



On the Estimation of the SHARP Parameter MEANALP from AIA Images Using Deep Neural Networks

B. Benson^{1,2} · W.D. Pan^{1,2} · A. Prasad^{1,3} · G.A. Gary^{1,3} · Q. Hu^{1,3}

Received: 30 April 2021 / Accepted: 13 October 2021 / Published online: 11 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Space-weather HMI Active Region Patches (SHARPs) data from the Helioseismic and Magnetic Imager (HMI) on board the Solar Dynamics Observatory (SDO) provides high cadence data from the full-disk photospheric magnetic field. The SHARP's MEANALP (α_m) parameter, which characterizes the twist, can provide a measure of nonpotentiality of an active region, which can be a condition for the occurrence of solar flares. The SDO/Atmospheric Imaging Assembly (AIA) captures images at a higher cadence (12 or 24 seconds) than the SDO/HMI. Hence, if the α_m can be inferred from the AIA data, we can estimate the magnetic field evolution of an active region at a higher temporal cadence. Shortly before a flare occurs, we observed a change in the α_m in some active regions that produced stronger (M- or X-class) flares. Therefore, we study the ability of neural networks to estimate the α_m parameter from SDO/AIA images. We propose a classification and regression scheme to train deep neural networks using AIA filtergrams of active regions with the objective to estimate the α_m of active regions outside our training set. Our results show a classification accuracy greater than 85% within two classes to identify the range of the α_m parameter. We also attempt to understand the nature of the solar images using variational autoencoders. Thus, this study opens a promising new application of neural networks which can be extended to other SHARP parameters in the future.

Keywords Deep neural networks · Active regions · SHARP data · SHARP MEANALP

✉ W.D. Pan
pand@uah.edu

B. Benson
bb0008@uah.edu

A. Prasad
ap0162@uah.edu

G.A. Gary
allen.gary@uah.edu

Q. Hu
qh0001@uah.edu

¹ The University of Alabama in Huntsville, Huntsville, AL, USA

² Department of Electrical and Computer Engineering, UAH, Huntsville, AL, USA

³ Center for Space Plasma and Aeronomic Research, UAH, Huntsville, AL, USA

1. Introduction

Solar coronal activity is driven by strong magnetic field which amount to several thousand gauss on the solar surface (Aschwanden, 2005). The internal rotation of the Sun creates complex coronal magnetic fields in active regions which are responsible for solar activity such as flares and coronal mass ejections (CMEs). Solar activity can be detrimental to the geospace environment and warrants a thorough understanding of the physical processes that cause it. The magnetic forces dominate other forces in the solar corona which causes plasma of different densities in the coronal loops to be confined and thus coronal loops can have different temperatures and hence are visible in different extreme ultraviolet (EUV) wavelengths. Solar activities such as flares and CMEs release part of the free magnetic energy (which is the difference in the magnetic energy of the coronal field to that of a potential field) present in the corona (Falconer, Moore, and Gary, 2002). Therefore, a high-fidelity magnetic field model is essential to understanding the dynamics of the solar corona. Certain assumptions about the nature of the magnetic field are necessary in order to model it accurately. Based on the observation that in the region between the photosphere and upper corona we can neglect all lower-order nonmagnetic forces (low plasma β), we arrive at a force-free approximation (Gary, 1989, 2001).

In our previous work we used a linear force-free magnetic field (LFFF) model to synthesize the coronal magnetic field of active regions of simple (Benson et al., 2017) and multi-dipolar magnetic fields (Benson et al., 2019). We also used loops generated from the LFFF based on a vector magnetogram of active region (AR) 11117, which was utilized to estimate the linear force-free parameter α_{ff} . In Benson et al. (2019), we generated datasets of synthetic coronal loop images by varying the value of α , which represents the top-view of 3D magnetic field flux tubes from the random footpoints of simple and multi-dipolar configurations. Figure 1 shows the vector magnetogram of AR 11117 and the synthetic coronal loop images for each of α_{ff} value. In the context of neural networks, each value of α represents a class of synthetic coronal loop images. These images were then used to train deep learning models to estimate the value of the force-free parameter α_{ff} and we found that we were able to predict the α_{ff} values with good accuracy.

Since large amounts of solar observational data from both ground- and space-based instruments have become readily available, machine learning algorithms have been increasingly applied for predictive studies. The ability of machine learning and deep learning models to analyze vast volumes of data quickly and efficiently, gives them an edge over traditional methods. Several studies implemented machine learning and deep learning models successfully in heliophysics for prediction and forecasting studies (Fernandez Borda et al., 2002; Qu et al., 2003; Li et al., 2007; Qahwaji and Colak, 2007; Wang et al., 2008; Yuan et al., 2010; Ahmed et al., 2013; Bobra and Couvidat, 2015; Nishizuka et al., 2018; Jonas et al., 2018; Huang et al., 2018; Pala and Atici, 2019; Benson et al., 2020). Camporeale (2019) provided an extensive review on the current state of machine learning research in heliophysics.

With the results in Benson et al. (2019), we showed that loops observed by the Atmospheric Imaging Assembly (AIA) can be useful in inferring some magnetic field parameters like α using deep neural networks. In this study, we generalize the method outlined earlier from synthetic data to actual observational data by training deep neural networks to estimate the SHARP's MEANALP α_m parameter, which is similar to the LFFF α_{ff} , from the high temporal cadence coronal EUV images without the need for photospheric magnetograms to infer the changing magnetic free energy. The objective is to determine a quick estimation of α_m given AIA filtergrams of active regions outside the training set. Since AIA images are

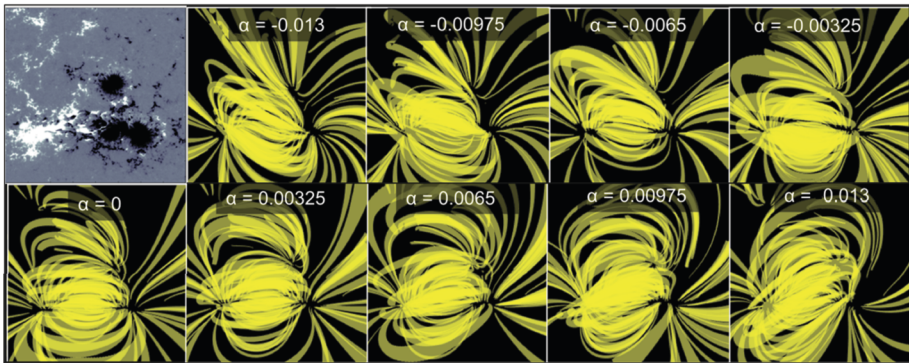


Figure 1 Vector magnetogram of AR 11117 and the pseudo-coronal loop simulations with varying values of the α_{ff} parameter. Each value in the panels corresponds to a class. $\alpha = -0.013$ represents class 0 and so on to $\alpha = 0.013$ representing class 9.

available at a higher (12 s) cadence as compared to the Helioseismic and Magnetic Imager (HMI) data (12 min), this method allows us to quickly predict magnetic field parameters, which can be of use in space weather studies.

