Using A.I. to Classify Lithology: Fact or Fiction?

17 September 2019

Cristal Palafox M.Sc., Spectral Geologist

Sam Scher M.Sc., Senior Geochemist

corescan

THE IMPORTANCE OF TERMINOLOGY

Thanks to machine-learning algorithms, the robot apocalypse was short-lived.

- Artificial Intelligence (A.I.) denotes machines that can think for themselves.
- Machine Learning is a subset of A.I. that is focused on systems that learn from their environment, e.g. data.
  - It is a method of data analysis that automates analytical model building by using algorithms that iteratively learn from data to find hidden insights and structure without being explicitly programmed where and how to look.

ADDRESSING THE LIMITATIONS IN MINING & EXPLORATION

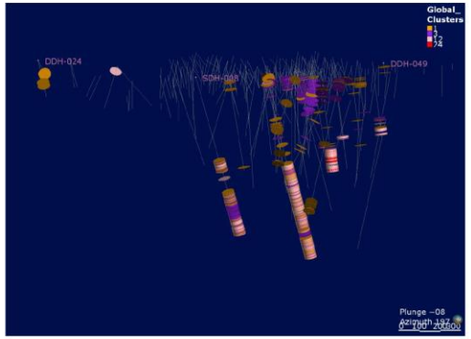
- A tale of two extremes
- Humans can see and do everything you might want to see and do
- Artificial Intelligence (AI) will save the world

- In reality, a machine is only as smart as what you put into it, and those who program it.

ADDRESSING THE LIMITATIONS IN MINING & EXPLORATION

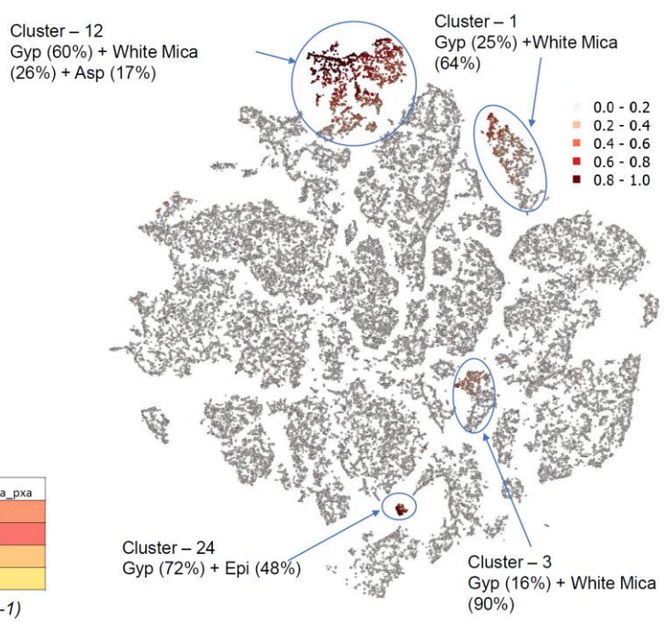
- A tale of two extremes
- Humans can see and do everything you might want to see and do
- Machine Learning will save the world

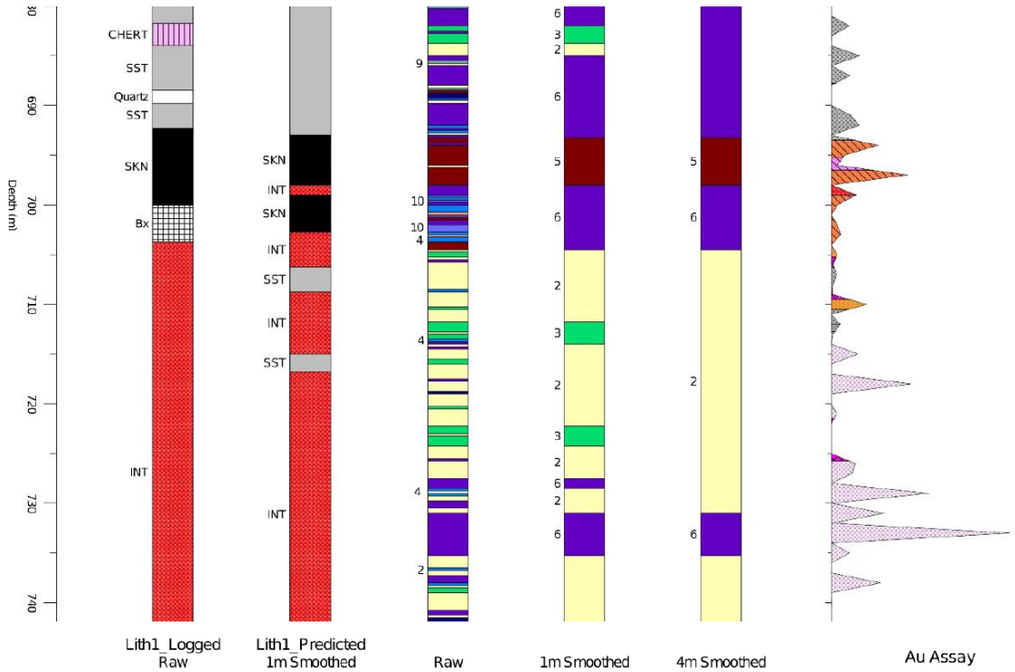
Gypsum



Cluster Number	aspectral_pxa	epidote_pxa	gypsum_pxa	white_mica_pxa
1	0.23	0.00	0.26	0.64
3	0.04	0.00	0.16	0.91
12	0.17	0.01	0.60	0.26
24	0.03	0.48	0.72	0.03

