FIRE: THE FIRST-YEAR
INNOVATION & RESEARCH
EXPERIENCE
CLOUD COMPUTING

Image Segmentation and Inpainting for Cloud Type Multilabel Classification

Anthony Louie

1. Introduction

One of the stream's goals was to find a way to identify amount obstructions in the GLOBE Cloud's data[1] for potential removal. Obstructions such as trees and buildings plague the data which is used for things like creating classification models. This could hurt the performance of our models. One approach from was to use otsu thresholding to try and identify those obstructions but the data is too complex to use non-learning methods. So, the purpose of this research is to use a learning based approach to gather this obstruction data and use it for things like training a multiclass and/or multilabel classifier to identify cloud types.

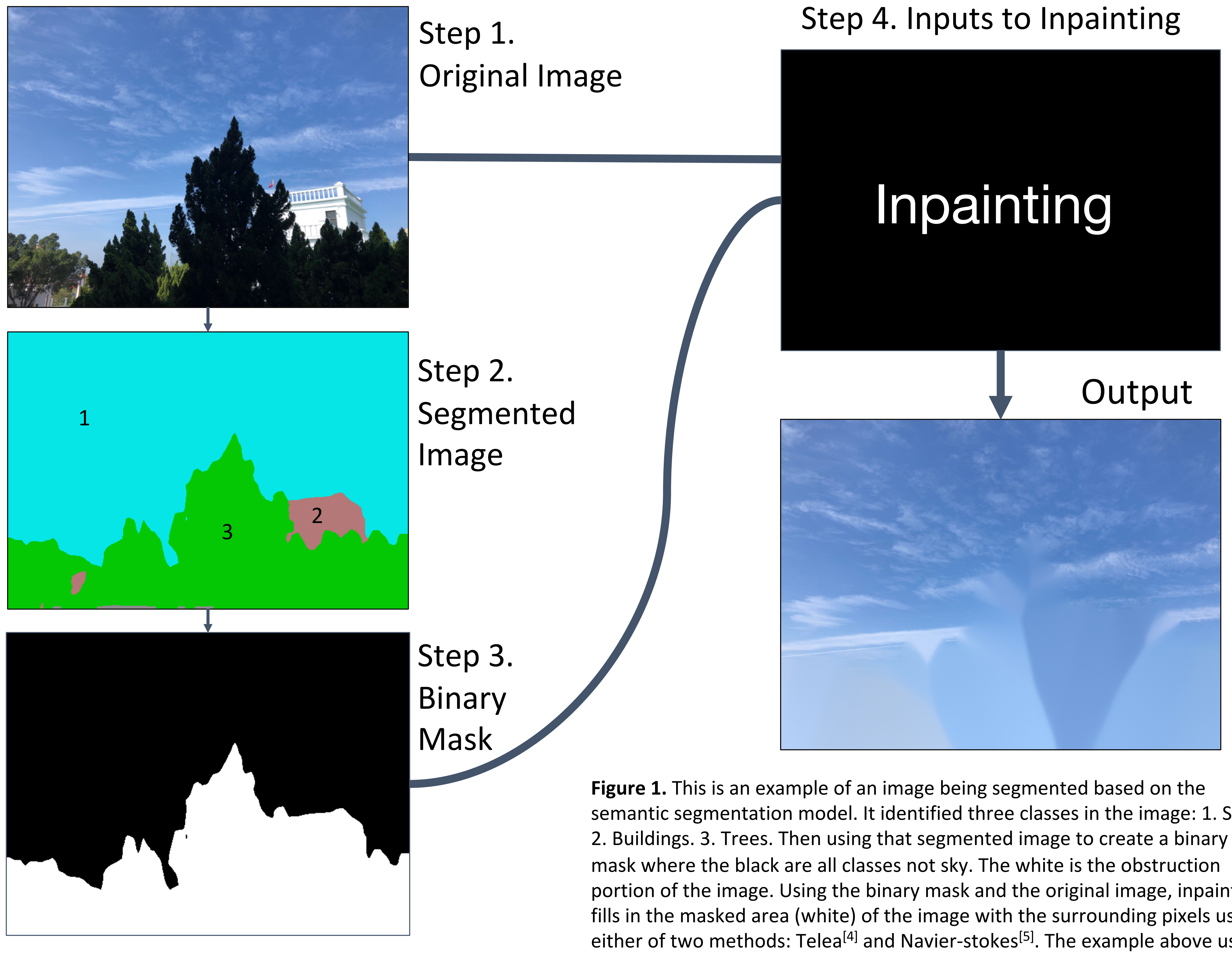
2. Tools & Data

- **NASA GLOBE Cloud:** Datasets from NASA's GLOBE program containing ground based images and metadata such as time and location
- **Semantic Segmentation Model**: A fully connected neural network to segment a single image into multiple identified classes.[2][3]
- **Inpainting**: Process of filling in parts of image identified by a binary mask based on surrounding pixels of image. There are two different methods.
- **Random Forest Classifier**: Classification model that uses many decision trees for an averaged prediction

6. Literature

