Rock Segmentation for Planetary Rovers

Prepared by: **Sam Mahoney** – smahoney229@gmail.com
Linkedin: **www.linkedin.com/in/sam-mahoney-7406401bb**
Supervised by: **Steven Kay** - skay@gmvnsl.com

ESA ViBEKO - GMV-NSL

(Vision Based Knowledge Extraction using Artificial Intelligence) This is an ESA led study with the aims of:

- Enabling extraction of operationally relevant information from visual sources acquired in space for enhancing the performance of operations.
- Capability to be implemented in ESA's mission operations data systems.

Lunar surface segmentation

My contribution to this project was to deliver a prototype Machine Learning based Computer Vision Rock Segmentation for Planetary Rovers

- Why?
  - Aids in obstacle avoidance, path planning / correction, sample identification and localisation.
- Challenges
  - Unlike segmentation in a man-made environment, many of the features on rocky planetary bodies are shared across objects. For example, for rock segmentation a small rock may share many features with a large rock, as well as sharing texture and colour with the ground, and shadow intensity matching the sky

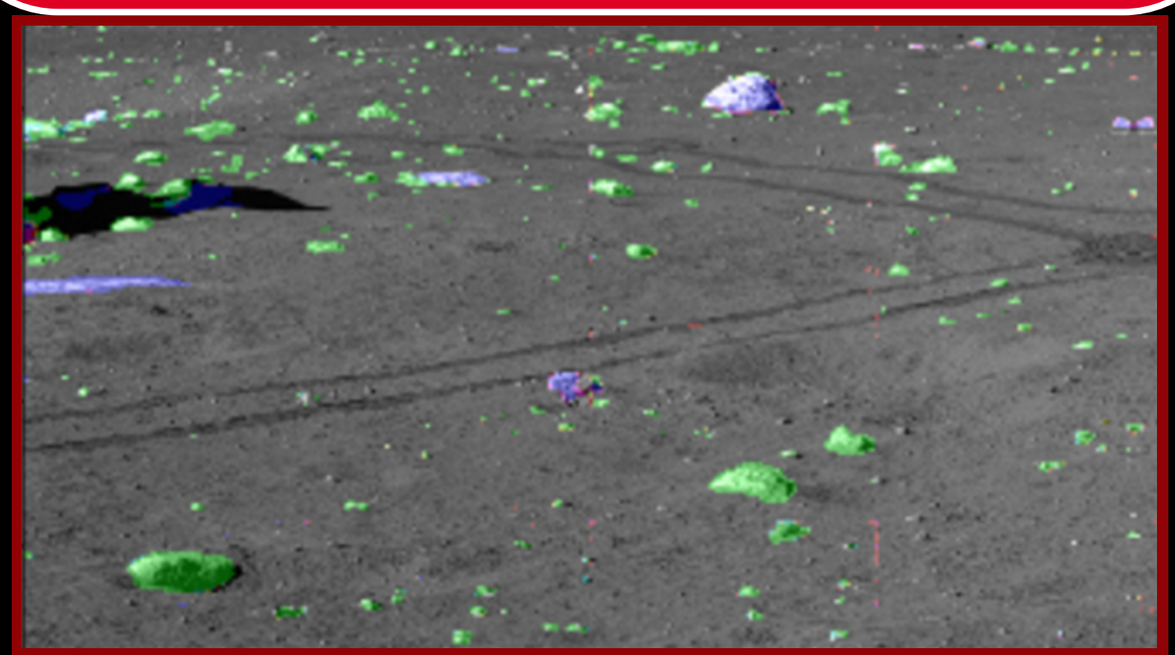
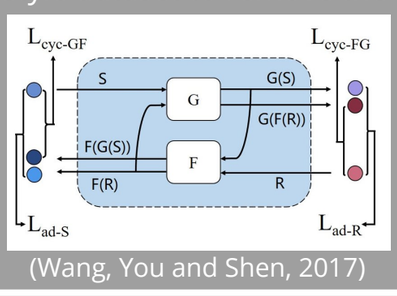
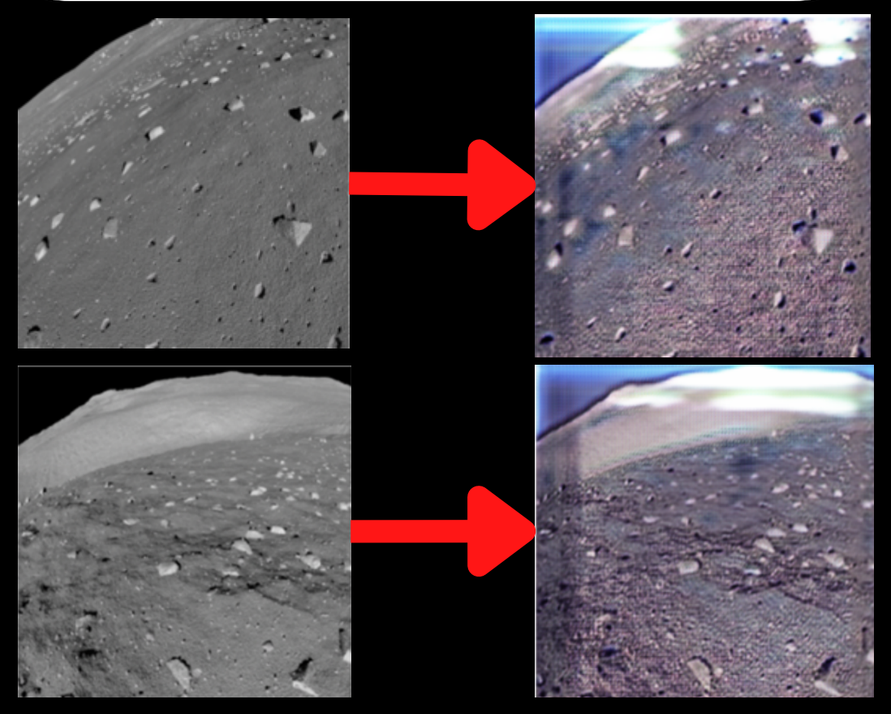


