Generation of High-resolution Solar Pseudo-magnetograms from Ca II K Images by Deep Learning

Gyungin Shin[1], Yong-Jae Moon[1], Eunsu Park[2], Hyunjin Jeong[1], Harim Lee[2], and Sung-Ho Bae[3]
[1] School of Space Research, Kyung Hee University Yongin, 17104, Republic of Korea; moonyj@khu.ac.kr
[2] Department of Astronomy and Space Science, College of Applied Science, Kyung Hee University, Yongin, 17104, Republic of Korea
[3] Department of Computer Science and Engineering, College of Software, Kyung Hee University, Yongin, 17104, Republic of Korea
Received 2020 January 29; revised 2020 April 11; accepted 2020 May 2; published 2020 May 21

Abstract

In this Letter, we generate realistic high-resolution (1024×1024 pixels) pseudo-magnetograms from Ca II K images using a deep learning model based on conditional generative adversarial networks. For this, we consider a model "pix2pixHD" that is specifically devised for high-resolution image translation tasks. We use Ca II K 393.3 nm images from the Precision Solar Photometric Telescope at the Rome Observatory and line-of-sight magnetograms from the Helioseismic and Magnetic Imager (HMI) at the Solar Dynamics Observatory from 2011 January to 2015 June. 2465 pairs of Ca II K and HMI are used for training except for January and July data. The remaining 436 pairs are used for an evaluation of the model. Our model shows that the mean correlation coefficient (CC) of total unsigned magnetic flux between AI-generated and real ones is 0.99 and the mean pixel-to-pixel CC after 8×8 binning over the full disk is 0.74. We find that the AI-generated absolute magnetic flux densities are highly consistent with real ones, even to the fine scale structures of quiet regions. On the other hand, the mean pixel-to-pixel correlations of magnetic flux densities strongly depend on a region of interest: 0.81 for active regions and 0.24 for quiet regions. Our results suggest a sufficient possibility that we can produce high-resolution solar magnetograms from historical Ca II data.

Unified Astronomy Thesaurus concepts: Solar physics (1476); The Sun (1693); Solar atmosphere (1477); Solar magnetic fields (1503); Astronomy data analysis (1858)

1. Introduction

Magnetic fields of the Sun are a fundamental component for its dynamic activities from small scales such as granules, to large scales such as flares and coronal mass ejections. To understand such solar activities, routine ground-based observations of full disk solar magnetic fields have been conducted since 1974 at the Kitt Peak Observatory (Livingston et al. 1976). The SOlar and Heliospheric Observatory (Domingo et al. 1995)/Michelson Doppler Imager (Scherrer et al. 1995) spacecraft started to offer line-of-sight full disk magnetograms of good quality in 1996 and the Solar Dynamics Observatory (SDO; Pesnell et al. 2012)/Helioseismic and Magnetic Imager (HMI; Scherrer et al. 2012; Schou et al. 2012) has offered better quality magnetograms since 2010. In fact, we have little information about how solar magnetic fields had evolved before 1974. A few attempts have been made to reconstruct past magnetic fields from Ca II full disk filtergrams, which is a well-known proxy for magnetic fields (Babcock & Babcock 1955). Pevtsov et al. (2016) reconstructed synoptic pseudo-magnetograms with Ca II 854.2 nm spectral line full disk images from the Vector Stokes Magnetograph of the Synoptic Optical Long-term Investigation of the Sun (Balasubramaniam & Pevtsov 2011) and sunspot magnetic field strength data from the Mount Wilson Observatory (Hale et al. 1919).

Convolutional Neural Networks (CNNs; LeCun & Bengio 1998) are one specific type of artificial intelligence to mediate drawbacks of formerly used multilayer perceptrons. CNNs are composed of several convolutional filters that are supposed to mimic a human visual system so that a CNN model can distinguish important parts of an input image from the rest. Due to its high performance in handling big data,

