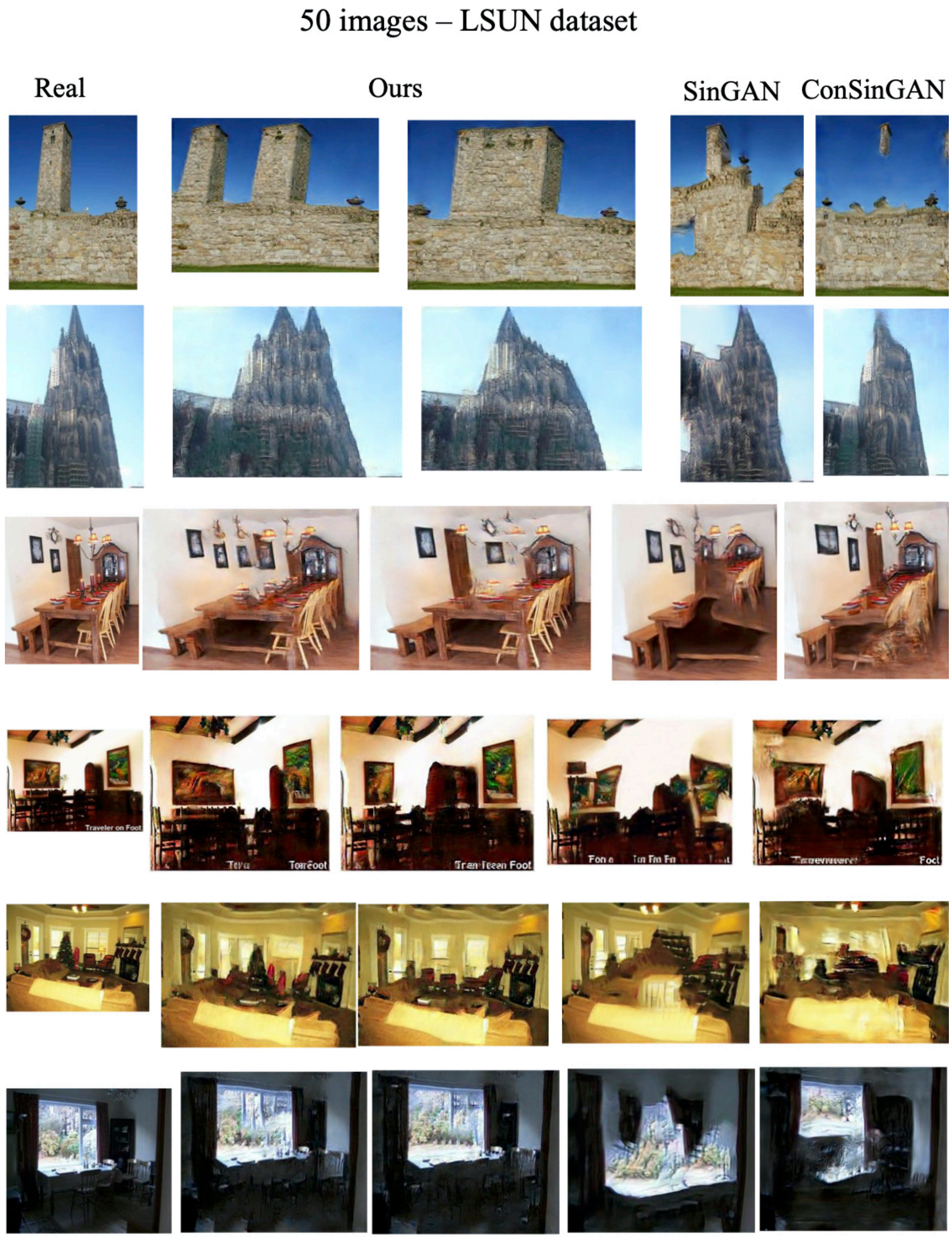
D.7 V2500 - Original images and random samples



