Piecing Together the Chaos Puzzle: A Machine Learning Model for Chaos Block Identification on Europa.

A. C. Mills[1,2], E. Duncan[2,3], C. A. Nixon[2], D. Trent[4], J. Santerre[3], I. Neidel[2,5], A. Larsen[6], and C. C. Walker[2,7], [1]Department of Geosciences, Baylor University, Waco, TX, alyssa_mills1@baylor.edu, [2]NASA Goddard Space Flight Center; [3]University of California, Berkeley; [4]NASA Langley Research Center; [5]Yale University; [6]Inntopia; [7]Woods Hole Oceanographic Institution.

Introduction: Of all the ocean worlds in our solar system, Europa has been the earliest and longest in the focus of the planetary community's search for the ingredients of life, and the quest to answer if such a world was or currently is habitable. To know if the ingredients for life exist/existed, areas where material exchange of the subsurface with the surface must be found. Such areas on Europa are 'chaos regions' which make up about a third of the moon's surface. Chaos regions are thought to be sites of current or recent geological activity, which means they can act as windows into internal processes and help us determine how material exchange occurs from the subsurface to surface. Thus chaos regions are amongst the best sites for investigations into the habitability of Europa.

Although chaos are known to be geologically active sites on Europa and important in investigating habitability, the formation mechanism(s) for these terrain has yet to be pinpointed. [1] and [2] previously posited that the observational constraints of chaos cannot be met by a single formation mechanism model, e.g. melt through, brine mobilization, partial melt, lens collapse [3, 4, 5, 6]. This means we do not know the true formation mechanism of chaos and inherently cannot fully know the interior processes involved to bring material from the ocean to the icy crust.

In order to begin to understand the enigmatic question of chaos formation, observations may provide vital clues as to what interior processes are involved, specifically considering the morphology, appearances and orientations of blocks inside chaos. [7] reconstructed the translations and rotations of chaos blocks inside Conamara Chaos working backward from observed to their probable original locations, formulating the conclusion that the surface must be warmed from some internal process for the movement to occur. [2] mapped various chaos regions in the Galileo SSI RegMap images to compare morphology, perimeter, block areas, and chaos area to each other to find trends between the different parameters.

To extend upon previous work, our goal is to map and analyze chaos regions all over Europa from varying missions, eventually utilizing future mission imagery from ESA JUICE and NASA Europa Clipper. Such a task can be accomplished by traditional methods where one or more scientists work manually to identify and locate features, however such methods have inherent biases introduced by each mapper and in addition are very time consuming. As an alternative, here we present a machine learning model that we devised to automatically identify blocks throughout Europa's chaos regions with currently existing imagery but could be extended to also ingest any future imagery. With the machine learning model identification of chaos blocks and their attributes like location, perimeter, and area, we can then easily compare any movement of blocks between past and future mission images. We will also be able in future to extend the geophysical interpretation of chaos based on the blocks in a similar fashion as [2] in an attempt to better understand how chaos may have formed.

Methods: *Training Dataset:* The training dataset was constructed by mapping various chaos regions and their blocks in the Galileo SSI RegMaps. These images were obtained from the USGS Astrogeology Science Center and chosen due to their latitudinal distribution along the trailing and leading hemispheres as well as their relative similar resolution of 250 m/pixel. Within the 92 photogrammetrically-controlled global mosaics, we mapped an additional 11 chaos regions from the original set mapped by [2], using the same definition for a chaos region. Within each chaos region are the chaos blocks which are defined to have: positive topographic relief, polygonal shape with clear boundary (distinctive shadow difference from matrix), and at least 2 km in length. Within this training dataset are attributes about the chaos such as morphology, but for the preliminary machine learning model we focus on identification of a specific chaos morphology, platy chaos (larger fragments with distinctive surface markings).

Machine Learning Model: With a robust fully labeled dataset of blocks within chaos regions in the Galileo SSI RegMap images, we started matching data from the training dataset to the input layer of our machine learning model (Mask R-CNN). In order to start the matching process, we preprocessed the input data from .tiff image files and .shp files to 750 dpi resolution images as well as cropped them to 50×50 km images to avoid jagged image edges and gaps, so they could be loaded onto the Google Cloud Platform. The preprocessed images were then split into 80% training, 10% testing, and 10% validation sets as per standard industry practice.

For the training process, the Mask R-CNN proceeded through the PyTorch architecture which involved batching. We augmented the training dataset into 4 batches of images with a binary cross-entropy loss function. Loss converged after 20 epochs. The model infers a label of "ice block" in yellow and "background" in purple for each pixel of a given image as shown in **Fig. 1**. Within the training process, reweighting occurs automatically by the model through gradient descent.