CNNs have been broadly adopted as a research tool across various scientific fields. On the other hand, Generative Adversarial Networks (GANs; Goodfellow et al. 2014) are a type of generative model in neural networks. GANs consist of two major components: a generator and a discriminator that are both multilayer perceptrons. Given 1D random values, the generator tries to produce a realistic image that is regarded as real data. On the contrary, the discriminator attempts to distinguish generated images from images in the real data set. In fact, this game between the two reaches at a point where the discriminator cannot tell apart the generated images and the real ones. To accommodate a better performance of GANs, Mirza & Osindero (2014) suggested inputting 1D values related to an output image that act as a condition, which is called conditional GAN (cGAN). Going a step further, Isola et al. (2016) proposed inputting an image in a generator instead of relevant 1D values and called their model "pix2pix." This pix2pix model translates an input image in one domain to the other image in a different domain that shares the same structure with the input data but has a different style. For instance, the pix2pix model can be trained to generate a color image from a black/white image. Because of its generality and moderate performance, the pix2pix model has been successfully used for several applications in the field of astronomy (Kim et al. 2019; Park et al. 2019). On the other hand, it has been well noted that the pix2pix model produces apparent artifacts and fails to capture fine details from time to time for high-resolution image translation tasks, e.g., larger than 1024×1024 pixels. Wang et al. (2017) proposed networks that resolve this issue and named them "pix2pixHD." Unlike a pix2pix model, it is empirically proven that a pix2pixHD model can manufacture even small features almost without artifacts for high-resolution data sets.

In this Letter, we reconstruct HMI-like high-resolution pseudo-magnetograms with Ca II 393.3 nm spectral line full disk intensity images by training a deep learning model called "pix2pixHD." For this, we use Ca II K images from the Precision Solar Photometric Telescope at the Rome Observatory (Rome/PSPT, Ermolli et al. 1998, 2007) and line-of-sight magnetograms from SDO/HMI. We train the model to generate high-quality magnetograms without loss of physical information during translation from Ca II intensity images to HMI-like magnetograms. We explain data sets in Section 2. Details in our deep learning method are delivered in Section 3. Results and discussion are present in Section 4. Finally, a brief summary and conclusion are given in Section 5.

2. Data

We use Ca II K 393.3 nm spectroheliograms for input data and line-of-sight magnetograms for target data. As mentioned in Section 1, this is due to high correspondences between bright features of Ca II K images and magnetic features on magnetograms (Leighton 1959; Chatzistergos et al. 2019a). For training, we use 2465 Ca II K image/magnetogram pairs from 2011 January to 2015 June except for every January and July. For evaluating our model, the remaining 436 pairs are used. To facilitate a training and test process with our limited computational resources, we set the size of all data as a 1024 by 1024 pixel.

2.1. Ca II 393.3 nm Full Disk Filtergram

We use Ca II 393.3 nm full disk intensity images that were obtained from the Rome/PSPT. We align Ca II K Images with SDO/HMI ones. In order to select relatively clear Ca II K images, we make Ca II contrast images by computing a local absolute maximum gradient within a small disk at every pixel[4] and discard those with mean contrast values lower than their average. Images containing obvious artifacts, e.g., deformed shape of the Sun or thick black line across the whole image, are manually removed.

2.2. HMI Magnetogram

We use line-of-sight magnetograms that were taken from SDO/HMI. Preprocessing contains aligning rotational axis to north up, locating solar center at the center of an image, and cleaning up. Magnetic flux densities are described within ± 1400 G. We adopt HMI data, which were observed within ± 12 minutes from observation time of each Ca II image, in order to minimize the effect of solar rotation. It is noted that 12 minutes correspond to one pixel movement by solar rotation at the disk center.

3. Method

As our purpose is to translate high-resolution (1024×1024 pixels) Ca II K intensity images to the corresponding magnetograms, we adopt a deep learning model called "pix2pixHD" based on cGAN, which are widely acknowledged for its great performance in high-resolution image translation tasks. We train our model for 200 epochs as in Wang et al. (2017), which correspond to about 490,000 iterations.

"Iteration" means the number of showing a Ca II/magnetogram pair to the model for training. An epoch refers to the count of showing the whole training data pairs, i.e., 2465 pairs. After training, we evaluate the trained model by comparing generated magnetograms and the corresponding SDO/HMI magnetograms. All the processes of training and testing our model are implemented on PyTorch (Paszke et al. 2019) developed by Facebook's Artificial Intelligence Research group, one of the popular deep learning frameworks. The official code used for the model, training, and test processes is publicly available at https://github.com/NoelShin/Ca2Mag. In the following, we describe the model architecture of pix2pixHD (see Figure 1) mainly in comparison with that of pix2pix. For more implementation details, please refer to Appendices B and C.

