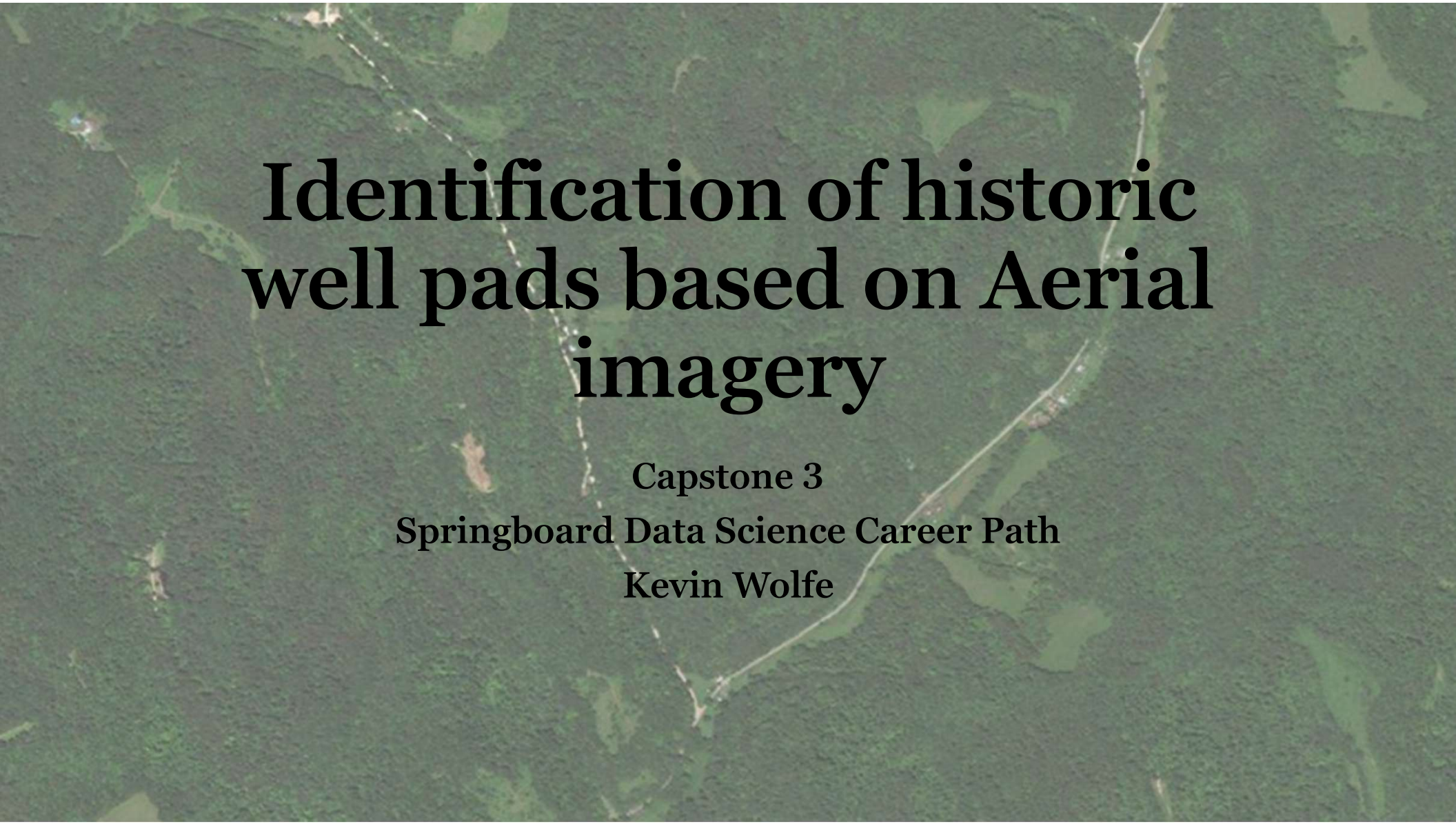


Identification of historic well pads based on Aerial imagery

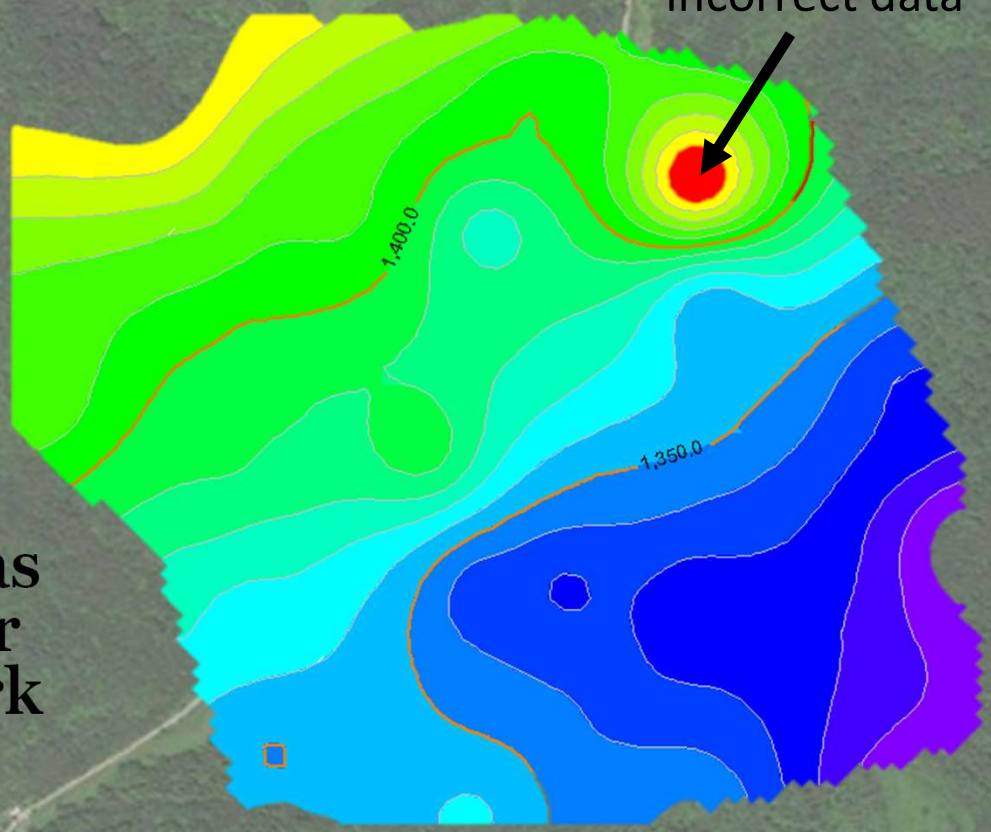
Capstone 3

Springboard Data Science Career Path

Kevin Wolfe

The Problem:

- Oil and gas mapping is hindered by poor ability to validate well locations for historic vertical wells
- Most mislocated wells are identified through “data busts” and bullseyes
- Each group mapping an areas must discover issues on their own and then redo their work after issues are corrected



Can computer vision provide a method for validating well locations?

A potential solution?

