Mid-Project Presentation

QIO Geo-Detection Project

Jingyun Chen, Socorro Dominguez, Milos Milic, Aditya Sharma

Table of Contents

Business Problem

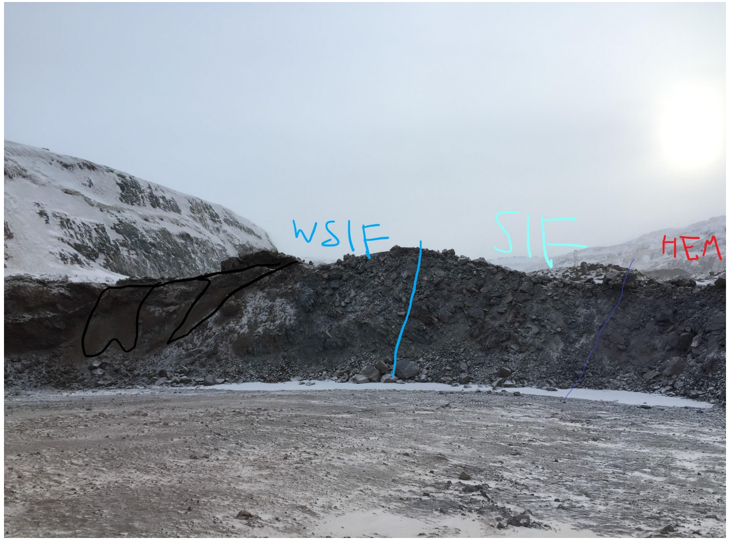
- Current operations need a trained geologist to identify rock types.
- Long term objective would be a tool that segments and identifies rock types from an image in real time, so that operators do not always have to depend on a geologist.
- We will deliver a prototype (proof of concept) that will help:
  - assist operations in identifying boundaries between blasted rock types at an active mining face.
  - identify which category the rocks fall into (8 types of rock that later classify into ore or waste).

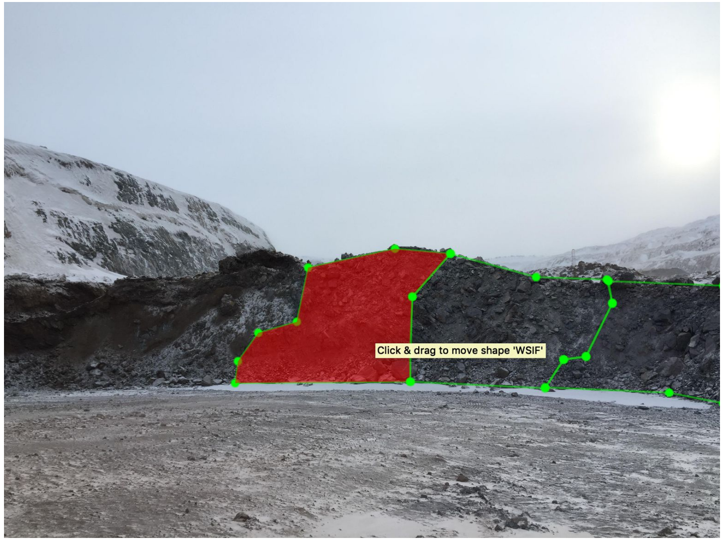
MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

Preprocessing the Data

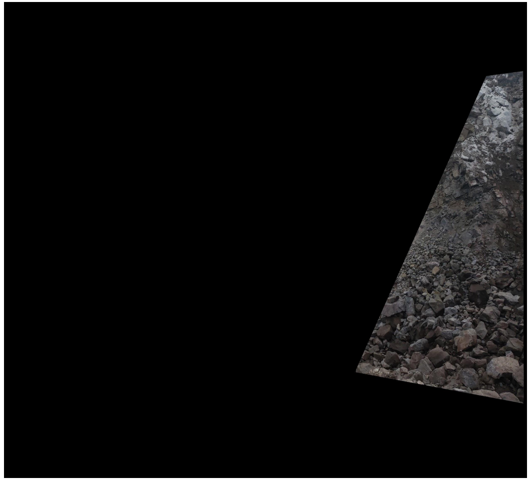
- The first problem we encountered, was the way data was labeled.
    - We changed the labeling method to LabelMe and QIO has adopted this change too!