The rest of the article is outlined as follows. Section 2 presents the motivation to use the α_m parameter and the data assimilation of active regions used for training deep neural network models. Section 3 discusses the initial tests and results that motivate our future experiments. In Section 4, we describe our phase-I and phase-II experiments in detail. We present our results and discussion in Section 5. Finally, we conclude the study in Section 6 and point out some avenues for future work.

2. Data Assimilation

2.1. SHARP MEANALP as a Flare Indicator

Photospheric magnetic field measurements are known to provide insights into the mechanism of solar flares and coronal mass ejections. Therefore, studying the correlations or the lack thereof between parameters of the photospheric magnetic field and eruptive solar activity is an ongoing research topic (Bobra et al., 2014).

As a starting point in this study, in order to verify the accuracy of the synthetic coronal loop simulations to those of actual EUV images, we test AIA image data from AR 11117 with the convolutional neural network (CNN) (LeCun et al., 1998) model trained on their corresponding synthetic coronal loop images using the method outlined in Benson et al. (2019). We select a time series of AIA EUV image data with a duration of 6 hours at 12 s cadence from 94 Å, 131 Å, 171 Å, 193 Å, and 211 Å wavelengths while centering the peak time of the flare. These images are passed through the SolarNet model (Benson et al., 2019) trained on the AR 11117 pseudo-coronal loops dataset to test the output prediction. Figure 2 shows a sharp change in the value of α during the time of a flare aligned against the GOES X-ray flux for AR 11117. The value of α changes at the same time as the flare occurs which signifies that this parameter can be a good indicator of flare activity. The blank values in the figure represent the lowest value of α belonging to class 0. As explainability is not a strong suite in deep learning models, it is difficult to evaluate the model predictions of the lowest class randomly before the flare peak time. To validate these results, we repeat the

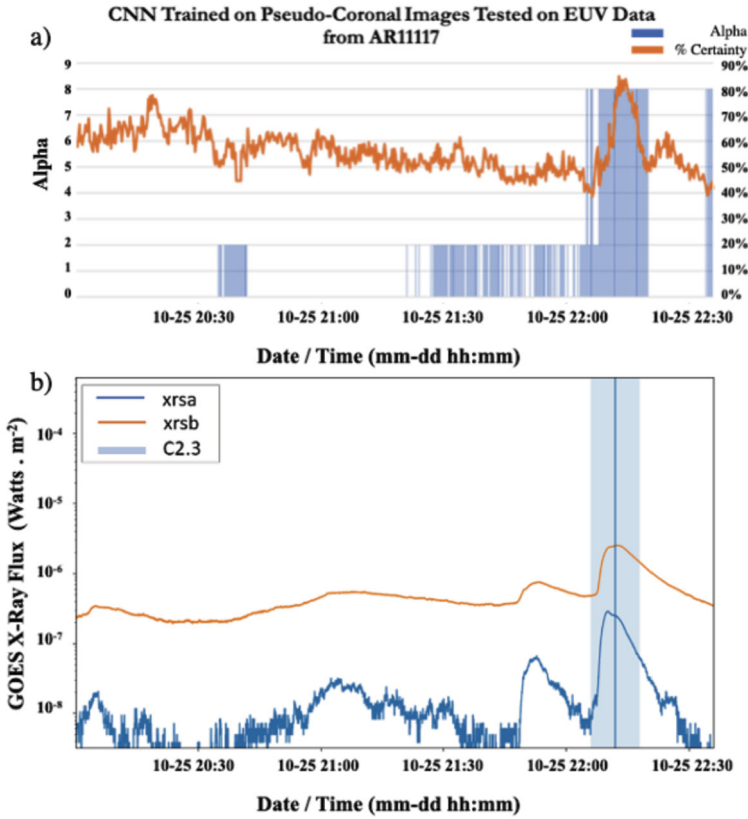


Figure 2 a) CNN model response while comparing pseudo-coronal loop images of AR 11117 (which occurred on 25 October 2010) to their corresponding sequence of AIA images. The height of the blue lines indicates the class value of α . The orange line shows the percentage of certainty that the CNN model associates with the prediction. The gaps in data represent the lowest value of alpha. b) The GOES X-ray flux data for AR 11117 aligned in time with the corresponding time series of AIA images. In the inset xrsa (xrsb) corresponds to GOES X-ray sensor a (b) and C2.3 indicates the flare X-ray class.

same process by generating coronal loop images for a different active region (AR 11283) and test them with their corresponding AIA loop images. These results as shown in Figure 3 also indicate a sharp change in α during the time of the flare.

To investigate the change in α during a flare using observed coronal images, in our study, we use the SHARP MEANALP parameter (α_m), which is the average characteristic twist of the magnetic field lines in the entire active region patch. The α_m parameter is defined by the equation

$$\alpha_m \propto \frac{\sum J_z B_z}{\sum B_z^2} \text{ Mm}^{-1}, \tag{1}$$

where J_z is the vertical electric current, and B_z is the vertical component of the magnetic field. The B_z^2 -weighted method proposed by Hagino and Sakurai (Hagino and Sakurai, 2004) was used to calculate α_m by computing the sum of the product $J_z B_z$ for the pixels in the active region patch, divided by the sum of B_z^2 . Dividing the pixels by the sum of

