Draft version April 11, 2023
Typeset using LaTeX default style in AASTeX631

Generating Space-based SDO/HMI-like Solar Magnetogram from Ground-based Hα Image by Deep Learning

Fei Gao,[1,2] Tie Liu,[3,4] WenQing Sun,[1,2] and Long Xu[1,5]

[1] State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] School of Astronomy and Space Science, Nanjing University, Nanjing 210023, China
[4] Key Laboratory for Modern Astronomy and Astrophysics (Nanjing University), Ministry of Education, Nanjing 210023, China
[5] Peng Cheng National Laboratory, Shenzhen 518000, China

ABSTRACT

Recently, the method of estimating magnetic field through monochromatic images by deep learning has been proposed, demonstrating good morphological similarity but somewhat poor magnetic polarity consistency relative to real observation. In this paper, we propose to estimate magnetic field from Hα images by using conditional Generative Adversarial Network (cGAN) as the basic framework. The Hα images from the Global Oscillation Network Group (GONG) are used as the inputs and the line-of-sight (LOS) magnetograms of the Helioseismic Magnetic Imager (HMI) are used as the targets. First, we train a cGAN model (Model A) with shuffling training data. However, the estimated magnetic polarities are not well consistent with real observations. Second, to improve the accuracy of estimated magnetic polarities, we train a cGAN model (Model B) with the chronological Hα and HMI images, which can implicitly exploit the magnetic polarity constraint of time series observation to generate more accurate magnetic polarities. We compare the generated magnetograms with the target HMI magnetograms to evaluate the two models. It can be observed that Model B has better magnetic polarity consistency than Model A. To quantitatively measure this consistency, we propose a new metric called pixel-to-pixel polarity accuracy (PPA). With respect to PPA, Model B is superior to Model A. This work gives us an insight that the time series constraint can be implicitly exploited through organizing training data chronologically, and this conclusion also can be applied to other similar tasks related to time series data.

Keywords: Solar magnetic fields (1503) — Space weather (2037) — Convolutional neural networks (1938) — generative adversarial network (GAN) — The Sun (1693)

1. INTRODUCTION

The solar magnetic field is the determining factor that controls solar activities in solar atmosphere. Measuring the solar magnetic field is crucial for forecasting solar activities, which can reduce or avoid the risks brought by catastrophic space weather. Nowadays, photospheric magnetic fields can be observed by ground-based telescopes and space-based telescopes, both of which have their advantages and disadvantages. Ground-based telescopes are susceptible to atmospheric effects and suffer from aberrations due to gravity. These issues can be avoided by using space-based telescopes. However, space-based telescopes need to be launched into space, which entails the higher costs as well as tougher scientific and technical requirements. In contrast to observations of magnetic fields, as the brightest hydrogen line in the visible spectral range, high-resolution Hα images can be directly observed by ground-based telescopes and have been routinely observed for decades. Liu et al. (2021) explored the mapping of Hα images to eight different wavelengths (94, 131, 171, 193, 211, 304, 335 and 1600 Å) by Solar Dynamics Observatory (SDO; Pesnell et al. (2011))/Atmospheric Imaging Assembly (AIA; Lemen et al. (2011)). Meanwhile, Veeder & Zirin (1970) suggested that

Corresponding author: Long Xu (lxu@nao.cas.cn); Fei Gao and Tie Liu are co-first authors.

a set of rules could be developed by comparing the photoelectric magnetograms with the high-resolution Hα images and then deriving the magnetic field directly from the filter. Inspired by these works, we consider the inference of photospheric magnetic fields from the Hα images.

Deep learning (LeCun et al. 2015) has been rapidly developed and has achieved great success in many fields in recent years. As the most popular and widely used networks in deep learning, convolutional neural networks (CNNs; LeCun et al. (1998)) are trained by defining a specific loss function and then optimizing the parameters using gradient descent and its improvement algorithms. In other words, we still need to design an effective loss function adapted to the task. However, in many cases, it is very difficult to find a reasonably effective loss function. Fortunately, this problem has been solved by the proposal of Generative Adversarial Networks (GANs; Goodfellow et al. (2020)), which are networks that can train a model using data with a certain distribution and automatically generate data with a similar distribution to the original data after the training is completed. The training of GANs is dynamic and involves an adversarial game between the generator G and the discriminator D. In this game process, GANs can learn a loss function and a generative model to discriminate whether the results are true or false. Conditional GANs (cGANs; Mirza & Osindero (2014)) are networks that learn a conditional generative model by adding additional conditional information to the generators and discriminators of the original GAN. cGANs are well suited for image-to-image translation tasks, where the goal is to learn the mapping between an input image and an output image. Further, based on cGANs, Isola et al. (2017) proposes a generic architecture for image-to-image translation, Pix2Pix, which yields convincing results in many application scenarios.

