Generative Adversarial Networks for automatic detection of mounds in Mars Arabia Terra.

Sahib Julka [1] Michael Granitzer [1] Barbara De Toffoli [2] Luca Penasa [3] Riccardo Pozzobon [3] Ute Amerstorfer [4]

[1]Chair of Data Science, University of Passau, Germany

[2]Deutsches Zentrum für Luft- und Raumfahrt, Institute for Planetary Research, Berlin, Germany

[3]Department of Geosciences, University of Padova, Italy

[4]Space Research Institute, Austrian Academy of Sciences, Graz, Austria

30/04/2021

Semantic Segmentation

- Mars Arabia Terra is region on mars with signs of past occurences of water beneath morphological structures such as mounds.
- Problem Statement: Given an elevation model of a terrain, automatically split the pixel space into disjoint regions of mound and non-mound segments.

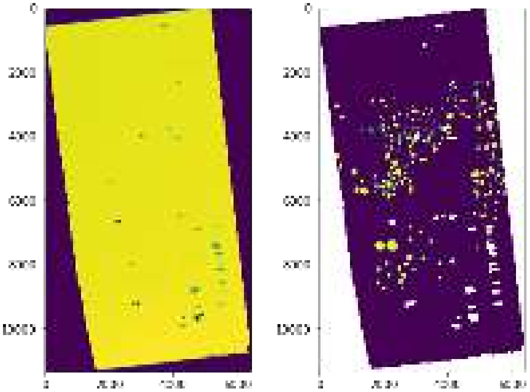

Figure: Left: HiRISE DTM sampled at 1m (as raster). Right: Hand-labelled mounds

Conditional GANs

- pix2pix [2] based supervised semantic segmentation with U-Net [3] backbone.
- Input: Tiled image, Labels: Ground-truth mask.
- Dice Score: 72 %
- Training loss: cGAN loss + L1 reconstruction loss expressed as:

$$\mathbb{E}_{x,y}[\log D(x,y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x,z)))] + \lambda \mathcal{L}_{L_1}(G)$$

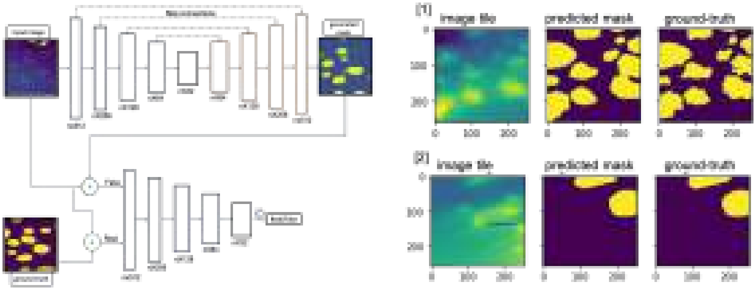

Figure: Left: Architecture of the conditional GAN, Right: Sample tile-wise segmentation results.

Mask-guided data augmentation

To estimate a distribution of variation

- **Experiment 1:** Ground truth mask guided generation: reverse mapping mask $\Rightarrow$ mound
- **Experiment 2:** Random mask based generation, using parameterised cubic bezier curve, with N (number of points) and α(smoothness of the curve), as parameters.
- Data augmentation induces variability and improves detection accuracy.

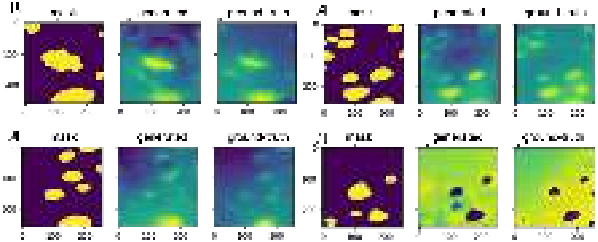

Figure: Example outputs of the mask-guided terrain simulation. The network learns to hallucinate mound-like features within the given mask shapes.

Unsupervised Disentanglement

- Unsupervised learning of latent codes corresponding to factors of variation via maximising mutual information [1].
- Latent factor variation enforced with Context discriminator via causal interventions.
- Training of G regularised with additional constraints.
- Training Loss: $\min_{G,H,Q} \max_D \mathcal{L}_{adv}(G, D) - \lambda\mathcal{L}_{info}(G, Q) - \alpha\mathcal{L}_c(G, H)$

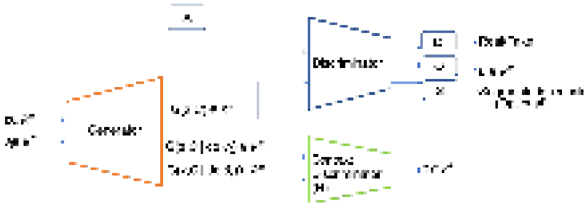

Figure: Overview of the Simulator. Generator receives a fixed number of latent codes (continuous or discrete) in addition to the noise. D predicts whether generated sample is real or fake. Q predicts the latent codes. H takes a pair of images generated with a fixed latent code, and predicts the dimension of the fixed code.

Effect of latent traversal

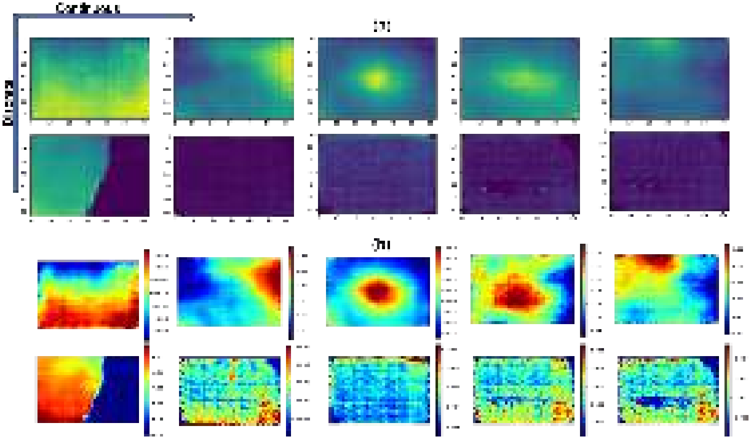

Figure: Results with using 1 latent (along Y axis) and 1 continuous variable (along X axis) (a). The discrete code appears to separate mound regions from non-mounds, while the continuous code appears to control slope (b).

Why terrain simulation?

- Labelling is expensive.
- Controlled generation.
- Implicit learning of factors of variation.
- Better searchability and domain adaptation.

References

Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel.
Infogan: Interpretable representation learning by information maximizing generative adversarial nets.
arXiv preprint arXiv:1606.03657, 2016.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros.
Image-to-image translation with conditional adversarial networks.
In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox.
U-net: Convolutional networks for biomedical image segmentation.
In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

Acknowledgement

Europlanet 2024 RI has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871149.