

# Proposal Presentation

## QIO Geo-Detection Project

*Aditya Sharma, Jingyun Chen,*

*Milos Milic, Socorro Dominguez*



## Table of Contents

- Introduction
- Business problem
- Objective
- Final product
- Workflow
- Raw data
- Data science techniques
- Challenges
- Timeline
- Summary

## Minerai de fer du Québec (Quebec Iron Ore)

- Mining company based in Quebec that specializes in iron ore extraction and processing.
- Around 500 specialists in various fields of iron ore extraction.
- Collect 100GB of production data per day.
- Want to employ machine learning to:
  - increase efficiency (6M tonnes to 7M tonnes of iron)
  - use stored data

## Business Problem

- Current operations need a trained geologist to identify rock types.
- Long term goal would be a tool that identifies rock types from an image in real time, so that operators do not always have to depend on a geologist.
- For now, they need a prototype 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).
- The types of rock would come from the training data images that have been previously classified by the experts.
- If prototype is successful, they would move forward to other ML projects.

# Business Problem



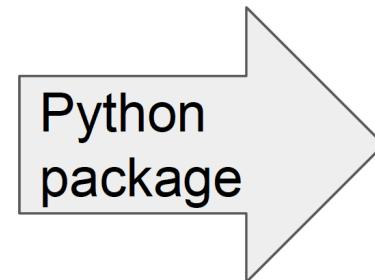
# Objective

- Automatic detection of 8 different rock types through image segmentation and classification.



# Final Data Product

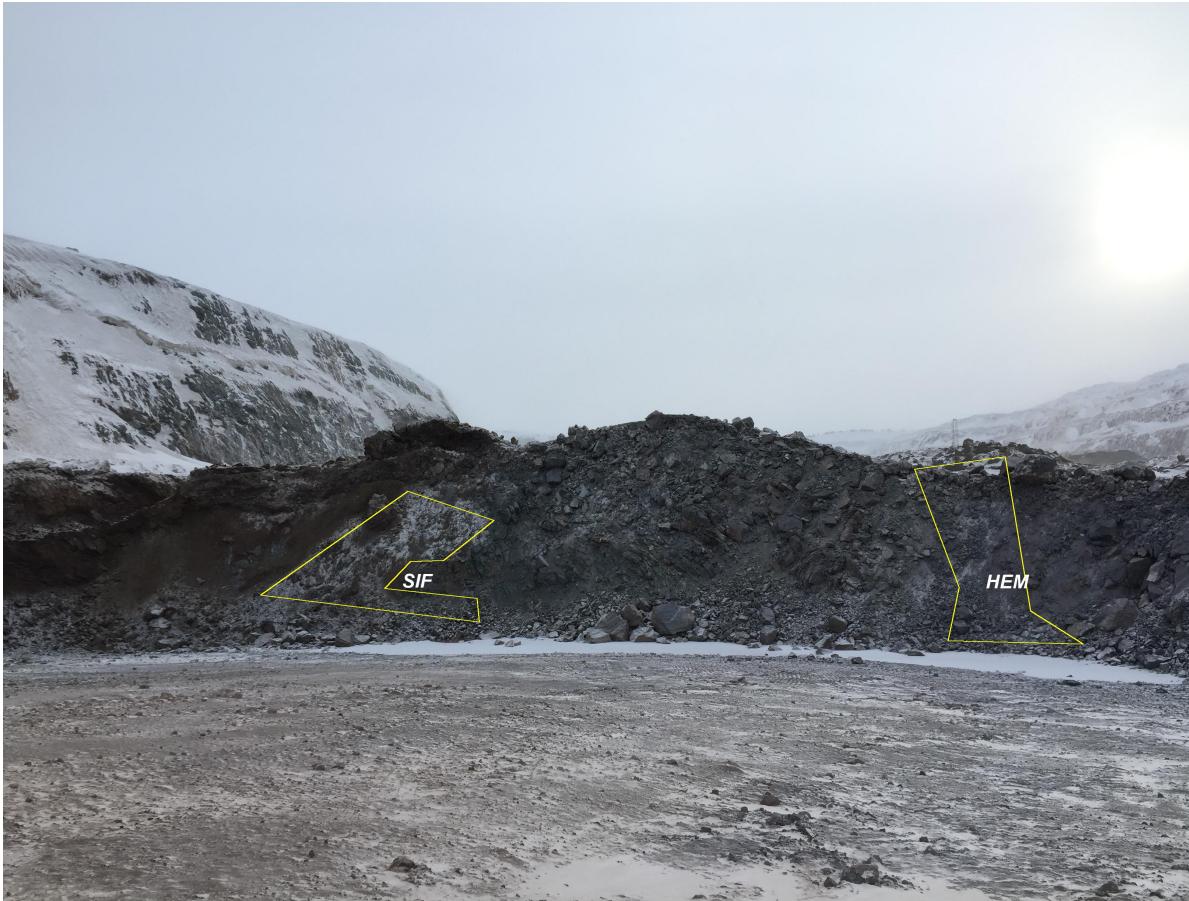
- A python package (prototype)
  - Input: images of blasted rocks taken at the mining field.
  - Output: images with rock type annotations (ideally, annotated rock types using polygons).