Several works have attempted to apply deep learning methods to translations between solar images. Kim et al. (2019) applied deep learning models based on cGANs to generate solar farside magnetograms from Solar Terrestrial Relations Observatory (STEREO; Kaiser et al. (2008))/Extreme UltraViolet Imager (EUVI; Howard et al. (2008)) 304Å images. Jeong et al. (2020) successfully used the Pix2PixHD model to generate the farside magnetograms from three-channel farside observations of STEREO Ahead (A) and Behind (B), and combined with the synchronic map to achieve coronal magnetic field extrapolation. Further, they improved the farside magnetograms using data sets of STEREO and SDO (Jeong et al. 2022). For the magnetogram fluctuation represented by sudden magnetic polarity reversal and drifting of magnetic field distribution in Kim et al. (2019), Sun et al. (2022) suggested a dynamic deep learning model which can generate a sequence of magnetogram with a smooth transition. In addition, Park et al. (2019) proposed a Pix2Pix-based deep learning method to generate SDO/AIA-like solar UV and extreme UV (EUV) images from SDO/Helioseismic Magnetic Imager (HMI; Schou et al. (2012)) line-of-sight (LOS) magnetograms. Lawrance et al. (2022) utilized pix2pix to generate the ground-based Mauna Loa Solar Observatory (MLSO; de Wijn et al. (2012)) K-Coronagraph like white-light images from SDO/AIA images. These works demonstrate that it is not only feasible but also effective to apply deep learning models based on cGANs to translate between solar images. Generating SDO/HMI magnetograms from Hα images is also a typical image-to-image translation task.

In this paper, we consider using a deep learning method based on cGANs to find out the mappings between Hα images and SDO/HMI magnetograms. In addition, we explore the impact of the different permutations of training samples on magnetogram generation. The organization of this paper is as follows. We introduce the deep learning models used in this paper in Section 2. The data are described in detail in Section 3. In Section 4, the experimental results and the analysis are presented. Section 5 summarizes our conclusions and prospects.

2. METHODS

As a typical image-to-image translation, we choose to utilize a general architecture for this task, Pix2Pix, to establish the mappings between HMI magnetograms and Hα observations. Pix2Pix consists of two vital components like the original GANs: a generative network (generator) G and a discriminative network (discriminator) D. The purpose of the generator is to generate target-like data from the input, while the discriminator is designed to distinguish between the image generated by the generator and the real one. Unlike the original GANs, cGANs augment the input with conditional information, which can be category labels or other auxiliary information. The condition information is specified as the observed image in Pix2Pix.

During the training process, the generator generates more realistic images to fool the discriminator, while the discriminator tries to separate the images generated by the network G from the real ones. This constant adversarial process between the generator and the discriminator allows the model to continuously optimize the adversarial loss function $\mathcal{L}_{cGAN}(G, D)$, which is given by

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{(x,y)}[\log D(x, y)] + \mathbb{E}[\log(1 - D(x, G(x, z)))], \quad (1)$$

