BRAIN2GAN; RECONSTRUCTING PERCEIVED FACES FROM THE PRIMATE BRAIN VIA STYLEGAN3

Anonymous authors
Paper under double-blind review

A APPENDIX

A.1 RECONSTRUCTION VIA Z- VERSUS W-LATENT SPACE

Figure 1 depicts the stimuli and their reconstructions from z- and $w-$latent space. Perceptually, it is obvious that the results were significantly enhanced by decoding from w-latent space.

Figure 1: **Qualitative results from z- and w-latent space.** The 100 test set stimuli (top row) and their reconstructions from brain activity via the w-latent space (middle row) and z-latent space (bottom row).

A.2 Permutation test analysis

The quantitative results were verified with a permutation test. In Figure 2, we plotted the six similarity metrics over 200 iterations. We discovered that the random samples were never better than our predictions from brain activity.

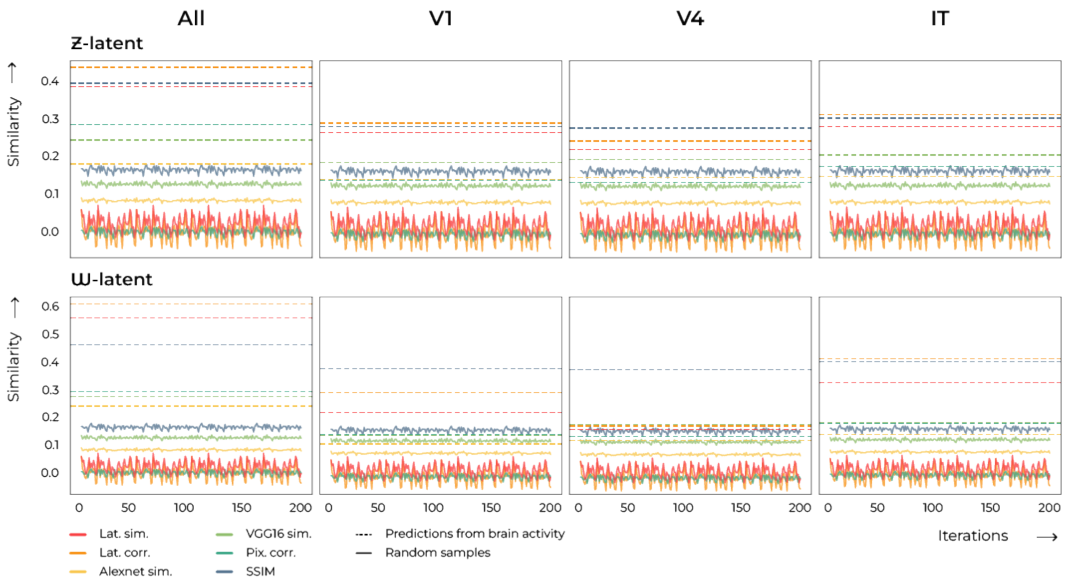

Figure 2: **Permutation test analysis.** In every iteration, one hundred latents/faces were randomly sampled to evaluate their similarity to the ground truth latents/faces in terms of the six similarity metrics of latent similarity, latent correlation, alexnet perceptual similarity, VGG16 perceptual similarity, pixel correlation and structural similarity index measure (SSIM). The eight plots denote similarity (Y-axis) over 200 iterations (X-axis) for the random samples as well as our predictions from brain activity (dashed lines) via z-space (top row) and w-space (bottom row). To speed up computation, the 1024^2 pixel images where resized to 224^2 pixels for pixel correlation and SSIM.

A.3 Second macaque, results

We repeated the experiment using MUA responses from V1, V2, V3 and V4 in a second macaque (male, 9 years old) with silicone-based electrodes (Figure 3, Table 1). Note that subject has no electrode arrays implanted in IT.

