

Cycle-consistent Generative Adversarial Networks for Neural Style Transfer using data from Chang'E-4

J. de Curtó and R. Duvall.
Iris Lunar Rover. Carnegie Mellon.

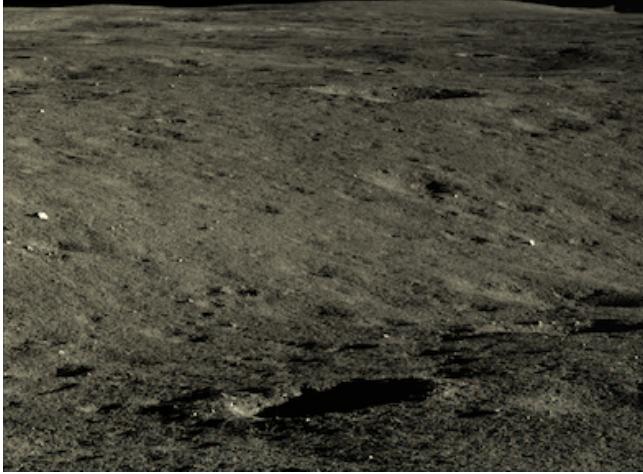


Figure 1. Images from the Moon. Panoramic camera of the rover. Chang'E-4.

Generative Adversarial Networks (GANs) [Goodfellow et al. 2014; Radford et al. 2016] have had tremendous applications in Computer Vision. Yet, in the context of space science and planetary exploration the door is open for major advances. We introduce tools to handle planetary data from the mission Chang'E-4¹ and present a framework for Neural Style Transfer using Cycle-consistency [Zhu et al. 2017] from rendered images. The experiments are conducted in the context of the Iris Lunar Rover, a nano-rover that will be deployed in lunar terrain in 2021 as the flagship of Carnegie Mellon, being the first unmanned rover of America to be on the Moon.

CCS Concepts: • Robotics → Perception.

Additional Key Words and Phrases: GAN, Perception, Deep Learning, Computer Vision, Robotics, Lunar Rover

1 Introduction

Generative Adversarial Networks (GAN) [Goodfellow et al. 2014] are able to produce good quality high-resolution samples from images, both in the unconstrained and conditional setting [Bousmalis et al. 2017; Chen et al. 2019; Li et al. 2017; Lombardi et al. 2018; Portenier et al. 2018; Romero et al. 2018; Sankaranarayanan et al. 2018; Wang et al. 2018a,b; Wu et al. 2016; Yang et al. 2017; Yu et al. 2018; Zhu et al. 2017]. Nonetheless, applications in the context of NASA missions and space exploration are scarce.

Given the difficulty to handle planetary data we provide downloadable files in PNG format from the missions Chang'E-3 and

Chang'E-4². In addition to a set of scripts to do the conversion given a different PDS4 Dataset. Example samples from the dataset can be seen in Figure 1. We also provide the corresponding labels, where localization information is present. We run extensive experiments to train a model able to be used as a hyperrealistic feature of the current simulator used in the Iris Lunar Rover.

2 Overall System

Following the design principles and the perception pipeline proposed in [Allan et al. 2019] in the context of the NASA Mission Resource Prospector, we intend to design a simulator with hyperrealistic characteristics of the Moon that helps us deploy VIO/SLAM in a rover of the same characteristics. The intention is also that helps us address object detection and segmentation in this unmapped environment, where training data is very difficult and costly to obtain. Although at the present time data from the Moon is scarce, there are already some open datasets available in analogue environments such as the POLAR Stereo Dataset [Wong et al. 2017] that includes stereo pairs and LiDAR information or [Vayugundla et al. 2018], that contains IMU, stereo pairs and odometry plus some additional localization data, all obtained on Mount Etna. Our intention is to provide downloadable files from the mission Chang'E-4 [Zhang et al. 2019] that could be easily used in CV and ML pipelines. We also provide scripts to handle alternate PDS4 Datasets. The context where these tools are being used is our specific sensor suite, that will be on-board the Iris Lunar Rover, a project led by Carnegie Mellon that intends to put forward a four pound rover into the

¹Data and more information will be available at github.com/decurtoidiaz/ce4/

decurtoidiaz@ieee.org, rduvall@andrew.cmu.edu.