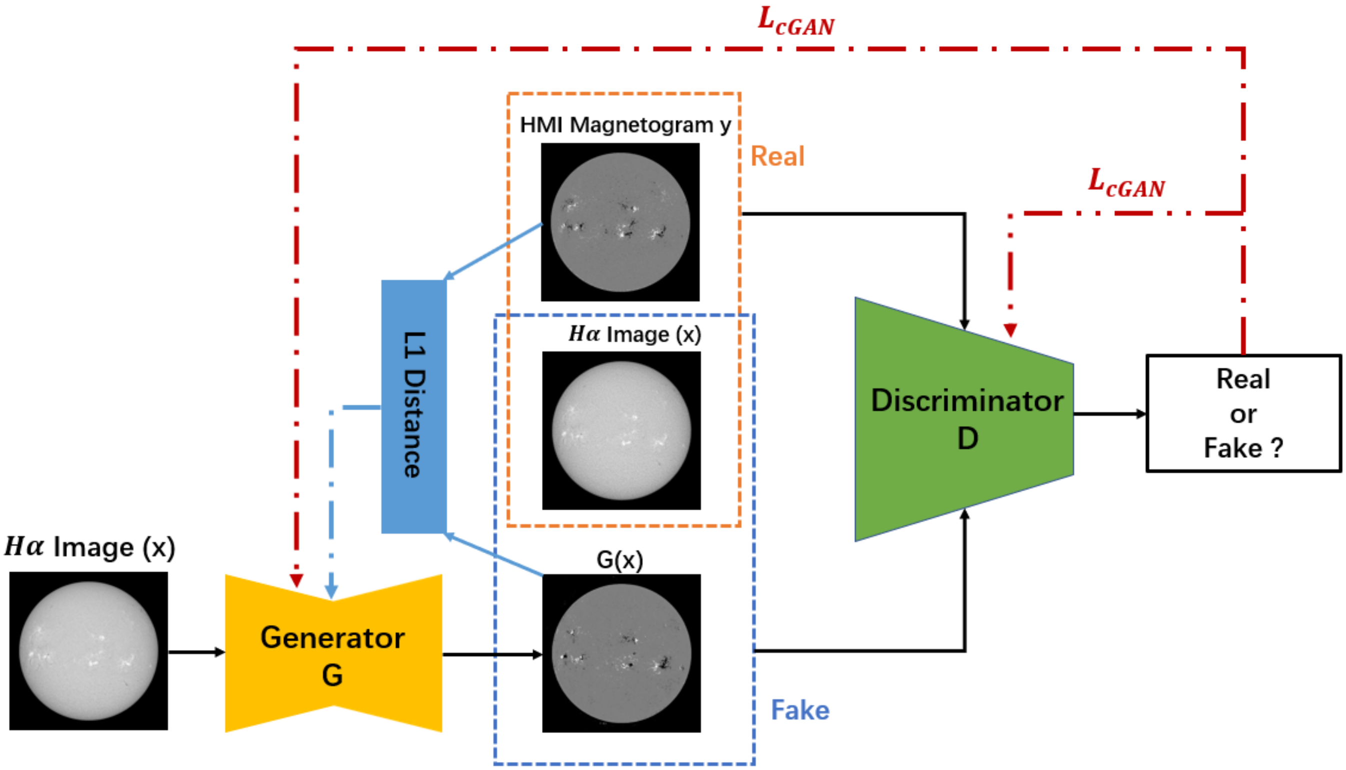


Figure 1. The architecture of Pix2Pix model used in the paper. G is the generator, D is the discriminator. x is an Hα image, y is the corresponding HMI magnetogram of x, and the $G(x)$ is the generated magnetogram of x.

where G is the generator, D is the discriminator, z is a random noise vector. x, y, $G(x, z)$ denote a real input (the Hα observed image), a real output (HMI magnetogram), and an image generated by the generator G, respectively. G tries to minimize the target in order to fool D, while D tries to maximize it to distinguish more accurately between the real and fake images. The objective of conditional GANs can be expressed by:

$$\min_G \max_D \mathcal{L}_{cGAN}(G, D). \tag{2}$$

Not only does the generator need to fool the discriminator, but it also needs to be close to the real output in the sense of $L2$. In Isola et al. (2017), experiments show that using $L1$ distance encourages less blurring compared to $L2$, so we choose to use $L1$ distance instead of $L2$.

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1]. \tag{3}$$

Therefore, the final objective is

$$G^* = \arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G), \tag{4}$$

where the λ is a hyperparameter which stands for the weight of the $L1$ loss term.

Figure 1 shows the main architecture of our model. In this work, we feed the model the Hα observed images and SDO/HMI magnetograms as input. The generator tries to generate SDO/HMI-like solar magnetograms from the Hα images. The discriminator tries to distinguish between the generated magnetograms and the real ones. We choose a U-Net with added skip connections between each layer i and layer $n - i$ as the structure of the generative network, where n denotes the total number of layers. Such a structure allows the large amount of low-level information shared between the input and output of the generator to circumvent the bottleneck. $L1$ can accurately capture low-frequency information in many cases, which prompts us to consider only model high-frequency information in the choice of the discriminator structure (Isola et al. (2017)). Therefore, we choose PatchGAN as our discriminator architecture, whose output is a $N \times N$ array rather than a single scalar. PatchGAN allows us to restrict our attention to the structure in local image patches and thus achieve high-frequency modeling. Each element of the output matrix of this discriminator corresponds to a patch of the input image, which means that the discriminator can distinguish whether $N \times N$ patches are real or fake, respectively, instead of the whole image. Additionally, the PatchGAN has fewer parameters and runs faster than classifying the whole image. For more details of the model architecture, please refer to Isola et al. (2017). The size of the receptive field of PatchGAN used in this paper is 70×70.

3. DATA

The Hα images and solar magnetograms come from the Global Oscillation Network Group (GONG; Hill et al. (1994)) and HMI onboard SDO, respectively. GONG includes six stations (the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias and Cerro Tololo Interamerican Observatory) located around the world and have the ability to observe the Sun with continuous observations. GONG provides full-disk line-center Hα solar images (2048×2048) with time cadences about 1 minute. The LOS magnetic field is measured by HMI which observes the solar photospheric magnetic field with high temporary and spatial resolutions. HMI provides full-disk LOS magnetograms with time cadences of 45 s.

We download 2048×2048 Hα images and 4096×4096 HMI LOS magnetograms by *sunpy.net.Fido* (SunPy; The SunPy Community et al. (2020)). In the end, 654 paired Hα and HMI LOS images from years 2013 and 2014 are collected and we process them as described below:

1. The Hα and HMI LOS images of the years 2013 and 2014 are downloaded with time cadences of one image a day. In order to keep the two images in one day consistent with each other, we choose one image in the time period 00:00-00:01 of that day according to universal time respectively.
2. Removing the damaged observations and picking out the paired images with the best image quality.
3. Processing the HMI LOS images to level 1.5 by the SolarSoft routine *hmi_prep*. And aligning the field of view of HMI and GONG images after normalizing them by exposure time.
4. The spatial resolution of HMI LOS images is 0.5 arcsec per pixel and that of GONG Hα images is about 1 arcsec per pixel. HMI images are downsampled to 1024×1024 images by averaging neighboring 4×4 pixels and GONG images are also compressed with neighboring 2×2 pixels to 1024×1024 images.