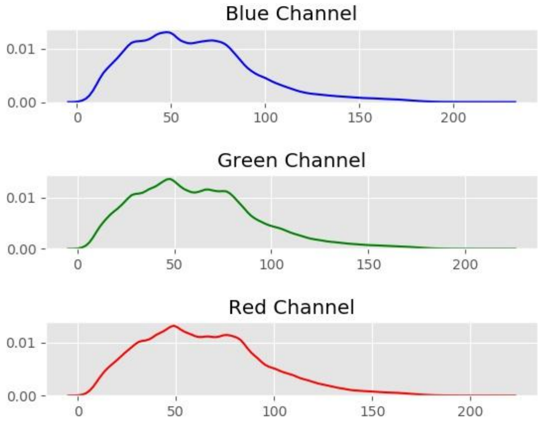




Exploratory Data Analysis - Colour Analysis

- Analyzed Colour Features for each image based on RGB channels.





- Nine different features (3 for each channel):

  - Skewness
  - Kurtosis
  - Mean Pixel

MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

EDA - Colour Analysis

- Our results:

	AvgMeanPixelBlue	SdMeanPixelBlue	AvgMeanPixelGreen	SdMeanPixelGreen	AvgMeanPixelRed	SdMeanPixelRed	MaxValueColour
AMP	84.359	33.439	81.638	31.002	79.146	30.009	Blue
BS	87.411	22.441	92.456	23.376	96.608	23.555	Red
GN	90.740	21.287	91.651	20.923	92.825	21.040	Red
HEM	87.529	36.227	85.068	33.933	87.778	32.411	Red
IFG	70.510	29.482	78.864	29.396	90.286	31.485	Red
LIM1	64.734	23.063	80.127	28.666	98.564	34.433	Red
LIM1-2	87.918	0.000	127.950	0.000	172.283	0.000	Red
MAG	127.611	35.239	126.241	35.817	126.894	38.738	Blue
MS	143.913	3.558	149.430	4.879	145.732	7.221	Green
QR	94.358	33.206	99.399	30.686	109.655	30.146	Red
QRIF	61.421	7.843	61.153	8.761	65.515	9.838	Red
SIF	91.682	29.066	91.071	29.683	89.173	29.894	Blue
WSIF	113.874	28.930	110.418	30.104	106.916	29.806	Blue

MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

EDA - Colour Analysis

- We also did an ANOVA test pairing each rock to all other rocks.
- This allowed us to see which rocks could be “easily differentiable” according to the 9 different features.

group1	group2	Kurtosis B	Kurtosis G	Kurtosis R	Skew Blu	Skew Gre	Skew Red	Mean Blu	Mean Gre	Mean Red	Total	Group1in	Group2in
HEM	QR	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	7	Ore	Dwaste
LIM2	QR	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	7	Ore	Dwaste
GN	LIM2	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	5	Cwaste	Ore
LIM2	WSIF	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	5	Ore	Cwaste
AMP	QR	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	4	Cwaste	Dwaste
AMP	WSIF	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	3	Cwaste	Cwaste
HEM	WSIF	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	3	Ore	Cwaste
AMP	GN	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	1	Cwaste	Cwaste
AMP	LIM1	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	1	Cwaste	Ore
AMP	LIM2	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	1	Cwaste	Ore
AMP	MAG	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	1	Cwaste	Ore
BS	LIM2	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	1	Cwaste	Ore
QR	SIF	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	1	Dwaste	Ore
HEM	LIM1	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	1	Ore	Ore

MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

EDA - Colour Analysis

- Most rocks are actually the colour they are supposed to be.
- Errors could be because there is insufficient data for some rocks or because of noise in the image: e.g. brightness or the existence of a window between the camera and the rock face.

Rock Characteristics and their Shades of Gray

	Red	Orange	Brown	Yellow	Green	Purple	White	Pale gray	Dark Gray	Dark Gray - Green
Blocky	HEM/IFH, QRIF		IFG/LIM2/LIM3	QR	WSIF, SIF		QRMS	GN	AMP, MAG/IFM	AMP
Sandy	QRIF	LIMO/LIM1	IFG/LIM2/LIM3		MS		MS			
Shiny	HEM/IFH				MS	HEM/IFH, QRMS	QRMS, MS			
Laminated and Fissile					MS		QRMS, MS	GN	GN	
Stratified					GN			GN	GN,MAG/IFM	

EDA - EXIF Data

- EXIF Data (exchangeable image file format) gives us information about the image that could correct colour imbalances.
- Examples of data fields: phone type, megapixels size, exposure time, lens aperture, focal length.
- If any of these extra features is what is affecting the images, we could normalize the data and strengthen our model.
- From the size of the pictures, we can identify if a picture was SMS-ed or taken from the device.

MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

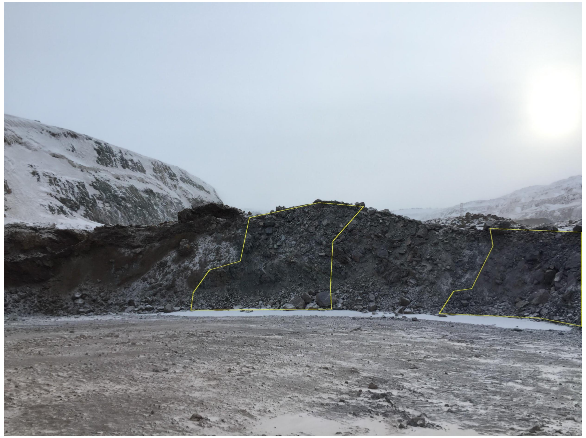
UBC Master of Data Science

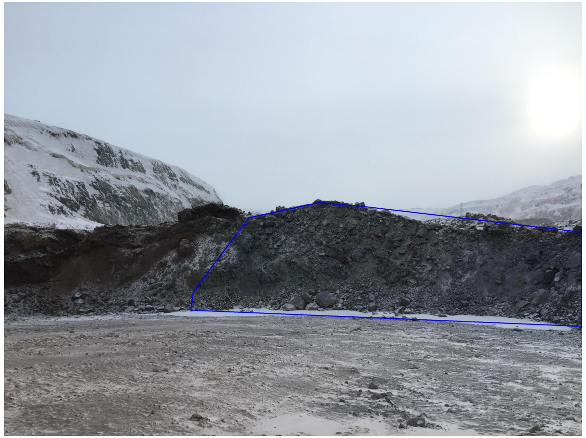
IRONN's First Steps

Image Region Optimizing Neural Network

- Once the labeling was done, we first created Convex Hulls for polygons.

- We call this the "Blasted Face"



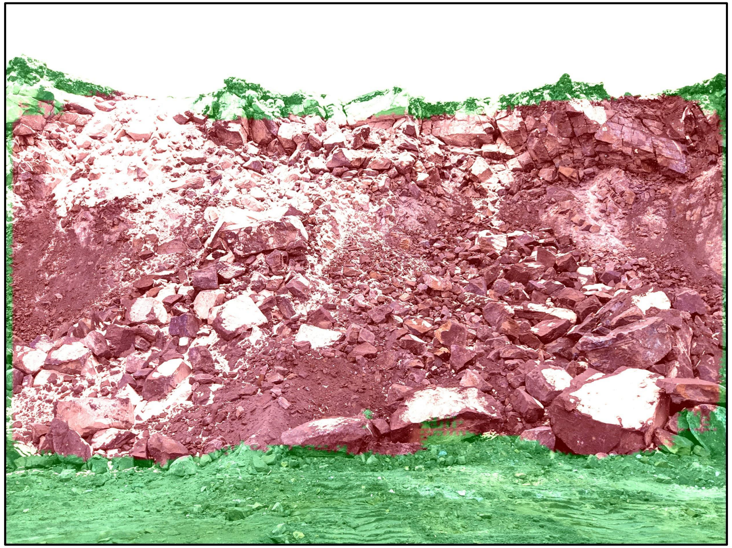


MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

Our Current Model

- With the ‘Blasted Face’ masks set, we trained a CNN on around 430 images.
- Original image and the model output for an example image where:
  - Red: blasted face
  - Green: rest of the image/background



