FIRE: THE FIRST-YEAR INNOVATION & RESEARCH EXPERIENCE

Engaging Freshmen Undergraduates in AI using Cloud Imagery and Model Output

Alexandra L. Jones (aljones4@umd.edu) at University of Maryland

Summary

The First-year Innovation & Research Experience (FIRE) program provides first-year University of Maryland students a 3-semester faculty-mentored research experience that drives accelerated professional development. Students use authentic tools of the trade to attack broadly relevant research objectives. The uncertainty of authentic research allows students to thrive through challenge and develop career competencies that employers are seeking. Students in the 2020 cohort of the "Cloud Computing" research group worked on several projects with the goal of applying machine learning to our datasets of cloud images, satellite data, and model output. Student teams applied either supervised (Google's Teachable Machine) or unsupervised learning tools (k-means clustering or watershed-ing) to one of these datasets to learn something about the black box algorithms they were employing as well as the underlying physical patterns in the data. Through this work students gain hands on experience with atmospheric datasets as well and machine learning tools and the transferrable skills of project management, problem solving, collaboration, and communication.

Cloud Photographs

One team worked to cluster feature vectors extracted from GLOBE Clouds photos using the Convolution Neural Network VGG16. Fig. 1 shows the 5 most similar images to each of 3 query images. A second team trained Google's Teachable Machine to automatically categorize GLOBE Clouds photos as one of 10 basic cloud types. The confusion matrix for their trained model shows predictability (Fig. 2).

Satellite imagery of Polar Lows

Using a dataset of satellite imagery of polar lows collected from NASA Worldview one team extracted feature vectors from the polar low images using VGG16 to see if they could identify distinct clusters relating to polar low characteristics. One of those clusters is shown in Fig. 3. The entire dataset projected into 2 dimensions using TSNE shows some clustering by image color (Fig. 4). Another team trained Google's Teachable Machine to classify images as either high confidence, low confidence or not a polar low (Fig. 5). The model's confusion matrix shows some difficulty in distinguishing between high and low confidence images.

Model Output of Convective Cores

One team visualized vertical velocities corresponding to convective cores in 3D (Fig. 6) and attempted to automate the identification of unique cores in space and time using watershed-ing (Fig. 7)