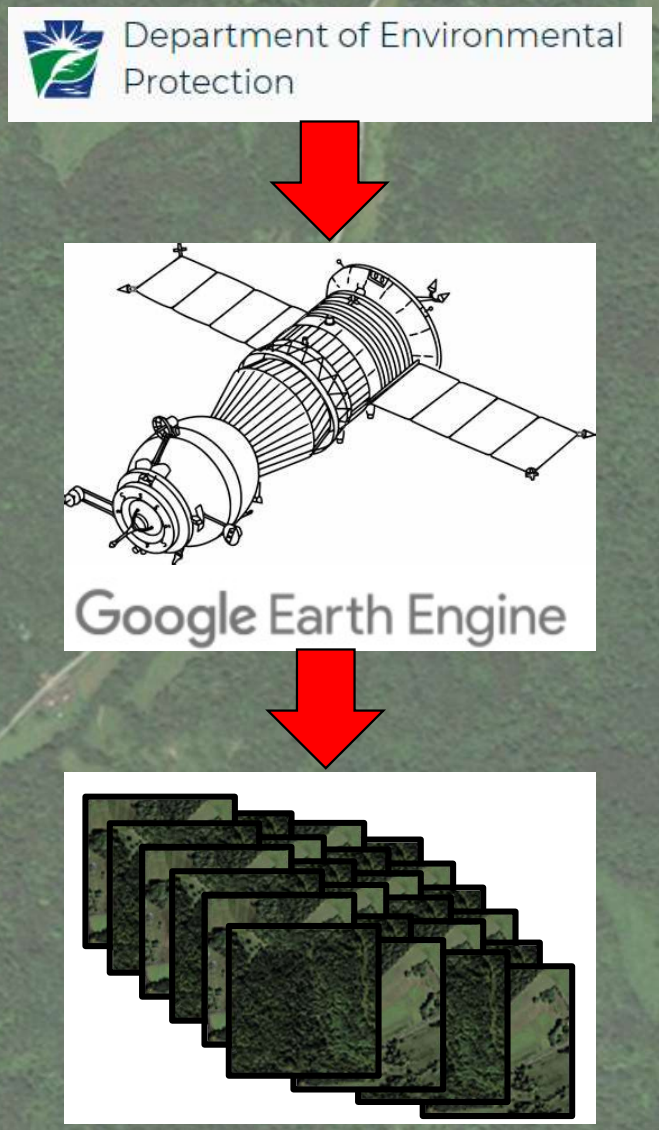
- Aerial photography is freely available from the US National Agricultural Imagery Program (NAIP)
- Goggle has provided a platform for viewing/exporting imagery data through Earth Engine
- Using well location data from state agencies NAIP data can be selected for further analysis

Location Not Visible

Location Visible

Image Data collection

- Well locations (lat/lon) for wells in Westmoreland Co. PA were selected for the base data in this experiment
- Locations were used to select a 500'x500' area around each wellhead
- NAIP imagery was exported for the selected area for each well



Reasons for a not visible location:

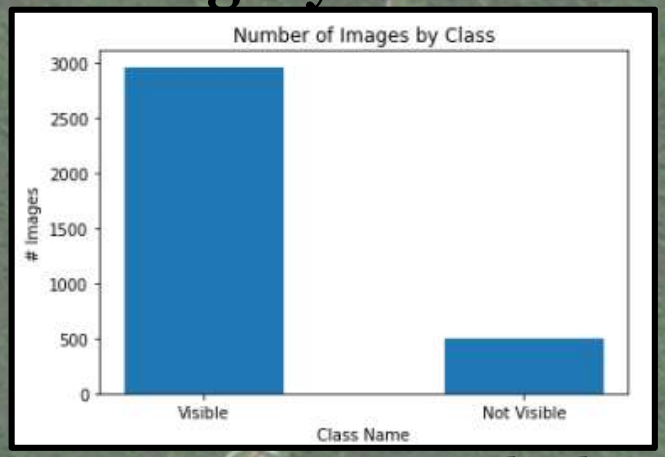
Well after plugging, will not be visible from the air

- Obscured –pipelines and lease roads should be visible even if surface is not
- Wrong location – the well was drilled somewhere else
- Plugged and Abandoned – the well used to be there but the operator has decommissioned it