²Original PDS4 and PDS3 images and labels from missions to the Moon Chang'E can be obtained at moon.bao.ac.cn.

surface of the Moon by 2021 and that will be America’s first rover to explore the surface of the planet, consists on IMU, two high-fidelity cameras and odometry sensors. Furthermore, it also has a UWB module [Alarifi et al. 2016; Ledergerber et al. 2015; Mueller et al. 2015; Xu et al. 2020] on-board to localize the rover with respect to the lander.

3 Approach, Long-term Goal and Prior Work

Generative image generation is a key problem in Computer Vision and Computer Graphics. Variational Autoencoders (VAE) [Kingma and Welling 2014; Lombardi et al. 2018] try to solve the problem with an approach that builds on probabilistic graphical models. Autoregressive models (for instance PIXELRNN [van den Oord et al. 2016]) have also achieved relative success generating synthetic images. In the past few years, Generative Adversarial Networks (GANs) [Antoniou et al. 2018; Goodfellow et al. 2014; Odena et al. 2017; Portenier et al. 2018; Radford et al. 2016; Wang and Gupta 2016; Zhu et al. 2016] have shown strong performance in image generation. Some works on the topic pinpoint the specific problem of scaling up to high-resolution samples [Zhang et al. 2017], where conditional image generation is also studied while some recent techniques focus on stabilizing the training procedure [Brock et al. 2019; Chen and Koltun 2017; Dosovitskiy and Brox 2016; Karras et al. 2018; Mescheder et al. 2017, 2018; Salimans et al. 2016; Wei et al. 2018; Zhao et al. 2017]. Other promising novel approaches include score matching with Langevin sampling [Song and Ermon 2019, 2020] and the use of sequence transformers for image generation [Parmar et al. 2018].

The use of these techniques though has seen little or no applications in space exploration and planetary research. We propose here a framework that could be used to generate abundant data of the Moon, Mars and other celestial bodies, so that learning algorithms could be trained on Earth and studied in simulation before being deployed in the real missions.

The proposed approach consists on using a technique of Neural Style Transfer or Generative Image Generation, such as the criteria of cycle-consistency, together with an augmentation of the given dataset (in our case using data from the lunar missions Chang’E-3 and Chang’E-4, but the same applies to Mars or other planets) using GANs in the setting of unconstrained image generation.

4 Cycle-consistent Generative Adversarial Networks

Our focus here is on Cycle-consistent Generative Adversarial Networks [Zhu et al. 2017], where we work on unpaired image-to-image translation [Park et al. 2020].

Image-to-image translation is a type of problem in Computer Vision and Computer Graphics where the objective is to learn a correspondence function between an input sample and an output sample, using a training set of aligned or non-aligned image pairs.

More precisely, our goal is to learn a function

$$G : X \rightarrow Y, \quad (1)$$

in a way that the distribution of samples $G(X)$ is as close as possible to the distribution Y . To accomplish this we are going to use an adversarial loss. Therefore, we couple it with the inverse correspondence

$$F : Y \rightarrow X, \quad (2)$$

and use a criteria of cycle-consistency to address the fact that the problem is highly under constrained

$$F(G(X)) \approx X \quad \text{and} \quad G(F(Y)) \approx Y. \quad (3)$$

When we talk about paired training data, we refer to the fact that the training data consists of training examples $\{x_c, y_c\}_{c=1}^N$, where the correspondences between x_c and y_c are given. Instead, we say that we are using unpaired training data, when the set consists of two training sets $\{x_c\}_{c=1}^N$ and $\{y_a\}_{a=1}^N$, where there is not explicitly given a correspondence between which x_c corresponds to which y_a .

