Exploring the Re-Creation of the McIntosh Archive with Deep

Ryan Nguyen(1), Jack Ziegler(1), Ryan Kelly(1), Ian Hewins(1), David Webb(2), Michael Kirk(3), Geoff Crowley(1)

(1) Orion Space Solutions, Louisville, CO; (2) Boston College; (3) NASA Goddard Spac

ORION SPACE SOLUTIONS
An Arcfield Company

Overview

The goal of the Machine Learning (ML) portion of this project is to automate the extraction of valuable solar features from H-alpha images, 195A (or 193A) images and Magnetograms to create Carrington Rotation (CR) maps of the sun over a full period of rotation. Currently, human experts must hand label and process individual images to capture Coronal Holes (CHs) Polarity Inversion Lines (PILs), filaments and other features from each of the three different types of images. Then the expert needs to repeat this process for several days throughout the rotational period of the sun and finally perform several augmentations using the IDL programming language to arrive at a final CR map. This process is time intensive. Ideally, we would like to have an AI model that can recreate CR maps as accurately as a human expert so that the process can be completed quickly and regularly.

We explored a reinforcement learning (RL) model and tried initializing the environment using an unsupervised model. Additionally, we explored various supervised learning models and saw incredible success. Our findings are shown here, and Figure 1 gives an overview of our approaches.

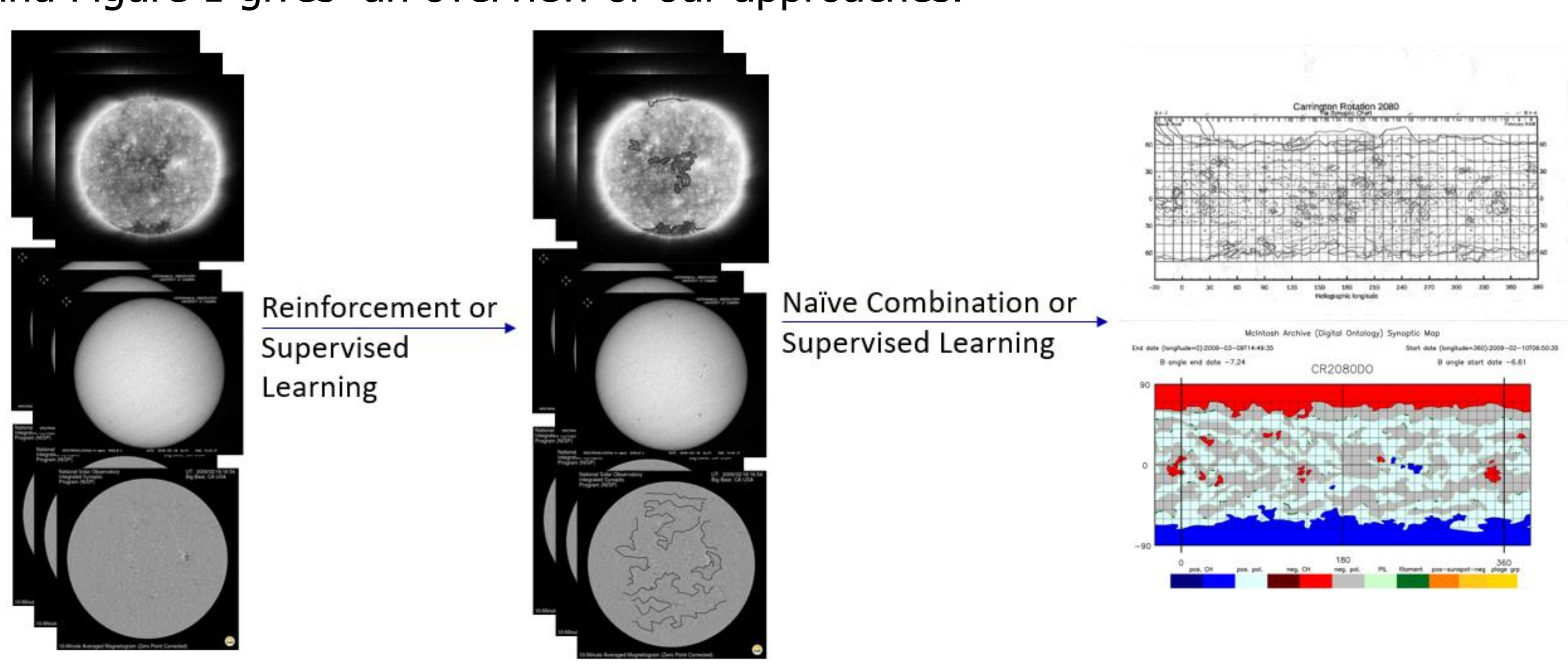

Figure 1: Mapping from images of the sun in various wavelengths to a synoptic map

