# Learning Images Across Scales Using Adversarial Training – Supplemental Document

KRZYSZTOF WOLSKI, Max-Planck-Institut für Informatik, Germany
ADARSH DJEACOUMAR, Max-Planck-Institut für Informatik, Germany
ALIREZA JAVANMARDI, Max-Planck-Institut für Informatik, Germany
HANS-PETER SEIDEL, Max-Planck-Institut für Informatik, Germany
CHRISTIAN THEOBALT, Max-Planck-Institut für Informatik, Germany
GUILLAUME CORDONNIER, Inria, Université Côte d'Azur, France
KAROL MYSZKOWSKI, Max-Planck-Institut für Informatik, Germany
GEORGE DRETTAKIS, Inria, Université Côte d'Azur, France
XINGANG PAN, Nanyang Technological University, Singapore
THOMAS LEIMKÜHLER, Max-Planck-Institut für Informatik, Germany

**ACM Reference Format:**
Krzysztof Wolski, Adarsh Djeacoumar, Alireza Javanmardi, Hans-Peter Seidel, Christian Theobalt, Guillaume Cordonnier, Karol Myszkowski, George Drettakis, Xingang Pan, and Thomas Leimkühler. 2024. Learning Images Across Scales Using Adversarial Training – Supplemental Document. *ACM Trans. Graph.* 43, 4, Article 131 (July 2024), 3 pages. https://doi.org/10.1145/3658190

## 1 STITCHING WITH PHOTOSHOP

In Fig. 1, we show the best result we could obtain using automatic image stitching with Adobe Photoshop. For this to work, we had to restrict the data to 420 patches at $s = 2$, which constitutes only a tiny fraction of our data and is four scales coarser than what our approach can deliver.

## 2 BASELINES FOR GENERATION

### 2.1 AnyresGAN [Chai et al. 2022]

As this method, different from our setting, requires input images at different resolutions, we modify our dataset extraction procedure. Specifically, we synthesize significantly larger images patches, from which the original implementation can extract the patches needed for training. Notice how this provides the method with more context than our method requires.

Authors' addresses: Krzysztof Wolski, Max-Planck-Institut für Informatik, Saarbrücken, Germany, kwolski@mpi-inf.mpg.de; Adarsh Djeacoumar, Max-Planck-Institut für Informatik, Saarbrücken, Germany, adjeacou@mpi-inf.mpg.de; Alireza Javanmardi, Max-Planck-Institut für Informatik, Saarbrücken, Germany, alireza.javanmardi@dfki.de; Hans-Peter Seidel, Max-Planck-Institut für Informatik, Saarbrücken, Germany, hpseidel@mpi-sb.mpg.de; Christian Theobalt, Max-Planck-Institut für Informatik, Saarbrücken, Germany, theobalt@mpi-inf.mpg.de; Guillaume Cordonnier, Inria, Université Côte d'Azur, Sophia-Antipolis, France, guillaume.cordonnier@inria.fr; Karol Myszkowski, Max-Planck-Institut für Informatik, Saarbrücken, Germany, karol@mpi-inf.mpg.de; George Drettakis, Inria, Université Côte d'Azur, Sophia-Antipolis, France, george.drettakis@inria.fr; Xingang Pan, Nanyang Technological University, Singapore, Singapore, xingang.pan@ntu.edu.sg; Thomas Leimkühler, Max-Planck-Institut für Informatik, Saarbrücken, Germany, thomas.leimkuehler@mpi-inf.mpg.de.



### 2.2 PULSE [Menon et al. 2020]

We first train a vanilla StyleGAN3 [Karras et al. 2021] model on all data patches. To synthesize a zooming-in sequence, we first generate a random sample from this model. We then perform GAN inversion to create a new sample that corresponds to a 2x zoomed-in image using the PULSE [Menon et al. 2020] framework on the central crop of the original sample. We repeat this process, until the required scale range is obtained. Continuous zooming is achieved using scaling and linear blending of the obtained sample sequence.

