Draft version November 7, 2022
Typeset using LaTeX default style in AASTeX631

A Deep Learning Approach to Generating Photospheric Vector Magnetograms of Solar Active Regions for SOHO/MDI Using SDO/HMI and BBSO Data

Haodi Jiang,[1,2,3] Qin Li,[1,4] Zhihang Hu,[1] Nian Liu,[1,4] Yasser Abduallah,[1] Ju Jing,[1,4,5] Genwei Zhang,[1] Yan Xu,[1,4,5] Wynne Hsu,[6,7] Jason T. L. Wang,[1,2] and Haimin Wang[1,4,5]

[1] *Institute for Space Weather Sciences, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA; hj78@njit.edu, wangj@njit.edu, haimin.wang@njit.edu*

[2] *Department of Computer Science, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA*

[3] *Department of Computer Science, Sam Houston State University, Huntsville, TX 77341, USA*

[4] *Center for Solar-Terrestrial Research, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA*

[5] *Big Bear Solar Observatory, New Jersey Institute of Technology, 40386 North Shore Lane, Big Bear City, CA 92314, USA*

[6] *Institute of Data Science, National University of Singapore, Singapore 119077*

[7] *Department of Computer Science, School of Computing, National University of Singapore, Singapore 119077*

ABSTRACT

Solar activity is usually caused by the evolution of solar magnetic fields. Magnetic field parameters derived from photospheric vector magnetograms of solar active regions have been used to analyze and forecast eruptive events such as solar flares and coronal mass ejections. Unfortunately, the most recent solar cycle 24 was relatively weak with few large flares, though it is the only solar cycle in which consistent time-sequence vector magnetograms have been available through the Helioseismic and Magnetic Imager (HMI) on board the Solar Dynamics Observatory (SDO) since its launch in 2010. In this paper, we look into another major instrument, namely the Michelson Doppler Imager (MDI) on board the Solar and Heliospheric Observatory (SOHO) from 1996 to 2010. The data archive of SOHO/MDI covers more active solar cycle 23 with many large flares. However, SOHO/MDI data only has line-of-sight (LOS) magnetograms. We propose a new deep learning method, named MagNet, to learn from combined LOS magnetograms, B_x and B_y taken by SDO/HMI along with Hα observations collected by the Big Bear Solar Observatory (BBSO), and to generate vector components B'_x and B'_y, which would form vector magnetograms with observed LOS data. In this way, we can expand the availability of vector magnetograms to the period from 1996 to present. Experimental results demonstrate the good performance of the proposed method. To our knowledge, this is the first time that deep learning has been used to generate photospheric vector magnetograms of solar active regions for SOHO/MDI using SDO/HMI and Hα data.

Keywords: Solar atmosphere; Solar magnetic fields; Convolutional neural networks

1. INTRODUCTION

Photospheric vector magnetograms of solar active regions (ARs) collected by the Helioseismic and Magnetic Imager (HMI) on board the Solar Dynamics Observatory (SDO; Scherrer et al. 2012) play an important role in solar physics. They are used by nonlinear force-free field extrapolation methods (Wheatland et al. 2000; Wiegelmann 2004; Metcalf et al. 2005; Schrijver et al. 2008; Sun et al. 2012) to calculate magnetic energy, which provides crucial information concerning an AR's capability of producing eruptive events including solar flares and coronal mass ejections (CMEs; Aschwanden et al. 2014). They are also used to derive magnetic field parameters such as those of Space-weather HMI Active Region Patches (SHARP; Bobra et al. 2014), which have been used in machine learning-based flare and CME forecasting (Bobra & Couvidat 2015; Chen et al. 2019; Liu et al. 2019, 2020; Wang et al. 2020; Abduallah et al. 2021).

Since 2010, consistent time-sequence full disk photospheric vector magnetograms have been available through SDO/HMI. This data set covers a relatively weak solar cycle, namely cycle 24, which had fewer large eruptive events, and therefore the sampling is not sufficient to understand the initiation of solar eruptions and to make accurate predictions of eruptive events. Prior to 2010, vector magnetograms were available sporadically for certain ARs from some

observatories, such as the Big Bear Solar Observatory (BBSO; Denker et al. 1999), the Synoptic Optical Long-term Investigations of the Sun (SOLIS; Keller et al. 2003) of the National Solar Observatory (NSO), the Imaging Vector Magnetograph (IVM; Mickey et al. 1996) and Haleakala Stokes Polarimeter (HSP; Mickey 1985) at the Mees Solar Observatory (MSO), the National Astronomical Observatory of Japan (NAOJ; Ichimoto et al. 1991), and the Solar Optical Telescope Spectro-Polarimeter (SOT-SP; Tsuneta et al. 2008) on board the Hinode spacecraft (Kosugi et al. 2007). However, these vector magnetograms were not consistent in the sense that they were not collected on a regular basis. On the other hand, prior to 2010, full disk line-of-sight (LOS) magnetograms have been consistently available from another source: the Michelson Doppler Imager (MDI) on board the Solar and Heliospheric Observatory (SOHO; Scherrer et al. 1995). This motivates us to develop a deep learning method, named MagNet, to generate consistent time-sequence vector magnetograms of all ARs from 1996 to 2010, which covered a more active solar cycle 23. Our MagNet model is trained by LOS magnetograms, B_x, B_y from SDO/HMI combined with Hα observations from BBSO. The validation of our approach uses the overlapping data from 2010-05-01 to 2011-04-11 when MDI and HMI obtained data simultaneously. There is a good physical reason to use Hα as the additional constraint for this research: in the solar atmosphere, magnetic fields and flows are at a frozen-in condition, and therefore the Hα fibril structure can provide the direction of magnetic fields in X and Y dimensions (Wiegelmann et al. 2008; Jiang et al. 2021).

Deep learning has been widely used in heliophysics (Galvez et al. 2019; Liu et al. 2019; Jiang et al. 2020, 2021) and astronomy (Kim et al. 2019; Liu et al. 2021). More recently, deep learning was employed to construct magnetograms and estimate magnetic fields. Kim et al. (2019) generated farside solar magnetograms from STEREO/Extreme UltraViolet Imager (EUVI) 304-Å images using a conditional generative adversarial network (cGAN). The authors trained their cGAN model using pairs of SDO/Atmospheric Imaging Assembly (AIA) 304-Å images and SDO/HMI magnetograms. Bai et al. (2021) estimated the unsigned radial component of the magnetic field from photospheric continuum images using a convolutional neural network (CNN). Higgins et al. (2021, 2022) performed Stokes inversion to generate vector magnetograms using deep neural networks. The authors considered both SDO and Hinode missions. Different from the above approaches, our trained MagNet model takes as input LOS magnetograms from SOHO/MDI as well as Hα images from BBSO, and generates as output magnetic field components B'_x and B'_y. The generated (predicted) B'_x and B'_y components along with the LOS components of the magnetic field, which can be treated as B_z components, create vector magnetograms from 1996 to 2010.

The rest of this paper is organized as follows. Section 2 describes solar observations and data preparation used in this study. Section 3 presents details of our MagNet model. Section 4 reports experimental results. Section 5 presents a discussion and concludes the paper.

2. OBSERVATIONS AND DATA PREPARATION

MDI (Scherrer et al. 1995), which is part of the SOHO satellite, acquires a LOS magnetogram every 96 minutes during the period from 1995 to 2011, which covers more active solar cycle 23 with many large flares. The spatial resolution is $4''$ and the full disk images are collected on a 1024×1024 detector. As a successor of MDI, the HMI instrument, which is part of the SDO mission, provides continuous coverage of full disk Doppler velocities, LOS magnetograms, and continuum proxy images (Scherrer et al. 2012). HMI has been operational since May 1, 2010, covering solar cycle 24. HMI observes the full solar disk at 6173 Å every 12 minutes for a better signal-to-noise ratio. The spatial resolution is $1''$ and the full disk images are collected on a 4096×4096 detector. Photospheric vector magnetograms have been available since the launch of SDO/HMI. BBSO, which is a ground-based observatory, has provided Hα observations (images) since 1970s (Denker et al. 1999). BBSO's full disk Hα observations are taken every 1 minute, up to 9 hours for one observing day at the wavelength of 6563 Å. The spatial resolution is $2''$ and the full disk images are collected on a 2048×2048 detector. Unlike satellite-based instruments such as MDI and HMI, the ground-based full disk telescope at BBSO sometimes has seeing limitations due to unstable conditions of Earth's atmosphere and weather.