Current Status

- **RL-Model**
  - We approximated the unit square with a regular n-by-n grid and mapped each point on the grid into a given image. The model was then rewarded for finding a mapping by which the points on the interior of the unit square were mapped onto a darker region, and points on the boundary of the unit square were mapped to brighter regions. (shown in Figure 2)
  - The model's reward is calculated as the approximation of a line integral along the boundary of the unit square, minus the surface integral of the interior of the square, with a penalty applied to discontinuous mappings.
  - To train the model, we provide an initial guess for the mapping, and let the model make small alterations to this map at each iteration. In an ideal situation, the mapping would converge to an optimal covering of a CH, and then we would run the model again, to find other CHs in the same image.
  - We explored K-Means as a way of providing an initialized environment for the RL model (shown in Figure 3)

Current Status (Continued)

- **Supervised Learning Model**
  - We have implemented four separate supervised learning models, one for each image frequency and a fourth that utilizes all three image types and attempts to identify all features simultaneously. The models trained separately on CHs and filaments have outperformed our initial expectations, achieving pixelwise classification accuracies above 98%.
  - For each of the three image types, we leveraged an Autoencoder style model with convolutional layers to classify pixels within the image that are labeled with the respective features.
  - We utilize residual skip connections, inspired by the ResNet model. A redrawing autoencoder can do such things as remove noise and sharpen the drawing. Our model has 74465 total parameter (290.88 KB) and utilizes the Adam optimizer with a low learning rate of 0.0002 and binary cross entropy for the loss function. Also note that the input image pixel values are scaled to range from zero to one. Results are shown in Figures 4 and 5.

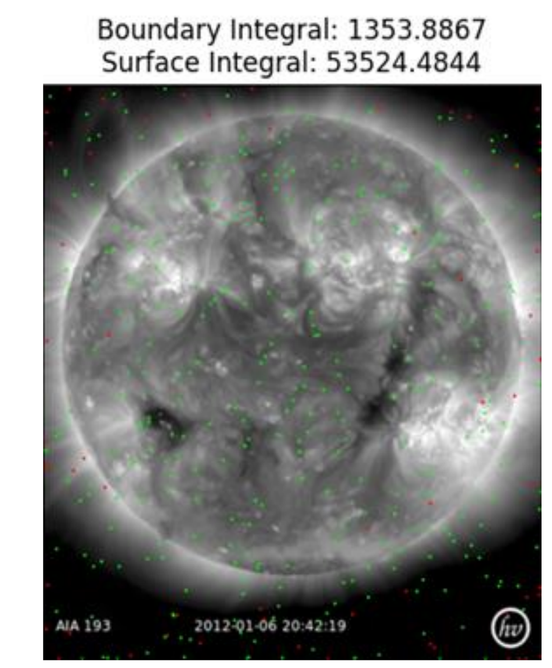

Figure 2: Visualization of PPO learned mapping.

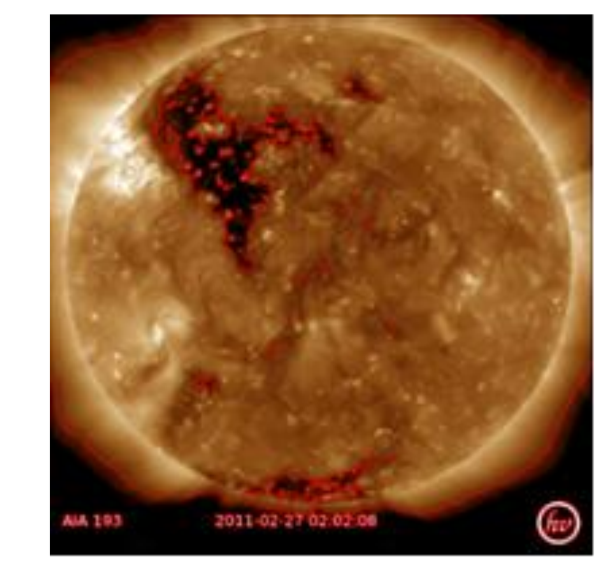

Figure 3: Visualization of K-means bounding of CHs to establish an initial environment for the RL model