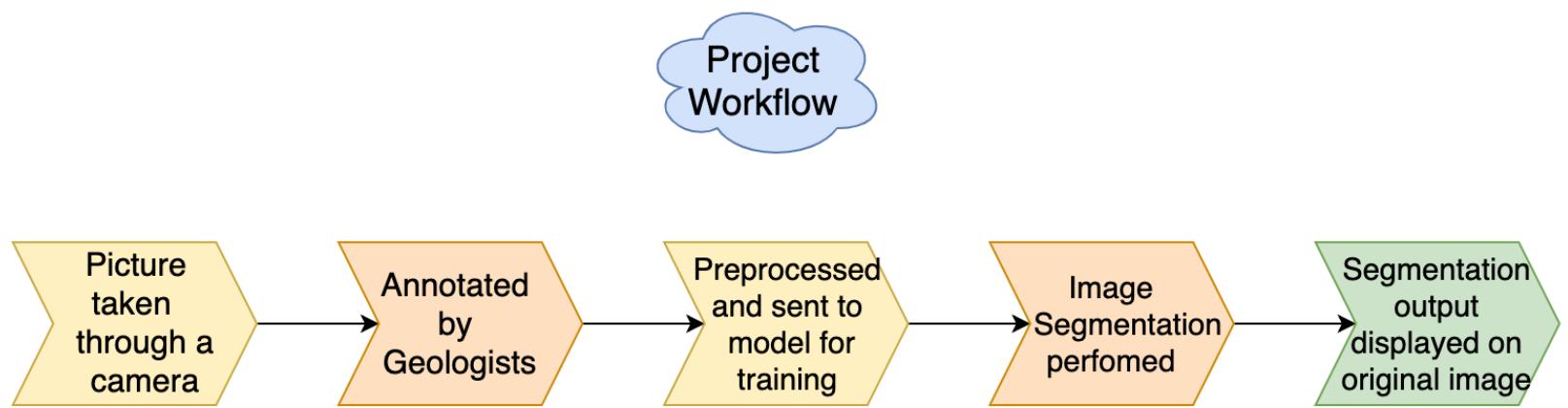
# Final Data Product (Input)