## 3 DATASET STATISTICS

In Tab. 1 we report detailed statistics on our datasets. We list the number of patches per scale interval, as well as the resulting density of patches. Density is reported as the expected number of patches that overlap a single pixel from an image that has the highest possible resolution in the scale interval.

## REFERENCES

Lucy Chai, Michaël Gharbi, Eli Shechtman, Phillip Isola, and Richard Zhang. 2022. Any-Resolution Training for High-Resolution Image Synthesis. In *ECCV*. 170–188.
Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2021. Alias-free generative adversarial networks. *NeurIPS* 34 (2021), 852–863.
Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, and Cynthia Rudin. 2020. PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models. In *CVPR*. 2437–2445.

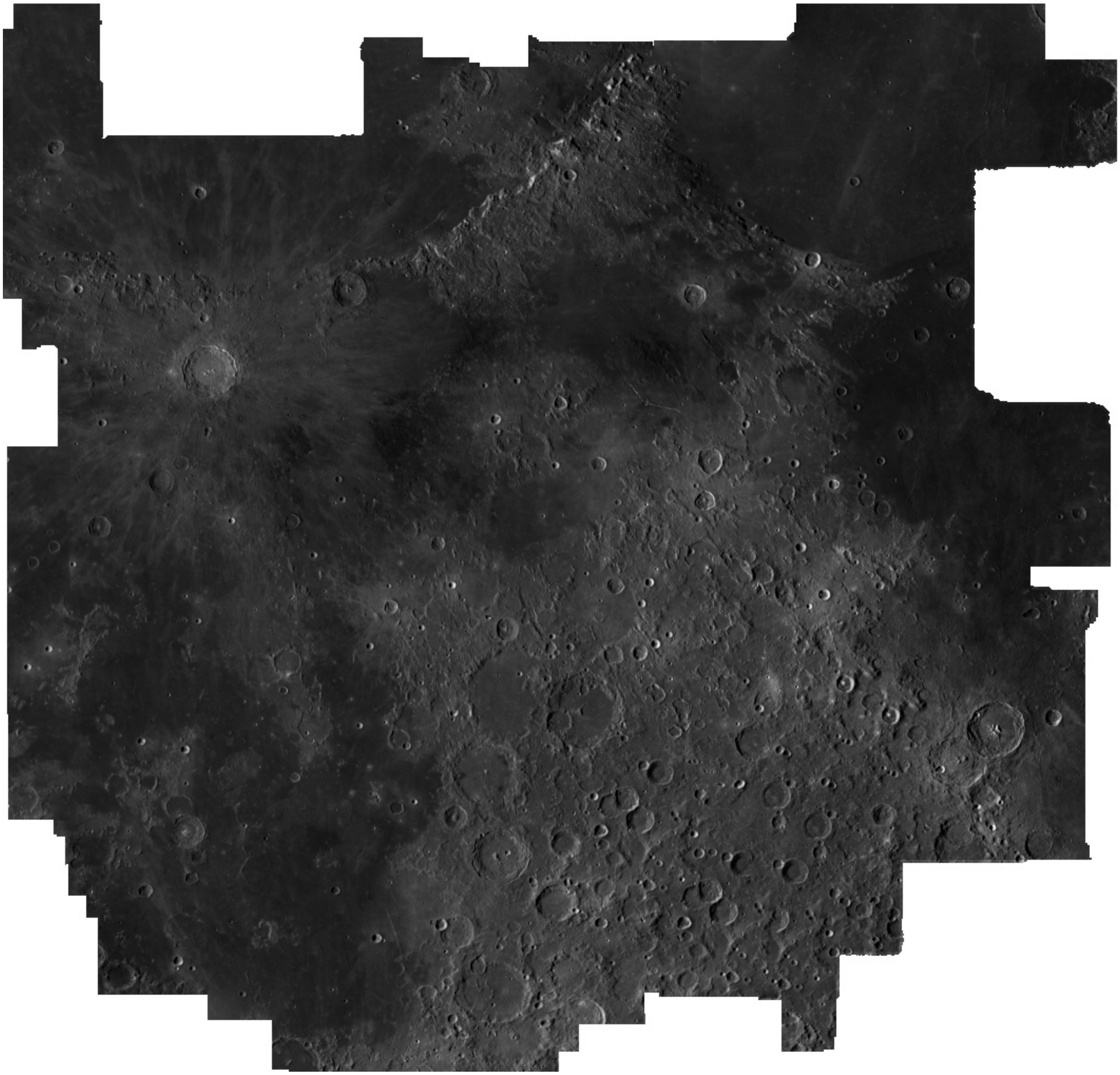

Fig. 1. Image stitching with Photoshop.

Table 1. Dataset details. For each scale interval, we list the number of patches (#Patch.) and the number of images per pixel (Dens.).

| Dataset | Scale 0 | | Scale 1 | | Scale 2 | | Scale 3 | | Scale 4 | | Scale 5 | | Scale 6 | | Scale 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. | #Patch. | Dens. |
| Milkyway | 0.5k | 276.23 | 1.5k | 204.62 | 4.k | 136.38 | 30k | 254.30 | 30k | 63.36 | 30k | 15.86 | - | - | - | - |
| Moon | 0.5k | 276.23 | 1.5k | 204.62 | 4.k | 136.38 | 30k | 254.30 | 30k | 63.36 | 30k | 15.86 | - | - | - | - |
