



River Delta Satellite Image Classification

Using a neural network to classify land and river delta images.

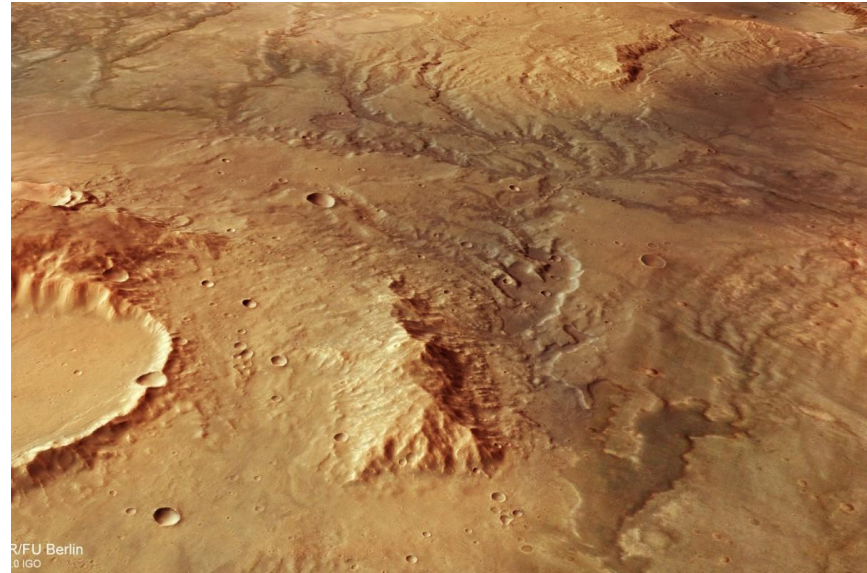
Alex Beat



Business Overview

Problem: A model is needed to help process the large amounts of raw images returned from Mars data.

Business Value: This model will speed up the processing of Mars image data to help classify river delta patterns and aid scientists in honing in on locations that contain traces of existing or pre-existing water on the planet.





Method

Images were scraped from Google Image Search to obtain a unique image dataset of 1,600+ Earth satellite images.

A convolutional neural network model (CNN) was trained on an expanded set of 2,232 images through use of data augmentation.

Phase 1: The model processes Earth image tiles classify land and river deltas.

Phase 2: Second model to process Mars image tiles.





Phase 1 Use Cases

Addition of time series into model processing will give an overview of changes to specific river deltas over time.

- Climate change
- Agriculture
- Commercial and residential real estate





Phase 2 Use Case: Water on Mars

Large amounts of raw data from Mars exists that needs to be processed as more and more of it continues to be collected.

Machine learning model will aid scientists in process.

Classify river delta patterns and return coordinates for scientists to follow up.

River delta patterns could provide potential traces for pre-existence of water.



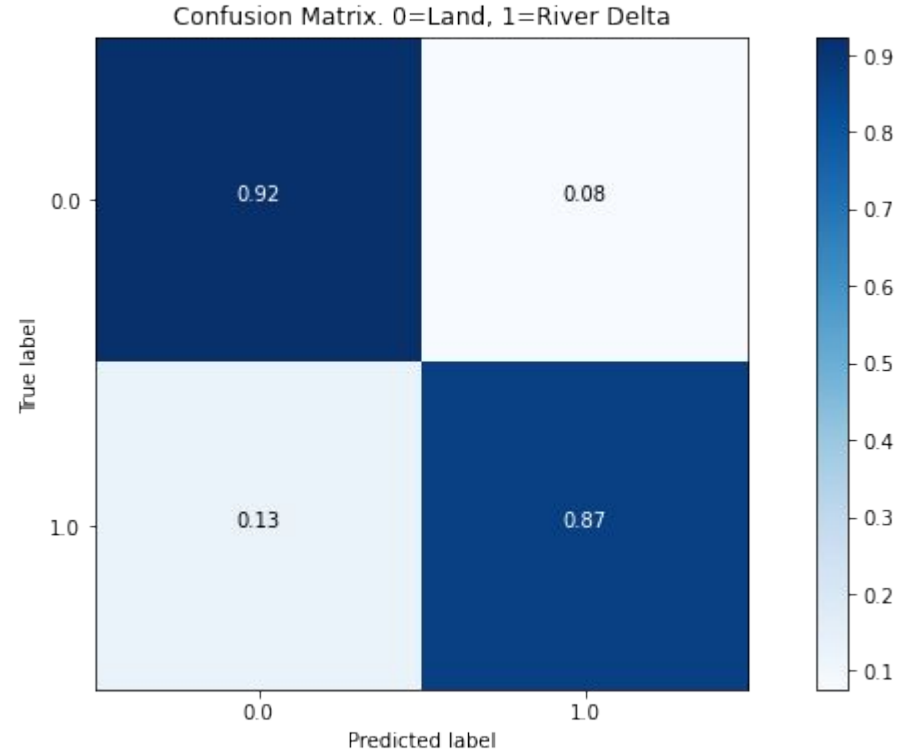
Best Model.