D.8 V5000 - Original images and random samples



D.9 LSUN-50



50 images – LSUN dataset
Real
Ours
SinGAN
ConSinGAN
Traveler on Foot

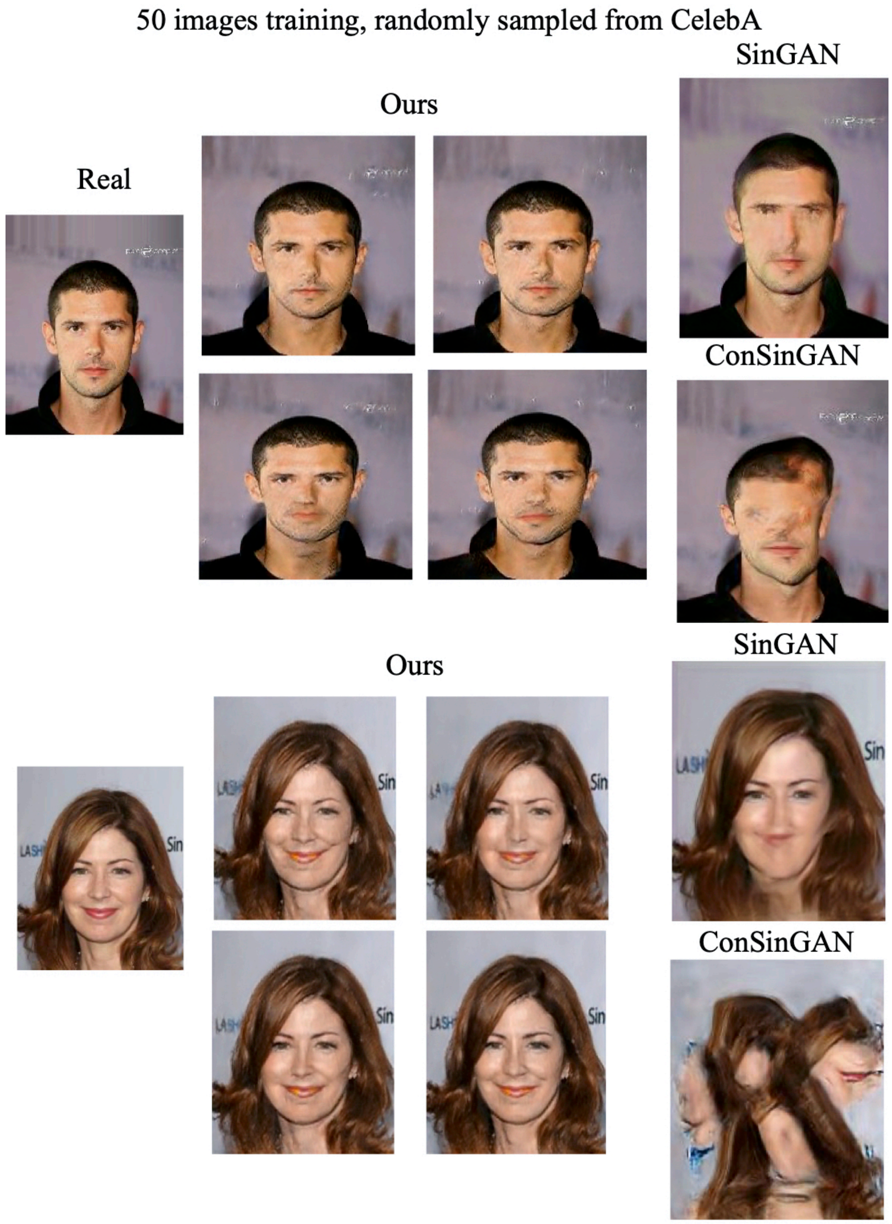
D.10 C250

250 images sampled from Churches Outdoor category - LSUN
SinGAN
Real
Ours
ConSinGAN