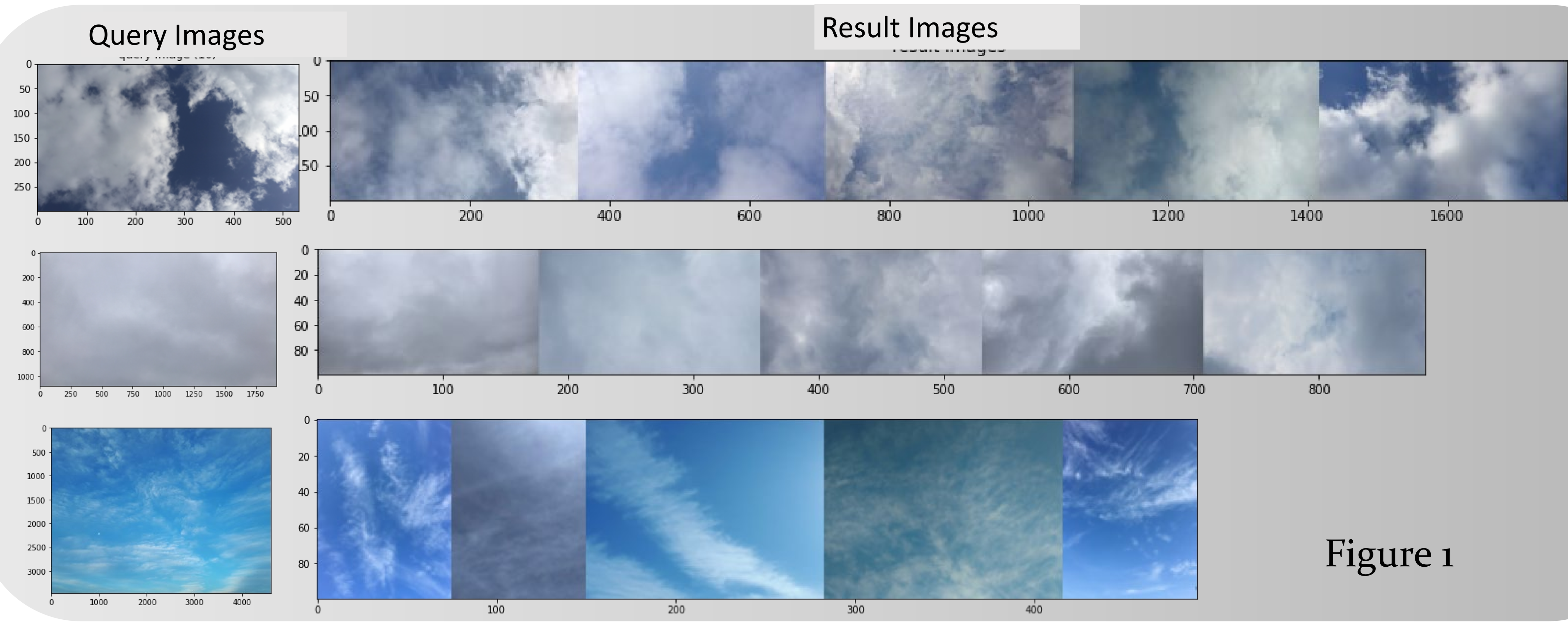

Figure 1

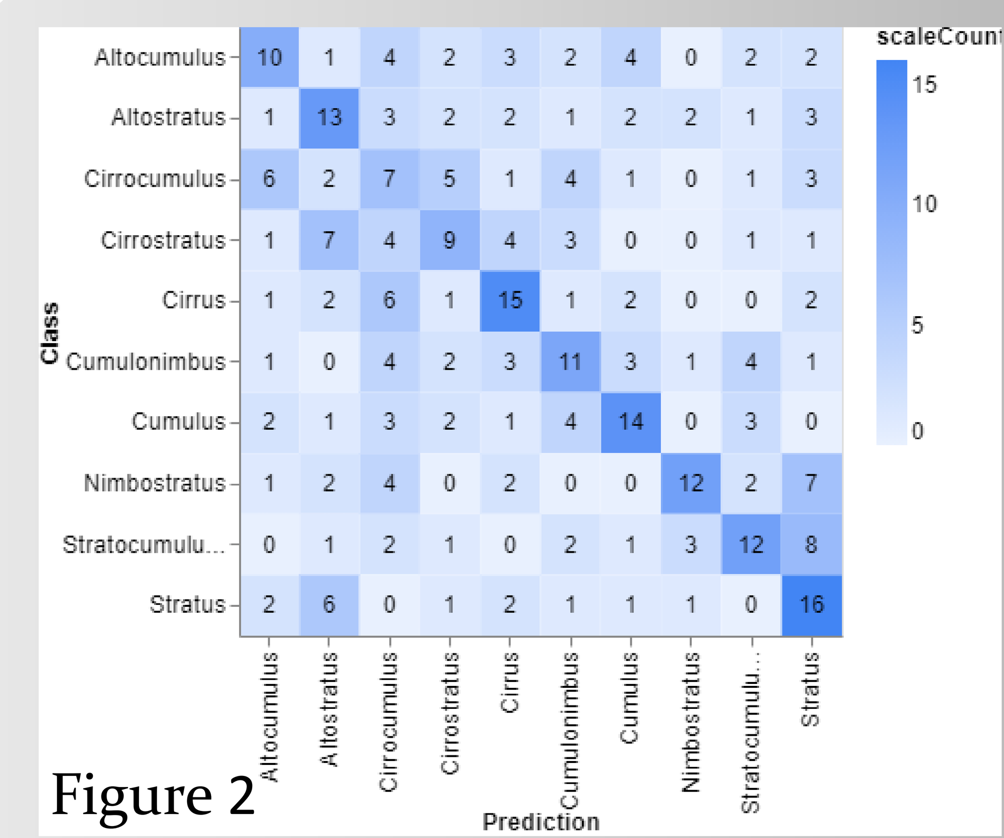

Figure 2

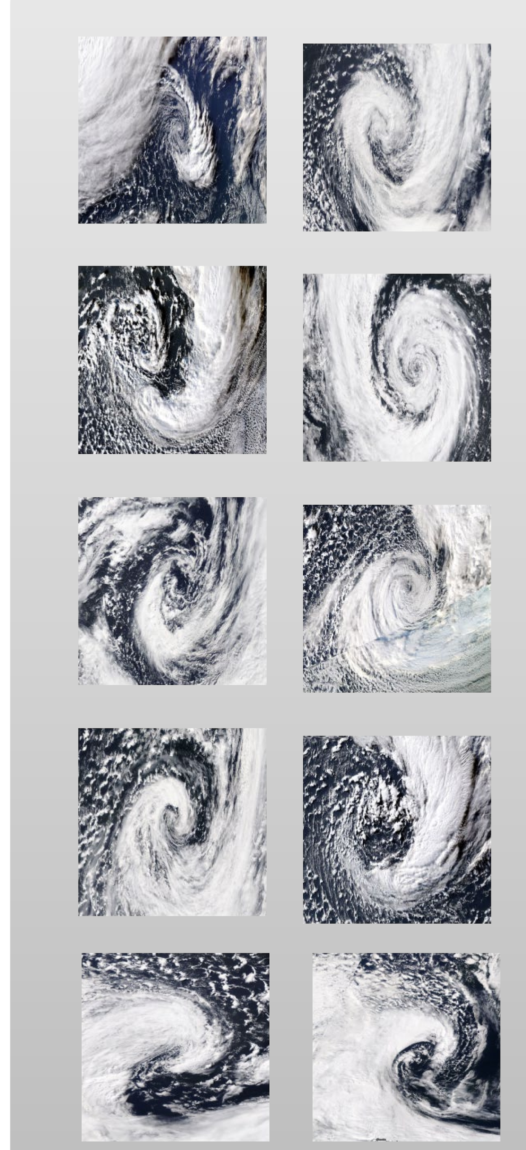