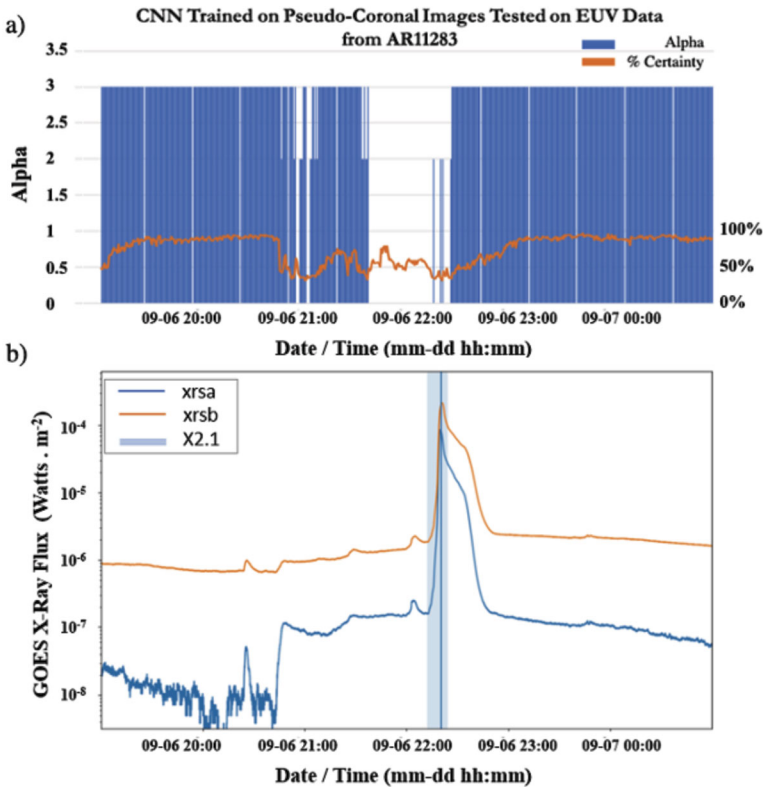


Figure 3 a) CNN model response while comparing pseudo-coronal loop images of AR 11283 (which occurred on 9 September 2011) to their corresponding sequence of AIA images. The height of the blue lines indicates the class value of α . The orange line shows the percentage of certainty that the CNN model associates with the prediction. The gaps in data represent the lowest value of alpha. b) The GOES X-ray flux data for AR 11283 aligned in time with the corresponding time series of AIA images. In the inset xrsa (xrsc) corresponds to GOES X-ray sensor a (b) and C2.3 indicates the flare X-ray class.

squares of the magnetic field ensures that only the regions with a strong magnetic field are represented in calculating the α_m parameter. The J_z values derived in the SHARP are dependent on some assumptions which we have supposed to be accurate in our study. The α_m considered here is averaged over the whole active region and might not show small scale changes. A study conducted by Bobra and Ionidis (2016) that predicts coronal mass ejections (CMEs) using SHARP parameters as features, rank them according to their F-score. The study found that the α_m parameter ranked third on a list of 19 parameters used to predict whether an X- or M-class flaring active region would produce a CME.

2.2. Active Region Selection

To check if the change in characteristic twist, α_m , of an active region during the occurrence of a flare is typical, active regions that appear within a maximum of 800 arcseconds from the center of the solar disk with a flare strength greater than M 1.0 were selected from the Hinode flare catalog (Watanabe, Masuda, and Segawa, 2012) using SunPy (Mumford et al., 2020). This yielded 379 X- and M-class flares that occurred between 2010 and 2015. We then plotted the α_m parameter for a duration of six hours centered around the peak time of

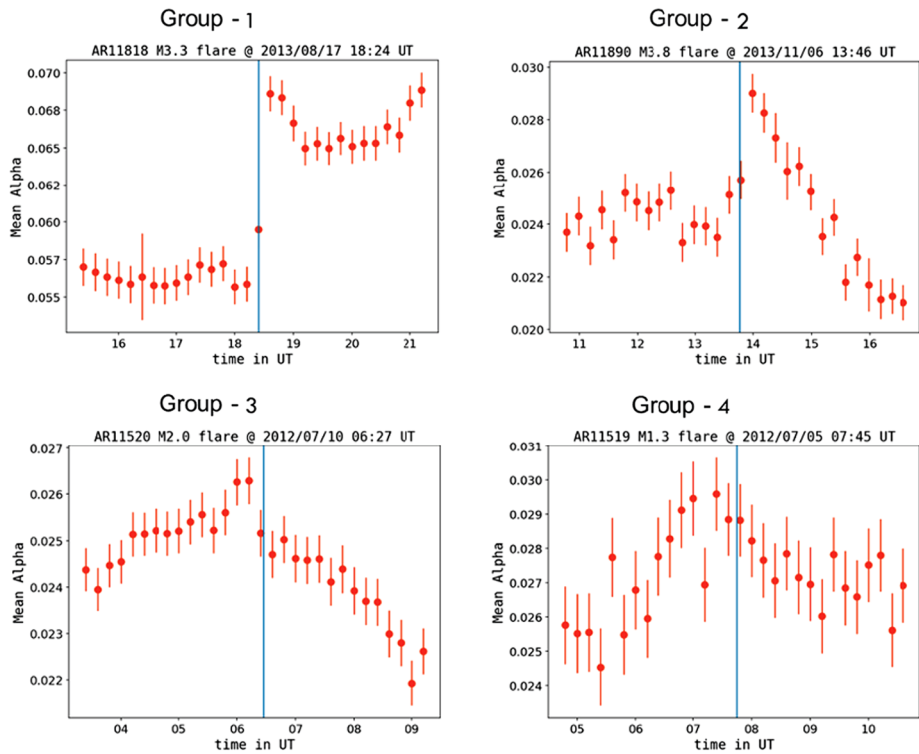


Figure 4 One active region example for each of the four characteristic groups of variation in the twist parameter α_m , plotted in time for 6 hours during a flare with the time of the flare centered. The blue line indicates the peak time of the flare. The labels Group-1 to Group-4 signify the amount of change in α_m .

the flare. As the photospheric data from the vector magnetogram is available with a cadence of 12 minutes, we could extract ≈ 30 time steps of the α_m parameter. These active regions were then binned into four groups based on the amount of change in the α_m parameter value during the flare.