Formally, the GAN objective [Goodfellow et al. 2014] involves finding a NASH equilibrium to the following two-player game:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}} [\log D(x)] + \quad (4)$$

$$+ \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] , \quad (5)$$

where x is a ground truth image sampled from the true distribution p_{data} , and z is a noise vector sampled from p_z (that is, uniform or normal distribution). G and D are parametric functions where $G : p_z \rightarrow p_{data}$ maps samples from noise distribution p_z to data distribution p_{data} .

5 Neural Style Transfer

Neural Style Transfer [Gatys et al. 2016a,b; Ulyanov et al. 2016] is based on the idea of synthesizing an original image by combining the content of one image together with the style of another sample. Here we will use cycle-consistent networks to attack this specific problem, with the aim of using a more general method that could help us solve concomitantly other tasks in the future. Moreover, the criteria of cycle-consistency assumes there is a bijection between the two domains, a constrain that could be often too restrictive, but that is very appropriate in our particular problem at-hand.

6 Unconstrained Image Generation

To tackle the problems that arise when training cycle-consistent networks with a dataset with few samples, i.e. mainly mode collapse and artifacts, we propose to use Unconstrained Image Generation using GANs to enlarge the original dataset with unseen examples, that is, as a way to generate additional training samples that will help the learning procedure converge to the desired solution. To achieve this we make use of the construction developed in [Brock et al. 2019].

7 Experiments

Extensive experiments using data from Chang’E-3 and Chang’E-4 have been conducted, in particular we are using images from the

panoramic camera of the rover and from the terrain camera of the lander. Some examples can be seen in Figures 2 and 3, model trained at image size 256×256 and 512×512 , respectively. As a source domain we are using samples from a rendered simulator of the Moon provided by Kaggle. The intention is to use the model in our actual renderer environment of the mission.

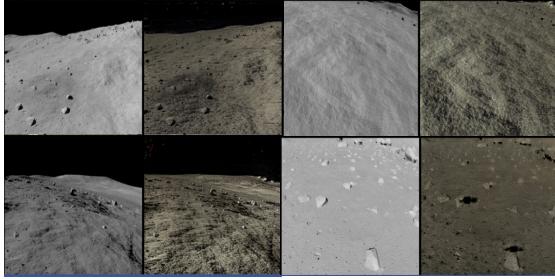


Figure 2. **Cycle-consistent gan.** Left: images from Kaggle, rendered simulator of the Moon. Right: style-Moon using our model. Trained at image size 256×256 .

8 Simulator

The context where this feature is being integrated is the actual simulator, see Figure 4, of the Iris Lunar Rover, the rover of Carnegie Mellon that will fly to the Moon onboard the Peregrine Lander of Astrobotic in 2021. Data from the simulator will be of the utmost importance to train and test localization algorithms such as SLAM/VIO [Schneider et al. 2018; Usenko et al. 2019]. The ability to have ample data to train will also amplify the capabilities of the modules designed for segmentation [Badrinarayanan et al. 2015; Chen et al. 2017; Chen and Koltun 2017; He et al. 2017] and object detection [Bolme et al. 2010; Girshick 2015; Redmon et al. 2016]. As well as to test the software design before the real mission.