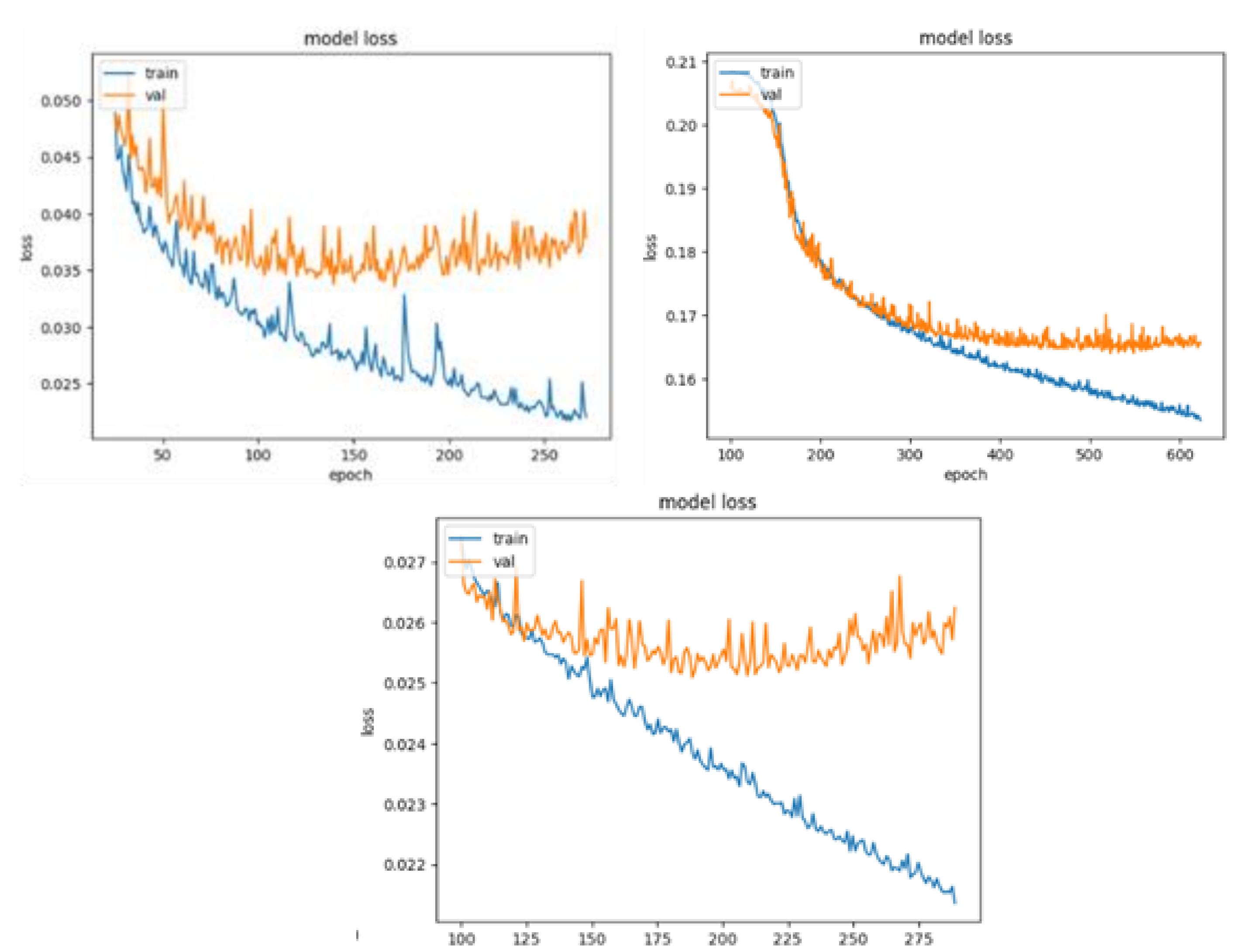

Figure 4:Loss vs. training time for CHs (top left), PILs (top right), and filaments (bottom)