Figure 3

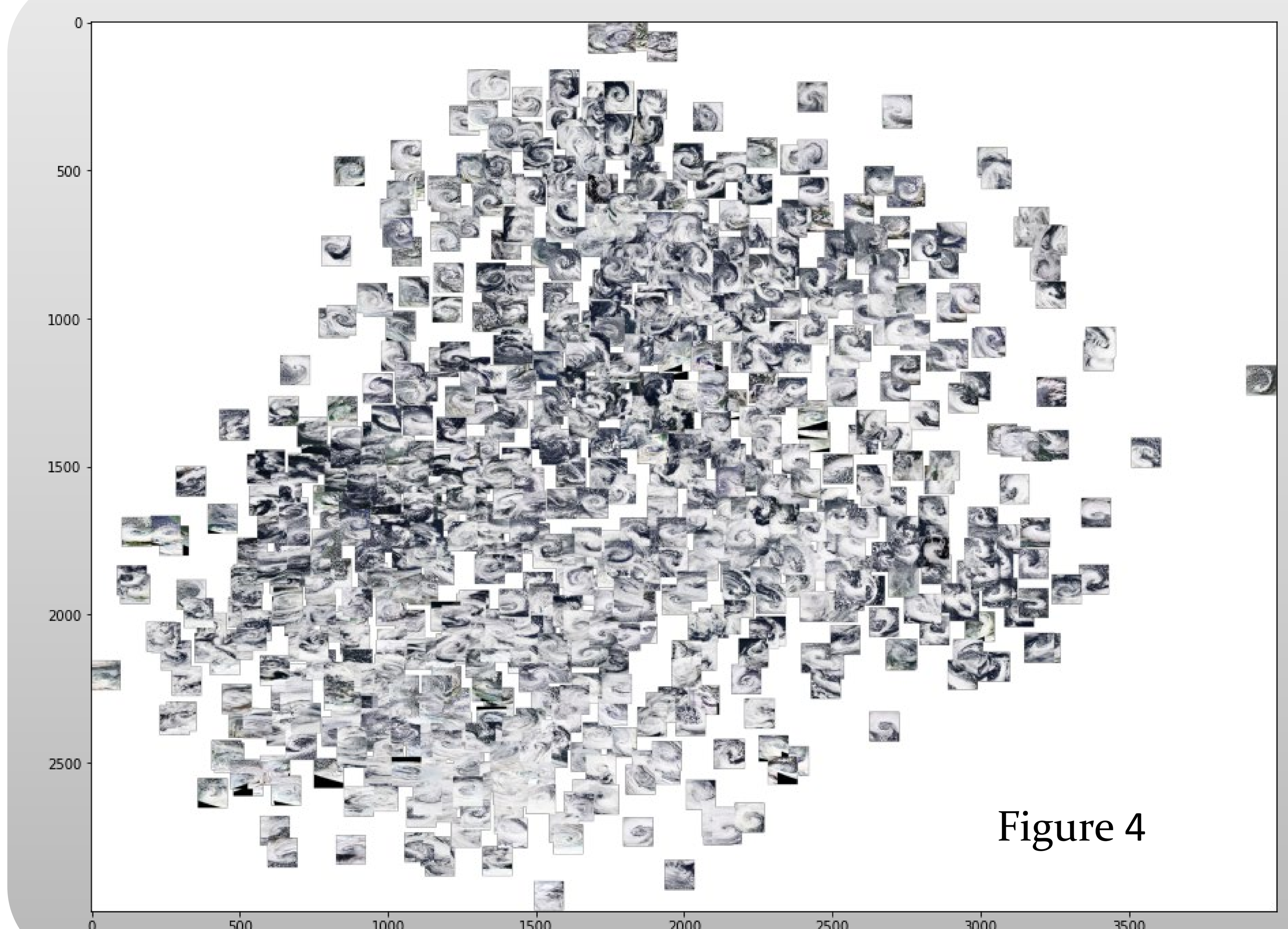
Figure 4

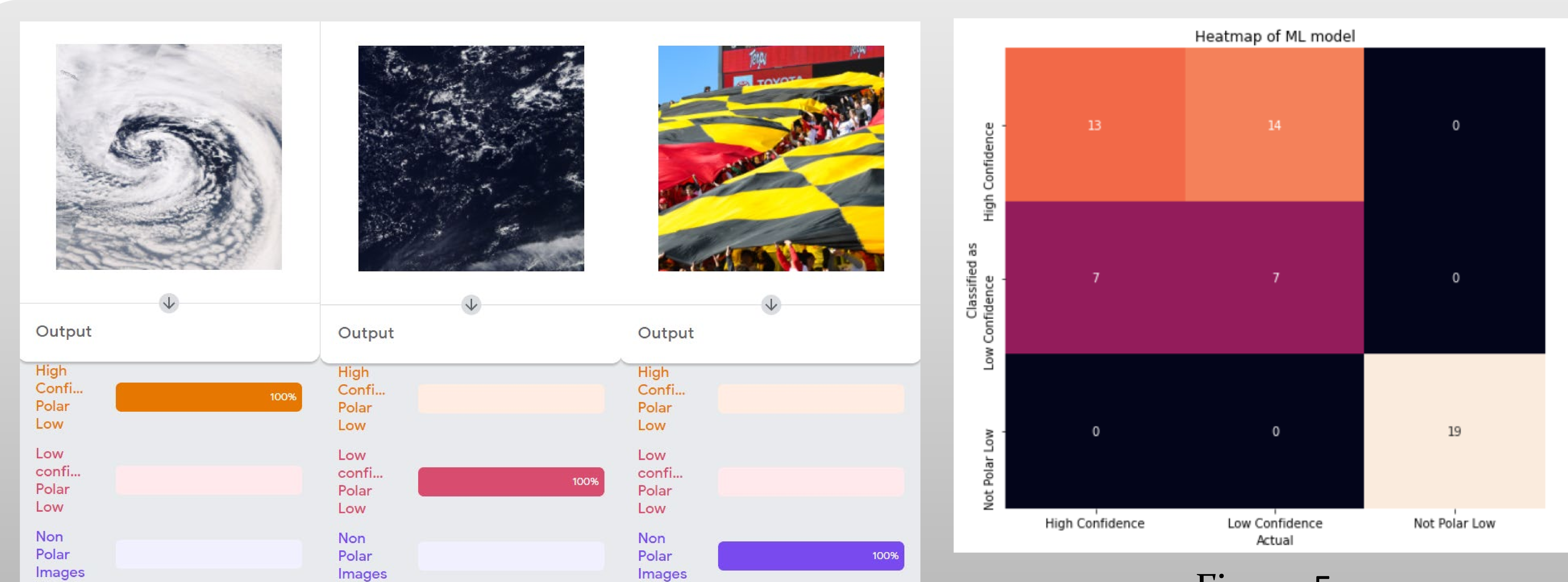

Figure 5

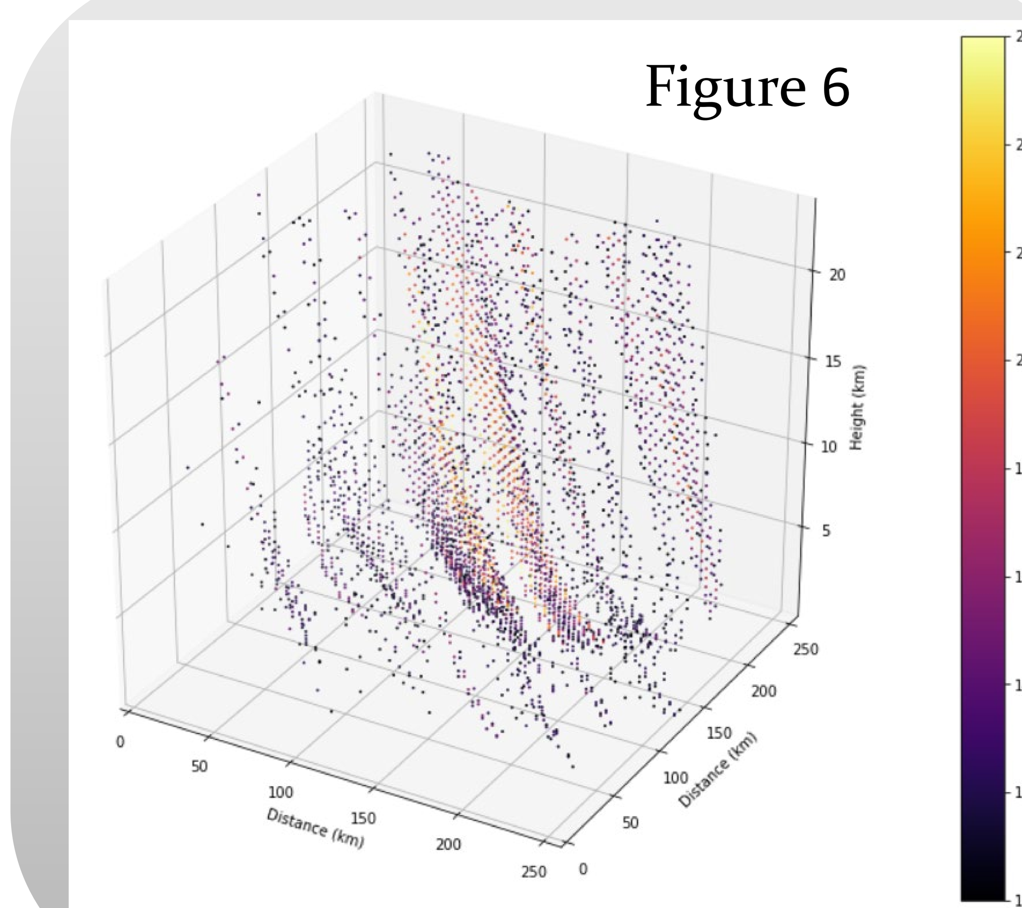