References

- A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafeesah, S. Al-Hadhrami, M. A. Al-Ammar, and H. S. Al-Khalifa. 2016. Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances. *Sensors* (2016).
- M. Allan, U. Wong, P. M. Furlong, A. Rogg, S. McMichael, T. Welsh, I. Chen, S. Peters, B. Gerkey, M. Quigley, M. Shirley, M. Deans, H. Cannon, and T. Fong. 2019. Planetary Rover Simulation for Lunar Exploration Missions. *IEEE Aerospace Conference* (2019).
- A. Antoniou, A. Storkey, and H. Edwards. 2018. Data Augmentation Generative Adversarial Networks. *ICLR* (2018).
- V. Badrinarayanan, A. Kendall, and R. Cipolla. 2015. SegNet: a Deep Convolutional Encoder-decoder Architecture for Image Segmentation. *IEEE International Conference on Computer Vision* (2015).
- D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui. 2010. Visual Object Tracking Using Adaptive Correlation Filters. *IEEE International Conference on Computer Vision* (2010).
- K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan. 2017. Unsupervised Pixel-level Domain Adaptation with Generative Adversarial Networks. *IEEE International Conference on Computer Vision* (2017).
- A. Brock, J. Donahue, and K. Simonyan. 2019. Large Scale GAN Training for High Fidelity Natural Image Synthesis. *ICLR* (2019).
- L. Chen, G. Papandreou, F. Schroff, and H. Adam. 2017. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv:1706.05587* (2017).
- Q. Chen and V. Koltun. 2017. Photographic Image Synthesis with Cascaded Refinement Networks. *IEEE International Conference on Computer Vision* (2017).
- Y. Chen, W. Li, X. Chen, and L. Gool. 2019. Learning Semantic Segmentation from Synthetic Data: a Geometrically Guided Input-output Adaptation Approach. *IEEE International Conference on Computer Vision* (2019).
- A. Dosovitskiy and T. Brox. 2016. Generating Images with Perceptual Similarity Metrics Based on Deep Networks. *NIPS* (2016).
- L. A. Gatys, M. Bethge, A. Hertzmann, and E. Shechtman. 2016a. Preserving Color in Neural Artistic Style Transfer. *arXiv:1606.05897* (2016).
- L. A. Gatys, A. S. Ecker, and M. Bethge. 2016b. Image Style Transfer Using Convolutional Neural Networks. *IEEE International Conference on Computer Vision* (2016).
- R. Girshick. 2015. FAST R-CNN. *IEEE International Conference on Computer Vision* (2015).
- I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. 2014. Generative Adversarial Networks. *NIPS* 27 (2014).
- K. He, G. Gkioxari, P. Dollár, and R. Girshick. 2017. MASK R-CNN. *IEEE International Conference on Computer Vision* (2017).
- T. Karras, T. Aila, S. Laine, and J. Lehtinen. 2018. Progressive Growing of GANs for Improved Quality, Stability, and Variation. *ICLR* (2018).
- D. P. Kingma and M. Welling. 2014. Auto-encoding Variational Bayes. *ICLR* (2014).
- A. Ledgerber, M. Hamer, and R. D'Andrea. 2015. A Robot Self-localization System Using One-way Ultra-wideband Communication. *IEEE International Conference on Intelligent Robots and Systems (IROS)* (2015).
- C. Li, K. Xu, J. Zhu, and B. Zhang. 2017. Triple Generative Adversarial Nets. *NIPS* 30 (2017).
- S. Lombardi, J. Saragih, T. Simon, and Y. Sheikh. 2018. Deep Appearance Models for Face Rendering. *ACM Transactions on Graphics (SIGGRAPH)* 37 (2018).
- L. Mescheder, S. Nowozin, and A. Geiger. 2017. The Numerics of GANs. *NIPS* (2017).