In our study, we selected full disk Hα images based on three criteria. First, the selected images must be intact. Some images are incomplete due to (i) operating problems such as incorrect angles of the telescope, loss of focus, sudden shaking of the telescope, etc., or (ii) external blocking entities such as birds, airplanes, etc. These incomplete images were excluded. Second, the selected images must show a clear appearance of the solar disk. Those images with unbalanced brightness caused by high and thin clouds were excluded. Third, among the images with a very close appearance where the identifiable items on the solar disk did not have observable changes, the clearest one was selected. When items like active regions and filaments are close to the limb, which is darker than the solar center, we manually check if the items are clear. The images with a very close appearance and less clear items were excluded.

In this way we selected full disk Hα images in the period from 2010-05-01 to 2017-12-31. For each selected full disk Hα image, we collected its temporally closest full disk MDI LOS magnetogram,[1] full disk HMI LOS magnetogram and full disk HMI vector magnetograms/components B_x, B_y where the difference between the time stamp of the full disk Hα image and the time stamp of a collected full disk magnetogram had to be less than 6 minutes. The full disk Hα image and the full disk MDI LOS magnetogram were linearly interpolated to the same spatial size of the full disk HMI LOS magnetogram and HMI vector magnetograms/components using the Interactive Data Language (IDL). These full disk MDI LOS magnetograms (HMI LOS magnetograms, HMI vector magnetograms/components B_x, B_y respectively) were taken directly from the mdi.fd_M_96m_lev182 (hmi.M_720s, hmi.B_720s respectively) series at the Joint Science Operations Center (JSOC).[2]

Because we were mainly interested in solar flares in ARs, we coaligned and cropped AR patches of 256 $\times$ 256 pixels that might produce flares from the full disk images using a two-step coaligning and cropping procedure written by IDL as follows.[3] First, we performed a full disk coalignment via region growing among the full disk images. Then, we cropped and coaligned the AR patches from the coaligned full disk images by maximizing the Pearson correlation coefficient (CC; Galton 1886; Pearson 1895; Sen & Srivastava 1990) among the AR patches. The coaligned AR patches with CC below a threshold were discarded. (In the study presented here, the threshold was set to 0.9.) The AR patch size of 256 $\times$ 256 was chosen for efficient model training purposes where the magnetic field strength at the center of each AR patch was greater than or equal to 1500 Gauss.[4] A large AR would be segmented into several patches, all of which were included and used for training.

In preparing the training data, we selected 2874 full disk Hα images from 2014-01-01 to 2017-08-04 and their temporally closest full disk HMI LOS magnetograms and full disk HMI vector magnetograms/components B_x, B_y. Each time we coaligned and cropped AR patches from four full disk images, namely a full disk HMI LOS magnetogram, a full disk HMI vector magnetogram/component B_x, a full disk HMI vector magnetogram/component B_y and their temporally closest full disk Hα image, using the two-step procedure written by IDL as described above. As a result of this coaligning and cropping process, we obtained 8442 coalignments each containing four AR patches.[5] These 8442 coalignments of AR patches were stored in the training set, denoted Train_HMI, which was used to train our MagNet model where all pixels in the AR patches were used to train the model. The 8442 AR patches of B_x, B_y in Train_HMI were used as labels when optimizing the model.

In preparing the test data, we selected and constructed two test sets to evaluate the performance of the trained MagNet model. The first test set contained 226 full disk HMI LOS magnetograms and full disk HMI vector magnetograms/components B_x, B_y from 2017-08-05 to 2017-12-31 and their temporally closest full disk Hα images. We coaligned and cropped AR patches of 256 $\times$ 256 pixels that might produce flares from these full disk images as done above. We obtained 261 coalignments each containing four AR patches, which were stored in the Test_HMI set. The 261 AR patches of B_x, B_y in Test_HMI were used to evaluate the AR patches of B'_x, B'_y that were predicted (generated) by the trained MagNet model on Test_HMI.

The second test set contained 115 full disk MDI LOS magnetograms, full disk HMI LOS magnetograms, full disk HMI vector magnetograms/components B_x and B_y from 2010-05-01 to 2011-04-11 and their temporally closest full disk Hα images. During this period in which both MDI's data and HMI's data were available, we were able to obtain all these five types of full disk images.[6] Each time we coaligned and cropped AR patches from five full disk images, namely a full disk MDI LOS magnetogram, a full disk HMI LOS magnetogram, a full disk HMI vector magnetogram/component B_x, a full disk HMI vector magnetogram/component B_y and their temporally closest full disk Hα image, using the two-step procedure written by IDL as described above. We obtained 26 coalignments each containing five AR patches, which were stored in the Test_MDI set. The 26 AR patches of B_x, B_y in Test_MDI were used to evaluate the AR patches of B'_x, B'_y that were predicted (generated) by the trained MagNet model on Test_MDI. Please note that, our

[1] The full disk MDI LOS magnetograms were only available from 2010-05-01 to 2011-04-11 in the period between 2010-05-01 and 2017-12-31.

[2] http://jsoc.stanford.edu/

[3] We cropped the AR patches (square regions) from the full disk images/magnetograms, and then all the cropped AR patches were re-sized to 256 $\times$ 256. Some of the AR patches for the training or testing of the MagNet model may be from the same active region on different days.

[4] Strictly speaking, the magnetic field strength of each pixel of the magnetograms collected by HMI and MDI is a pixel-area-averaged signal with a unit generally quoted as Mx/cm^2 (Centeno et al. 2014; Hoeksema et al. 2014). We use Gauss as done in the literature (Rahman et al. 2020; Bai et al. 2021) where Gauss is equivalent to Mx/cm^2 in units.

[5] Both the full disk HMI vector magnetograms B_x, B_y and the AR patches of B_x, B_y cropped from the full disk magnetograms were in the image coordinate system.

[6] Notice that these 115 collections of full disk images in the second test set were independent and disjoint from the 226 collections of full disk images in the first test set due to their totally different collecting periods. Specifically, the first test set covered the period from 2017-08-05 to 2017-12-31 whereas the second test set covered the period from 2010-05-01 to 2011-04-11. The training set, Train_HMI, covered the period from 2014-01-01 to 2017-08-04.