The 654 paired images are divided into training (394 pairs) validation (130 pairs) and test (130 pairs) datasets in chronological order.

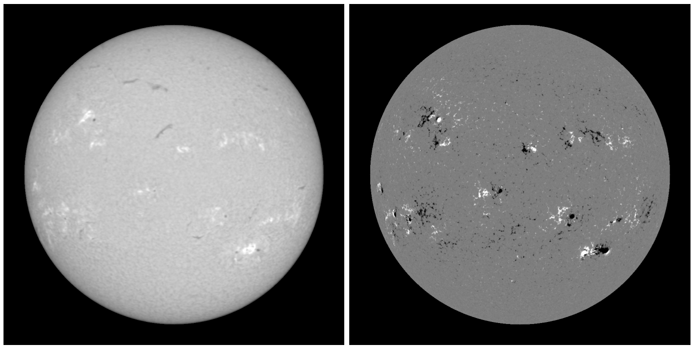


Figure 2. An example of the paired images in our data set. The left one is a GONG Hα image. The right one is the corresponding SDO/HMI magnetogram.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Our experiments are all implemented by Pytorch (Paszke et al. 2019). We train the models with 394 pairs of Hα images and SDO/HMI magnetograms from the training set. We expand the training set with data augmentation. The 1024×1024 images are flipped and randomly cropped to 512×512. The cropped images are then fed to the model for training. In order to optimize the training process of the models, several parameters need to be set. We use the adaptive momentum estimation optimizer (Adam; Kingma & Ba (2014)) for both the generator and the discriminator.

The initial learning rate is set to 0.0002 and decays linearly after 200 epochs. The batch size is set to 8. The λ in (4) is set to 100. In addition to the above parameters, we consider two different data input orders and experiment with them to obtain two models, one is Model A where the data are randomly input after being shuffled, and the other is Model B where the inputs are in chronological order. Except for the data input order and the number of epoch, we set other hyperparameters of both models to the same value or strategy. Based on the performance of the models on the validation set, the epoch of Model A is set to 400, and the epoch of Model B is set to 380. The performance of these optimal models on the test set is the final result of our model. The dataset and code used in this work can be accessed via gitHub https://github.com/filterbank/Ha2Mag.

In order to evaluate the model objectively, we calculate Peak Signal-to-Noise Ratio (PSNR), Structure Similarity Index Measure (SSIM), and pixel-to-pixel correlation coefficients (CCs) between the real observed magnetograms and the generated ones. For these three metrics, we only consider the pixels on the solar disk. PSNR is the ratio of the energy of the peak signal to the average energy of the noise, which is defined as

$$PSNR = 10 \log_{10} \frac{MaxValue^2}{MSE}, \tag{5}$$

where MaxValue is the maximum pixel value of the image. MSE is the mean square error. Compared with PSNR, SSIM is more in line with the evaluation results of the human visual system. We calculate SSIM to measure the structural similarity between images, which is defined as

$$SSIM\left(x, y\right) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_x y + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}, \tag{6}$$

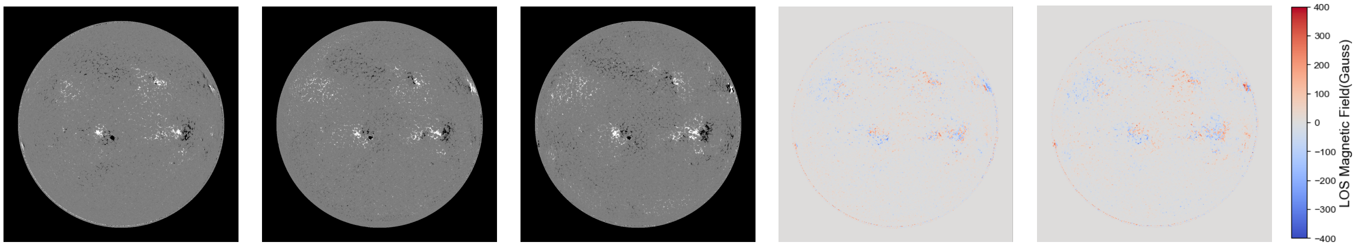
where x and y denote the real and generated images respectively, μ_x and μ_y denote the means of x and y respectively, σ_x^2 and σ_y^2 denote the variances of x and y respectively. C_1 and C_2 are two constants which are given by $C_i = \left(k_1 L\right)^2 \left(i = 1, 2\right)$, where L is the range of pixel values, k_1 and k_2 are 0.01 and 0.03, respectively. The Pearson correlation coefficient can reflect how linearly the images are related. By calculating pixel-to-pixel CCs, we are able to measure the correlation between the observed and generated images. Each notation has the same meaning as (6), and the Pearson correlation coefficient is defined as

$$PCC\left(x, y\right) = \frac{E\left[\left(x - \mu_x\right)\left(y - \mu_y\right)\right]}{\sigma_x \sigma_y}. \tag{7}$$