Figure 4 shows the α_m plots for one example in each of the four active region groups. The lines across the data points are error bars. Active regions in group 1 showed a stark change in α_m . In contrast, active regions in groups 2 and 3 showed a noticeable change, and active regions in group 4 showed no change or noisy data with big error bars. Tables 1 and 2 show the breakdown of the number of X- and M-class flares. Approximately 60% of the total X-class flares and approximately 27% of M-class flares show a noticeable change in the α_m parameter. Of the active regions that have flares stronger than M2.0 class, $\approx 40\%$ show a noticeable change in the α_m parameter. These statistics indicate that stronger flares are more likely to show a noticeable change in the characteristic twist parameter α_m since the magnetic field changes are greater.

3. Regression Tests

The estimation of the value of the α_m parameter for future active regions from EUV coronal images is formulated as a regression problem. Regression is a supervised learning method

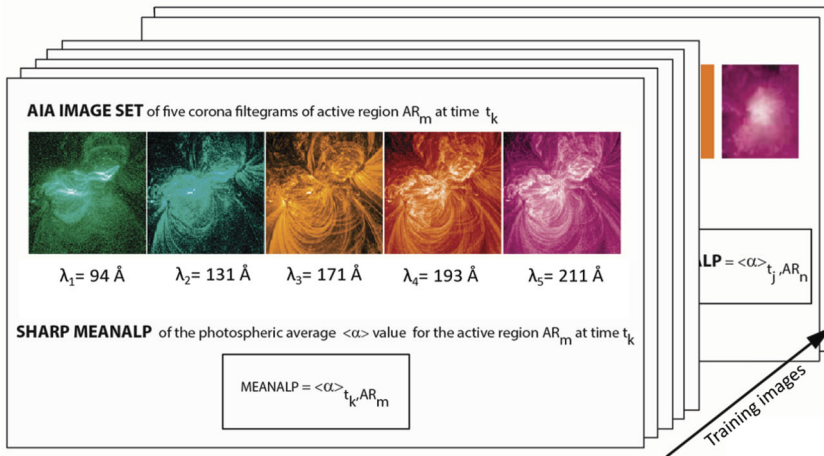


Figure 5 Collection of AIA training images over time t and over different active regions in 94, 131, 171, 193, 211 Å wavelengths and the associated α_m .

where the model learns to predict a continuous target value. To test the ability of the CNN-regression model to handle AIA images, initially, we use a single active region (AR 11283) with its corresponding α_m values as labels. AIA images are available at a 12-second cadence and the α_m is available at a 12-minute cadence. This allows us to use more AIA data per α_m value. We use 60 AIA images (obtained during a 12-minute duration and five wavelengths) that are labeled with a single α_m value which is obtained from the header of the HMI SHARP data. The 12-minute time window for the AIA images is centered around the SHARP data timestamp which ensures that no image is less than 6 minutes away from the SHARP α_m label. This data is taken for 6 hours forming 1800 images per active region. These images are cropped to only represent the active region patch.

We treat these images with a multi-scale Gaussian normalization filter (MGN) (Morgan and Druckmüller, 2014) which enhances the features in the image while flattening the noisy regions. Figure 5 shows the collection of the AIA training images over time t in different wavelengths used to train the CNN-regression model. The α_m values for our training data come from the HMI SHARP header and thus are our labels, labels do not change with a change in saturation. The CNN learns the same α_m label for both saturated and unsaturated AIA images belonging to a single α_m . As the same label is applied to a 12-minute dataset, the effects of fluctuations in saturation during flares are also averaged out by the CNN. Finally, the percentage of over-saturated images in the dataset is less than 3%, therefore, we do not consider any separate treatment for over-saturated images.

In the first experiment, we train the model on five wavelengths of AIA data for AR 12297, which produced an X 2.1 flare, and test it with AIA data from AR 11261, which produced an M 6.0 flare. Figure 6 shows the box plot of the α_m ranges for the training and testing data, respectively. The box plot shows that the ranges for α_m values about the mean are similar. Figure 7 shows¹ the results of the regression test. It can be seen that the predicted values follow the rise in α_m value. This gives us the framework to group active regions by their α_m values and train the models over a broad range of α_m values.

¹The data needed to reproduce this figure is missing.

Figure 6 Box plot showing the ranges of α_m value for AR 12297 (X2.1) (train set) and AR 11261 (M6.0) (test set). The box denotes the first and third quartiles of the entire range of α_m values and the green line is the median. The circle denotes an outlier value.

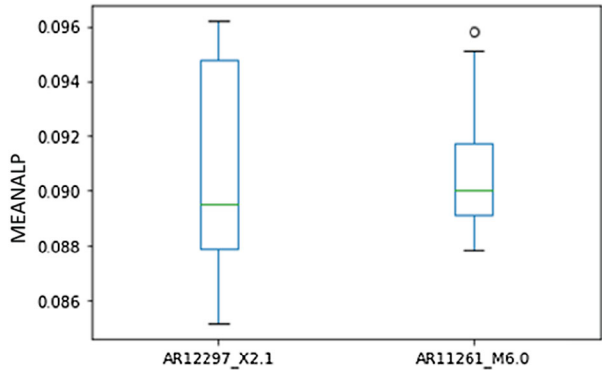


Figure 7 Actual vs. predicted values of α_m for all five wavelengths for the initial regression test. The blue lines represent the actual α_m values of the test images and the orange lines represent the predicted values. The model was trained on AIA images from AR 12297 (X2.1) and tested with AIA images from AR 11261 (M6.0).

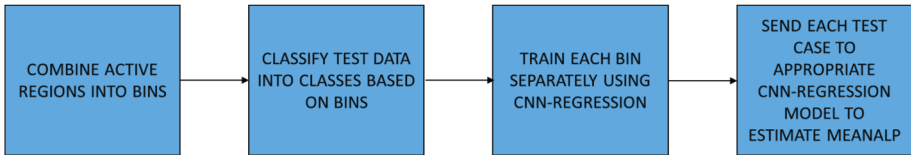
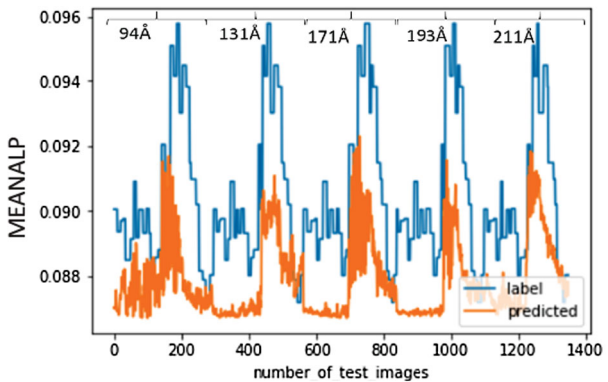


Figure 8 Flow diagram showing the training scheme to estimate the α_m value of an unseen active region image.