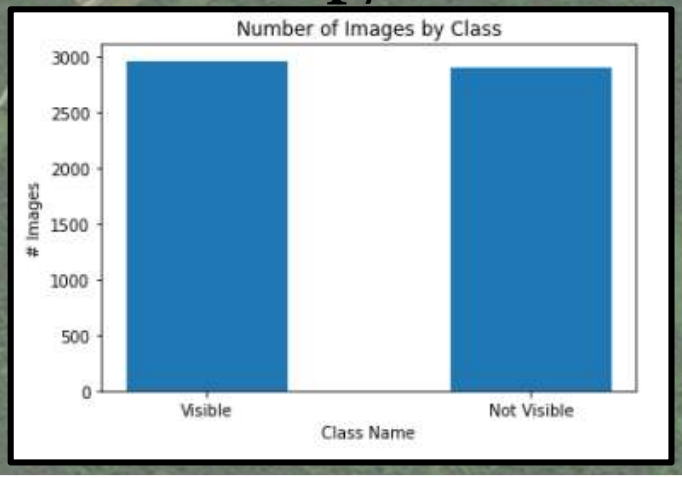
Image Dataset

- Well locations were utilized to locate and crop 3457 images.
- In 499 of the images, I could not clearly identify a well location or the auxiliary features of a well location
- These 499 images were rotated and flipped to create a balanced dataset of visible/not visible images