Cluster Legend showing average abundance (between 0-1)





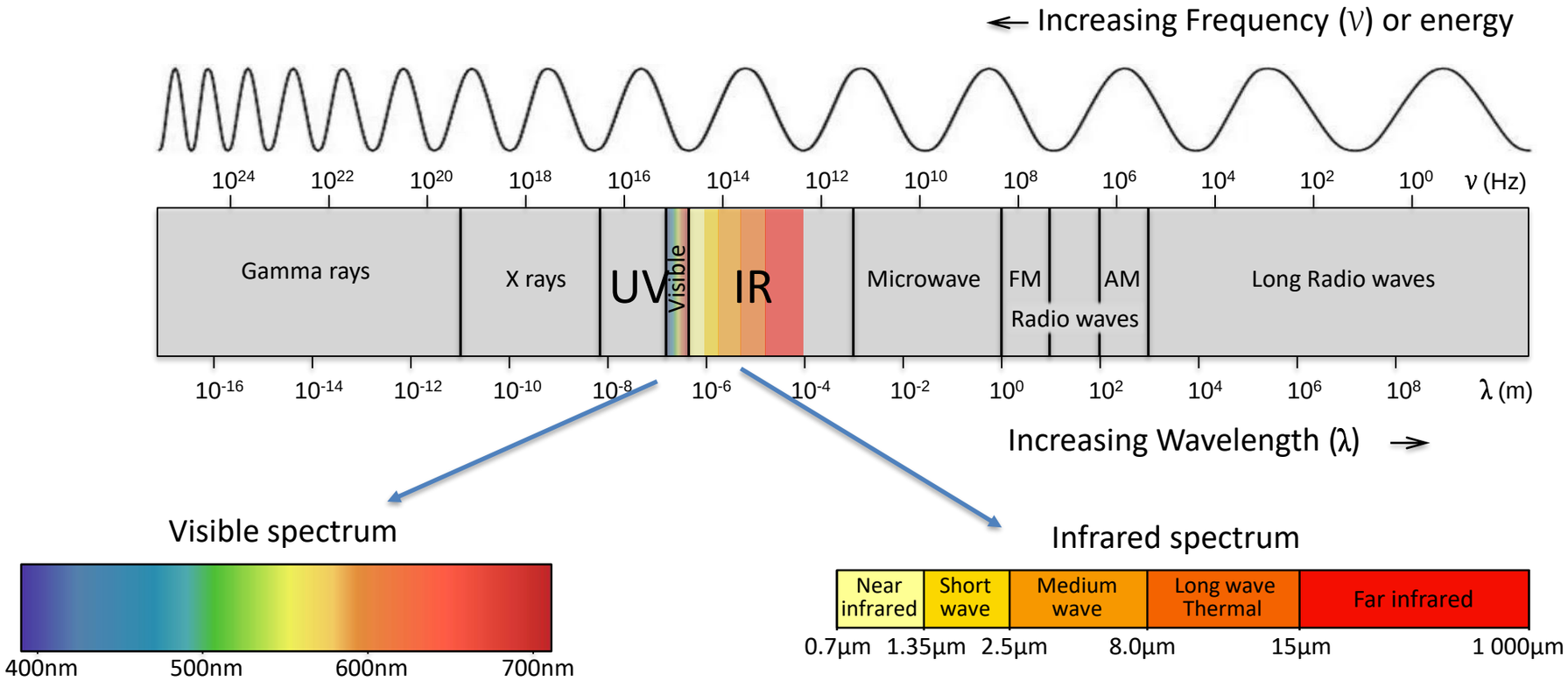
- In reality, a machine is only as smart as what you put into it, and those who program it…and ***what*** you do with it