| Himalayas | 0.5k | 276.23 | 1.5k | 204.62 | 4.k | 136.38 | 30k | 254.30 | 30k | 63.36 | 30k | 15.86 | 30k | 3.96 | 30k | 0.99 |
| Spain | 0.5k | 276.23 | 1.5k | 204.62 | 4.k | 136.38 | 30k | 254.30 | 30k | 63.36 | 30k | 15.86 | 30k | 3.96 | 30k | 0.99 |
| Rembrandt | 0.5k | 276.23 | 1.5k | 204.62 | 4.k | 136.38 | 30k | 254.30 | 30k | 63.36 | 30k | 15.86 | 30k | 3.96 | 30k | 0.99 |
| MilkywayGen | 30k | - | 30k | - | 30k | - | 30k | - | - | - | - | - | - | - | - | - |
| MoonGen | 30k | - | 30k | - | 30k | - | 30k | - | - | - | - | - | - | - | - | - |
| HimalayasGen | 30k | - | 30k | - | 30k | - | 30k | - | - | - | - | - | - | - | - | - |
| SpainGen | 30k | - | 30k | - | 30k | - | 30k | - | - | - | - | - | - | - | - | - |
| RembrandtGen | 30k | - | 30k | - | 30k | - | 30k | - | - | - | - | - | - | - | - | - |
| Sunflowers | 30k | - | 60k | - | 60k | - | 35k | - | - | - | - | - | - | - | - | - |
| Bricks | 25k | - | 67k | - | 82k | - | 60k | - | - | - | - | - | - | - | - | - |
| Milkyway 12k | 62 | 33.59 | 187 | 25.28 | 500 | 16.91 | 3750 | 31.70 | 3750 | 7.92 | 3750 | 1.98 | - | - | - | - |
| Milkyway 3k | 15 | 8.14 | 46 | 6.22 | 125 | 4.23 | 937 | 7.92 | 937 | 1.98 | 937 | 0.50 | - | - | - | - |
| Milkyway 1k | 5 | 2.70 | 15 | 2.02 | 41 | 1.39 | 312 | 2.64 | 312 | 0.66 | 312 | 0.16 | - | - | - | - |
| Milkyway 0.5k | 2 | 1.08 | 7 | 0.95 | 20 | 0.68 | 156 | 1.32 | 156 | 0.33 | 156 | 0.08 | - | - | - | - |
| Milkyway 0.25k | 3 | 1.62 | 3 | 0.41 | 10 | 0.34 | 78 | 0.66 | 78 | 0.16 | 78 | 0.04 | - | - | - | - |