



IceNet (Alpha): Labeled Image Dataset for Icy World Surface Autonomy

Principal Investigator: Hiro Ono (347F)

Marissa Cameron (398P), Yuliya Marchetti (398J), Cynthia Phillips (394D)

<http://goto/icenet>

Project Objective – 1) Curate analogue images, 2) Annotate images for training and V&V of autonomy algorithms, and 3) share the dataset via JPL network

Curated Images

World	Location/spacecraft	# of images	Analogue of		
			Icy surface	Lighting	Instrument
Earth	<ul style="list-style-type: none"> Borup Fiord Matanuska Glacier Glacier images from NSIDC: Chile, Colombia, Iceland, India, Norway, Peru, Spain 	133			
Moon	Apollo 16 and 17	99			
	Chang'e 3 (Chinese Moon rover)	16			
Mars	<ul style="list-style-type: none"> MSL Mastcam InSight IDC 	86			
Titan	Huygens	1	?		
Total		355			

Included Datasets

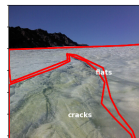


“Descriptive” Dataset

General characterization of dominant features and endmembers within an image

Labels:

- Texture: rough, smooth, fluffy, porous
- Size: big, small, coarse, fine
- Brightness: dark, bright, midtone
- Distinct Color



“Interpretative” Dataset

Higher order analysis of an image

Labels:

- Material: Vegetation, water, rocks, ice
- Terrain: Peaks, fractures/cracks, blades

POSTPONED



“Sampling” Dataset

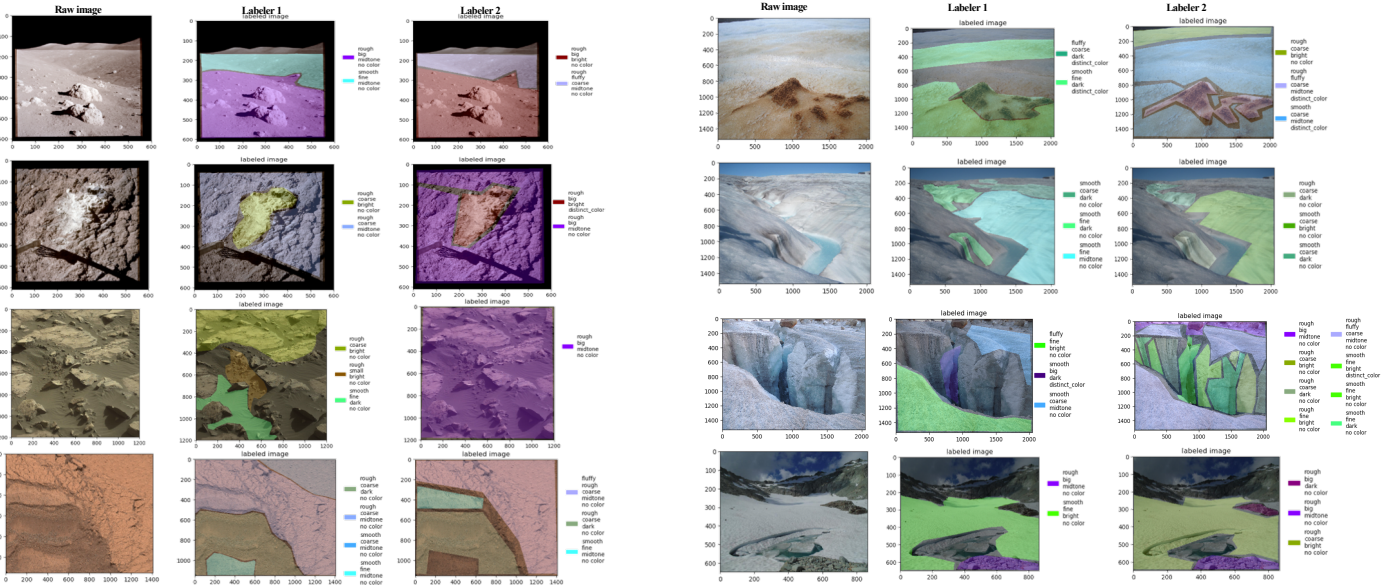
Assess scientifically interesting sampling locations

Labels:

- Fluffiness (inductive of newly deposited materials)
- Colored surface (inductive of non-ice materials)

Labels

- “Descriptive” dataset: 528 labels produced by 2 JPL scientists (173 images have two labels) – examples below
- “Sampling” dataset has 355 labels produced by a JPL scientist
- Collected through Zooniverse (web-based citizen science tool); required time: ~2min/label



Benefits to NASA and JPL (or significance of results):

- A necessity for vision-based autonomy on icy world surface
- First dataset of this kind, to the best of our knowledge
- Any JPL projects can download the dataset
- Currently datatize too small for supervised machine learning; need follow-on to scale

Classification results (Proof-of-concept)

- Semantic segmentation of size and texture
- Trained TextureCam (random forest) with the Descriptive Dataset
- Currently dataset too small to train accurate models; Results will improve with more labels