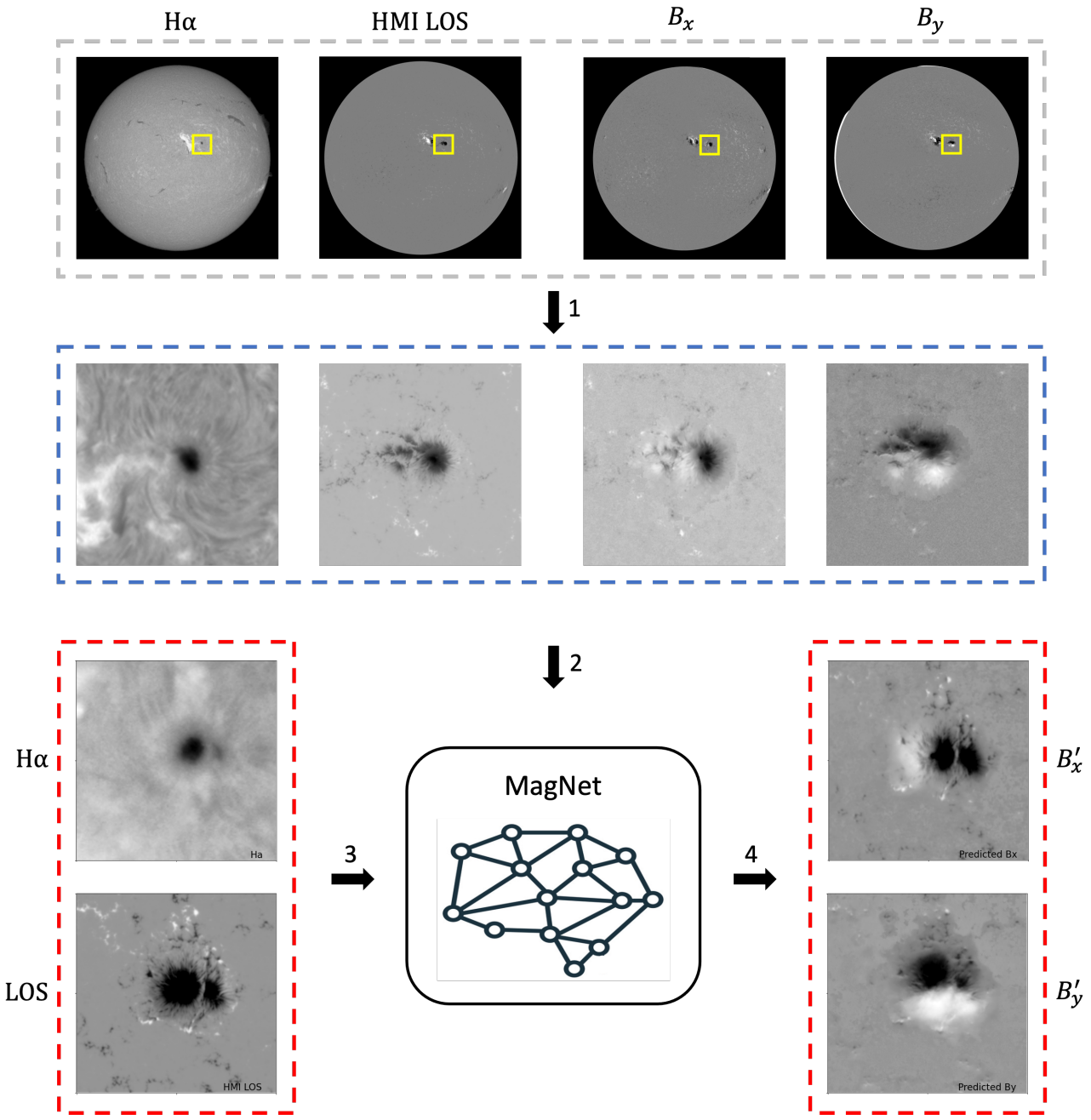

Figure 1. The workflow of MagNet. During training, we coalign and crop AR patches of 256 × 256 pixels from four full disk images (step 1) and use the coaligned AR patches to train the MagNet model (step 2). During testing/prediction, a pair of coaligned AR patches of Hα image and HMI/MDI LOS magnetogram is fed to the trained Magnet model (step 3), which generates the pair of AR patches of B'_x and B'_y that correspond to the input test data (step 4).

training set (Train_HMI) and test sets (Test_HMI and Test_MDI) covered different time periods, and therefore were disjoint. Thus, the trained MagNet model can make predictions on data that it has never seen before.

3. METHODOLOGY

3.1. *The Workflow of MagNet*

Figure 1 illustrates the workflow of MagNet. Collected full disk images including BBSO Hα images, HMI LOS magnetograms, and HMI vector components B_x and B_y are shown in the gray dashed box where an AR patch in a

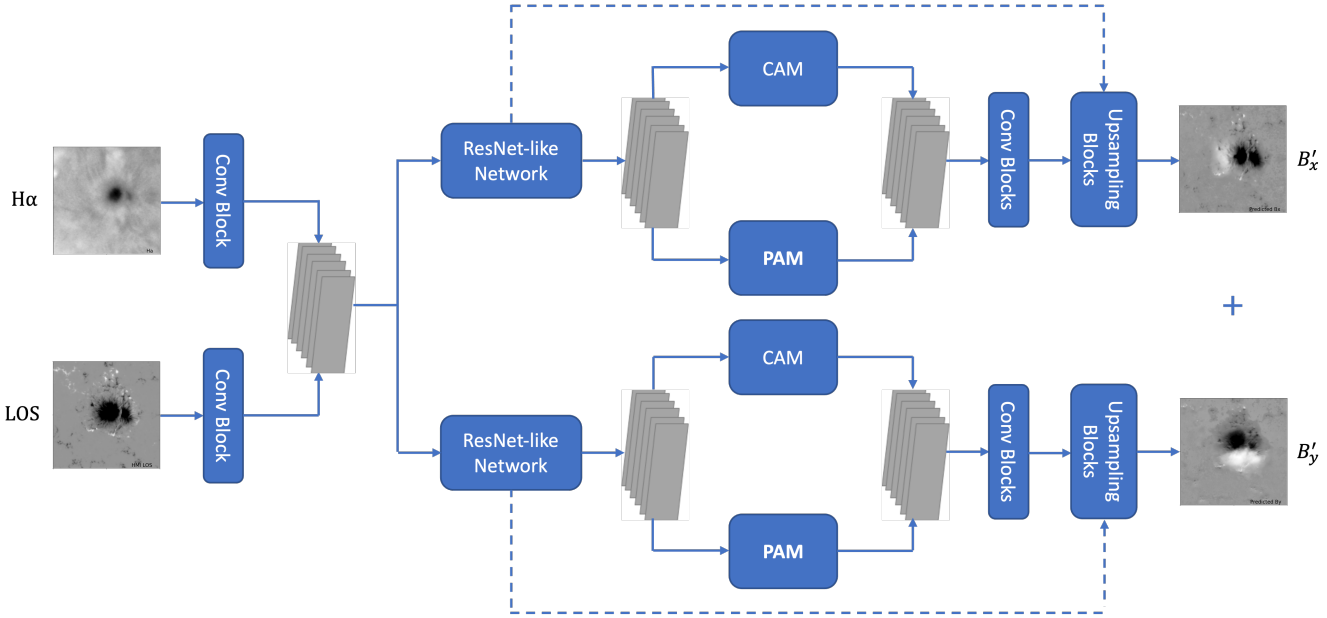

Figure 2. Illustration of the architecture of the MagNet model. The model takes as input a pair of coaligned AR patches of Hα image and HMI/MDI LOS magnetogram, and generates as output the pair of AR patches of B'_x and B'_y that correspond to the input data. See text for details of the model architecture.

full disk image is highlighted in a small yellow box. We process these full disk images by coaligning and cropping the AR patch out of each full disk image to produce coaligned AR patches of 256 × 256 pixels enclosed in the blue dashed box (step 1). During training, pairs of coaligned AR patches of BBSO Hα images and HMI LOS magnetograms are fed to the MagNet model where the corresponding AR patches of HMI vector components B_x and B_y, treated as true B_x and B_y, are used as labels (step 2). During testing/prediction, the test data containing a pair of coaligned AR patches of BBSO Hα image and HMI/MDI LOS magnetogram, shown in the left red dashed box, is fed to the trained MagNet model (step 3). The model predicts (generates) the pair of AR patches of B'_x and B'_y, shown in the right red dashed box, that correspond to the input test data (step 4).

3.2. *The MagNet Model*

Generating AR patches of B'_x and B'_y amounts to solving a regression problem because the output of the generating procedure consists of real numbers, i.e., magnetic field strengths. We employ a novel deep learning model containing an advanced CNN with self-attention (Zhao et al. 2020) to solve this regression problem. Self-attention was originally developed to learn global dependencies of input data and was used in machine translation (Vaswani et al. 2017). Later, self-attention was applied to many areas such as image recognition, scene segmentation and image synthesis, etc. Figure 2 presents the architecture of our deep learning model.

During training, each pair of coaligned AR patches of BBSO Hα image and HMI LOS magnetogram taken from the training set, Train_HMI, is in turn fed to the MagNet model. The input images are first processed by convolutional blocks. Each convolutional block consists of a convolution layer with batch normalization, followed by a parametric ReLU (PReLU) activation function (He et al. 2015). The feature maps produced by the convolutional blocks are concatenated and sent to two separate paths where one path generates the AR patch of B'_x and the other path generates the AR patch of B'_y. Each path starts with a ResNet-like network, specifically ResNet-101 (He et al. 2016) with a dilated convolution strategy (Zhang et al. 2018) as described in Fu et al. (2021). The output of the ResNet-like network is then sent to two modules: the channel attention module (CAM) and position attention module (PAM) (Fu et al. 2019). CAM and PAM leverage the self-attention mechanisms in the modules to better capture and transform a wider range of contextual information into local features, thus enhancing their representation capability. Both CAM and PAM are calculated in a similar way where CAM applies the self-attention mechanism to image channels while PAM focuses on the calculation of location information. The outputs of CAM and PAM are combined and sent to convolutional blocks, followed by upsampling blocks (i.e., convolutional blocks with upsampling and skip connection represented as dashed lines in Figure 2 (Falk et al. 2019)), to generate the AR patch of B'_x (B'_y, respectively). The

AR patches of generated B'_x and B'_y are compared with labels (i.e., the corresponding AR patches of true B_x and B_y). The weights of neurons in the MagNet model are then updated to minimize the error (loss) caused by the comparison. At the end of the training process, the weights are optimized.

During testing/prediction, the trained MagNet model takes as input a pair of coaligned AR patches of BBSO Hα image and HMI LOS magnetogram (MDI LOS magnetogram, respectively) taken from the Test_HMI (Test_MDI, respectively) set and generates as output the pair of AR patches of B'_x and B'_y that correspond to the input test data. The AR patches of generated B'_x and B'_y are evaluated by the corresponding AR patches of HMI vector components B_x and B_y, treated as true B_x and B_y, in Test_HMI (Test_MDI, respectively).