4. Experimental Setup

We carry out our experiments in two phases based on the initial regression test results. While we know the value of the α_m parameter of active regions known to the model, the goal is to estimate the value of α_m for an unseen active region. The CNN-regression model used here is shown in Figure 9, and contains 4.7 million parameters and consists of eleven convolution layers, two fully connected layers, and a linear layer to output the single prediction. The same architecture is used for the classification model by replacing the linear layer with a softmax layer. The results from the initial regression tests indicated that the CNN-regression model was able to estimate the value of an unseen active region within the same α_m range. Therefore, we propose a scheme to first classify active regions into bins before sending them to the CNN-regression models trained on separate bins. Figure 8 shows the flow diagram for

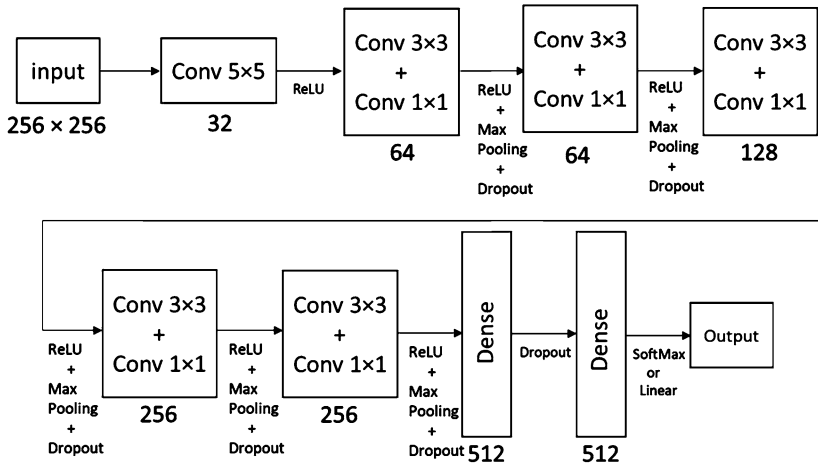


Figure 9 Convolutional neural network architecture used for training. The output layer provides a classification result with a softmax activation or a regression output with a linear activation.

Table 1 Count of X-class and M-class flares for the groups 1–4 seen in Figure 3.

	Group-1	Group-2	Group-3	Group-4	Total
X-class	11	1	1	9	22
M-class	15	41	40	261	357

our scheme. First, we group known active regions into groups/bins. Each of these bins is trained separately on a CNN-regression model and saved. Then, a classification model is developed to classify new data into these bins. Based on the classification results, the new data is sent to the appropriate CNN-regression model to estimate the value of α_m . We train all our datasets on a NVIDIA DIGITS™ DevBox workstation using Keras (Chollet, 2015) layers with a Tensorflow backend (Abadi et al., 2015).

4.1. Phase-I Experiments

In phase-I experiments, we train the model to classify active regions into bins using 20 active regions at specific times and test the model using an additional 11 active regions. Based on the α_m values, the active regions are grouped into eight bins. Figure 10 shows the ranges of the α_m parameter for the selected active regions. The data consists of five wavelengths of AIA images belonging to active regions producing greater than M2.0-class flares. The training was performed with 70% of the images from the dataset, while 15% of the images were each allocated for cross-validation and testing purposes.

4.2. Phase-II Experiments

For the phase-II experiments, we implement the full proposed scheme of classification of active regions into bins and implement the α_m parameter estimation using the CNN-regression models. We choose only the active regions belonging to groups 1 and 2 of Table 2 due to our limited data handling capabilities for these sets of experiments. Of the 68 active regions that belong to groups 1 and 2 from Table 1, we could only use 57 active regions as the remaining

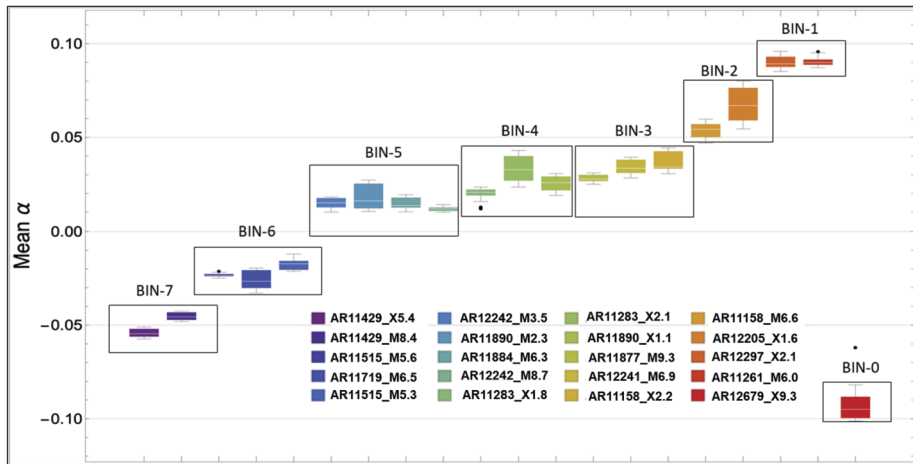


Figure 10 Box plot of α_m ranges of active regions from phase-I experiments grouped into eight bins based on their average value.

Table 2 Breakdown of flaring active regions into four characteristic groups.

Class	Group-1	Group-2	Group-3	Group-4	Total
$\geq M2.0$	13	23	32	89	157
$\geq M3.0$	11	13	12	46	82
$\geq M4.0$	10	8	5	32	55
$\geq M5.0$	8	7	4	21	40

were located on the solar limb during the time of their flares. This dataset contains approximately 102,600 images from 57 active regions, 5 wavelengths, 30 α_m values, 12 images per α_m value. For the classification tasks, the active regions are grouped into six bins based on the α_m parameter range. Each bin is trained separately using a CNN-regression model for estimating the α_m parameter. Figure 11 shows the ranges of α_m for all the active regions in the dataset and their grouping into bins. It clearly shows that the data has a class imbalance problem. To counter this, we assign weights to each class based on the number of images per each class compared to the total number of images in the dataset. This reduces the tendency of the classification model from favoring the class with the highest number of images.