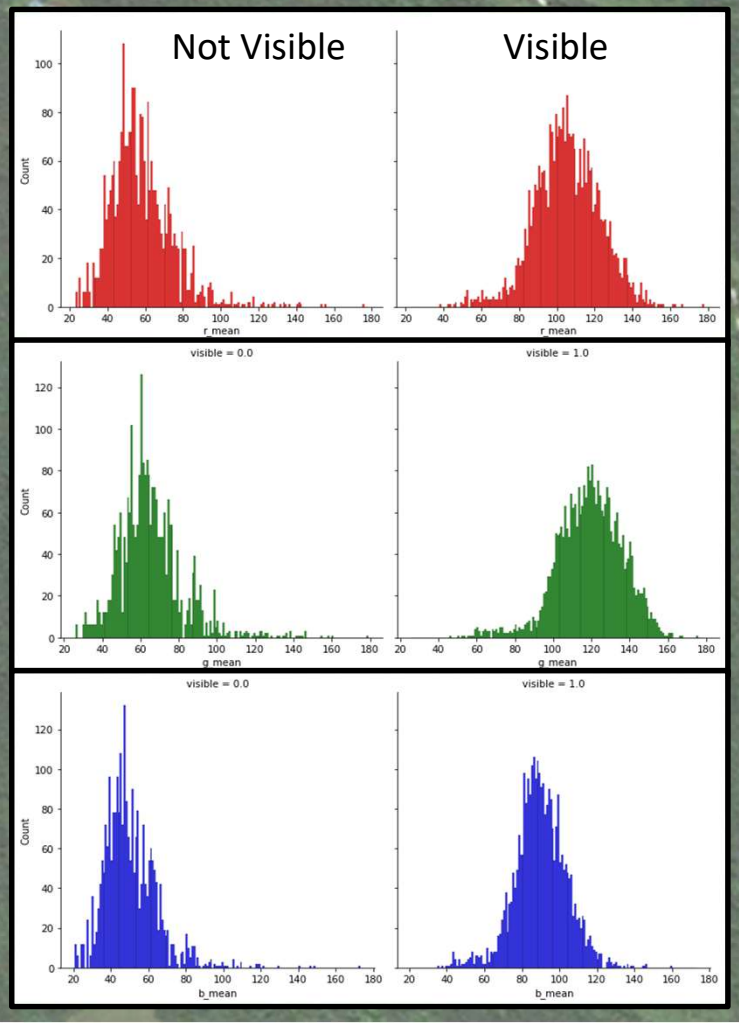
Initial dataset category balance



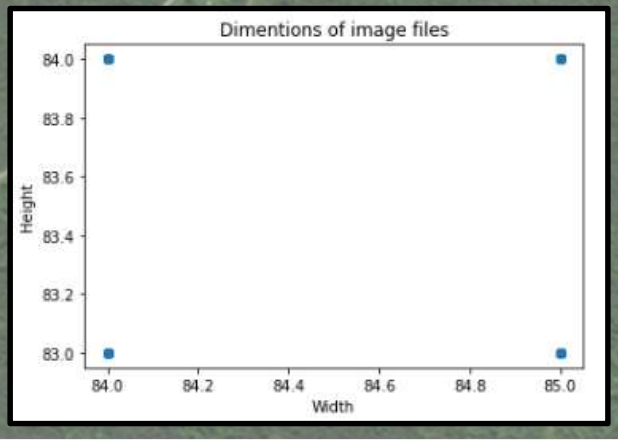
Dataset category balance after flip/rotation



Imagery data characterization



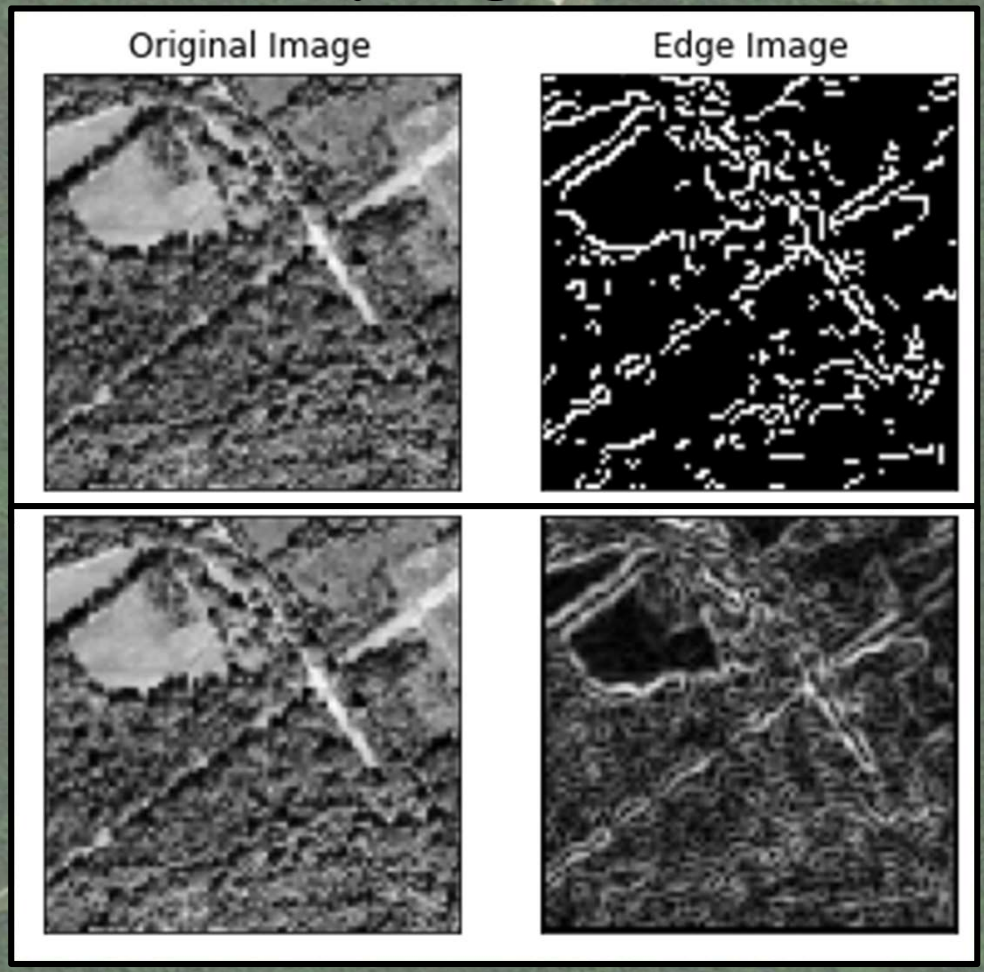
- Mean color band distributions are nearly normally distributed
- There is a distinct decrease in mean color intensity in the “Not Visible” images
- The shape of the image files are nearly constant