Table 1: **Quantitative results.** The t row decoded the z-latent whereas the lower row decoded the w-latent from brain activity in V1, V2, V3 and V4 in a second macaque. Decoding performance is quantified in terms of six metrics: latent cosine similarity, latent correlation (Student's t-test), perceptual cosine similarity using alexnet and VGGFace, pixel-wise correlation (Student's t-test) and structural similarity index (SSIM) in pixel space between stimuli and their reconstructions ($mean \pm std.error$).

	Lat. sim.	Lat. corr.	Alexnet sim.	VGG16 sim.	Pixel corr.	SSIM
z	0.3292 ± 0.0072	0.2065 ± 0.0054	0.2190 ± 0.0022	0.1568 ± 0.0012	0.4251 ± 0.0001	0.3633
w	0.4206 ± 0.0081	0.2204 ± 0.0051	0.2009 ± 0.0026	0.1754 ± 0.0015	0.5412 ± 0.0001	0.4750

A.4 Visual guide

Six metrics were used to evaluate decoding performance: latent similarity, latent correlation, alexnet perceptual similarity, VGG16 perceptual similarity and structural similarity index measure (SSIM). In Figure 5, we present the stimulus-reconstruction pairs with the highest and lowest similarity for each metric.

Figure 3: **Qualitative results:** The 100 test set stimuli (top row) and their reconstructions from brain activity in V1, V2, V3 and V4 in a second macaque (bottom row).

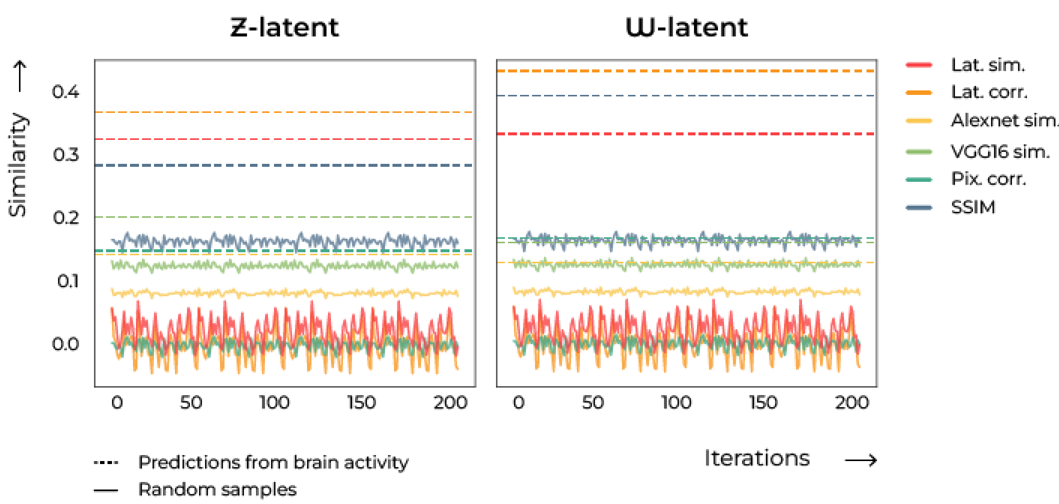

Figure 4: **Permutation test analysis.** In every iteration, one hundred latents/faces were randomly sampled to evaluate their similarity to the ground truth latents/faces in terms of the six similarity metrics of latent similarity, latent correlation, alexnet perceptual similarity, VGG16 perceptual similarity, pixel correlation and structural similarity index measure. The two plots denote similarity (Y-axis) over 200 iterations (X-axis) for the random samples as well as our predictions from brain activity (dashed lines) via z-space (left plot) and w-space (right plot). To speed up computation, the 1024^2 pixel images where resized to 224^2 pixels for pixel correlation and SSIM.