D.11 CelebA

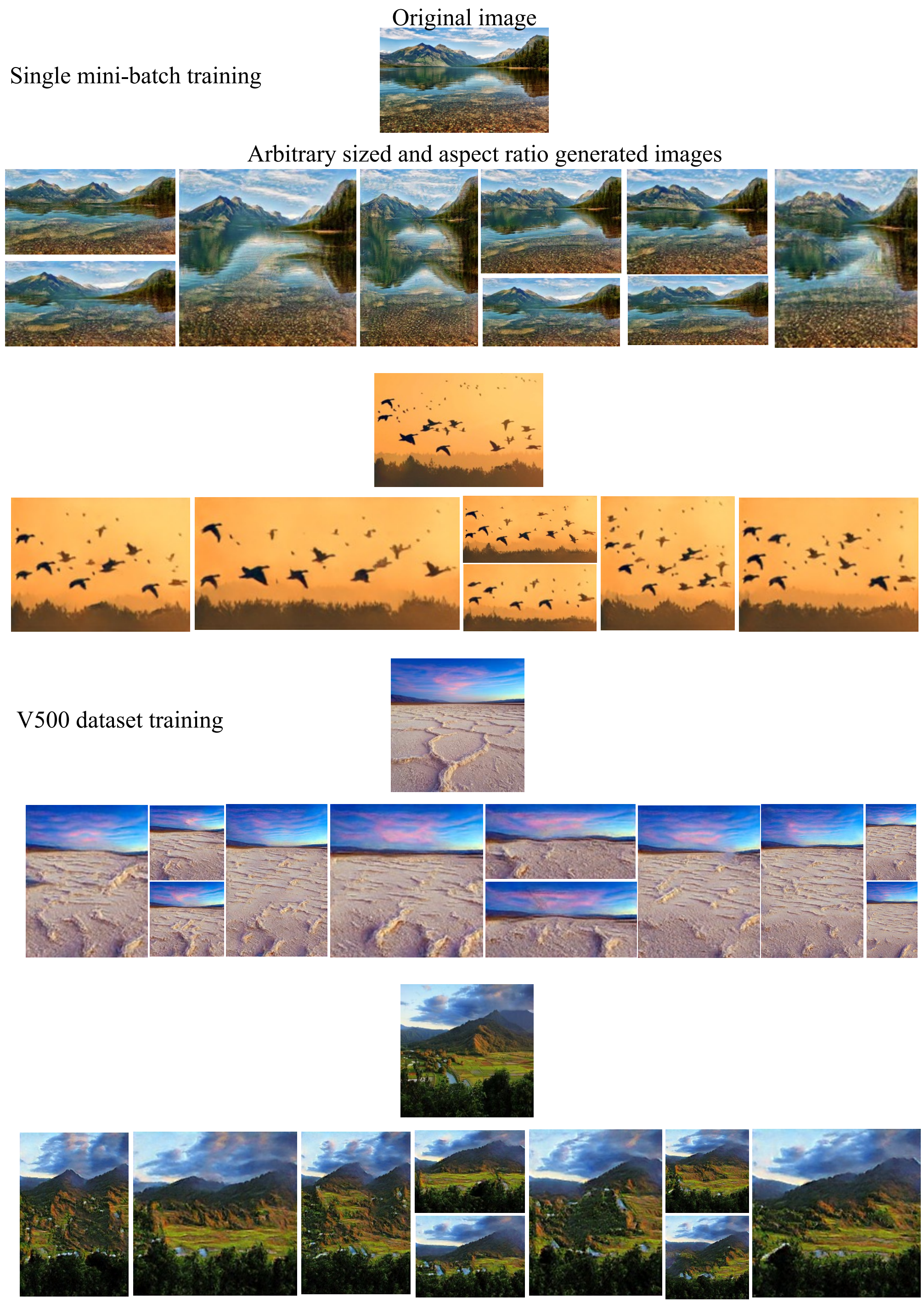
We have tested the method on 50 randomly sampled face images from the CelebA dataset. We attach side-by-side results along with the baselines, where we used the same initial noise size (of width 22) for all of the methods to allow for a fair comparison.