Table 1. Objective Measures of Comparison between SDO/HMI Magnetograms and AI-generated Ones

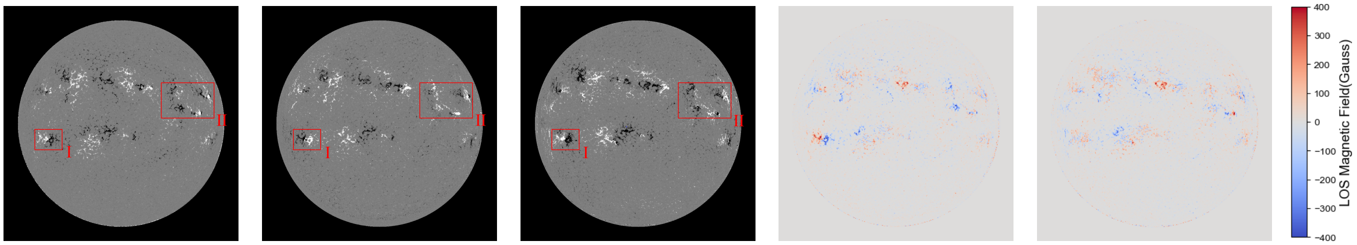
	Full-disk		
	PSNR	SSIM	Pixel-to-Pixel CCs (8 × 8 binning)
Model A	22.21 ± 1.06	0.76 ± 0.01	0.85 ± 0.06
Model B	21.77 ± 0.91	0.75 ± 0.01	0.85 ± 0.05
	Active Regions		
	PSNR	SSIM	Pixel-to-Pixel CCs
Model A	14.83 ± 2.13	0.38 ± 0.08	0.65 ± 0.13
Model B	14.51 ± 1.90	0.34 ± 0.08	0.67 ± 0.11

Table 1 shows the average values and standard deviations of three metrics for the full-disk and active regions (ARs) between the magnetograms generated by Model A and Model B and the corresponding real ones in the test set. For the results calculated on the full-disk, the average PSNR values of Model A and Model B are 22.21 and 21.77, respectively. The average values of SSIM of the two models are 0.76 and 0.75, respectively. The average pixel-to-pixel CCs of the two models after 8 × 8 binning are both 0.85. The results of these metrics of Model A are mostly comparable or slightly better than those of Model B. For the results calculated on the ARs, the average PSNR of Model A and Model B are 14.83 and 14.51, respectively. The average values of SSIM of the two models are 0.38 and 0.34, respectively.



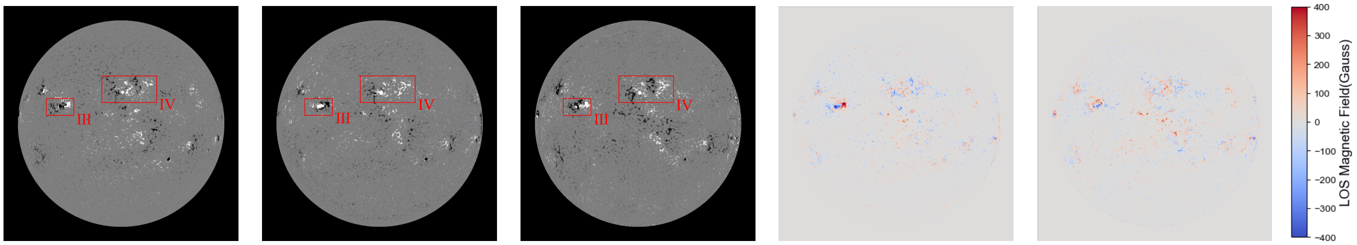
(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(a) 2014-08-11



(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(b) 2014-10-07



(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(c) 2014-12-30

Figure 3. Comparison between HMI magnetograms and generated ones on August 11, October 7, and December 30, 2014. The first column represents SDO/HMI magnetograms. The second and third columns represent generated ones by Model A and Model B, respectively. The last two columns represent the difference maps between SDO/HMI magnetograms and generated ones by Model A and Model B, respectively. The boxes in red color mark the ARs highlighted in Figure 4.

The pixel-to-pixel CCs of the two models without binning are 0.65 and 0.67, respectively. Compared with the results calculated on the full-disk, the difference between the results calculated on ARs of the two models is more obvious. It can be seen that the results of PSNR and SSIM of Model A are better than those of Model B, similar to the results calculated for the full-disk. However, the average pixel-to-pixel CCs of Model B is slightly better than that of Model A, which indicates that there is a better correlation between the ARs of the magnetograms generated by Model B and those of the real ones. Combining the results of these three metrics, it seems that the magnetograms generated by the two models are at the same level. The values of these three metrics calculated on both full-disk and ARs indicate that the two models have the ability to generate magnetograms that are well-consistent with the real magnetograms. Therefore, we further compare the generated magnetograms visually.