GETTING TO THE ROOT OF THE ISSUE

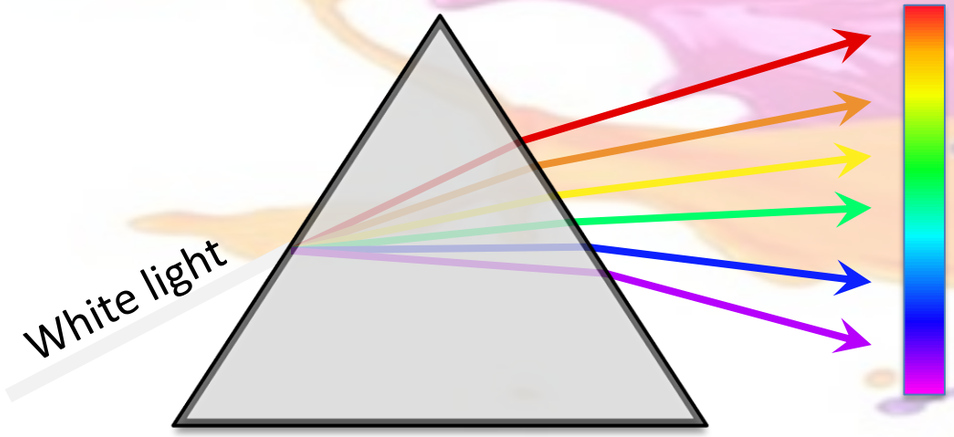
- 8 geologists pictured at left
- 15 opinions of mineralogy present, which in turn affects:
  - Interpretation of lithology
  - Interpretation of alteration
  - Construction of paragenesis
  - Exploration model recommendations
  - Ultimate 3D representation used for mining

How to combine valuable human interpretations with quantitative data, e.g. mineral identification…?

THE PHYSICS OF LIGHT: THE ELECTROMAGNETIC SPECTRUM

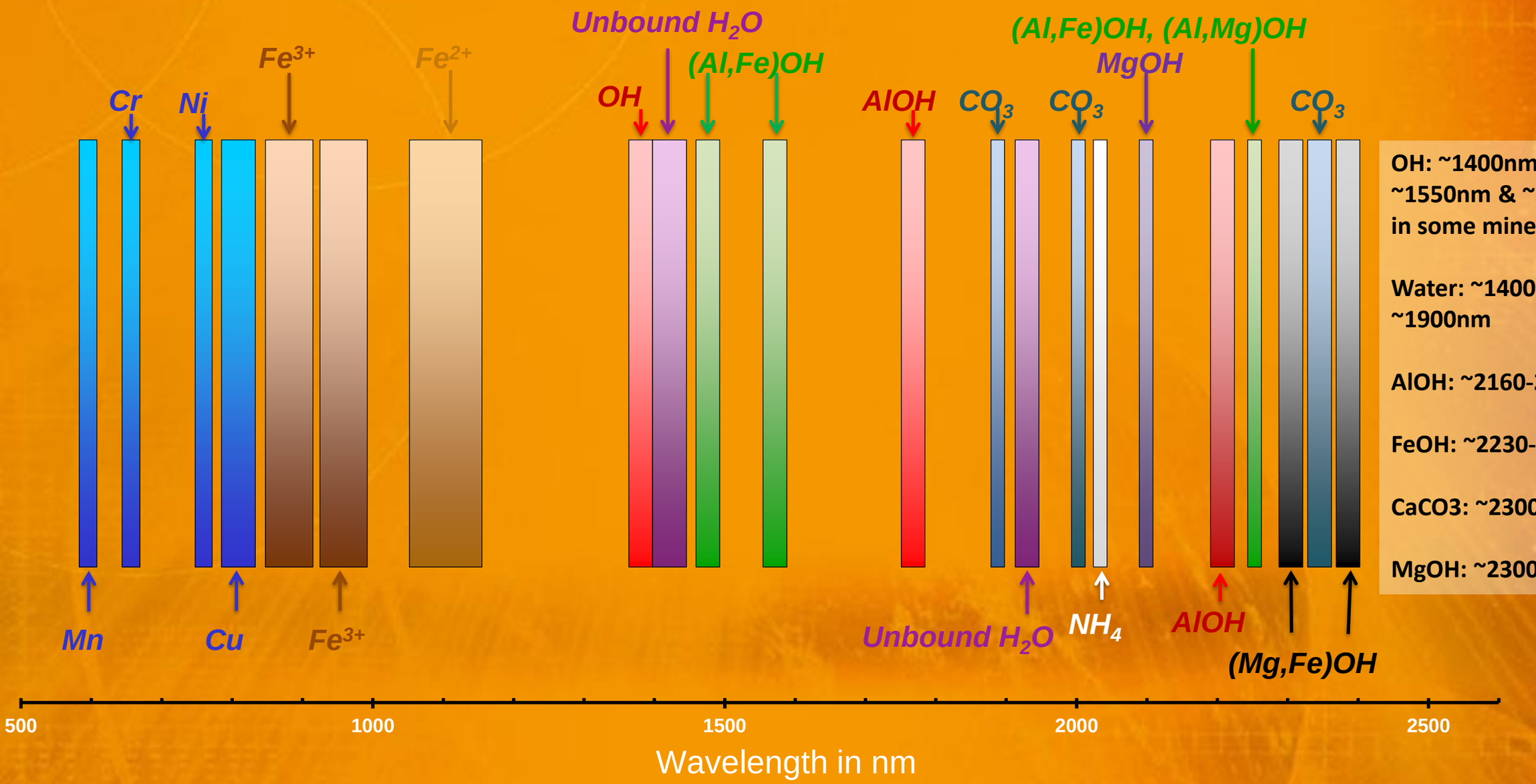


IDENTIFYING MINERALS WITH THE VISIBLE COLOR SPECTRUM



Violet:	**380nm to 424nm**
Blue:	**424nm to 491nm**
Green:	**491nm to 575nm**
Yellow:	**575nm to 585nm**
Orange:	**585nm to 647nm**
Red:	**647nm to 760nm**