Recall for river deltas: 87%

Precision: 88%

Overall accuracy 90%

Note: The land tile images have a higher score due to a class imbalance issue of having more images contained within the training and test sets compared to the delta class. This can be improved by removing some of the land images in future work.



*Additional model reports in the appendix.



Baseline vs. Dropout.

baseline model					
	precision	recall	f1-score	support	
0.0	0.93	0.94	0.93	212	
1.0	0.90	0.88	0.89	137	
accuracy			0.92	349	
macro avg	0.91	0.91	0.91	349	
weighted avg	0.92	0.92	0.92	349	
baseline model with dropout					
	precision	recall	f1-score	support	
0.0	0.92	0.92	0.92	212	
1.0	0.88	0.87	0.87	137	
accuracy			0.90	349	
macro avg	0.90	0.90	0.90	349	
weighted avg	0.90	0.90	0.90	349	



Summary & Recommendations

Implement this model to aid in faster processing of large quantities of raw Mars data.

Additional use cases on Earth:

- Climate change
- Agriculture
- Commercial and residential real estate

Future Work

Obtain cleaner images.

Note the displayed images and the text/image artifacts they contain.

Result of basic Google Image Search algorithm webscrape.

Land



Delta





Future Work cont.

Grid search for better hyperparameter tuning.

Obtain a bigger set of images and address class imbalance issue.

Obtain Mars API images that have paired location coordinates.

Time series adaptation of Phase 1 for Earth river delta change observation.

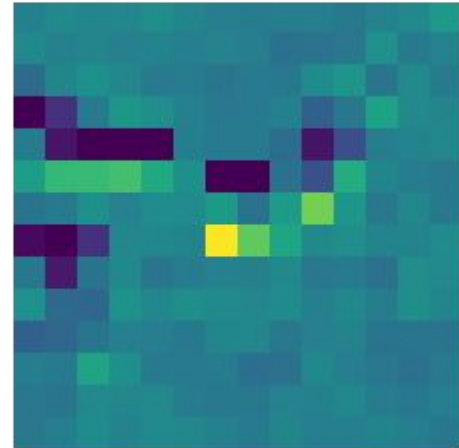
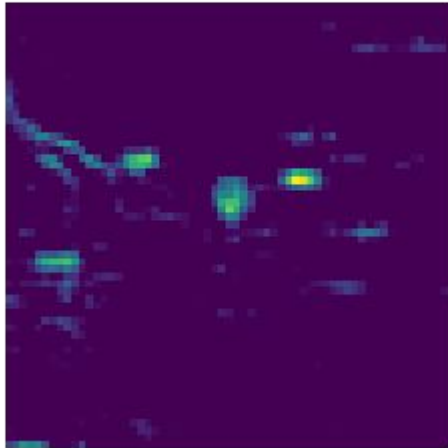
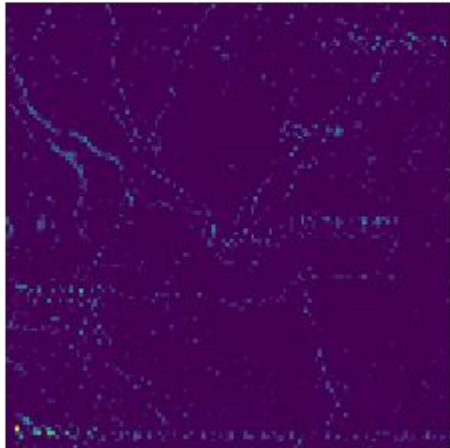


Thank you.



Appendix.

Feature map of CNN pattern
recognition within hidden layers.



Appendix.

CNN model layers.

Each layer develops more patterns and more channel dimensions of the images, going from 32 up to 128 feature mapped dimensions.

Dense layers at the end connect all of the parameters as final learning step for output.

Model: "sequential_16"

Layer (type)	Output Shape	Param #
=====		
conv2d_48 (Conv2D)	(None, 126, 126, 32)	896
max_pooling2d_48 (MaxPooling)	(None, 63, 63, 32)	0
conv2d_49 (Conv2D)	(None, 61, 61, 64)	18496
max_pooling2d_49 (MaxPooling)	(None, 30, 30, 64)	0
conv2d_50 (Conv2D)	(None, 28, 28, 128)	73856
max_pooling2d_50 (MaxPooling)	(None, 14, 14, 128)	0
flatten_16 (Flatten)	(None, 25088)	0
dense_40 (Dense)	(None, 128)	3211392
dropout_8 (Dropout)	(None, 128)	0
dense_41 (Dense)	(None, 512)	66048
dense_42 (Dense)	(None, 1)	513
=====		

Total params: 3,371,201
 Trainable params: 3,371,201
 Non-trainable params: 0



Appendix.

Training and testing accuracy shown through model training process.

X-axis steps are known as epochs.

Loss shows how well the model is predicting during training.