- L. Mescheder, S. Nowozin, and A. Geiger. 2018. Which Training Methods for GANs Do Actually Converge? *ICML* (2018).
- M. W. Mueller, M. Hamer, and R. D'Andrea. 2015. Fusing Ultra-wideband Range Measurements with Accelerometers and Rate Gyroscopes for Quadcopter State Estimation. *ICRA* (2015).
- A. Odena, C. Olah, and J. Shlens. 2017. Conditional Image Synthesis with Auxiliary Classifier GANs. *ICML* (2017).
- T. Park, A. A. Efros, R. Zhang, and J. Zhu. 2020. Contrastive Learning for Unpaired Image-to-Image Translation. *EUROPEAN Conference on Computer Vision* (2020).
- N. Parmar, A. Vaswani, J. Uszkoreit, L. Kaiser, N. Shazeer, A. Ku, and D. Tran. 2018. Image Transformer. *ICML* (2018).
- T. Portenier, Q. Hu, A. Szabó, S. A. Bigdeli, P. Favaro, and M. Zwicker. 2018. FaceShop: Deep Sketch-based Face Image Editing. *ACM Transactions on Graphics (SIGGRAPH)* 37 (2018).
- A. Radford, M. Metz, and S. Chintala. 2016. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *ICLR* (2016).
- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. 2016. You Only Look Once: Unified, Real-time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition* (2016).
- A. Romero, P. Arbeláez, L. Gool, and R. Timofte. 2018. SMIT: Stochastic Multi-label Image-to-image Translation. *arXiv:1812.03704* (2018).
- T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. 2016. Improved Techniques for Training GANs. *NIPS* (2016).
- S. Sankaranarayanan, Y. Balaji, A. Jain, S. N. Lim, and R. Chellappa. 2018. Learning from Synthetic Data: Addressing Domain Shift for Semantic Segmentation. *IEEE International Conference on Computer Vision* (2018).
- T. Schneider, M. Dymczyk, M. Fehr, K. Egger, S. Lynen, I. Gilitschenski, and R. Siegwart. 2018. Maplab: an Open Framework for Research in Visual-inertial Mapping and Localization. *IEEE Robotics and Automation Letters* (2018).
- Y. Song and S. Ermon. 2019. Generative Modeling by Estimating Gradients of the Data Distribution. *NIPS* (2019).
- Y. Song and S. Ermon. 2020. Improved Techniques for Training Score-based Generative Models. *NIPS* (2020).
- D. Ulyanov, V. Lebedev, A. Vedaldi, and V. Lempitsky. 2016. Texture Networks: Feed-forward Synthesis of Textures and Stylized Images. *ICML* (2016).
- V. Usenko, N. Demmel, D. Schubert, J. Stückler, and D. Cremers. 2019. Visual-inertial Mapping with Non-linear Factor Recovery. *IEEE Robotics and Automation Letters* (2019).
- A. van den Oord, N. Kalchbrenner, and K. Kavukcuoglu. 2016. Pixel Recurrent Neural Networks. *ICML* (2016).
- M. Vayugundla, F. Steidle, M. Smisek, M. J. Schuster, K. Bussmann, and A. Wedler. 2018. Datasets of Long Range Navigation Experiments in a Moon Analogue Environment on Mount Etna. *International Symposium on Robotics* (2018).
- T. Wang, M. Liu, J. Zhu, G. Liu, A. Tao, J. Kautz, and B. Catanzaro. 2018a. Video-to-video Synthesis. *NIPS* (2018).
- T. Wang, M. Liu, J. Zhu, A. Tao, J. Kautz, and B. Catanzaro. 2018b. High-resolution Image Synthesis and Semantic Manipulation with Conditional GANs. *IEEE International Conference on Computer Vision* (2018).
- X. Wang and A. Gupta. 2016. Generative Image Modeling Using Style and Structure Adversarial Networks. *EUROPEAN Conference on Computer Vision* (2016).
- X. Wei, B. Gong, Z. Liu, W. Lu, and L. Wang. 2018. Improving the Improved Training of WASSERSTEIN GANs: a Consistency Term and its Dual Effect. *ICLR* (2018).