Figure 6

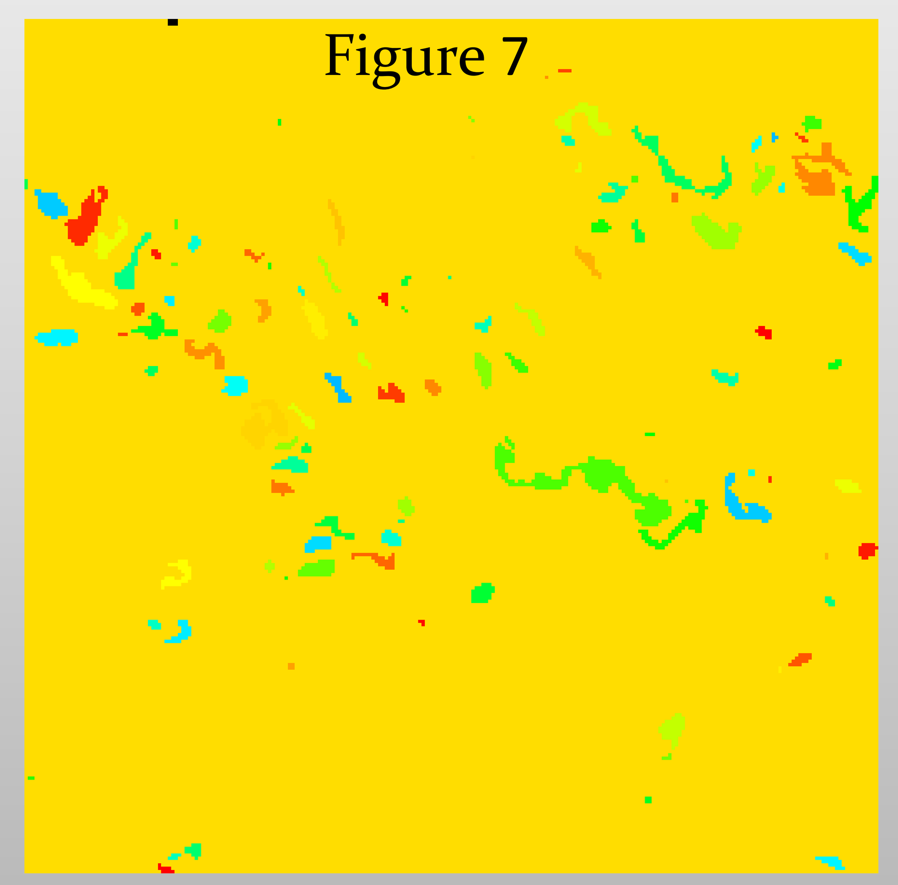
Figure 7

Acknowledgements

This research was conducted by undergraduate members of the 2020-cohort of FIRE-CC. We would like to thank our collaborators Tianle Yuan and Xiaowen Li for project inspiration and data.