Edge Detection

- Canny and Sobel edge detection were tested to investigate the most reasonable window size model parameters
- Aperture sizes above 3 produced images that seemed to be random noise due to the size of the images
- The edge detection seems to highlight the lease road and pipelines associated with the well locations

Canny Edge Detection



Sobel Edge Detection

Well location Modeling

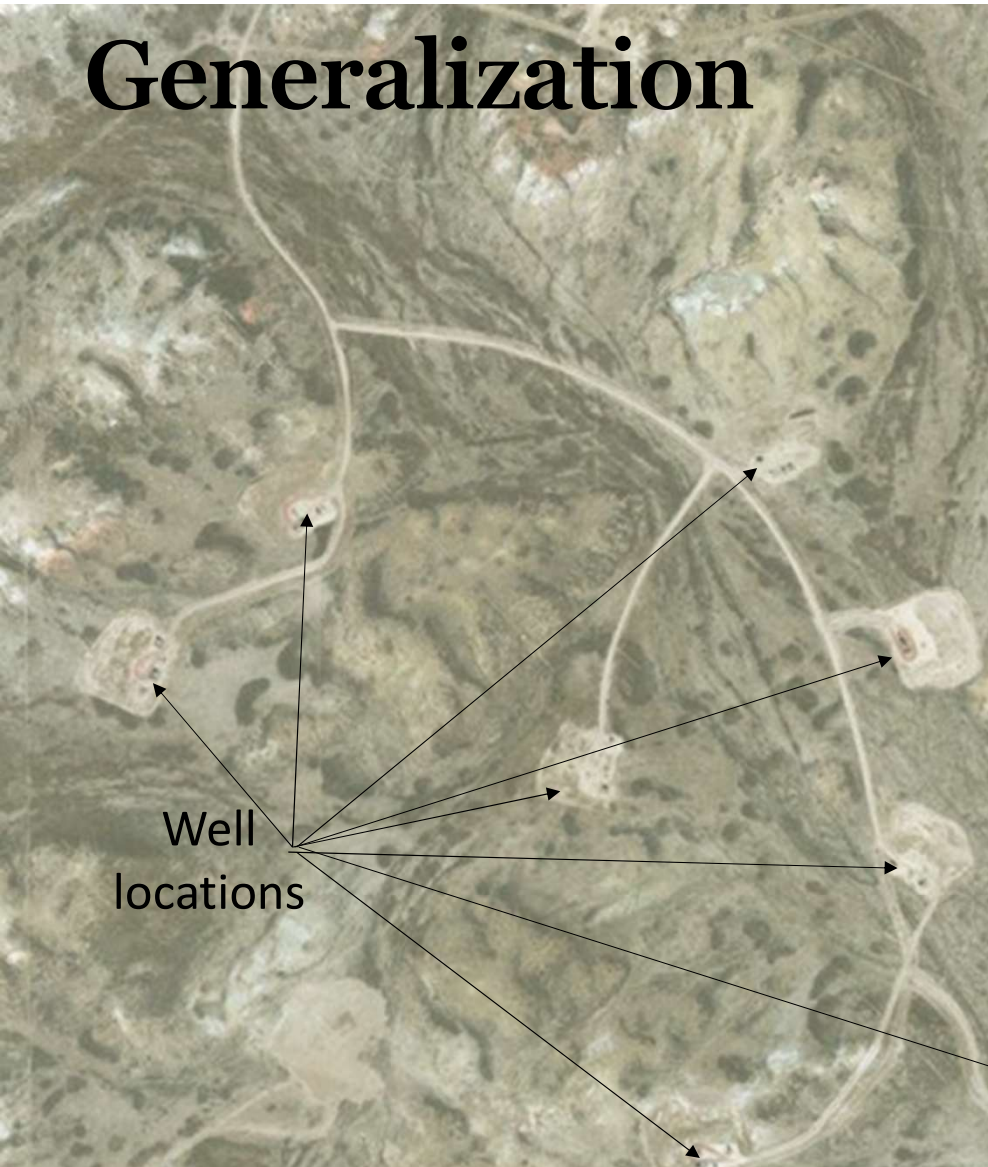
3 models:

- Dummy model – 51% accuracy
- Single Convolution – 97% acc.
- Dual Convolution - 95% acc.