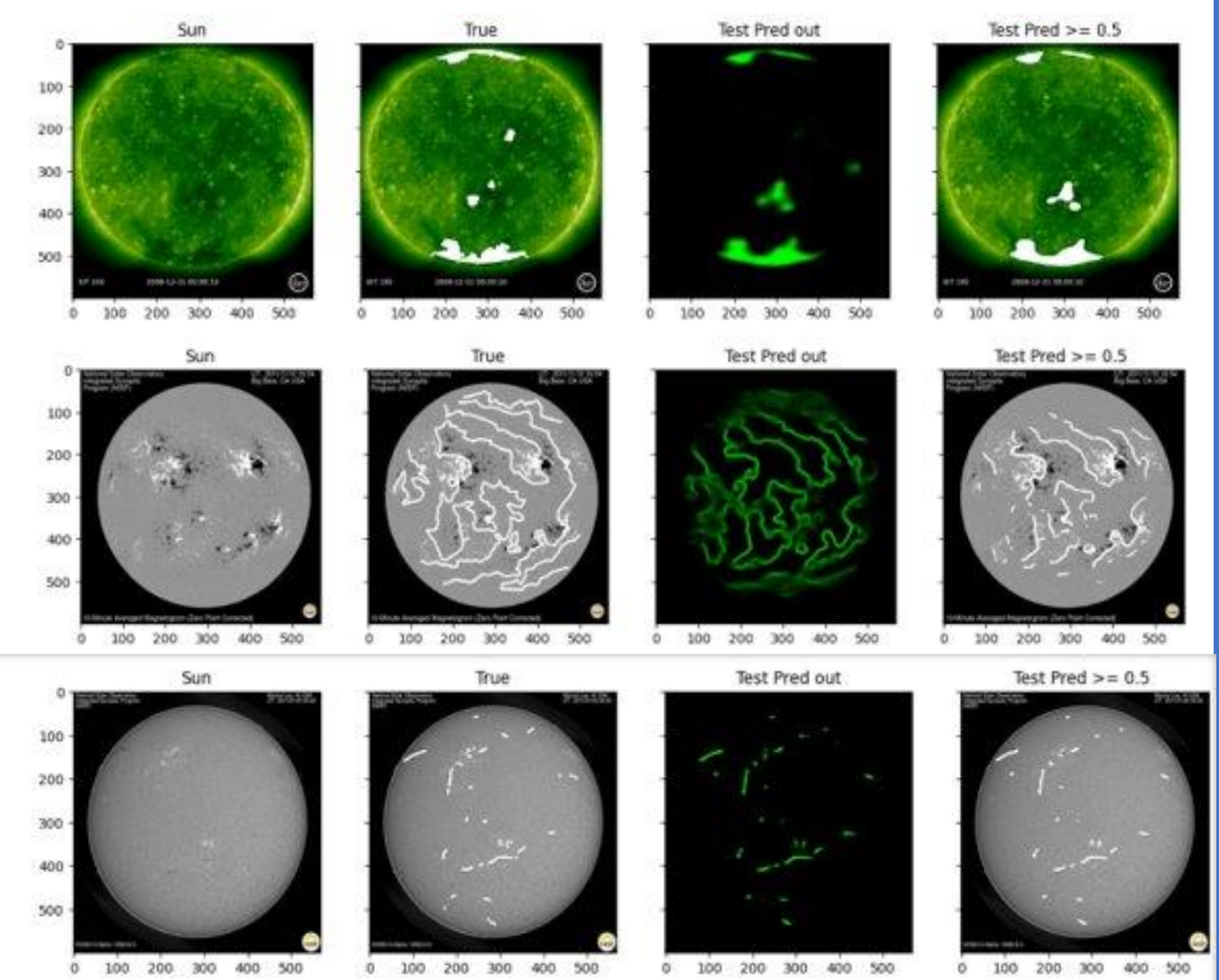

Figure 5:Visualization of the supervised learning model capturing the various features. CHs (top), PILs (middle), and filaments (bottom)

Future Work

- Improve the generalizability and overall accuracy of our supervised learning models
- Develop a supervised model to combine a full rotation's worth of labeled images into a completed CR map or Synoptic map
- Integrate the RL model and the supervised approach for optimal performance

Ryan Nguyen, Ph.D., Research Scientist: ryan.nguyen@arcfield.com

This work was funded under NSF Award 2035710