For many of the AR patches used in our study, a large portion of each of them has small magnetic field strengths ($\leq$ 200 Gauss). Relatively few pixels in an AR patch have large magnetic field strengths ($>$ 200 Gauss). To tackle this imbalanced problem in our datasets, we employ a novel weighted L_w^p loss of a pixel p, defined as:

$$L_w^p(s', s) = |\frac{s}{c}||s' - s|. \tag{1}$$

Here, s' represents the MagNet-generated magnetic field strength at p, s represents the true magnetic field strength at p, and c is a threshold. (In the study presented here, the threshold was set to 0.95.) The absolute difference between s' and s at p, usually reflected by the L_1 loss, is multiplied by a weight, $|\frac{s}{c}|$. This suggests that a pixel p with a larger (smaller, respectively) magnetic field strength yield a larger (smaller, respectively) L_w^p loss.

The weighted loss between an AR patch, A, of generated B'_x and the corresponding AR patch of true B_x, denoted $L_w(B'_x, B_x)$, is defined as:

$$L_w(B'_x, B_x) = \frac{1}{N}\sum_{p\in A} L_w^p(s', s), \tag{2}$$

where $N = 256 \times 256 = 65536$ is the total number of pixels in A. The weighted loss between an AR patch of generated B'_y and the corresponding AR patch of true B_y, denoted $L_w(B'_y, B_y)$, is defined similarly.

The total loss, denoted L_{MagNet}, is then defined as the sum of $L_w(B'_x, B_x)$ and $L_w(B'_y, B_y)$, as shown in Equation (3) below:

$$L_{MagNet} = L_w(B'_x, B_x) + L_w(B'_y, B_y). \tag{3}$$

The training of MagNet was done by applying the adaptive moment estimation (Adam) optimizer (Balles & Hennig 2018; Zou et al. 2019) to minimize L_{MagNet} with 100 epochs on an NVIDIA A100 GPU.[7]

4. EXPERIMENTS AND RESULTS

4.1. *Evaluation Metrics*

We adopt three metrics, namely the mean absolute error (MAE; Sen & Srivastava 1990), Pearson correlation coefficient (CC; Galton 1886; Pearson 1895; Sen & Srivastava 1990) and % Within t (Higgins et al. 2021) to quantitatively evaluate the performance of MagNet. The first metric is defined as:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |s'_i - s_i|, \tag{4}$$

where $N = 65536$ is the total number of pixels in an AR patch of a vector component, and s_i (s'_i, respectively) denotes the true (generated, respectively) magnetic field strength for the ith pixel, $1 \leq i \leq 65536$, in the vector component. This metric has a unit of Gauss, which is used to quantitatively assess the dissimilarity (distance) between the true magnetic field strengths and generated magnetic field strengths in the vector component.

The second metric is defined as:

$$\text{CC} = \frac{\text{E}[(T - \mu_T)(G - \mu_G)]}{\sigma_T \sigma_G}, \tag{5}$$

where $E(\cdot)$ is the expectation. T and G denote the true magnetic field strengths and generated magnetic field strengths respectively in the vector component. μ_T and μ_G denote the mean of T and G respectively. σ_T and σ_G denote the

[7] Our MagNet model is coded by Python and TensorFlow. The code and pretrained model are available at https://nature.njit.edu/solardb/magnet. The full disk Hα images can be downloaded from http://www.bbso.njit.edu/Research/FDHA/. There are two types of images, *fl* and *fr*, from the above BBSO site. In this study, we use the *fl* images.

Table 1. Evaluation Metric Values of MagNet Based on the Test Data

Test Sets	MAE	CC	% Within t
Test_HMI(B'_x)	(38.55, 63.11, 295.92)	(0.7149, 0.9105, 0.9703)	(36.50%, 84.19%, 95.03%)
Test_HMI(B'_y)	(41.76, 62.62, 216.13)	(0.7043, 0.8641, 0.9704)	(34.39%, 82.03%, 93.90%)
Test_MDI(B'_x)	(51.98, 87.35, 174.85)	(0.6683, 0.7510, 0.8384)	(56.52%, 74.30%, 88.18%)
Test_MDI(B'_y)	(48.73, 83.63, 130.58)	(0.6737, 0.7956, 0.8808)	(55.40%, 72.01%, 89.93%)

standard deviation of T and G respectively. CC, which does not have units, has a value between -1 and 1. When CC is -1 or 1, there is an exact linear relationship between T and G. When CC is 0, there is no linear dependency between T and G.

The third metric is the percentage count of well-estimated pixels (Higgins et al. 2022), which is defined as:

$$\%\ \text{Within t} = \frac{M}{N} \times 100\%, \tag{6}$$

where M denotes the total number of agreement pixels in the vector component. We say the ith pixel in the vector component is an agreement pixel if $|s_i - s'_i|$ is smaller than a user-specified threshold t. (In the study presented here, the threshold t was set to 100 Gauss.) This metric is used to quantitatively assess the similarity between the true magnetic field strengths and generated magnetic field strengths in the vector component.

4.2. *Quantitative Evaluation of the MagNet Model on HMI and MDI Data*

Table 1 presents the three evaluation metric values of MagNet based on the data in the Test_HMI and Test_MDI sets. Each triple in the table consists of three values where the left one is the minimum value, the middle one is the average value, and the right one is the maximum value. The row of Test_HMI(B'_x) (Test_HMI(B'_y), respectively) shows the metric values obtained by generating B'_x (B'_y, respectively) based on the test data in Test_HMI. The row of Test_MDI(B'_x) (Test_MDI(B'_y), respectively) shows the metric values obtained by generating B'_x (B'_y, respectively) based on the test data in Test_MDI. The training data were from Train_HMI.

It can be seen from Table 1 that the generated B'_x (B'_y, respectively) components are close to the true B_x (B_y, respectively) components with the average MAE being less than 100 Gauss for both of the HMI and MDI test data. The average CC is approximately 0.9 for the HMI test data and close to 0.8 for the MDI test data. The average % Within t is approximately 83% for the HMI test data and 73% for the MDI test data. MagNet generally performs better on the HMI test data than on the MDI test data. This happens due to several reasons. First, the MagNet model is trained by HMI data, not MDI data. The spatial resolution of MDI images is $4''$, which is lower than the resolution, $1''$, of HMI images. Furthermore, due to the larger cadence (96 minutes) of MDI compared to the cadence (12 minutes) of HMI, the time gaps between the MDI images and their coaligned Hα images are often larger than the time gaps between the HMI images and their coaligned Hα images. As a result, the quality of the coalignments of AR patches for MDI is lower than the quality of the coalignments of AR patches for HMI. In what follows we present some sample predictions for HMI and MDI.

4.3. *Case Studies of the MagNet Model on HMI and MDI Data*

We follow the Mount Wilson (or Hale) classification system for sunspot groups (Hale et al. 1919; Jaeggli & Norton 2016) to indicate whether an active region is complex or not. Specifically, we consider active regions in the α and β classes to be simple ARs, and active regions in the other classes to be complex ARs.

Generating vector components of AR 12683 based on BBSO Hα and HMI LOS data. Figure 3 presents generated B'_x and B'_y components for AR 12683 on 2017 October 2 21:00:00 UT where training data were from Train_HMI. Figure 3(A) shows a pair of coaligned 256×256 patches of BBSO Hα image and HMI LOS magnetogram from AR 12683. This pair of test images is used as input to the trained MagNet model. Figure 3(B) presents results produced by MagNet with respect to the test data in Figure 3(A). In Figure 3(B), the first column shows 2D histograms which illustrate the density of the pixel distribution (Rahman et al. 2020), the second column shows the AR patches of B'_x and B'_y components generated/predicted by MagNet, and the third column shows the corresponding AR patches of true B_x and B_y components. In each 2D histogram, the X-axis represents true B_x (B_y, respectively) and the Y-axis represents generated/predicted B'_x (B'_y, respectively). The diagonal line in the 2D histogram corresponds to pixels