Diagnostic absorption features of minerals in the Visible-near (VNIR) to shortwave infrared (SWIR) range



OH: ~1400nm (also ~1550nm & ~1750-1850 nm in some minerals)

Water: ~1400nm & ~1900nm

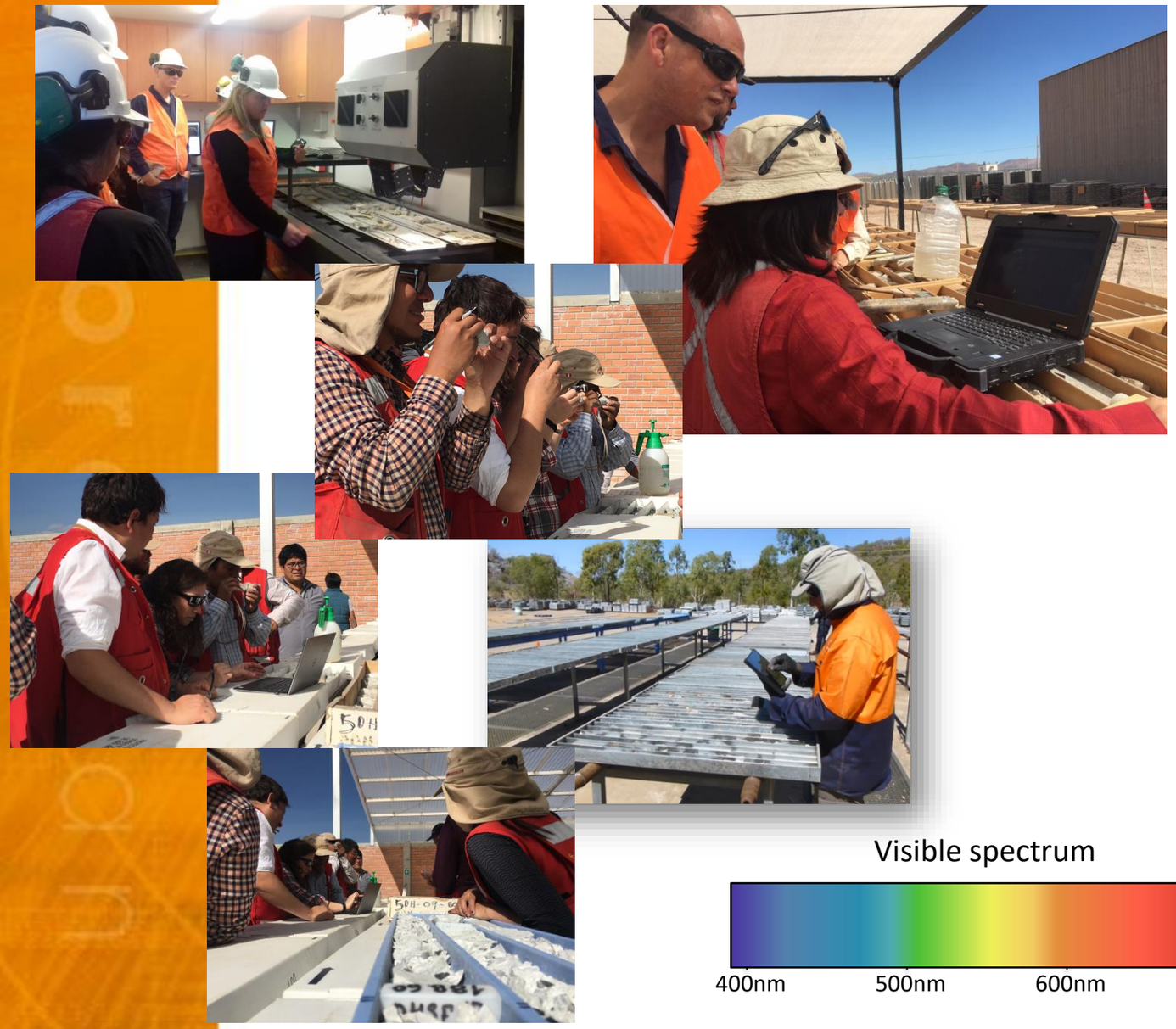
AlOH: ~2160-2228nm

FeOH: ~2230-2295nm

CaCO3: ~2300-2370nm

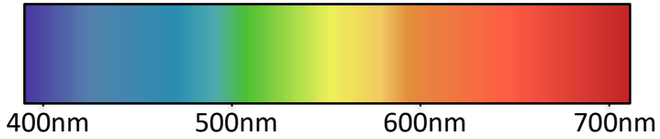
MgOH: ~2300-2370nm

ADVANTAGES OF INTEGRATING HYPERSPECTRAL CORE IMAGING IN LOGGING

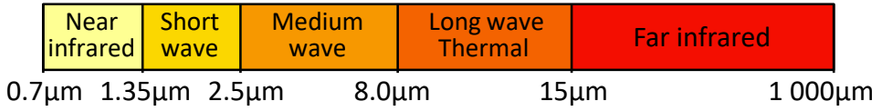


- Geologists bring knowledge, context and hard-earned wisdom
- ‘Domain experts’ are indispensable
- The synthesis of humans and consistent, precise analytical data is best-practice