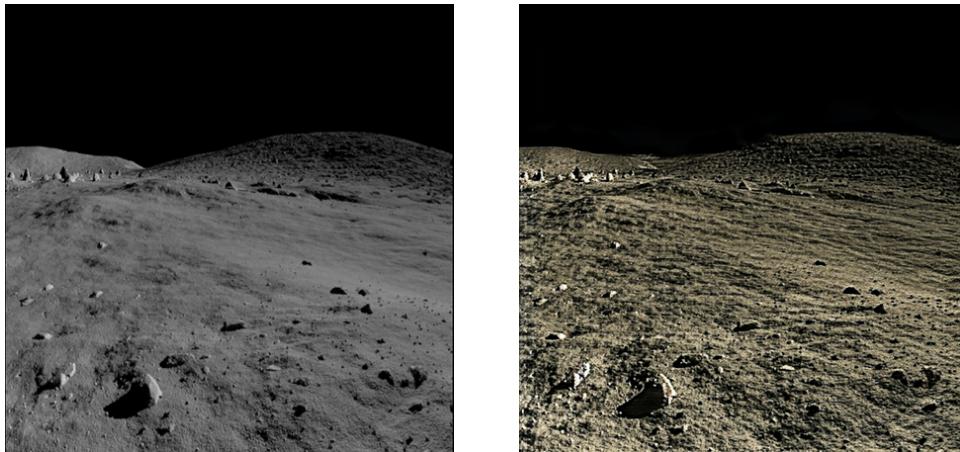


Figure 3. **Cycle-consistent gan.** Left: image from Kaggle, rendered simulator of the Moon. Right: style-Moon using our model. Trained at image size 512×512.

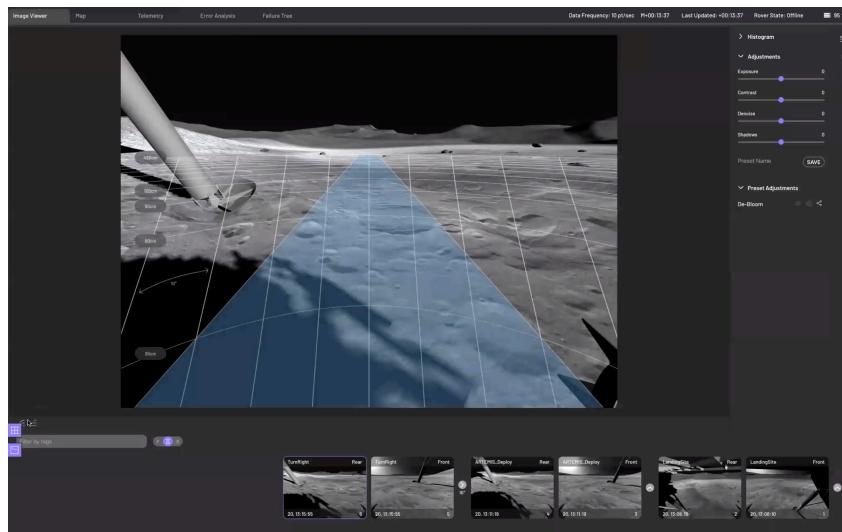


Figure 4. **Iris Lunar Rover.** Simulator used in the actual mission.

- U. Wong, A. Nefian, L. Edwards, X. Buoyssounouse, P. M. Furlong, M. Deans, and T. Fong. 2017. POLAR (Polar Optical Lunar Analog Reconstruction) Stereo Dataset. *NASA Ames Research Center* (2017).
- J. Wu, C. Zhang, T. Xue, W. T. Freeman, and J. B. Tenenbaum. 2016. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-adversarial Modeling. *NIPS* 29 (2016).
- H. Xu, L. Wang, Y. Zhang, K. Qiu, and S. Shen. 2020. Decentralized Visual-Inertial-UWB Fusion for Relative State Estimation of Aerial Swarm. *ICRA* (2020).
- C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li. 2017. High-resolution Image Inpainting Using Multi-scale Neural Patch Synthesis. *IEEE International Conference on Computer Vision* (2017).
- J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang. 2018. Generative Image Inpainting with Contextual Attention. *IEEE International Conference on Computer Vision* (2018).
- H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. Metaxas. 2017. StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE International Conference on Computer Vision* (2017).
- Z. B. Zhang, W. Zuo, X. G. Zeng, X. Y. Gao, and X. Ren. 2019. The Scientific Data and its Archiving from Chang'E 4 Mission. *4th Planetary Data Workshop* (2019).
- J. Zhao, M. Mathieu, and Y. LeCun. 2017. Energy-based Generative Adversarial Networks. *ICLR* (2017).
- J. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros. 2016. Generative Visual Manipulation on the Natural Image Manifold. *EUROPEAN Conference on Computer Vision* (2016).
- J. Zhu, T. Park, P. Isola, and A. A. Efros. 2017. Unpaired Image-to-Image Translation Using Cycle-consistent Adversarial Networks. *IEEE International Conference on Computer Vision* (2017).