pix2pixHD is composed of a generator and two discriminators A and B. The generator is a CNN whose objective is to produce realistic magnetograms, which the discriminators fail to distinguish from corresponding real SDO/HMI magnetograms. On the other hand, given a Ca II/generated HMI pair or a Ca II real HMI pair, the objective of the discriminators is to find whether or not the pair contains real HMI data. As an optimization of the generator relies on performances of the two discriminators, it is important to make sure that the discriminators are good enough to distinguish (Arjovsky et al. 2017). In pix2pix, it introduces a "patch discriminator" whose receptive fields[5] of a pixel in its final output correspond to a certain patch size of an input image. Isola et al. (2016) showed that a discriminator with a different size of a receptive field has divergent outcomes and that a discriminator with a 70×70 receptive field generally performs satisfactorily. Inspired by this, Wang et al. (2017) proposed the use of several 70×70 patch discriminators and to give input pairs of different spatial sizes to the discriminators. A discriminator which gets a smaller (larger) input pairs has larger (smaller) a receptive field size. In our model, we use two 70×70 patch discriminators A and B. For the discriminator A, input pairs of the original 1024 by 1024 pixel size are given. For the discriminator B, input pairs that are downsampled by half are given, which makes its effective receptive field size 141×141. This setup encourages the generator to produce realistic magnetograms in two different view sizes. For more detailed information about objectives of the generator and discriminators, please refer to Appendix A.

4. Results and Discussions

Figure 2 shows a comparison between a generated HMI-like magnetogram from a Ca II K intensity image and an SDO/HMI magnetogram observed at 18:48 UT on 2012 January 31. The magnetic field polarities of the generated HMI-like magnetogram are consistent with the ones in the corresponding magnetogram, even to small-scale magnetic features. However, in the case of complex magnetic fields near the polarity inversion line, the model struggles to assign proper polarities to active regions.

To investigate whether generated magnetograms reconstruct physical meanings as well, we derived "Magnetic Range of Influence" (MRoI; McIntosh et al. 2006, 2007) from generated and SDO/HMI magnetograms. MRoI is a simple means for

---

[4] We use the scikit-image package (van der Walt et al. 2014) which is a Python image processing library for making the contrast maps. We set a radius of the disk as 2.

[5] A receptive field refers to a size that each pixel of the final output from a convolutional neural network perceives on the input image. For example, if a receptive field of a network is 70×70, it means that each pixel in the final output from the network considers a field of 70×70 size on its input image.
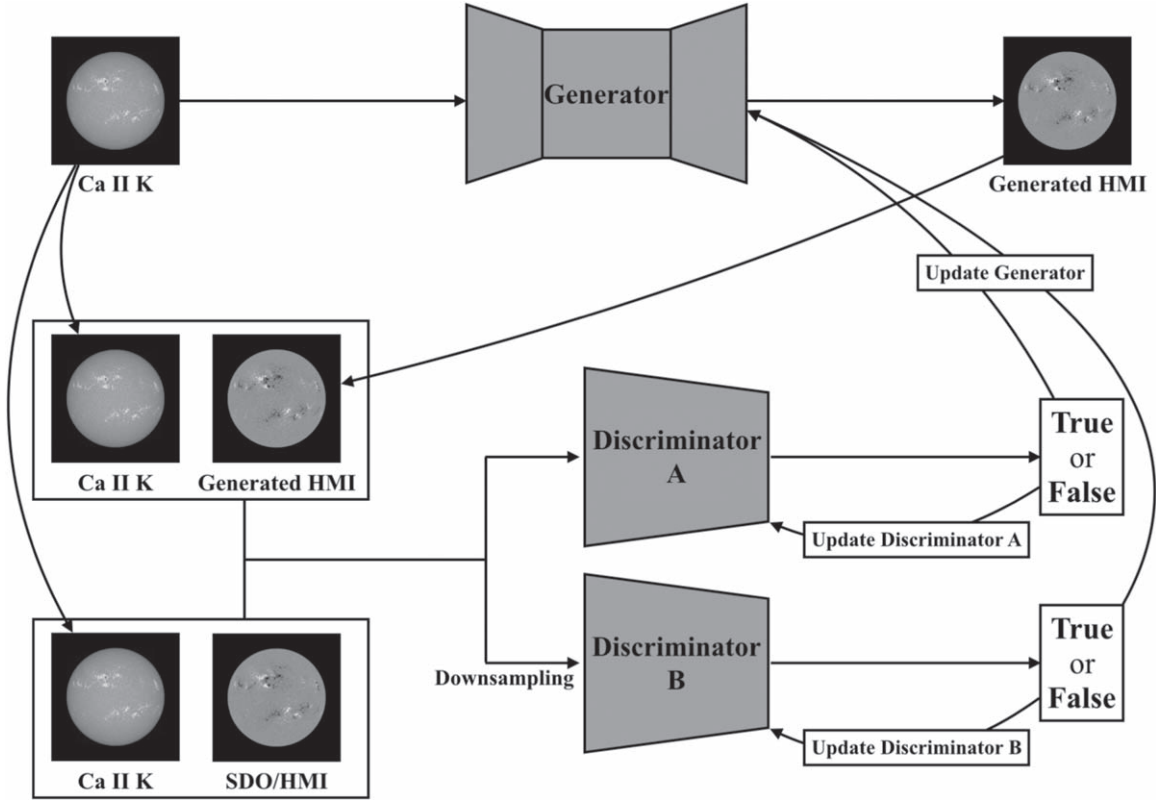
Figure 1. Overview of pix2pixHD. A Ca II K intensity image of 1024 × 1024 pixels is given to the generator and it produces a corresponding HMI-like magnetogram. The generated magnetogram is paired with the input Ca II K image and given to the discriminators A and B. Also, another pair of the same Ca II K image and the corresponding SDO/HMI is given to the discriminators A and B. When they are given to discriminator B, the size of the pairs is downsampled by half so that it can be determined with a different receptive field. Then, the discriminators strive to distinguish which pair contains the real HMI image. Their output values are close to 1 if they suppose that the input pair contains the real HMI image and 0 for the opposite case. These results affect the generator such that it can produce a more realistic HMI-like magnetogram afterward.