# Final Data Product (Output)



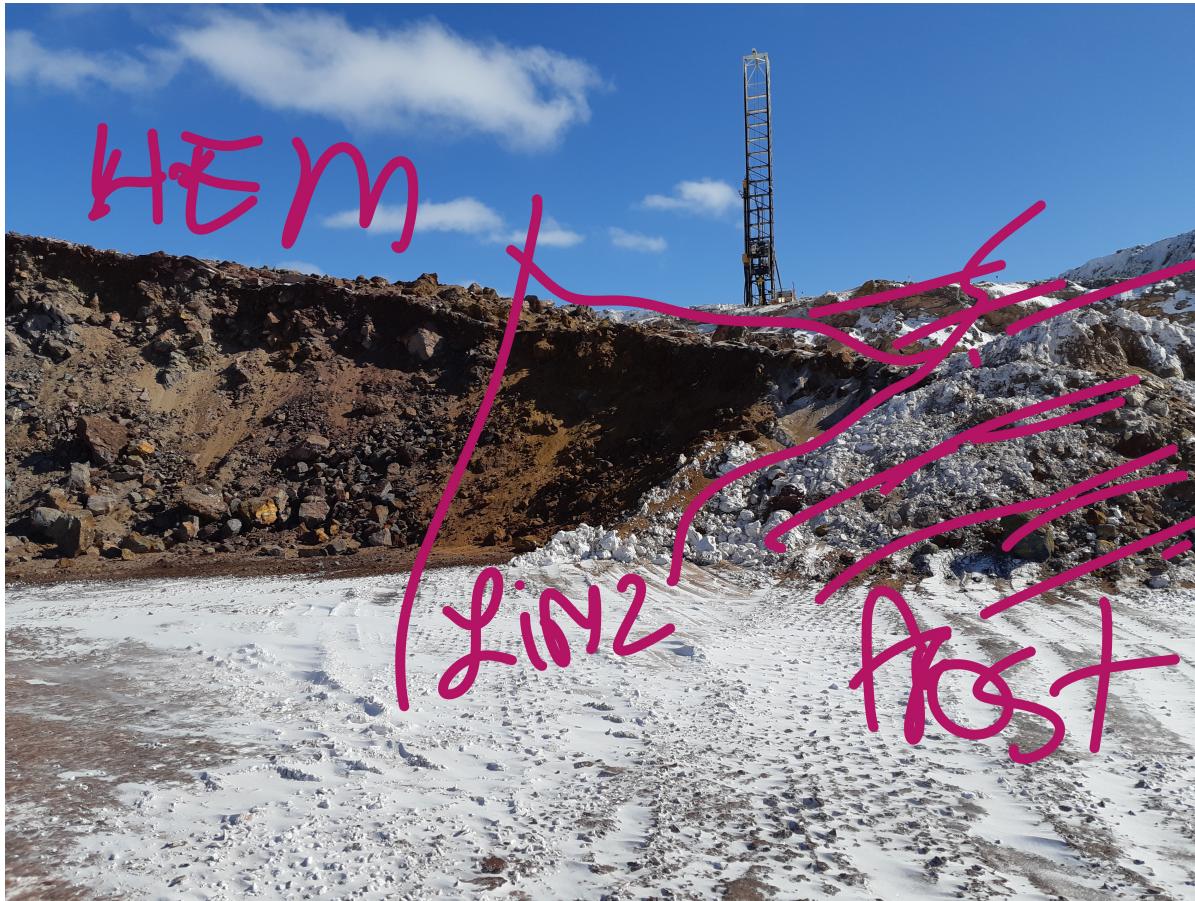
# Project Workflow



## Raw Data

- Currently: ~500 original images, ~500 hand-annotated reference images.
- Features: 8 rock types.
- Issues:
  - Hand-annotated images cannot be used for training a model.
  - Images contain background information (e.g., sky, ground, snow).

## Raw Data



# Data Science Techniques

## Preprocessing of the data

- Image labelling software (LabelMe: open source Polygonal Annotation tool)
- Image processing (OpenCV: open source Computer Vision library)

## Model training

- CNN for object detection for now (looking into other techniques)
- Semantic Segmentation using CNNs (to be further researched)

## Difficulties & Challenges

- Segmentation algorithm may remove parts of the ore face and may keep parts of the ground.
- Some training images may be unsuitable for training (bad lighting/angles).
- May not have enough training data for 8 different rock types.
- Model may have a poor accuracy.
- Thinking about the input and output format.
- Handling fine-grained differences from textures.
- Thinking about what is possible and what is not in the given time period.

# **Timeline**

Week 1: EDA and Proposal Presentation (April 26th)

Week 2: End of Proposal Document (May 3rd)

Week 3: Preprocess original data

Week 4 - 7: Create/train/evaluate/update model

Week 8: Finish data product (Presentation on June 17/18)

Week 9: Polish data product

Week 10: End of project

# Summary

- Quebec Iron Ore currently collects huge amounts of production data (100GB per day) and would like to leverage data science solutions in order to optimize different stages in the mining process.
- We need a way to classify rock types using object detection and segmentation.
- The final product would be a python package with:
  - Input: an image of the blasted rock face.
  - Output: an image with annotated boundaries labeling rock types.
- There are some unique challenges due to similarity in the texture, color and other properties in different rock types.

**Thank you!!**