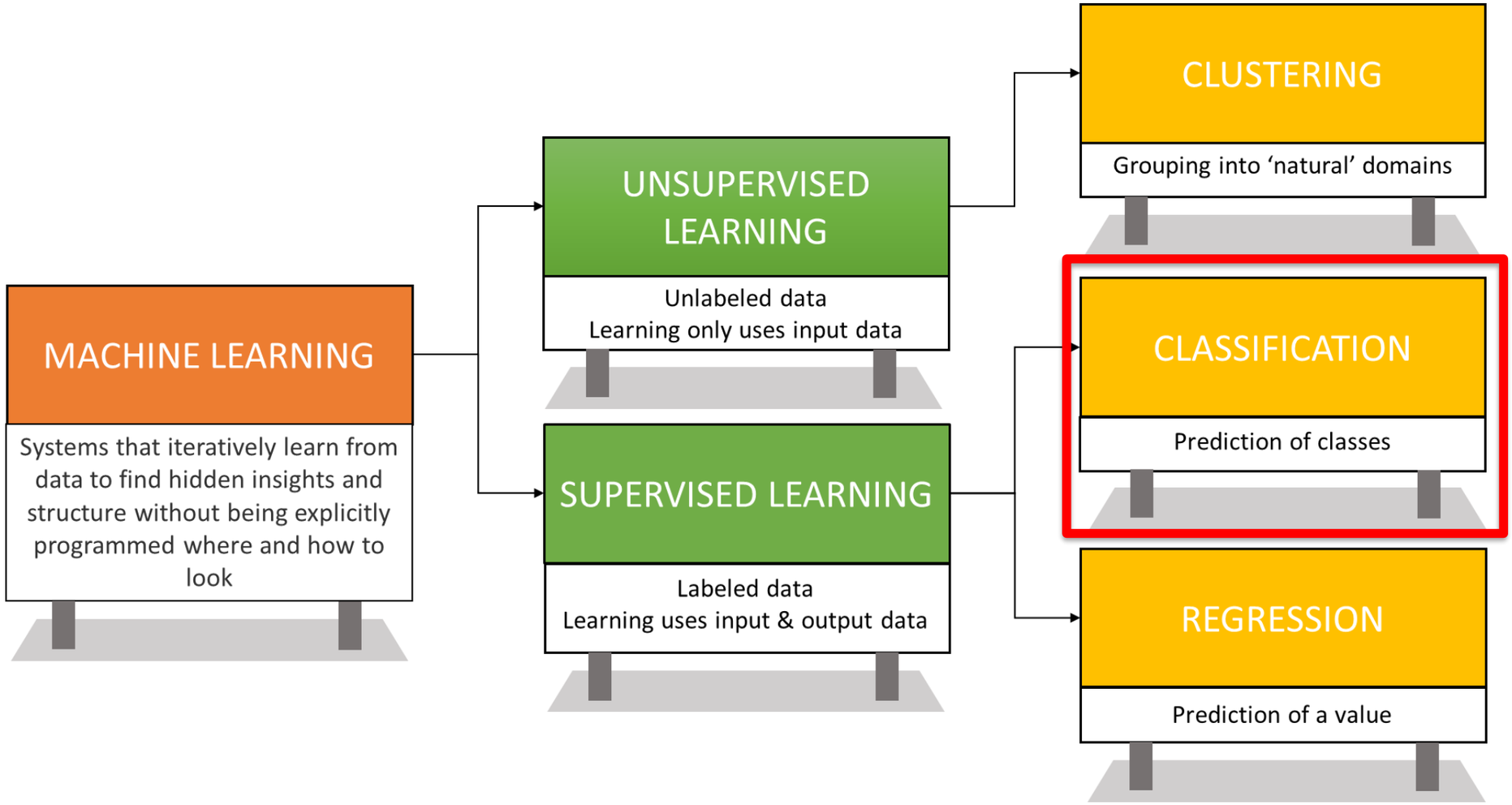
Visible spectrum



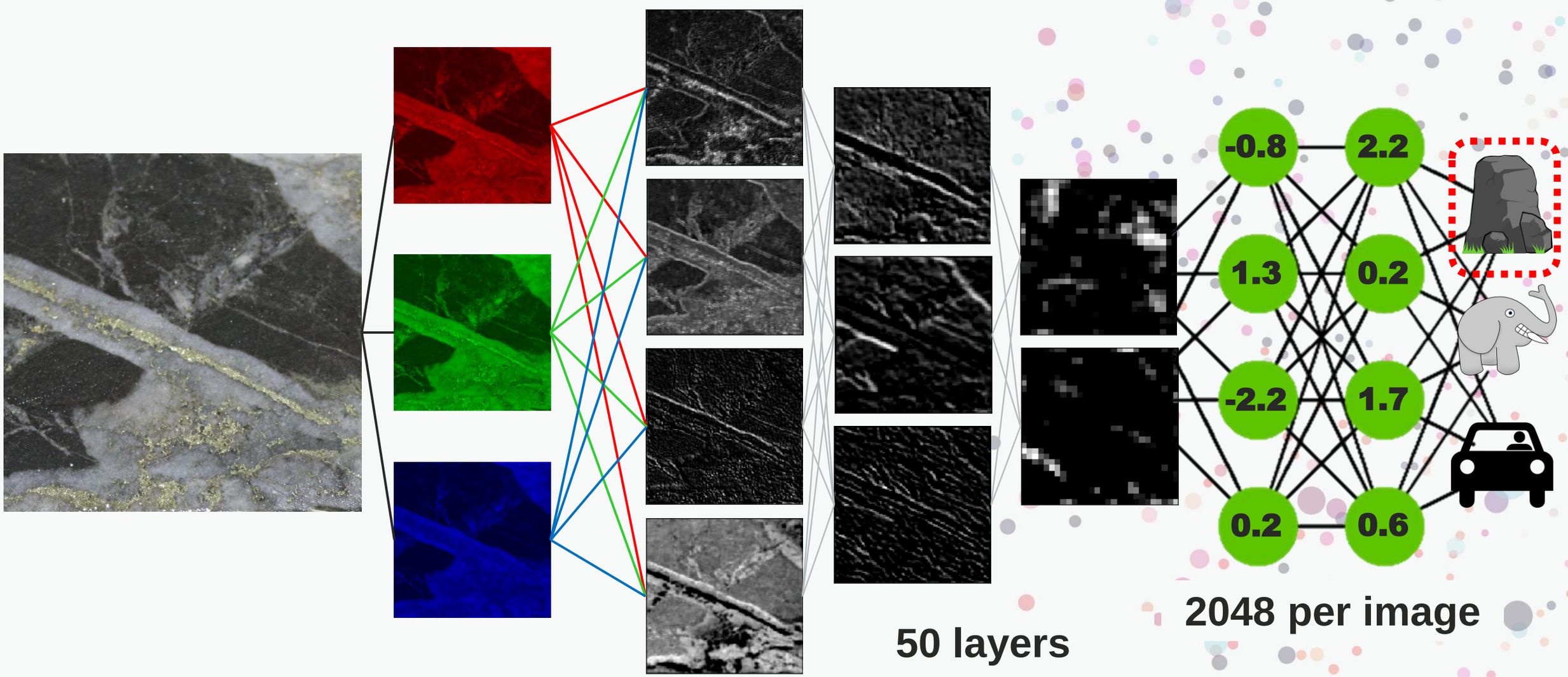
Infrared spectrum



THE BASICS OF CREATING A MACHINE LEARNING MODEL TO ASSIST LOGGING

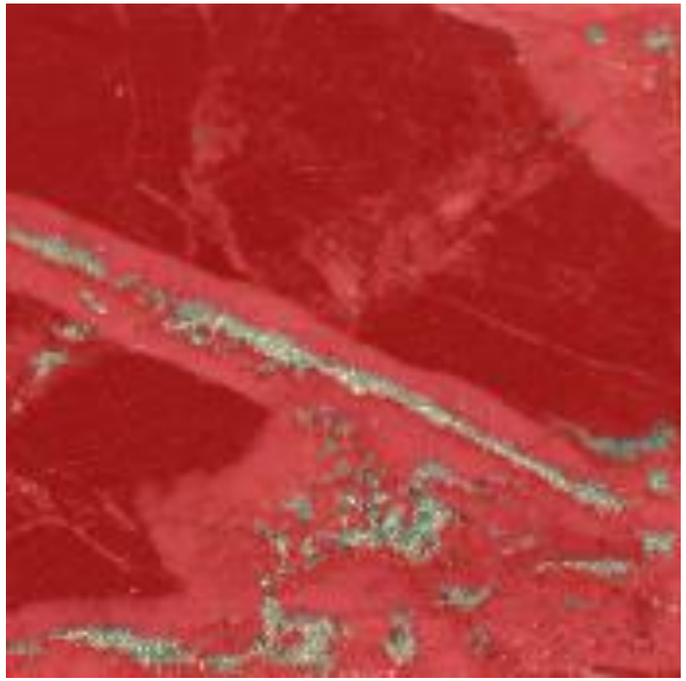


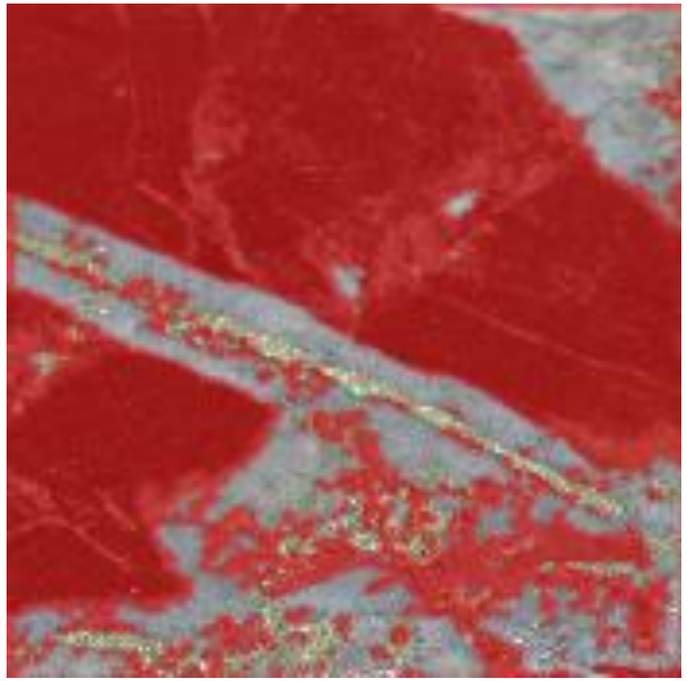