Figure 3 shows the full-disk SDO/HMI magnetograms, the corresponding AI-generated magnetograms by Model A and Model B, and the difference maps between them on August 11, October 7, and December 30, 2014. These three cases are all from the test set. The results demonstrate the ability of the two models to generate HMI magnetograms with good visual perceptibility (Figure 3(a)). It can be seen that there is a strong correlation between the overall magnetic field distributions of the magnetograms generated by both Model A and Model B and those of the real ones. As shown in the color maps of the difference maps, some details of the magnetic fields in the ARs of the generated magnetograms, especially those generated by Model A, do not match well with the HMI magnetograms. There is a part of the magnetic field of ARs in the magnetograms generated by Model A with the opposite polarity of those in the real ones, while Model B generates most of the magnetic field polarities correctly (Figures 3(b) and (c)).

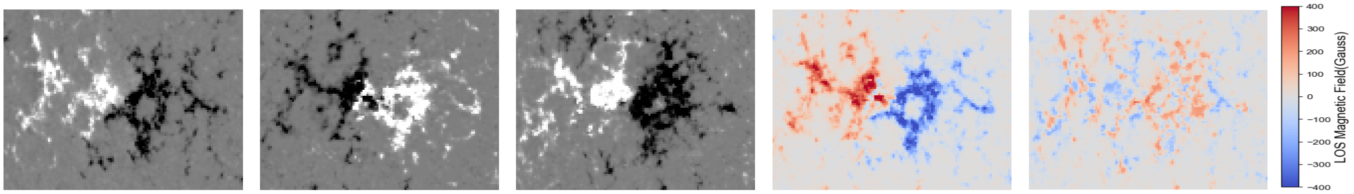
According to Figure 3, we find that the effect of magnetic field generation in some ARs of the models is not satisfactory. Therefore, We mark these ARs on each observed magnetogram and the same region on its corresponding generated ones in Figure 3(b) and (c). We then experiment on them to compare the generation performance of the two models. In Figure 4, the HMI magnetogram, the magnetograms generated by the two compared models, and the difference maps between them are illustrated for the selected ARs mentioned above. For the AR(I) on October 7, 2014 (Figure 4(a)), it can be seen that Model A generates the magnetic polarity of AR(I) completely opposite to that in the HMI magnetogram, while Model B generated it roughly correctly. Depending on the difference map, Model A generates a part of the negative magnetic field of AR(II) into a positive one, while Model B amends this problem partially (Figure 4(b)). For December 30, it is similar to the situation on October 7. It is clear that the magnetic polarity in the real magnetogram for AR(III) is opposite to that of the magnetogram generated by Model A and the same as that of the magnetogram generated by Model B (Figure 4(c)). For AR(IV), Model A generates a part of the negative magnetic field as positive, while the magnetogram generated by Model B is in good agreement with the real one (Figure 4(d)).

We further examine the ability to generate the magnetic field polarity distribution of our models through the experiments below. For objective evaluation, we propose a metric that can measure the accuracy of the magnetic field polarity of the generated image, called pixel-to-pixel polarity accuracy (PPA). First, we set a threshold to filter out the pixels whose magnetic field strength is greater than the threshold in the real magnetogram. The number of these pixels is denoted as N. Then, N pixels in the generated image corresponding to these pixels are found. Each pixel that is filtered is then individually assigned a value based on its sign, which is defined as $\hat{x}_{ij} = sgn(x_{ij})$, where x_{ij} is the pixel value of the row i and column j on the image, $\hat{x}_{ij}$ is either 1 or -1. After processing a pair of the real and the generated magnetograms as described above, We subtract the corresponding pixels of the two images to obtain the number of pixels with a value of 0, which is denoted as n. Finally, we calculate the proportion of the number of pixels with the same magnetic field polarity (pixels with a value of 0 after subtraction) among all pixels satisfying the threshold condition in the real and generated images, i.e. $\frac{n}{N}$. This final result is the PPA of the pair of magnetograms.

Table 2. The average and standard deviation of pixel-to-pixel polarity accuracies (PPAs) between SDO/HMI magnetograms and AI-generated ones for Model A and Model B at different thresholds.

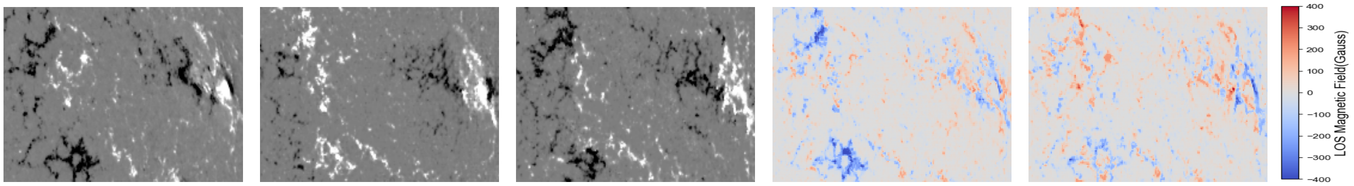
	Threshold						
	0	5	10	20	30	40	50
Model A	0.59 ± 0.07	0.62 ± 0.10	0.64 ± 0.12	0.66 ± 0.14	0.67 ± 0.15	0.68 ± 0.16	0.68 ± 0.16
Model B	0.61 ± 0.07	0.65 ± 0.10	0.67 ± 0.11	0.70 ± 0.13	0.71 ± 0.14	0.72 ± 0.14	0.72 ± 0.14