To evaluate the performance of the model, we used two metrics:

1. How well does the model identify a chaos block within a given bounding box (binary classification)?
2. Given a positive classification for the existence of a chaos block, how well does the model classify each pixel within the bounding box (instance segmentation)?

The first metric is based on accuracy, precision, and recall as defined by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Where TP is true positive identifications; TN is true negative identifications; FP is false positive identifications; and FN is false negative identifications. The second metric is evaluated with an intersection over union (IOU) score which assesses how well the model classifies each pixel:

$$IoU = \frac{True\ Label\ \cap Model\ Label}{True\ Label\ \cup Model\ Label}$$

Results/Conclusion: The accuracy of the model currently is 0.75, indicating that the model correctly identifies 75% of all pixels in the test set, regardless of their true label. The precision score indicates that, of all the positive classification made by the model, 68% were correct. The recall scored indicates that, of all the true positive labels in the test set, the model correctly identified 75% of them. These results show that the model missed smaller chaos blocks, often mistaking an entire image as an ice block. When the model was only looking at larger chaos blocks, typically platy blocks, the model has a much higher accuracy.

In the future, we plan to improve the metrics of the model by potentially combining previous iterations of the ResNet50 model which was able to more accurately identify smaller blocks. Once our metrics are at satisfactory numbers, we will finish the model testing of the training sets onto other chaos regions that were not mapped. To ensure the machine learning model is accurately identifying the chaos blocks, the lead author here will check the model labels incase further training needs to be done. Eventually, we would like to be able to identify all chaos blocks on Europa with the current imagery available, so we can use the model on future mission imagery to identify how chaos blocks have changed in size and movement. Once the machine learning model identification of chaos blocks is complete, we plan to complete a geophysical interpretation of chaos on Europa, running further analyses to build on previous chaos studies [2], but also running the same Spearman correlation tests to see if the earlier results are confirmed with a larger dataset of block size distributions resulting from our model.

References: [1] E.J. Leonard, R. T. Pappalardo, and A. Yin (2018) *Icarus*, 312, 100-120. [2] E. J. Leonard, S. M. Howell, A. C. Mills, D. A. Senske, D. A. Patthoff, H. C. F. C. Hay, and R. T. Pappalardo (2022) *GRL*, 49, 8. [3] R. Greenberg, G. Hoppa, R. Tufts, P. Geissler, J. Riley, and S. Kadel (1999) *Icarus*, 141, 263-286. [4] J. W. Head and R. T. Pappalardo (1999) *JGR*, 104, E11, 143-27, 155. [5] G.C. Collins, J. W. Head III, R. T. Papparaldo, and N. A. Spaun (2000) *JGR*, 105, 1709-1716. [6] B. E. Schmidt, D. D. Blankenship, and P. M. Schenk (2011) *Nature Let*, 479, 502-505. [7] N. A. Spaun, J. W. Head, G. C. Collins, L. M. Prockter, and R. T. Pappalardo (1998) *GRL*, 25, 23, 4277-4280.

Figures:

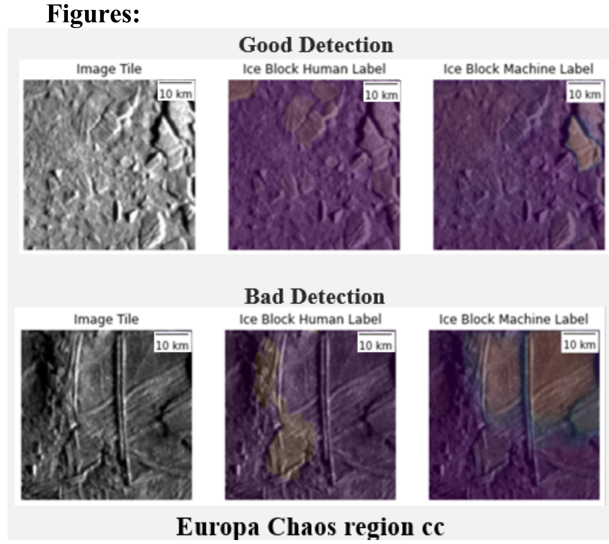

Fig. 1. Top: correct classification of blocks in Chaos cc region on Europa. Bottom: incorrect classification of blocks in Chaos cc region on Europa. Left: original Galileo SSI image. Center: human labeled blocks from training dataset. Right: Model labeled blocks.