The best performing model was the single convolution layer Keras model

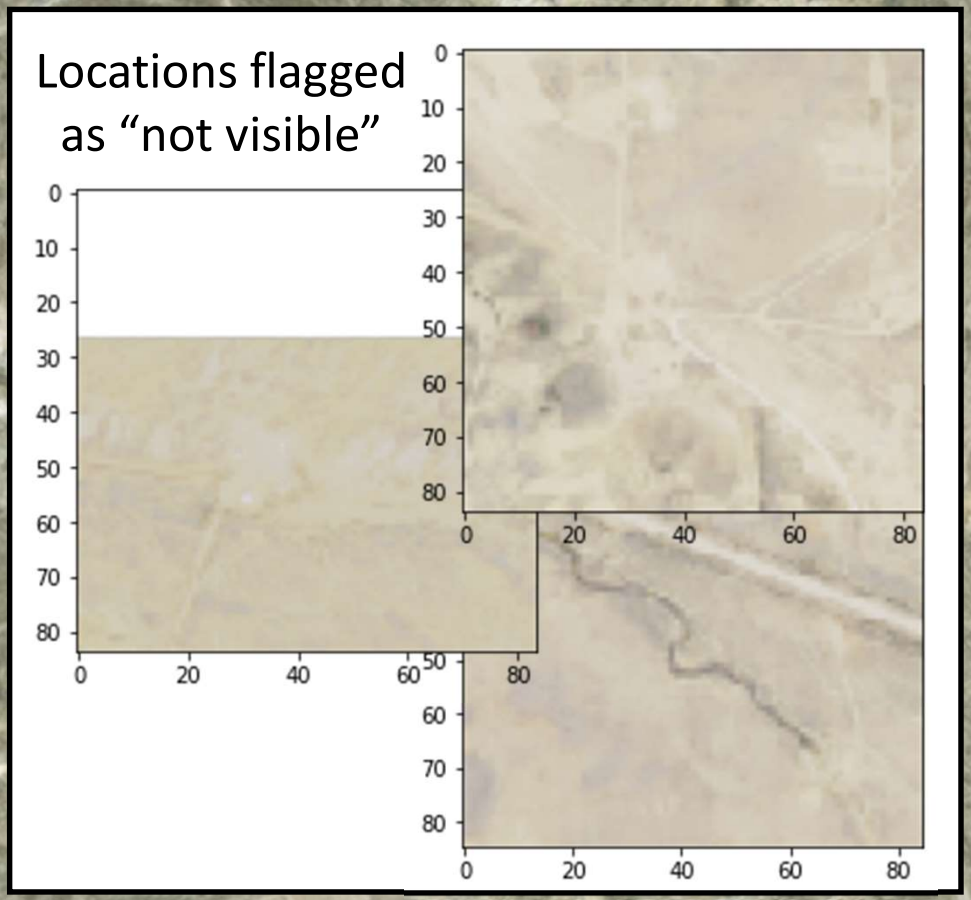
Model	Training Score	Testing Score	Precision 'visible'	Precision 'not visible'	Recall 'visible'	Recall 'not visible'
Dummy	0.516	0.511	0.557	0.456	0.548	0.465
Keras 1x Conv	**0.997**	**0.97**	**0.975**	**0.962**	**0.968**	**0.969**
Keras 2x Conv	0.972	0.955	0.995	0.905	0.915	0.994