5. Estimating SHARP MEANALP Parameter Results

5.1. Phase-I Results

Phase-I tests consist of testing the classification performance of the 20 active regions using SolarNet. We use an additional 11 active regions to test the generalization ability of the phase-I classification model.

Table 3 shows the summary for the phase-I classification results. Due to the limited number of active regions, none of the test images belongs to either Bin-0 or Bin-7. The Top-1 accuracy is 17.24%, while the Top-2 and Top-3 accuracies are 34.44% and 82.33%, respec-

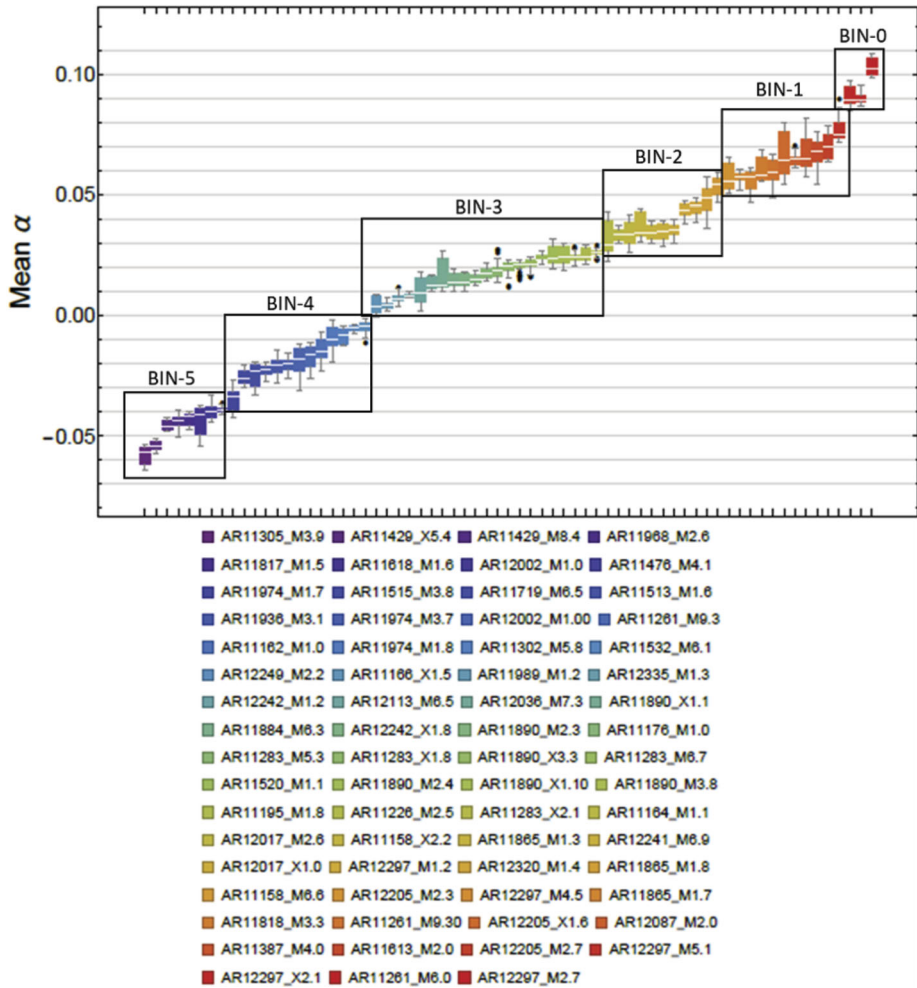


Figure 11 Box plot of SHARP MEANALP (α_m) ranges of active regions from phase-II experiments grouped into six bins based on their range. The bins are made based on the average value of the α_m parameter.

tively. The Top-1 accuracy measures the proportion of correctly classified examples compared to the target label. Top-2 and Top-3 accuracies indicate the second and third highest probabilities of the model that match the target label. The results indicate that over 80% of the guesses by the classifier are within two classes. The overall generalization ability of the classifier is not very clear from the phase-I tests due to the limited data. Therefore, we do not conduct any further regression tests on this dataset.

5.2. Phase-II Results

For phase-II tests, we select 57 active regions from our groups 1 and 2 that showed a noticeable change with respect to the characteristic twist parameter α_m . To maximize the generalization ability of our classification model, we separate our data into training (60%) and holdout sets (40%). We then divide the holdout set equally into validation and test sets. By

Table 3 Summary of classification performance on phase-I and phase-II experiments.

	Phase-I	Phase-II
Top-1% Accuracy	17.24%	29.43%
Top-2% Accuracy	34.44%	61.51%
Top-3% Accuracy	82.33%	87.76%

Table 4 CNN-regression model MSE of SHARP MEANALP (α_m) on individual bins for the phase-II active regions.

Bin	MSE
Bin-0	4.57×10^{-4}
Bin-1	1.53×10^{-4}
Bin-2	5.09×10^{-5}
Bin-3	1.49×10^{-5}
Bin-4	6.55×10^{-5}
Bin-5	7.68×10^{-5}

doing so, we ensure that the model does not see the images from the holdout set. We also ensure that active regions used for training are not present in the test sets. By making sure the validation and test sets come from the same distribution, we can optimize the classification model improving the validation performance, which guarantees that the test set will indicate the generalization ability of the trained model.

Table 3 shows the summary of the results. The Top-1 accuracy is 29.43%, while the Top-2 and Top-3 accuracies are 61.51% and 87.76%, respectively. This indicates that over 87% of the guesses by the classifier are within two classes. These results are better than the results from the phase-I classifier while using significantly more data, which adds much more complexity to identifying similar active regions. We can say that the generalization ability of the phase-II classifier is better than that of the the phase-I classifier. The regression bins are trained individually using the CNN-regression model with the mean square error (MSE) shown in Table 4. However, as the Top-1 accuracy is low, incorrect classification could lead to a higher error in the estimation of the α_m parameter.