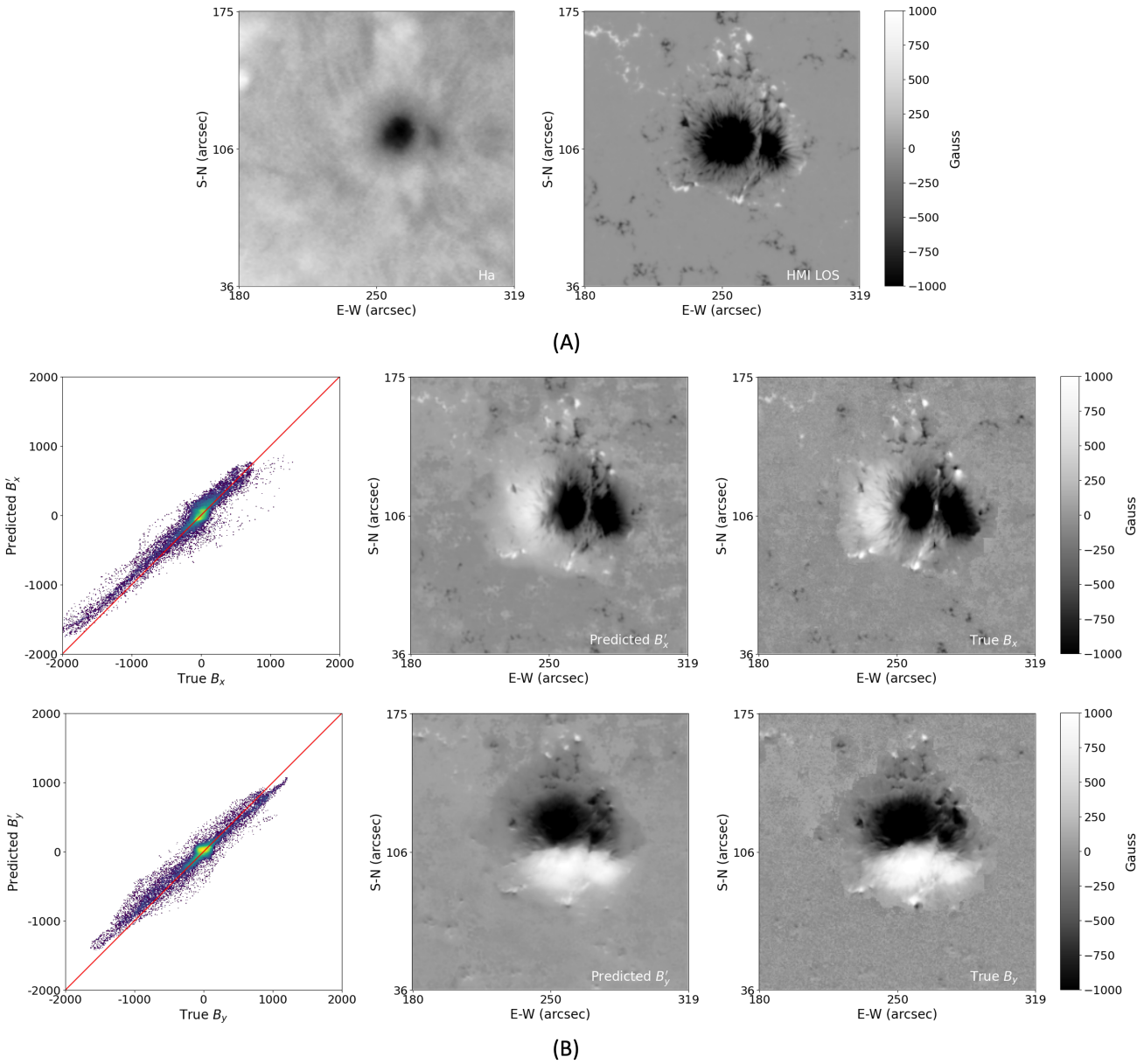

Figure 3. Comparison between MagNet-generated B'_x, B'_y and true B_x, B_y based on BBSO Hα and HMI LOS data of AR 12683 on 2017 October 2 21:00:00 UT. (A) The Hα and HMI LOS data inputted to MagNet. (B) The output results from MagNet. See text for details of the input data and output results.

whose generated B'_x (B'_y, respectively) values are identical to true B_x (B_y, respectively) values. AR 12683 was a relatively simple active region. It can be seen from Figure 3(B) that MagNet works extremely well on this simple AR, capable of generating vector components that are very close to true components with an MAE of 52.34 Gauss and 57.49 Gauss, a CC of 0.9581 and 0.9631, and a % Within t of 87.20% and 84.72% for B_x and B_y respectively.

Generating vector components of AR 12673 based on BBSO Hα and HMI LOS data. AR 12673 on 2017 September 6 was a very complex active region. It was the most flare-productive AR in solar cycle 24, showing strong magnetic fields in the light bridge and apparent photospheric twist (Wang et al. 2018), and produced four X-class flares including an X9.3 flare on 2017 September 06. This active region contained many pixels with extremely large magnetic field strengths (saturated at the value of 5000 Gauss or larger). The MagNet model trained by the 8442 coalignments of AR patches in the Train_HMI set described in Section 2 did not perform well in AR 12673 on 2017

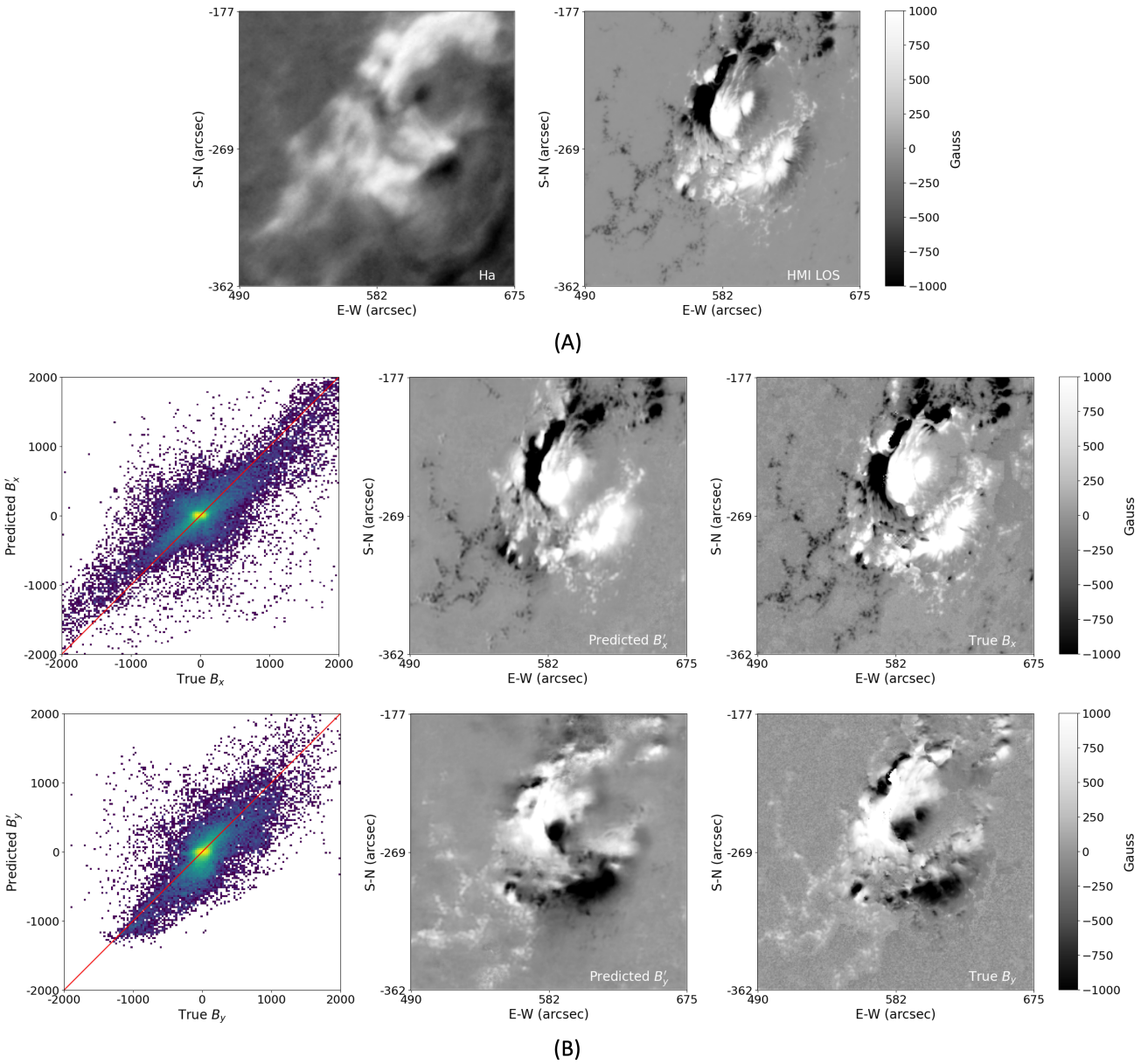

Figure 4. Comparison between MagNet-generated B_x', B_y' and true B_x, B_y based on BBSO Hα and HMI LOS data of AR 12673 on 2017 September 6 19:00:00 UT. (A) The Hα and HMI LOS data inputted to MagNet. (B) The output results from MagNet.

September 6, due to the lack of knowledge of very complex structures and extremely large magnetic field strengths such as those in AR 12673 on 2017 September 6.