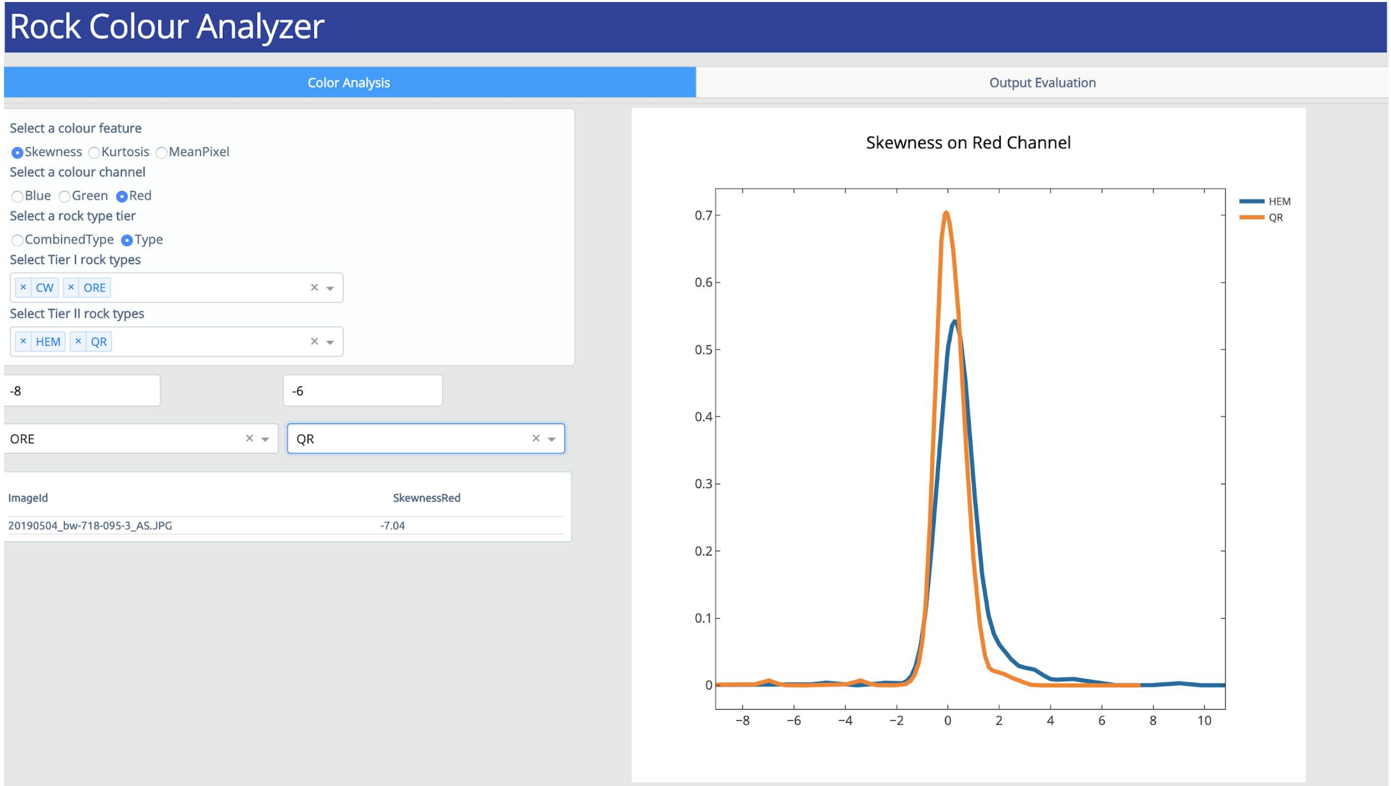
Our Current Model

- Currently using Fully Convolutional Networks for Image Segmentation
- Takes in images and spits out images
- Framework: PyTorch
- Next step:
  - Train a model with all the polygons
  - Predict and check performance
  - Improve performance

Our Current Product

Plotly Dashboard

http://127.0.0.1:8050/



MINERAI DE FER QUÉBEC
QUEBEC IRON ORE

UBC Master of Data Science

Our Current Product

- Features we are adding:
    - Quality assurance for pictures.
    - Quality assurance for masks.

Qualitive image analysis

Filename	How close is the rock face? (0 = Perfect, 1 = Too close, 2 = Too far)	Is image too light or dark? (0 = Perfect, 1 = Too dark, 2 = Too light)	Is image taken behind glass? (0 = Yes, 1 = No)	Is the mask good?(0 = Yes, 1 = No)
File1.jpg	2	2	0	0
File2.jpg	0	1	0	0
File3.jpg	1	0	0	0
File4.jpg	1	0	0	0
File5.jpg	0	0	0	1
File6.jpg	2	0	1	0

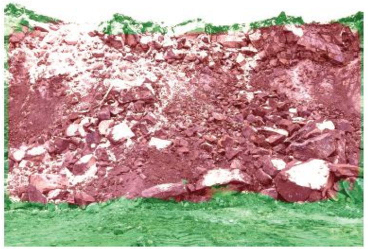
Reference_file.jpg

download

File1.jpg

Upload a file

File1_mask.jpg



Upload a file

Difficulties & Challenges

- Some training images are unsuitable for training (bad lighting, angles, and background noise).
- May not have enough training data due to a large number of rock types.
- Model’s current performance might not be the same when we introduce more classes.
- Model might not be able to handle fine-grained differences from textures between rock types.
- Available time frame

Next Steps

- Pipeline parts have been (mostly) created
- QIO would need to train mine operators to capture images in a standardized format
- Train a model to segment different rock types
- With a better and properly labeled training set, model performance will improve
- Develop a user friendly interface that can be used by the operators