5.3. Latent Space Visualization and Discussion

To understand the nature of the AIA image data and the reasons for the low Top-1 accuracy, we use a variational autoencoder (Kingma and Welling, 2013; Chollet, 2016) to map the latent representations of the images belonging to each active region bin. To compare the visualizations, we also map the latent representations of the Modified National Institute of Standards and Technology (MNIST) dataset (LeCun et al., 1998) and the phase-II active regions. Figure 12 shows the comparison of the latent space visualizations of the MNIST digits and the phase-II active regions dataset. The color bars on each side represent the different bins or classes. For the MNIST latent space visualization, it can be clearly seen that the digits that look different are separated in the latent space. For example, the red color representing digit 0 and the blue color representing digit 1 are on different corners in the latent space because there is no similarity in their physical representations. However, for the phase-II active region images, the samples that belong to each class are dispersed throughout

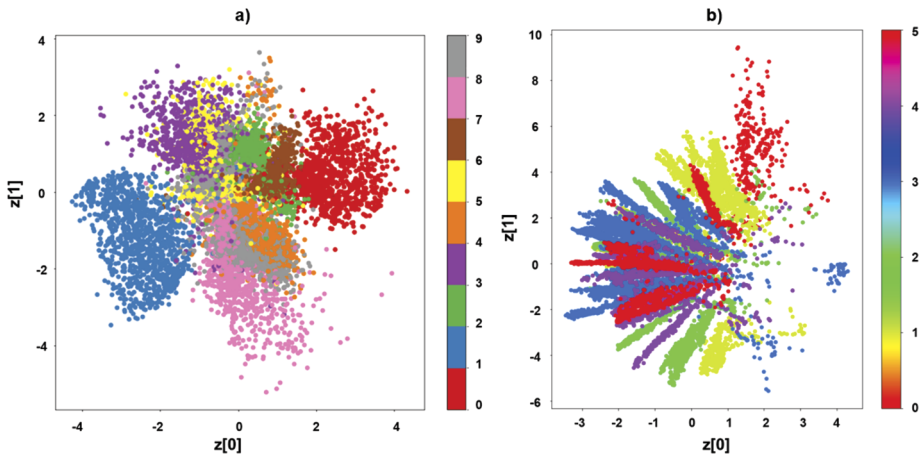
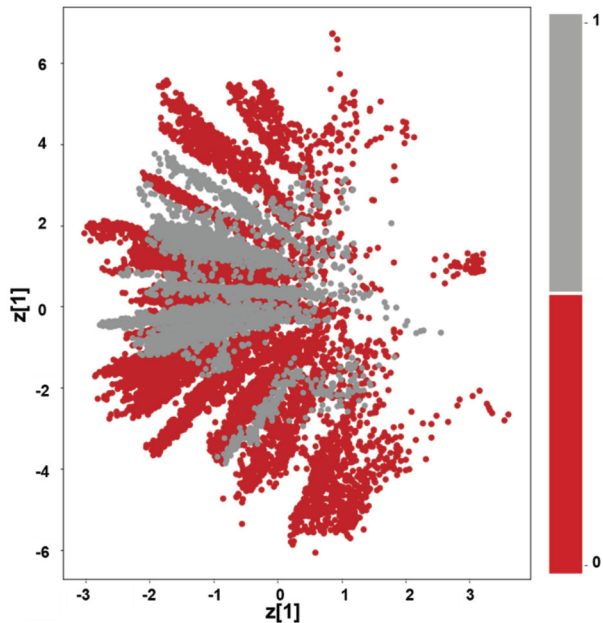


Figure 12 Comparison of the latent space visualizations using variational autoencoders for phase-II active regions. Here $z[0]$ and $z[1]$ are the two latent variables that remain after the encoding stage. The color bars on each side represent the different bins or classes. a) Visualization for MNIST dataset. b) Visualization for phase-II active regions.

Figure 13 Latent space visualization in two dimensions for the phase-II active regions grouped into two bins based on the sign of the SHARP MEANALP (α_m) parameters. Here $z[0]$ and $z[1]$ are the two latent variables that remain after the encoding stage.



the latent space and are not separable. This explains to a certain extent why immediate classes or bins cannot be sharply differentiated, thus causing the CNN to perform poorly on the Top-1 accuracy but perform far better within one or two groups or classes.

We bin the active regions into two groups based on the sign of the α_m parameters to see if the classifier improved its performance. Figure 13 shows the latent space visualization for the binary classes. It can be seen that the classes are better separable than the active

regions grouped into six bins. For training this model, class weights are assigned to penalize the over-represented positive class and favor the under-represented negative class. The data are divided into training and holdout sets as done with the 6-bin dataset. Validation and test sets are then formed from the holdout set. The classifier score an accuracy of 80.78% in being able to differentiate between the positive and negative classes. As this is a binary classification problem, the true skill score TSS which varies between -1 and 1 is calculated as 0.47.

The binary classification results reveal the conundrum with the classification-regression scheme. The classification would need fewer bins to reduce the classification error whereas narrower bins are required to keep the regression error low. Also, as the latent space visualizations shows, active regions are not well separable by the values of the MEANALP parameter.

6. Conclusion and Future Work

In this study, we build on our work with the pseudo-coronal loop image models (Benson et al., 2019), the response of AIA data to the SolarNet model. This derived from our observation of a change in the characteristic twist parameter α_m shortly before the peak flaring time for strong flares. This correlation between the change in the α_m parameter during a short time before flares can be used as a feature in flare prediction studies. We propose a framework to estimate the α_m parameter by using a classification and regression scheme. Our phase-II results show greater than 85% accuracy within two classes in identifying the range of the α_m parameter. We also analyze the nature of solar satellite image data from the SDO/AIA mission using variational autoencoders by plotting the latent space visualizations of AIA data. These experiments aid our understanding of the complex nature of the magnetic field configurations of solar active regions. Our effort to estimate the α_m parameter from AIA image data yields acceptable results despite the complex nature of the data and given the limited number of active regions used. This study also reveals some of the limitations of convolutional neural networks in handling data with high dynamic ranges.

In the future, the correlation between solar flares and the α_m parameter, including other SHARP parameters, can be further explored during short time periods before the eruption of solar flares. This study can also be extended to using the AIA data to estimate other SHARP parameters that are better indicators of flaring activity. The ability to process large amounts of data is key to such tasks, especially while using high-resolution and high dynamic range image data. The use of magnetogram images of active regions for estimating the α_m parameter can also be explored. A scheme where we can use magnetogram image data for classification and AIA images for regression can be considered given the issues with grouping active regions into very narrow bins.