1. Amos, H. M., Starke, M. J., Rogerson, T. M., Colón Robles, M., Andersen, T., Boger, R., Campbell, B. A., Low, R. D., Nelson, P., Overoye, D., Taylor, J. E., Weaver, K. L., Ferrell, T. M., Kohl, H., & Schwerin, T. G. (2020). Globe Observer Data: 2016–2019. *Earth and Space Science*, *7*(8). https://doi.org/10.1029/2020ea001175
2. Scene Parsing through ADE20K Dataset. B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso and A. Torralba. Computer Vision and Pattern Recognition (CVPR), 2017.
3. Semantic Understanding of Scenes through ADE20K Dataset. B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, A. Barriuso and A. Torralba. International Journal on Computer Vision (IJCV), 2018.
4. Telea, Alexandru. "An image inpainting technique based on the fast marching method." Journal of graphics tools 9.1 (2004): 23-34.
5. Bertalmio, Marcelo, Andrea L. Bertozzi, and Guillermo Sapiro. "Navier-stokes, fluid dynamics, and image and video inpainting." In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I-355. IEEE, 2001.

3. Inpainting



Step 1. Original Image

Step 2. Segmented Image

Step 3. Binary Mask

Step 4. Inputs to Inpainting

Inpainting

Output

Figure 1. This is an example of an image being segmented based on the semantic segmentation model. It identified three classes in the image: 1. Sky. 2. Buildings. 3. Trees. Then using that segmented image to create a binary mask where the black are all classes not sky. The white is the obstruction portion of the image. Using the binary mask and the original image, inpainting fills in the masked area (white) of the image with the surrounding pixels using either of two methods: Telea[4] and Navier-stokes[5]. The example above uses Telea.