It was observed that AR 12673 evolved very rapidly. There were dramatic changes and significant differences during the period between 2017 September 5 and September 7 in which AR 12673 was very complex. On 2017 September 4, AR 12673 was a very simple active region and on September 8, it was near the limb. To enhance the knowledge of MagNet, we created a new training set, denoted Train_HMI_New, by manually picking 835 HMI LOS magnetograms from complex ARs in 2015 together with 3 HMI LOS magnetograms and 18 HMI LOS magnetograms from AR 12673 on 2017 September 5 and September 7 respectively. Due to the rapid daily changes of AR 12673 as explained above, the information in the new training set, Train_HMI_New, which contained images from 2017 September 5 and September 7, was disjoint and independent from the information in the test images from 2017 September 6. This new training

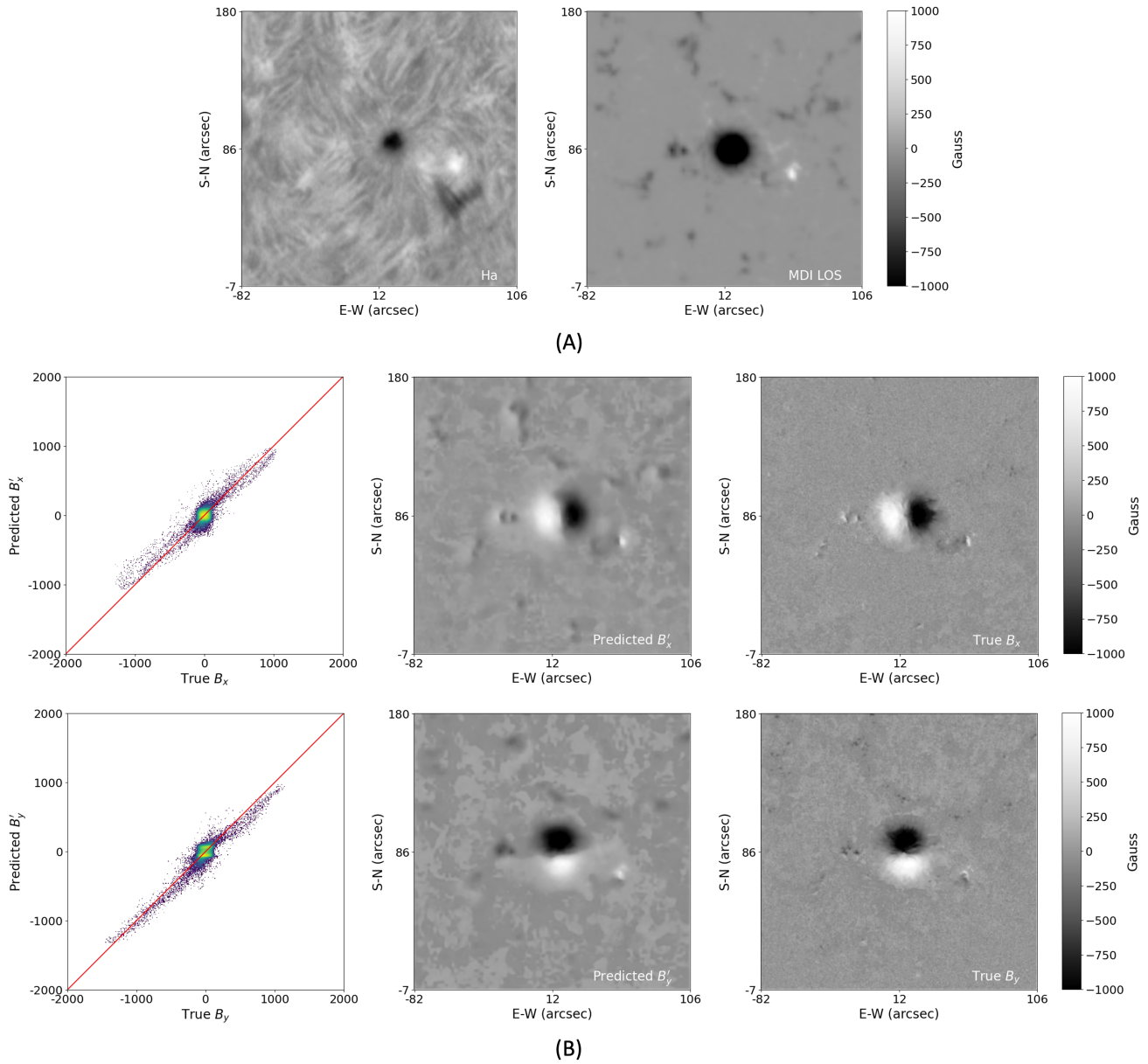

Figure 5. Comparison between MagNet-generated B'_x, B'_y and true B_x, B_y based on BBSO Hα and MDI LOS data of AR 11101 on 2010 August 30 17:36:00 UT. (A) The Hα and MDI LOS data inputted to MagNet. (B) The output results from MagNet.

set contained 835 + 21 = 856 coalignments each containing four AR patches from Hα image, HMI LOS magnetogram, HMI vector component B_x and HMI vector component B_y respectively.

Figure 4 presents the B'_x and B'_y components for AR 12673 on 2017 September 6 19:00:00 UT generated by the new MagNet model trained by Train_HMI_New. Figure 4(A) shows a pair of coaligned 256 × 256 patches of BBSO Hα image and HMI LOS magnetogram from AR 12673. This pair of test images is used as input to the new MagNet model. Figure 4(B) presents results produced by the new MagNet model with respect to the test data in Figure 4(A). It can be seen from Figure 4(B) that the new MagNet model works well on the very complex AR 12673 on 2017 September 6, capable of generating vector components that are close to true components with an MAE of 145.31 Gauss and 132.33 Gauss, a CC of 0.8150 and 0.7304, and a % Within t of 66.11% and 63.96% for B_x and B_y respectively.

Generating vector components of AR 11101 based on BBSO Hα and MDI LOS data. Figure 5 presents generated B'_x and B'_y components for AR 11101 on 2010 August 30 17:36:00 UT where training data were from