In our case, the minimum and maximum pixel value of the magnetogram are -200 and 200. We crop out the ARs of all the real magnetograms in the test set and their corresponding generated ones by the two models since we only consider the pixels in the ARs. We then calculate the PPA on the cropped images. Table 2 shows the average and standard deviation of PPAs between the real magnetograms and the generated ones by the two models at different thresholds. Model A illustrates that the PPAs of ARs are 0.59, 0.62, 0.64, 0.66, 0.67, 0.68, and 0.68 at the thresholds of 0, 5, 10, 20, 30, 40, and 50, respectively. The PPAs of ARs for Model B at the corresponding thresholds are up to 0.61, 0.65, 0.67, 0.70, 0.71, 0.72, and 0.72, respectively. These values indicate that while the Hα observation does not contain information on the magnetic field polarity distribution, our model can learn a part of it during training, which enables both models to correctly generate the magnetic field polarity for the majority of ARs. Model B outperforms Model A in reconstructing the polarity of ARs with varying magnetic field strengths, as shown by the fact that the PPA values of Model B are higher than that of Model A regardless of the choice of different thresholds (Table 2). These results are consistent with the visual results (Figure 3 and 4). The input order of the data is what distinguishes Model A from Model B. Therefore, the reason for the better performance of Model B may be that the observations of the solar magnetic field are time series, and our choice of chronological input is equivalent to adding the prior time constraint. The Hα images themselves do not contain information on the magnetic field polarity, which makes it very difficult to directly and accurately generate the distribution of polarities without any prior conditions. This prior



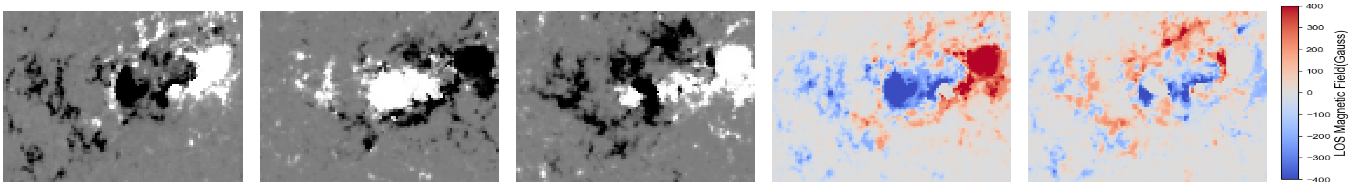
(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(a) 2014-10-07 AR(I)



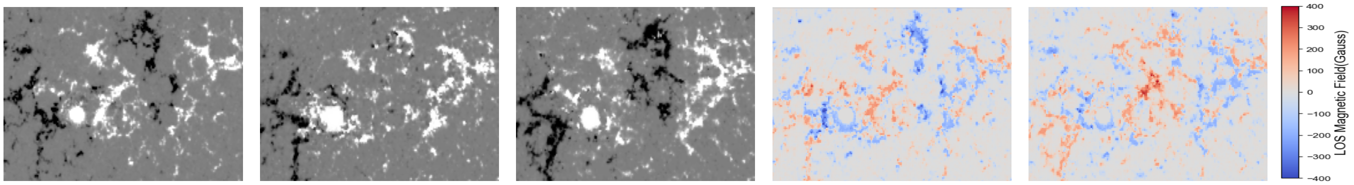
(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(b) 2014-10-07 AR(II)



(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(c) 2014-12-30 AR(III)



(a) HMI magnetograms (b) Model A (c) Model B (d) Model A - HMI (e) Model B - HMI

(d) 2014-12-30 AR(IV)

Figure 4. Comparisons of ARs between HMI magnetograms and generated ones by Model A and Model B on October 7 and December 30, 2014. The marked ARs in Figure 3 are zoomed. The first column represents the ARs of SDO/HMI magnetograms, the second column represents the ARs of the generated ones by Model A, the third column represents generated ones by Model B, and the last two columns represent the difference maps of the ARs between SDO/HMI magnetograms and generated ones by Model A and Model B, respectively.

constraint enables our model to obtain more information to confirm the magnetic field polarity through training, which is more conducive to generate SDO/HMI-like magnetograms. In addition, the results demonstrate that both Model A and Model B improve the accuracy of magnetic field polarity generation in the ARs with the increase of threshold. Although there may be more noise interference to reduce the accuracy when the threshold setting is low. Our results show that the accuracy is still improved when the threshold is increased to ignore the noise, which indicates that our models are more accurate in generating the polarity of the strong magnetic field. We would like to note that PPA values do not necessarily correspond strictly to model performance. PPA is an additional performance metric that complements other metrics such as PSNR, SSIM, and PCC.