4. Classification

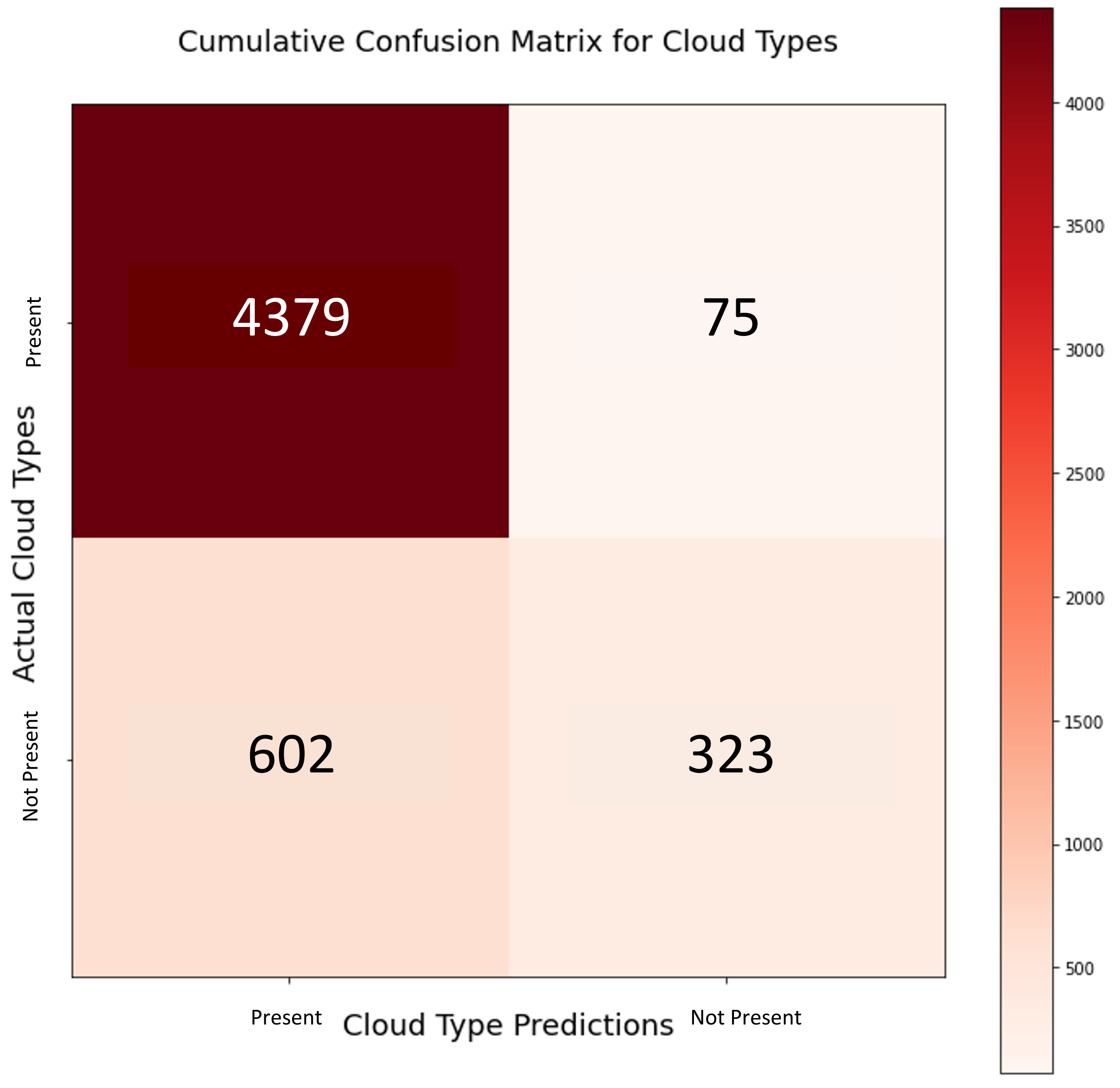
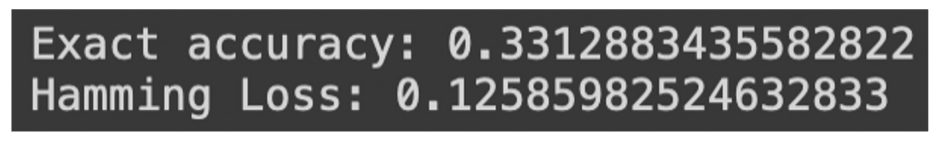


Figure 2. This confusion matrix shows the performance of a random forest classifier trained on the metadata from the GLOBE Clouds observations along with the obstruction ratios (white pixels) from the binary masks. Parameters max depth and max features were set to 14 and log2 respectively.



```
Exact accuracy: 0.3312883435582822
Hamming Loss: 0.12585982524632833
```

Because the classification problem is a multi output or multi label problem, meaning that the predictions can contain more than one class. Exact accuracy may not be the best metric to assess model performance. We have 11 different labels representing cloud types and if one single cloud type is misclassified that entire prediction is wrong. Hamming loss is a way to look at individual class predictions. It is the fraction of wrong individual labels to the total number of labels. Here we have 12% of the predicted labels being wrong. Which means our model predicts 12% of individual classes wrong or it has a 88% accuracy for the multiclass case.