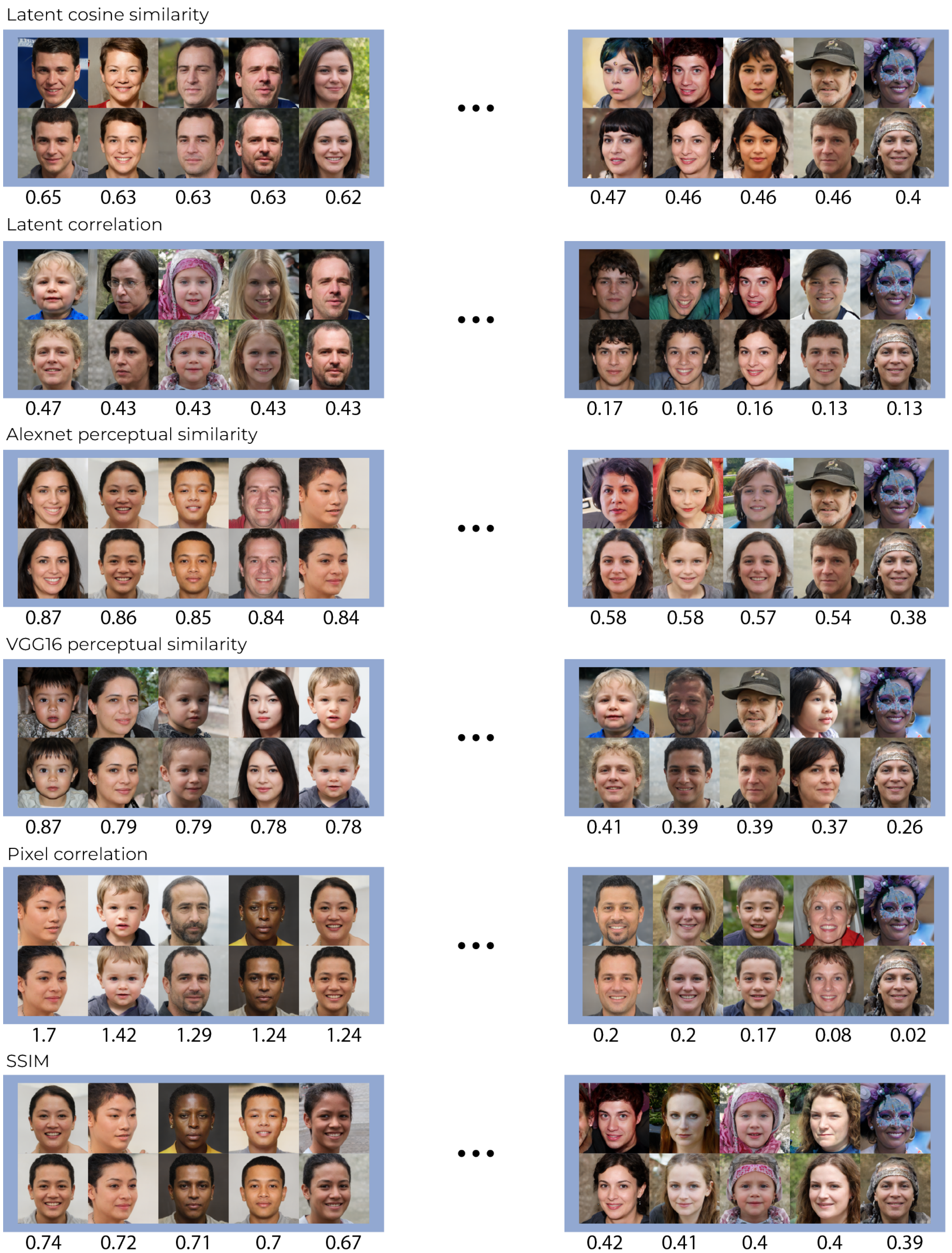

Figure 5: **Visual guide.** The five stimulus-reconstruction pairs with the highest similarity are displayed on the left whereas those with the lowest similarity are displayed on the right. The top row denotes the stimuli and the bottom row their reconstructions from brain activity.