5. CONCLUSION AND SUMMARY

In this paper, generating HMI-like magnetograms from Hα images is investigated, which is taken as an image-to-image translation task and a deep learning model based on cGAN is applied. For model training, a dataset containing

Hα image and HMI magnetogram pairs collected from GONG and SDO respectively is established. Two models with different data input orders are trained and examined the influence of input order on magnetogram generation.

The quantitative results show that we successfully generate SDO/HMI-like magnetograms. The generated magnetograms by our model have good correlations with the real ones. In addition, by adding a temporal constraint to train the model (Model B) with data input in chronological order, we can generate more consistent magnetic polarity. To evaluate magnetic polarity consistency quantitatively, we then proposed a new metric to measure the accuracy of magnetic polarity. The results show that Model B outperforms Model A in reconstructing the polarity of ARs with varying magnetic field strengths, which demonstrates that adding time series constraints to the model by organizing training set chronologically is effective for the generation of magnetic field polarity distribution.

Our work not only provides a feasible and effective way to generate SDO/HMI-like magnetograms but also demonstrates that it is possible to obtain space-based generated data from ground-based observations. Our models will be helpful to replace low-quality data such as noisy, degraded ones, and data observed in poor atmospheric circumstances to supplement the database. We hope that our models can be useful for solar activity forecasting using SDO/HMI magnetograms, such as solar flare forecasting. In addition, the reason why we consider the data input in chronological order is that the Hα images themselves do not contain information about the magnetic field polarity, which makes it very difficult to directly and accurately generate the distribution of polarities without any prior conditions. Both Hα observations and HMI magnetograms are time series data. Therefore, restricting input data in chronological order can implicitly embed some prior information into the model while shuffling input data would completely abandon this prior information. We would like to emphasize that this training strategy by adding time constraints through organizing input data in chronological order can be extended to other similar tasks concerning the modeling with time series data.

We appreciate the instructive comments and suggestions of the editor and reviewer to improve the manuscript. The SDO and GONG science team generously provided the data that were used in the study. We acknowledge all of the team members who have contributed to making these missions successful. We appreciate the community for their contributions to developing the open-source packages used in this work.

This work was supported by the National Key R&D Program of China (Nos. 2021YFA1600504, 2022YFE0133700), the National Natural Science Foundation of China (NSFC) (Nos. 11790305, 11963003, 12203023, 12273007).

REFERENCES

de Wijn, A. G., Burkepile, J. T., Tomczyk, S., et al. 2012, in Ground-based and Airborne Telescopes IV, Vol. 8444, SPIE, 1201–1208

Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. 2020, Communications of the ACM, 63, 139

Hill, F., Fischer, G., Forgach, S., et al. 1994, Solar Physics, 152, 351

Howard, R. A., Moses, J., Vourlidas, A., et al. 2008, Space Science Reviews, 136, 67

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. 2017, in Proceedings of the IEEE conference on computer vision and pattern recognition, 1125–1134

Jeong, H.-J., Moon, Y.-J., Park, E., & Lee, H. 2020, The Astrophysical Journal Letters, 903, L25

Jeong, H.-J., Moon, Y.-J., Park, E., Lee, H., & Baek, J.-H. 2022, The Astrophysical Journal Supplement Series, 262, 50

Kaiser, M. L., Kucera, T., Davila, J., et al. 2008, Space Science Reviews, 136, 5

Kim, T., Park, E., Lee, H., et al. 2019, Nature Astronomy, 3, 397

Kingma, D. P., & Ba, J. 2014, arXiv preprint arXiv:1412.6980

Lawrance, B., Lee, H., Park, E., et al. 2022, The Astrophysical Journal, 937, 111

LeCun, Y., Bengio, Y., & Hinton, G. 2015, nature, 521, 436

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, Proceedings of the IEEE, 86, 2278

Lemen, J. R., Akin, D. J., Boerner, P. F., et al. 2011, in The solar dynamics observatory (Springer), 17–40

Liu, T., Su, Y.-N., Xu, L.-M., & Ji, H.-S. 2021, Research in Astronomy and Astrophysics, 21, 135

Mirza, M., & Osindero, S. 2014, arXiv preprint arXiv:1411.1784

Park, E., Moon, Y.-J., Lee, J.-Y., et al. 2019, The Astrophysical Journal Letters, 884, L23

Paszke, A., Gross, S., Massa, F., et al. 2019, Advances in neural information processing systems, 32

Pesnell, W. D., Thompson, B. J., & Chamberlin, P. 2011, in The solar dynamics observatory (Springer), 3–15

Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, Solar Physics, 275, 229

Sun, W., Xu, L., Ma, S., et al. 2022, The Astrophysical Journal Supplement Series, 262, 45

The SunPy Community, Barnes, W. T., Bobra, M. G., et al. 2020, The Astrophysical Journal, 890, 68, doi: 10.3847/1538-4357/ab4f7a

Veeder, G. J., & Zirin, H. 1970, Solar Physics, 12, 391