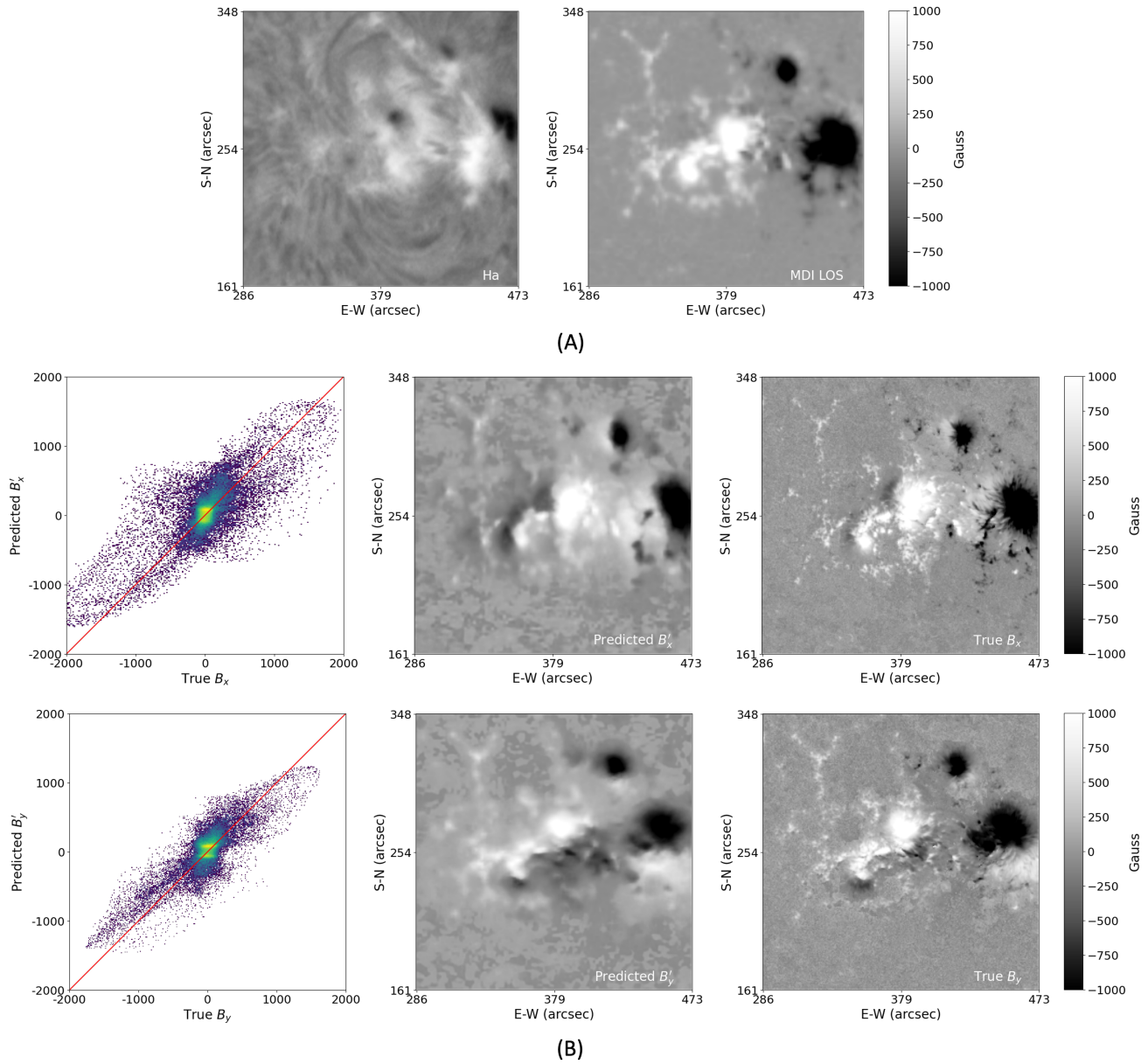

Figure 6. Comparison between MagNet-generated B_x', B_y' and true B_x, B_y based on BBSO Hα and MDI LOS data of AR 11117 on 2010 October 27 22:24:00 UT. (A) The Hα and MDI LOS data inputted to MagNet. (B) The output results from MagNet.

Train_HMI. Figure 5(A) shows a pair of coaligned 256 × 256 patches of BBSO Hα image and MDI LOS magnetogram of AR 11101. This pair of images is used as input to the trained MagNet model. Figure 5(B) presents results produced by MagNet with respect to the test data in Figure 5(A). AR 11101 was a relatively simple active region. It can be seen from Figure 5(B) that MagNet works pretty well on this simple AR, capable of generating vector components that are similar to true components with an MAE of 52.53 Gauss and 53.16 Gauss, a CC of 0.8264 and 0.8459, and a % Within t of 87.66% and 86.33% for B_x and B_y respectively. We note both AR 11101 and the AR 12683 described above have similar evaluation metric values. MagNet is trained by HMI data and the vector components of AR 12683 are generated based on test data also from HMI (though the training set and test set are disjoint). On the other hand, the vector components of AR 11101 are generated based on test data from MDI. The fact that AR 11101 and AR 12683 have similar evaluation metric values demonstrates the good learning and inference capability of MagNet.

Generating vector components of AR 11117 based on BBSO Hα and MDI LOS data. Figure 6 presents generated B'_x and B'_y components for AR 11117 on 2010 October 27 22:24:00 UT where training data were from Train_HMI. Figure 6(A) shows a pair of coaligned 256 × 256 patches of BBSO Hα image and MDI LOS magnetogram of AR 11117. This pair of images is used as input to the trained MagNet model. Figure 6(B) presents results produced by MagNet with respect to the test data in Figure 6(A). AR 11117 was a relatively complex active region. It can be seen from Figure 6(B) that MagNet works reasonably well on this complex AR, capable of generating vector components that are similar to true components with an MAE of 129.00 Gauss and 114.68 Gauss, a CC of 0.7514 and 0.8050, and a % Within t of 61.22% and 58.63% for B_x and B_y respectively. Compared to the evaluation metric values for the relatively simple AR 11101 described above, the evaluation metric values for the more complex AR 11117 are lower.

5. DISCUSSION AND CONCLUSIONS

We present a new deep learning method (MagNet) for generating photospheric vector magnetograms of solar active regions for SOHO/MDI using SDO/HMI and BBSO data. This method allows us to expand the availability of photospheric vector magnetograms to the period from 1996 to present, covering solar cycles 23 and 24. The vector magnetograms can be used by physics-based methods to calculate magnetic energy and magnetic field parameters useful for predicting solar flare activity (Liu et al. 2017, 2019, 2020).

Our main results are summarized as follows.

1. The experimental results obtained by using BBSO Hα observations and HMI magnetograms in the period between 2014-01-01 and 2017-08-04 as training data demonstrated the good performance of the proposed method. Specifically, when using the trained MagNet model to generate vector components B'_x and B'_y based on BBSO Hα and HMI LOS data and evaluated by HMI/true vector components B_x and B_y in the Test_HMI set during the period between 2017-08-05 and 2017-12-31, we obtained an average MAE of ∼63 Gauss, CC of ∼0.9, and % Within t of ∼83%. When using the trained MagNet model to generate vector components B'_x and B'_y based on BBSO Hα and MDI LOS data and evaluated by the HMI/true vector components B_x and B_y in the Test_MDI set from the overlapping period of MDI and HMI between 2010-05-01 and 2011-04-11, we obtained an average MAE of ∼85 Gauss, CC of ∼0.8, and % Within t of ∼73%.

2. We reported case studies of the MagNet model on HMI and MDI data, which quantitatively and visually demonstrated how our model worked on these data. Like other machine learning methods, the performance of MagNet depends on training data. For example, in handling the very complex active region AR 12673, our original model trained by the Train_HMI set does not perform well. We have to include complex images in the training set in order to generate satisfactory vector magnetograms. Most of the AR patches in the Train_HMI set are relatively simple; they have small magnetic field strengths with the maximum magnetic field strength being much smaller than 5000 Gauss. As a consequence, the model trained by the Train_HMI set performs well in generating vector components with small magnetic field strengths. It suffers when generating vector components having complex structures with very large magnetic field strengths. On the other hand, with the new training data in the Train_HMI_New set, which contains selected complex structures with very large magnetic field strengths, MagNet exhibits good performance as shown in Figure 4.

3. When training and testing the MagNet model, we did not impose any threshold on the data. In other words, all pixels in the original images from the instruments (MDI, HMI, BBSO) were included. Imposing thresholds on the data has an impact on the performance of MagNet. For example, suppose that, during training, we only consider pixels whose magnetic field strengths range from −2000 Gauss and +2000 Gauss (Rahman et al. 2020). Due to the lack of knowledge of pixels with large magnetic field strengths (e.g., 5000 Gauss) in a complex active region such as AR 12673, MagNet suffers when predicting the test images from the complex AR with large magnetic field strengths.

4. MagNet is trained by HMI data and tested on MDI data. However, HMI and MDI are two different instruments on different observatories with different temporal and spatial resolutions. One could use deep learning techniques (Menon et al. 2020; Danier et al. 2022; Rombach et al. 2022) to improve the temporal and spatial resolutions of the images from the different sources to enhance the performance of MagNet.

Transfer learning could also be helpful in obtaining knowledge from one source and applying the knowledge to another source (Zhuang et al. 2021).

5. Our approach heavily depends on the coalignments of AR patches, which are produced by a two-step coaligning and cropping procedure written by IDL. Coaligning images from the different instruments is a challenge. Developing more accurate coalignment algorithms implemented in OpenCV or IDL may further improve the performance of MagNet.

6. Due to the operation time and condition limitations, BBSO did not collect full disk Hα images for all the active regions (ARs) in solar cycle 23. As a consequence, vector components of those missing ARs would not be generated. To get full coverage of the ARs in solar cycle 23, one would need to resort to other sources (e.g., Kanzelhöhe Solar Observatory (KSO; Otruba 1999; Otruba et al. 2008)) that provide complementary full disk Hα images.

To our knowledge, MagNet is the first method capable of generating photospheric vector magnetograms of solar active regions for SOHO/MDI using SDO/HMI and Hα data. According to our findings and results, the vector components generated by MagNet are reasonably close to ground truths (observed data). More research is needed to further improve the quality of the generated data.

We thank the referee and scientific editor for very helpful and thoughtful comments. We also thank the BBSO team for providing the data used in this study. The BBSO operation is supported by the New Jersey Institute of Technology and U.S. NSF grant AGS-1821294. This work was supported by U.S. NSF grants AGS-1927578, AGS-1954737, AGS-2149748 and AGS-2228996. Q.L., J.J., Y.X. and H.W. acknowledge the support of NASA under grants 80NSSC18K1705, 80NSSC19K0068 and 80NSSC20K1282.