Generalization



- To evaluate the ability of this model to be utilized in a larger area imagery data was selected from 2 new areas
- Sweetwater Co. WY was used for an oil and gas productive area with a significantly different landscape
- The area surrounding Gooding, ID was used as an area with no oil and gas activity (all locations should be “not visible”)

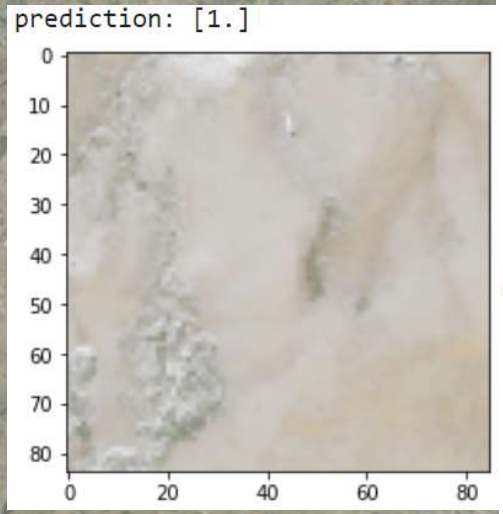
Generalization – Sweetwater Co. WY

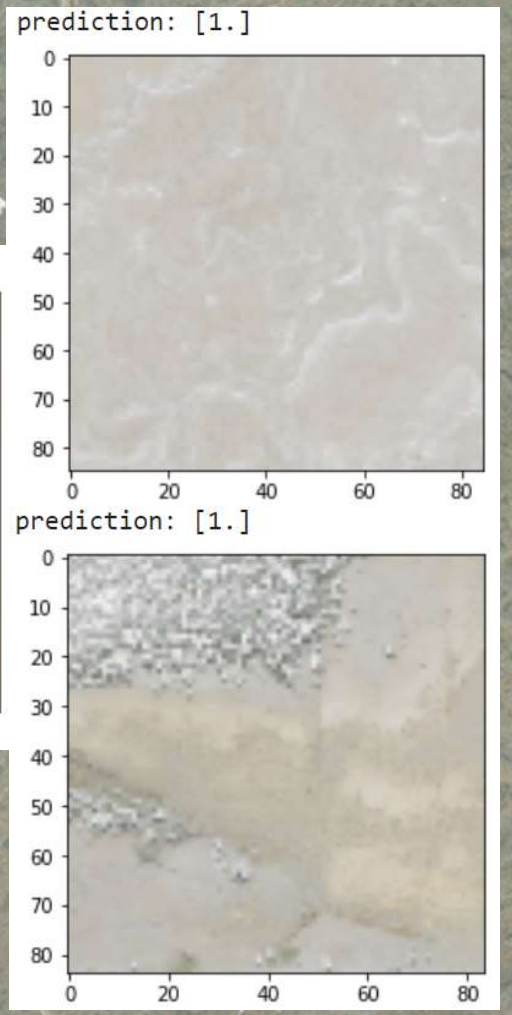


- Due to the desert conditions and slow growing vegetation well locations are visible for 30-50 years after reclamation
- Westmorland Co. model predicted only 75% of the well locations were in fact a location
- 25% of locations being mislocated or not visible is highly improbable in this landscape

Generalization – Gooding, ID

- There are no well locations in this area
- Apart from irrigated areas the vegetation is similar to Sweetwater Co. WY
- The model predicted 75% of the randomly selected locations were well locations.
- This is a significant failure





Conclusions and Recommendations

- While the model performed very well against test data in Westmorland Co. PA the poor performance in different landscapes indicates additional work will be needed to utilize this model in areas outside of western PA
- To enhance the ability of models like this to predict the presence of a well location, a more varied landscape would be required in the training data
- My recommendation for additional work would be to expand this project to utilize locations from as many vegetation, climate, and topographic landscapes as reasonable