Even though our method generates more realistic images than baselines by a notable margin, our results are still non-comparable to classic face generation (by standard GAN and Flow-based models). Thus, we consider face datasets as a limitation of our approach.



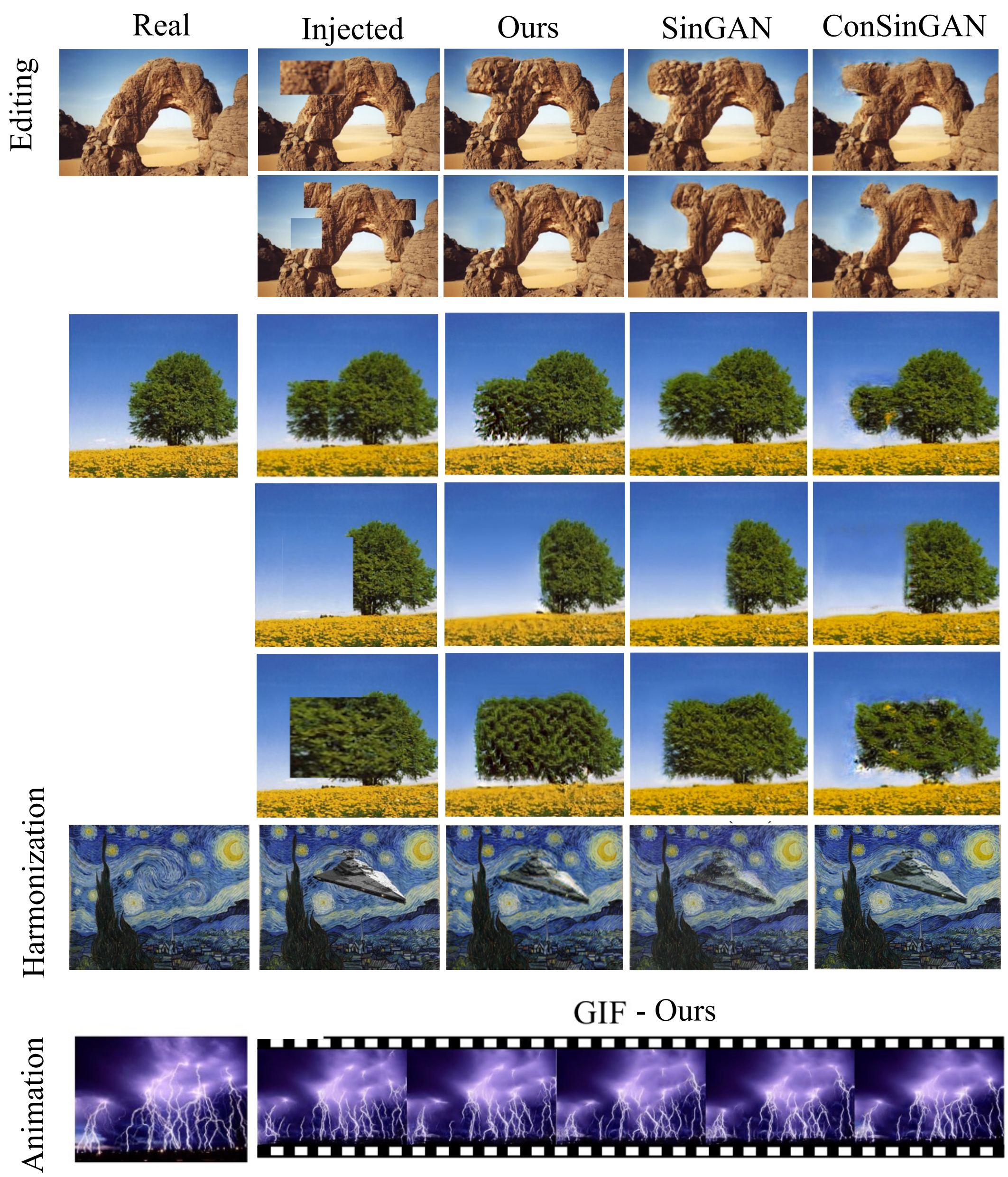
D.12 Arbitrary sized and aspect-ratio image generation