Computer vision: Convolutional Neural Networks



MAPPING LITHOLOGIES WITH RGB

- **Limitations with ML using RGB photography**
  - Masking, problems with photography (sun/angle)
  - RGB photography can only see ~380-760nm; information is limited to what a human eye can see!
  - The machine has all the same difficulties of differentiating complex mineralogy and lithologies as humans.
- Importantly, **consistency** is improved by using ML.





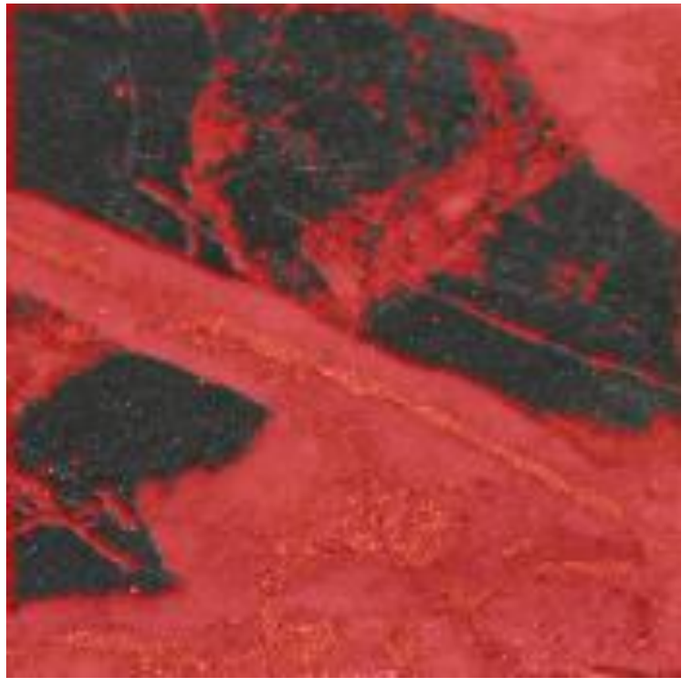
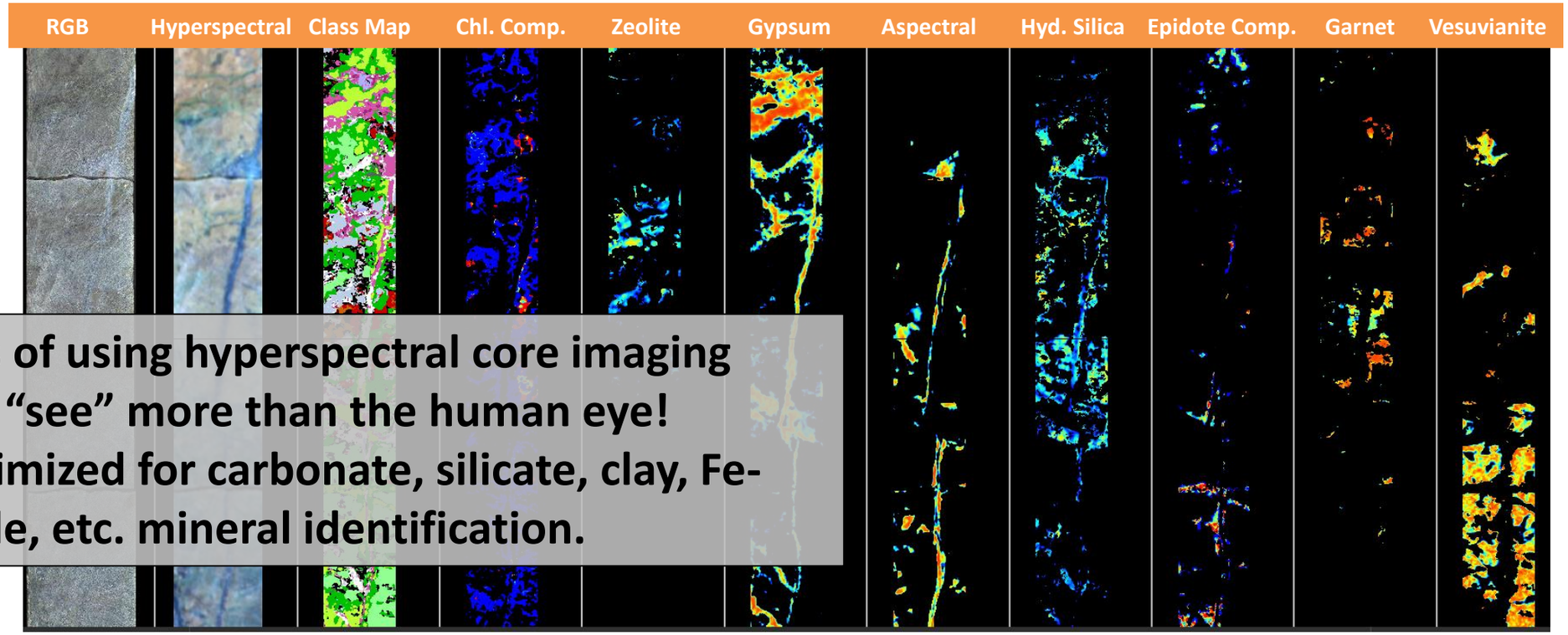