Image Augmentation - CycleGAN

- A Generative Adversarial Network designed to transfer features or 'styles' between two datasets whilst retaining high level features.
- This model was used to apply more realistic texture to the synthetic dataset. With a larger dataset the results would be further improved.
- Further improvement to this could be achieved via pre-processing the style dataset to remove the sky.



(Wang, You and Shen, 2017)



UNET

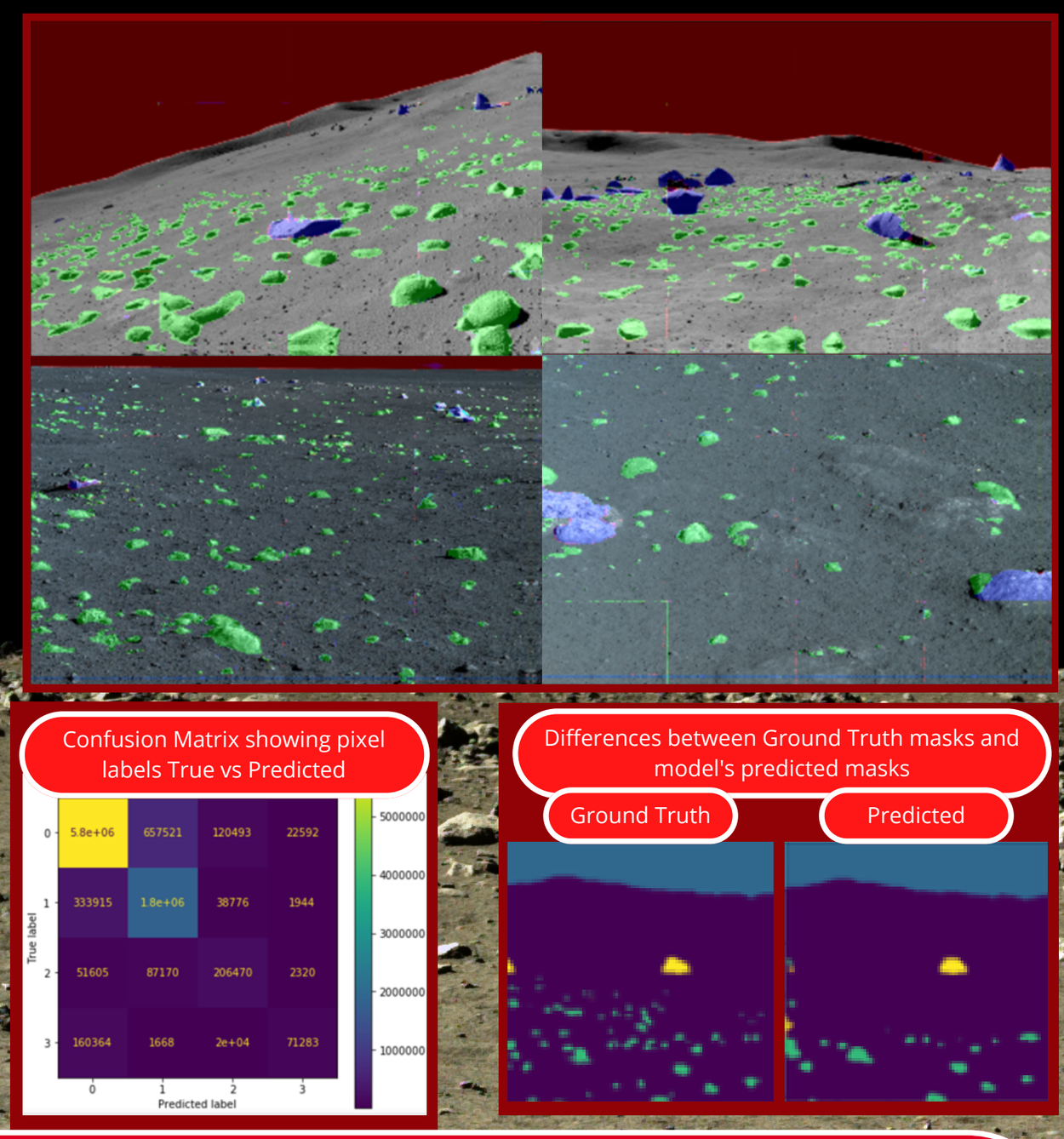
- A 352 layer UNET with trainable InceptionV3 backbone was used.
- Custom weights determined by comparing pixel quantity in mask ground truth dataset
- Dice Loss function - widely used in ML to calculate similarity between two images.

$$DL(y,\hat{p}) = 1 - \frac{2y\hat{p}+1}{y+\hat{p}+1}$$

- Dataset of synthetic images : 100 for training and 30 for testing, converted using patching into 5100 for training and 480 for testing.
- 30 epochs with batch size of 32
- Trains in approx 38mins on a single Nvidia RTX 2080-Ti

Datasets

- **Artificial Lunar Landscape** - (Pessia, 2019) https://www.kaggle.com/datasets/romainpessia/artificial-lunar-rocky-landscape-dataset
- **Chang'e 3 Yutu Rover PCAM images** - (Chinese Academy of Science, n.d.) https://moon.bao.ac.cn/
- **Autonomous Space Robotics Lab: Devon Island Rover Navigation Dataset.** - http://asrl.utias.utoronto.ca/datasets/devon-island-rover-navigation/



Confusion Matrix showing pixel labels True vs Predicted

Differences between Ground Truth masks and model's predicted masks

Segmentation results

- 84% pixel accuracy on Evaluation set when comparing predicted mask vs ground truth mask
- Performs well on both synthetic lunar images and real images taken by Chang'e 3 Yutu Rover
- Each 1000x1000 image takes 0.19 seconds to segment.

References

- Wang, X., You, M. and Shen, C. (2017) Fig. 1. (a) The architecture of the CycleGAN model, which is essential ResearchGate.
- Ronneberger, O., Fischer, P. and Brox, T. (2015) *U-Net: Convolutional Networks for Biomedical Image Segmentation*.
- Jadon, S. (no date) *shruti-jadon - Overview* [online]. Github. Available from: https://github.com/shruti-jadon