Due to the fully convolutional architecture adopted, all of our models are able to generate an image with an arbitrary size of aspect ratio by simply changing the dimensions of the noise maps used. Below are some examples obtained during single mini-batch training and V500 dataset training.



E Editing, Harmonization and Animation

Following results were obtained using a single model, trained on the 50 image dataset merged with these 4 images (a total of 54 images). The applications are performed in the same exact way [31] did.



F Interpolation

We conducted an experiment to study the smoothness of our interpolations at different scales. We estimated the slope of the generated images $H_i(\alpha)$, for a fixed set of random seeds, on a discrete set of values $\alpha \in \{0.1j\}_{j=1}^{9}$ as follows: $s_{i,j} := \frac{\|H_i(\alpha_{j+1}) - H_i(\alpha_j)\|_1}{h \times w \cdot (\alpha_{j+1} - \alpha_j)}$, where $h \times w$ is the size of the images. As can be seen in Fig. 8, the interpolations at higher scales tend to be significantly smoother than the interpolations at lower scales.

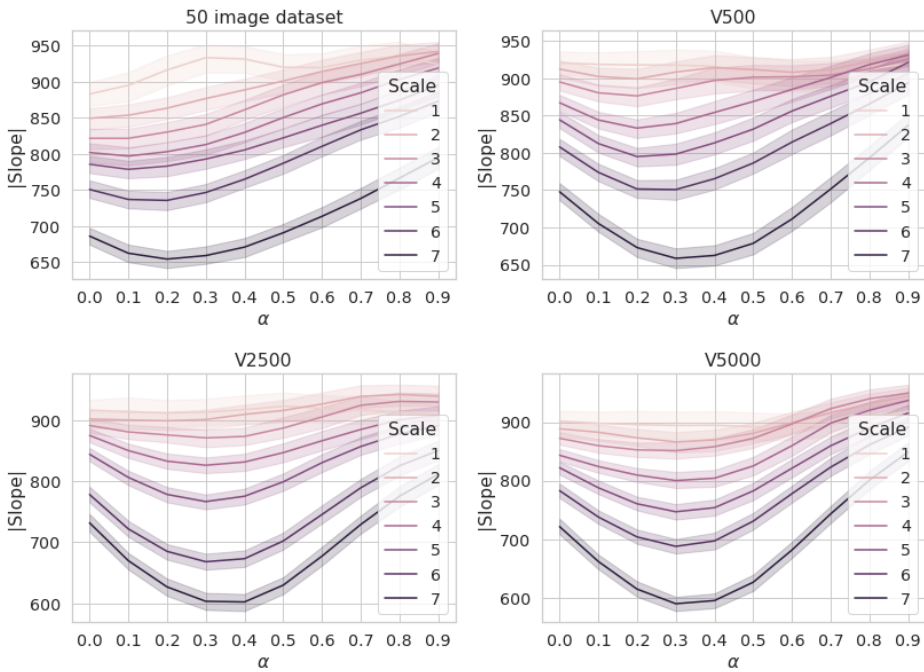
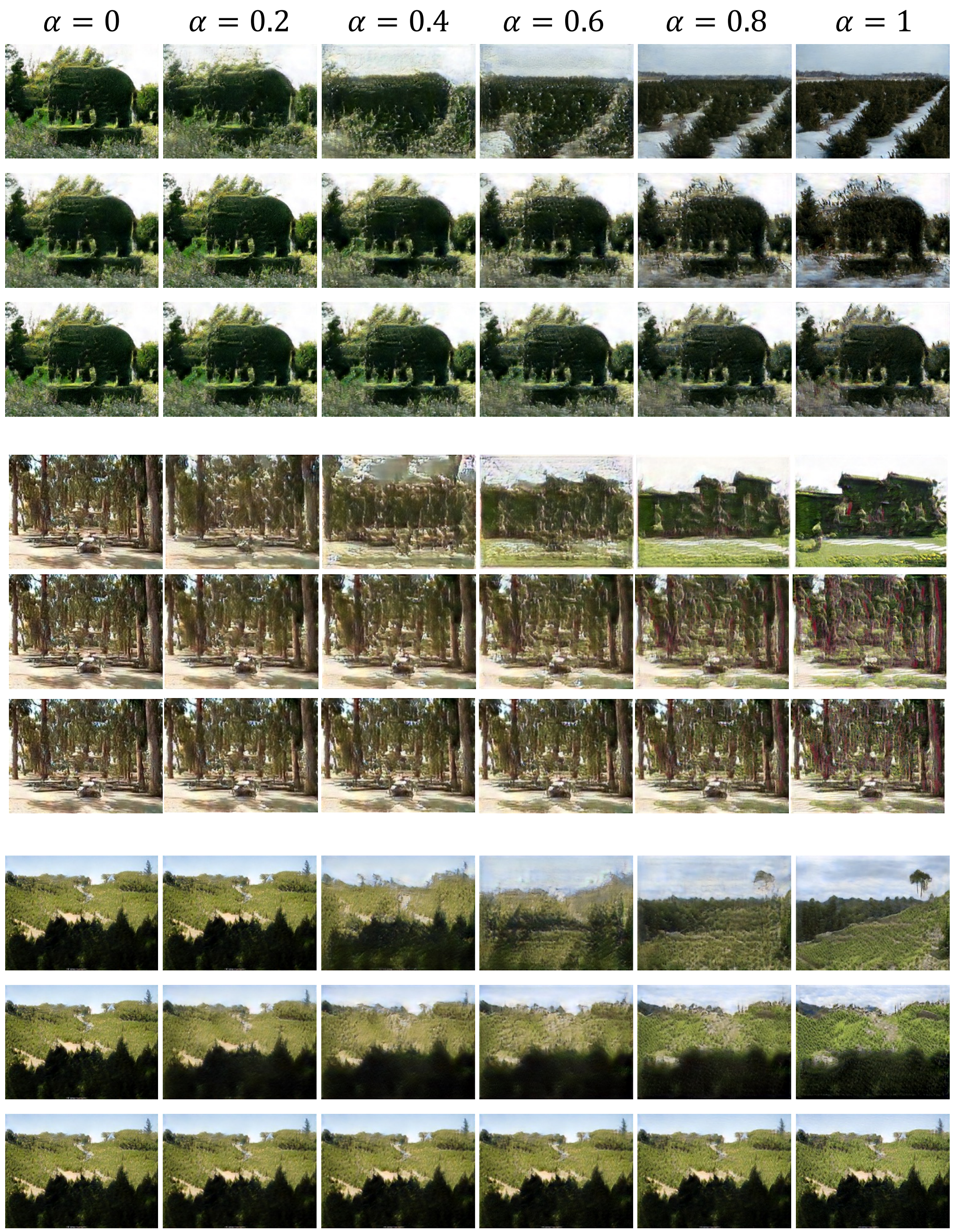
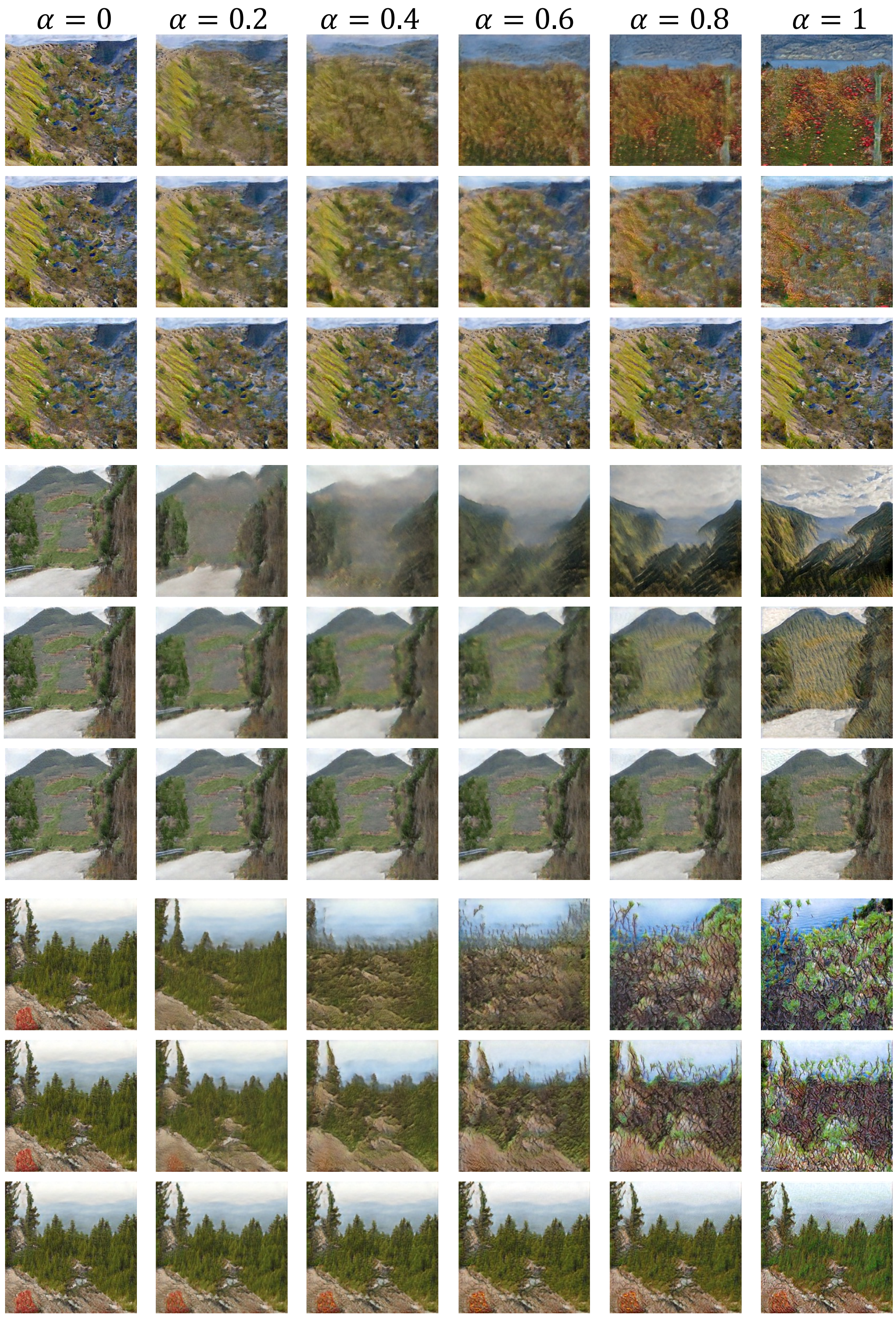