IMAGE CLASSIFICATION AND CORESCAN DATA

- Most machine learning workflows require large amount of high resolution and **consistent data**.
- Corescan collects approximately 200,000 pixels of data per meter. The HCI-4 will produce 800,000 pixels per meter.
- Significantly, Corescan data provides spatial relationships between the pixels in the form of an image; therefore we know which minerals exist as a given and the position of every other mineral.

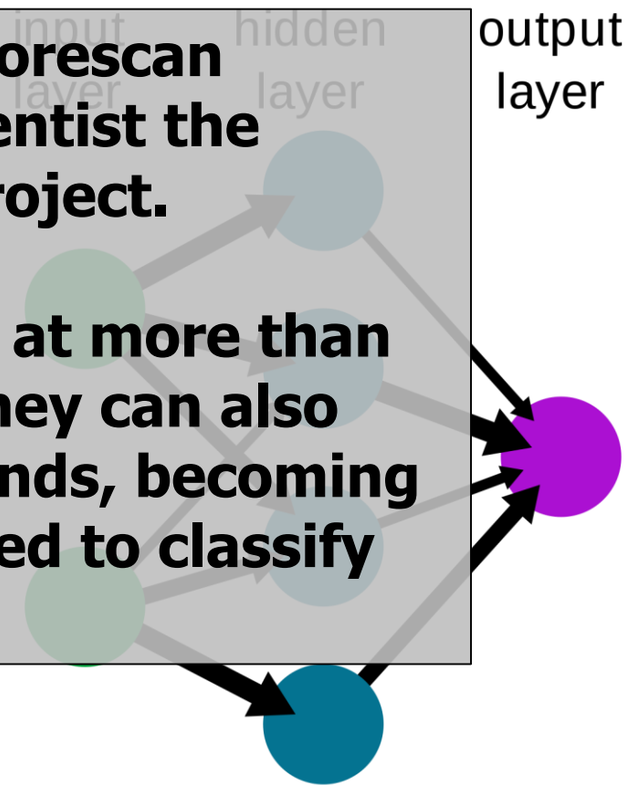


Advantages of using hyperspectral core imaging

- **Can "see" more than the human eye!**
- **Optimized for carbonate, silicate, clay, Fe-oxide, etc. mineral identification.**

MAPPING LITHOLOGIES WITH CORESCAN RGB + 512 BANDS + LASER PROFILER

A simple neural network



Project geologists in conjunction with a Corescan spectral geologist can advise the data scientist the optimal bands to use as inputs for each project.

This will create neural networks that look at more than the individual Corescan bands of data… they can also look at the interactions between these bands, becoming similar to ratios and gradients that are used to classify spectra!

USING MACHINE LEARNING TO LOG CORE

- **Warning:**
  - **High degree of geological knowledge required!**
  - **Will require an extreme amount of effort on the part of the staff geologists!**

Observational Logging

Current practice

Large amount of underutilized data

Augmented Logging

Taking the observations of the geologist and integrating them with quantitative data

First step is to identify major classes for automation (e.g. lithology, alteration)

Automated Logging

Driven by quantitative data that measures the physical properties of the rock

Trained by geologists to recognize everything that is deemed important

- **Will need to train models to recognize all classes that are of importance:**
  - **Lithology and alteration are the tip of the iceberg!**
  - **Veins, breccias, faults… whatever you log and is important to model will need training!**
  - **Re-evaluation of what is important may be required!**

Corescan would like to acknowledge Solve Geosolutions for their input in this talk.

Thank you all for attending.

SOLVE
GEOSOLUTIONS