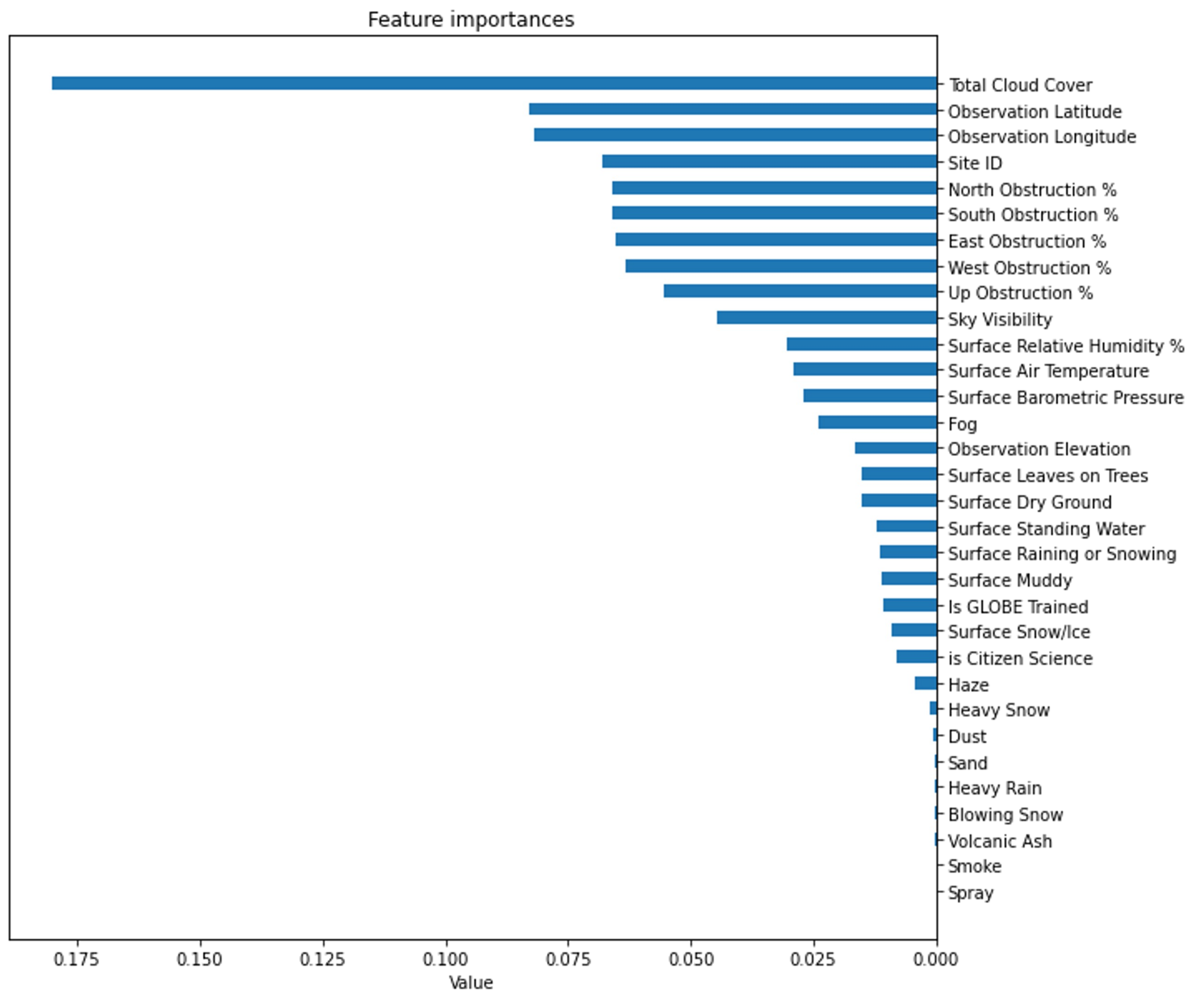


Figure 3. This plot is the feature importance for the random forest classifier. The higher the value, the more weight that feature holds. Meaning the better that feature is at being a split parameter for the decision trees.

5. Discussion

The idea of using obstructions as a part of the feature vectors for cloud images doesn't make much sense because obstructions are unrelated to cloud types. They fall under human error and are not a part of the data distribution. This is further supported by the fact that when training a random forest classifier without the obstruction ratios in the data, we actually get higher exact accuracy and lower hamming loss.

So instead of using the obstruction ratios from the raw binary mask information, we can use the mask to inpaint the image and gather direct image data such as image contrast and red-green mean difference, which is the difference between the average number of red pixels and the average number of green pixels. This should work because the inpainted version of the image does not have the obstructions like trees so it won't affect green average and so on. The performance of this classifier is better than both of the others, but only by 1%.

A potential future direction for this project is to use a deep learning approach, such as neural networks to explore this data. Techniques such as convolutions may also be impactful for image data. But now we have a process to clean up the existing GLOBE Cloud data to be used for other research.