Figure 8: **Smoothness rate of the interpolations.** We plot the smoothness rate $s_{i,j}$ (y-axis) as a function of α (x-axis), averaged over 500 pairs of images A, B along with their standard deviations.

F.1 Places-50

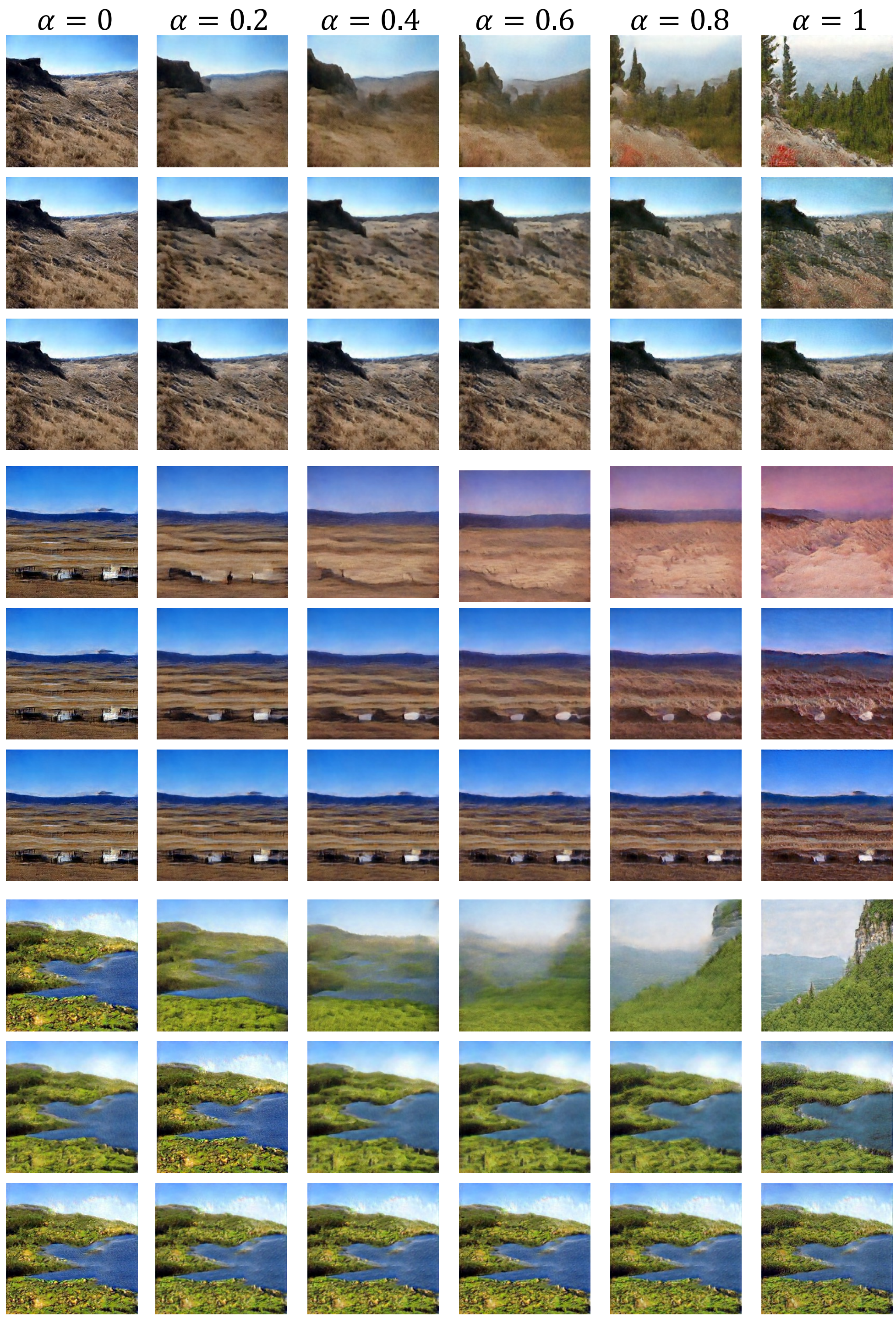


F.2 V500



F.3 V2500

F.4 V5000



G Feedforward modeling