Facilities: Big Bear Solar Observatory, Solar Dynamics Observatory, Solar and Heliospheric Observatory

REFERENCES

Abduallah, Y., Wang, J. T. L., Nie, Y., Liu, C., & Wang, H. 2021, Research in Astronomy and Astrophysics, 21, 160, doi: 10.1088/1674-4527/21/7/160

Aschwanden, M. J., Xu, Y., & Jing, J. 2014, ApJ, 797, 50, doi: 10.1088/0004-637X/797/1/50

Bai, X., Liu, H., Deng, Y., et al. 2021, A&A, 652, A143, doi: 10.1051/0004-6361/202140374

Balles, L., & Hennig, P. 2018, in Proceedings of Machine Learning Research, Vol. 80, Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018, ed. J. G. Dy & A. Krause (PMLR), 413–422. http://proceedings.mlr.press/v80/balles18a.html

Bobra, M. G., & Couvidat, S. 2015, ApJ, 798, 135, doi: 10.1088/0004-637X/798/2/135

Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549, doi: 10.1007/s11207-014-0529-3

Centeno, R., Schou, J., Hayashi, K., et al. 2014, SoPh, 289, 3531, doi: 10.1007/s11207-014-0497-7

Chen, Y., Manchester, W. B., Hero, A. O., et al. 2019, Space Weather, 17, 1404, doi: 10.1029/2019SW002214

Danier, D., Zhang, F., & Bull, D. R. 2022, in IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022 (IEEE), 3511–3521, doi: 10.1109/CVPR52688.2022.00351

Denker, C., Johannesson, A., Marquette, W., et al. 1999, SoPh, 184, 87, doi: 10.1023/A:1005047906097

Falk, T., Mai, D., Bensch, R., et al. 2019, Nature Methods, 16, 67, doi: 10.1038/s41592-018-0261-2

Fu, J., Liu, J., Jiang, J., et al. 2021, IEEE Transactions on Neural Networks and Learning Systems, 32, 2547, doi: 10.1109/TNNLS.2020.3006524

Fu, J., Liu, J., Tian, H., et al. 2019, in IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019 (Computer Vision Foundation / IEEE), 3146–3154, doi: 10.1109/CVPR.2019.00326

Galton, F. 1886, The Journal of the Anthropological Institute of Great Britain and Ireland, 15, 246. https://doi.org/10.2307%2F2841583

Galvez, R., Fouhey, D. F., Jin, M., et al. 2019, ApJS, 242, 7, doi: 10.3847/1538-4365/ab1005

Hale, G. E., Ellerman, F., Nicholson, S. B., & Joy, A. H. 1919, ApJ, 49, 153, doi: 10.1086/142452

He, K., Zhang, X., Ren, S., & Sun, J. 2015, in 2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015, 1026–1034. https://doi.org/10.1109/ICCV.2015.123

He, K., Zhang, X., Ren, S., & Sun, J. 2016, in 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016 (IEEE Computer Society), 770–778, doi: 10.1109/CVPR.2016.90

Higgins, R. E. L., Fouhey, D. F., Zhang, D., et al. 2021, ApJ, 911, 130, doi: 10.3847/1538-4357/abd7fe

Higgins, R. E. L., Fouhey, D. F., Antiochos, S. K., et al. 2022, ApJS, 259, 24, doi: 10.3847/1538-4365/ac42d5

Hoeksema, J. T., Liu, Y., Hayashi, K., et al. 2014, SoPh, 289, 3483, doi: 10.1007/s11207-014-0516-8

Ichimoto, K., Sakurai, T., Yamaguchi, A., et al. 1991, in Flare Physics in Solar Activity Maximum 22, ed. Y. Uchida, R. C. Canfield, T. Watanabe, & E. Hiei, Vol. 387, 320, doi: 10.1007/BFb0032657

Jaeggli, S. A., & Norton, A. A. 2016, ApJL, 820, L11, doi: 10.3847/2041-8205/820/1/L11

Jiang, H., Jing, J., Wang, J., et al. 2021, ApJS, 256, 20, doi: 10.3847/1538-4365/ac14b7

Jiang, H., Wang, J., Liu, C., et al. 2020, ApJS, 250, 5, doi: 10.3847/1538-4365/aba4aa

Keller, C. U., Harvey, J. W., & Giampapa, M. S. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4853, Innovative Telescopes and Instrumentation for Solar Astrophysics, ed. S. L. Keil & S. V. Avakyan, 194–204, doi: 10.1117/12.460373

Kim, T., Park, E., Lee, H., et al. 2019, Nature Astronomy, 3, 397, doi: 10.1038/s41550-019-0711-5

Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, SoPh, 243, 3, doi: 10.1007/s11207-007-9014-6

Liu, C., Deng, N., Wang, J. T. L., & Wang, H. 2017, ApJ, 843, 104, doi: 10.3847/1538-4357/aa789b

Liu, H., Liu, C., Wang, J. T. L., & Wang, H. 2019, ApJ, 877, 121, doi: 10.3847/1538-4357/ab1b3c

—. 2020, ApJ, 890, 12, doi: 10.3847/1538-4357/ab6850

Liu, J., Wang, Y., Huang, X., et al. 2021, Nature Astronomy, 5, 108, doi: 10.1038/s41550-021-01310-6

Menon, S., Damian, A., Hu, S., Ravi, N., & Rudin, C. 2020, in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020 (Computer Vision Foundation / IEEE), 2434–2442, doi: 10.1109/CVPR42600.2020.00251

Metcalf, T. R., Leka, K. D., & Mickey, D. L. 2005, ApJL, 623, L53, doi: 10.1086/429961

Mickey, D. L. 1985, SoPh, 97, 223, doi: 10.1007/BF00165987

Mickey, D. L., Canfield, R. C., Labonte, B. J., et al. 1996, SoPh, 168, 229, doi: 10.1007/BF00148052

Otruba, W. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 184, Third Advances in Solar Physics Euroconference: Magnetic Fields and Oscillations, ed. B. Schmieder, A. Hofmann, & J. Staude, 314–318. https://ui.adsabs.harvard.edu/abs/1999ASPC..184..314O

Otruba, W., Freislich, H., & Hanslmeier, A. 2008, Central European Astrophysical Bulletin, 32, 1. https://ui.adsabs.harvard.edu/abs/2008CEAB...32....1O

Pearson, K. 1895, Proceedings of the Royal Society of London, 58, 240. https://doi.org/10.1098%2Frspl.1895.0041

Rahman, S., Moon, Y.-J., Park, E., et al. 2020, ApJL, 897, L32, doi: 10.3847/2041-8213/ab9d79

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. 2022, in IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022 (IEEE), 10674–10685, doi: 10.1109/CVPR52688.2022.01042

Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129, doi: 10.1007/BF00733429

Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207, doi: 10.1007/s11207-011-9834-2

Schrijver, C. J., DeRosa, M. L., Metcalf, T., et al. 2008, ApJ, 675, 1637, doi: 10.1086/527413

Sen, A., & Srivastava, M. 1990, Regression Analysis (Springer-Verlag New York), doi: https://doi.org/10.1007/978-1-4612-4470-7

Sun, X., Hoeksema, J. T., Liu, Y., et al. 2012, ApJ, 748, 77, doi: 10.1088/0004-637X/748/2/77

Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. 2008, SoPh, 249, 167, doi: 10.1007/s11207-008-9174-z

Vaswani, A., Shazeer, N., Parmar, N., et al. 2017, in Advances in Neural Information Processing Systems, ed. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett, Vol. 30 (Curran Associates, Inc.). https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Wang, H., Yurchyshyn, V., Liu, C., et al. 2018, Research Notes of the AAS, 2, 8. https://doi.org/10.3847/2515-5172/aaa670

Wang, X., Chen, Y., Toth, G., et al. 2020, ApJ, 895, 3, doi: 10.3847/1538-4357/ab89ac

Wheatland, M. S., Sturrock, P. A., & Roumeliotis, G. 2000, ApJ, 540, 1150, doi: 10.1086/309355

Wiegelmann, T. 2004, SoPh, 219, 87, doi: 10.1023/B:SOLA.0000021799.39465.36

Wiegelmann, T., Thalmann, J. K., Schrijver, C. J., De Rosa, M. L., & Metcalf, T. R. 2008, SoPh, 247, 249, doi: 10.1007/s11207-008-9130-y

Zhang, H., Dana, K. J., Shi, J., et al. 2018, in 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018 (Computer Vision Foundation / IEEE Computer Society), 7151–7160, doi: 10.1109/CVPR.2018.00747

Zhao, H., Jia, J., & Koltun, V. 2020, in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020 (Computer Vision Foundation / IEEE), 10073–10082, doi: 10.1109/CVPR42600.2020.01009

Zhuang, F., Qi, Z., Duan, K., et al. 2021, Proc. IEEE, 109, 43, doi: 10.1109/JPROC.2020.3004555

Zou, F., Shen, L., Jie, Z., Zhang, W., & Liu, W. 2019, in IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019 (Computer Vision Foundation / IEEE), 11127–11135, doi: 10.1109/CVPR.2019.01138