A time series analysis of all SHARP parameters where solar activity can be classified based on the type of flares also merits consideration and thought. Studying the long term variability and evolution of the SHARP parameters over the lifetime of an active region and using forecasting methods can also be an area of further research. Forecasting models using univariate/multivariate time series data can be applied to SHARP parameters that display trends that correlate to solar activity.

Declarations

Disclosure of Potential Conflicts of Interest The authors declare that there are no conflicts of interest.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al.: 2015, *TensorFlow: Large-scale machine learning on heterogeneous systems*. Software available from [tensorflow.org](https://www.tensorflow.org).
- Ahmed, O.W., Qahwaji, R., Colak, T., Higgins, P.A., Gallagher, P.T., Bloomfield, D.S.: 2013, Solar flare prediction using advanced feature extraction, machine learning, and feature selection. *Solar Phys.* **283**(1), 157. DOI.
- Aschwanden, M.J.: 2005, *Physics of the Solar Corona. An Introduction with Problems and Solutions (2nd edition)*, Springer, Berlin. DOI.
- Benson, B., Jiang, Z., Pan, W.D., Gary, G.A., Hu, Q.: 2017, Determination of linear force-free magnetic field constant alpha using deep learning. In: *International Conference on Computational Science and Computational Intelligence*, 760.
- Benson, B., Pan, W.D., Gary, G.A., Hu, Q., Staudinger, T.: 2019, Determining the parameter for the linear force-free magnetic field model with multi-dipolar configurations using deep neural networks. *Astron. Comput.* **26**, 50. DOI.
- Benson, B., Pan, W.D., Prasad, A., Gary, G.A., Hu, Q.: 2020, Forecasting solar cycle 25 using deep neural networks. *Solar Phys.* **295**(5), 65. DOI.
- Bobra, M.G., Couvidat, S.: 2015, Solar flare prediction using SDO/HMI vector magnetic field data with a machine-learning algorithm. *Astrophys. J.* **798**(2), 135. DOI.
- Bobra, M.G., Ionidis, S.: 2016, *Astrophys. J.* **821**(2), 127. DOI.
- Bobra, M.G., Sun, X., Hoeksema, J.T., Turmon, M., Liu, Y., Hayashi, K., Barnes, G., Leka, K.D.: 2014, The helioseismic and magnetic imager (HMI) vector magnetic field pipeline: SHARPs – space-weather HMI active region patches. *Solar Phys.* **289**(9), 3549. DOI.
- Camporeale, E.: 2019, The challenge of machine learning in space weather: nowcasting and forecasting. *Space Weather* **17**(8), 1166. DOI.
- Chollet, F.: 2015, *Keras*, GitHub.
- Chollet, F.: 2016, *The Keras Blog: Building Autoencoders in Keras*. Accessed: 2019-08-31. <https://blog.keras.io/building-autoencoders-in-keras.html>.
- Falconer, D.A., Moore, R.L., Gary, G.A.: 2002, Correlation of the coronal mass ejection productivity of solar active regions with measures of their global nonpotentiality from vector magnetograms: baseline results. *Astrophys. J.* **569**, 1016.
- Fernandez Borda, R.A., Mininni, P.D., Mandrini, C.H., Gómez, D.O., Bauer, O.H., Rovira, M.G.: 2002, Automatic solar flare detection using neural network techniques. *Solar Phys.* **206**(2), 347. DOI.
- Gary, G.A.: 1989, Linear force-free magnetic fields for solar extrapolation and interpretation. *Astron. Astrophys. Suppl. Ser.* **69**, 323.
- Gary, G.A.: 2001, Plasma beta above a solar active region: rethinking the paradigm. *Solar Phys.* **203**(1), 71.
- Hagino, M., Sakurai, T.: 2004, Latitude variation of helicity in solar active regions. *Publ. Astron. Soc. Japan* **56**(5), 831. DOI.
- Huang, X., Wang, H., Xu, L., Liu, J., Li, R., Dai, X.: 2018, Deep learning based solar flare forecasting model. I. Results for line-of-sight magnetograms. *Astrophys. J.* **856**(1), 7. DOI.
- Jonas, E., Bobra, M., Shankar, V., Hoeksema, J.T., Recht, B.: 2018, Flare prediction using photospheric and coronal image data. *Solar Phys.* **293**(3), 48. DOI.
- Kingma, D.P., Welling, M.: 2013, Auto-encoding variational Bayes. [arXiv](https://arxiv.org/abs/1312.6114).
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: 1998, Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278.
- Li, R., Wang, H., He, H., Cui, Y., Du, Z.: 2007, Support vector machine combined with k-nearest neighbors for solar flare forecasting. *Chin. J. Astron. Astrophys.* **7**(3), 441. DOI.
- Morgan, H., Druckmüller, M.: 2014, Multi-scale Gaussian normalization for solar image processing. *Solar Phys.* **289**(8), 2945. DOI. ADS.
- Mumford, S.J., Freij, N., Christe, S., Ireland, J., Mayer, F., Hughitt, V.K., et al.: 2020, Sunpy: a Python package for solar physics *J. Open Sour. Softw.* **5**(46), 1832. DOI.
- Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., Ishii, M.: 2018, Deep flare net (DeFN) model for solar flare prediction. *Astrophys. J.* **858**(2), 113. DOI.
- Pala, Z., Atici, R.: 2019, Forecasting sunspot time series using deep learning methods. *Solar Phys.* **294**(5), 50. DOI.
- Qahwaji, R., Colak, T.: 2007, Automatic short-term solar flare prediction using machine learning and sunspot associations. *Solar Phys.* **241**(1), 195. DOI.
- Qu, M., Shih, F.Y., Jing, J., Wang, H.: 2003, Automatic solar flare detection using mlp, rbf, and svm. *Solar Phys.* **217**(1), 157. DOI.
- Wang, H.N., Cui, Y.M., Li, R., Zhang, L.Y., Han, H.: 2008, Solar flare forecasting model supported with artificial neural network techniques. *Adv. Space Res.* **42**(9), 1464. DOI.

Watanabe, K., Masuda, S., Segawa, T.: 2012, Hinode flare catalogue. *Solar Phys.* **279**(1), 317. [DOI](#).

Yuan, Y., Shih, F.Y., Jing, J., Wang, H.: 2010, Automated flare forecasting using a statistical learning technique. *Res. Astron. Astrophys.* **10**(8), 785. [DOI](#).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.