understanding a magnetic environment on the photosphere, delivering the radial distance required to balance magnetic flux densities on each pixel in a magnetogram (Leamon & McIntosh 2009). This is computed by increasing a radius of a circle by one pixel at a time until the contained fluxes are balanced. For a large MRoI, the magnetic field is interpreted to be unbalanced, and thus to be "open." We plot MRoI in Figure 3. The MRoI derived from the generated HMI is consistent with the one from SDO/HMI, describing that the model successfully reconstructs the physical meanings. Mean Pearson correlation coefficient between the MRoI from generated magnetograms and from the real ones is 0.82 after 8 by 8 binning.

To visualize how accurate absolute magnetic flux density distributions are alike, we plot three absolute magnetic flux maps: baseline, one from generated HMI, and one from SDO/HMI along with a difference map between ones from the generated and SDO/HMI magnetogram as shown in Figure 4. For baseline, we made a simple pixel to pixel calibration curve based on our training data set. We gathered pairs of Ca II K pixel intensity and unsigned magnetic flux in pixel level on the solar disk. We took a median value where there are multiple magnetic flux values for the same Ca II K intensity such that each intensity value has only one corresponding absolute magnetic flux density value.

As can be seen in the figure, predictions are highly reliable for quiet regions as well as active regions. This is because bright regions on Ca II K intensity images agree with magnetic

features on magnetograms (Leighton 1959; Chatzistergos et al. 2019a).

To examine how accurately magnetograms are reconstructed, we derive four different metrics: total unsigned magnetic fluxes (TUMF), pixel-to-pixel Pearson correlation coefficient (CC), relative error (R), and structural similarity index (SSIM; Wang et al. 2004) by comparing real SDO/HMI magnetograms and the corresponding generated magnetograms. The equation of R is given by

$$R_i = (\Phi_i^{\text{generated}} - \Phi_i^{\text{real}}) / \Phi_i^{\text{real}}, \qquad (1)$$

where Φ, i denotes total unsigned magnetic flux and the serial number of 436 test data. The equation of SSIM is given by

$$\text{SSIM}_i = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \qquad (2)$$

where x, y, μ_x, μ_y, σ_x^2, σ_y^2, σ_{xy}, i represent two different data, the average of x and y, the variance of x and y and the covariance of x and y, and the serial number, respectively. c_1 and c_2 are two constants for stabilizing the division. Here, we set c_1 and c_2 to 6.50 and 58.52 as default. SSIM produces a value between 0 and 1. A value close to 1 means that the two data share common structures and 0 means the opposite case.

For estimation we count pixels over +10 G for positive flux and under −10 G for negative flux on the solar disk by noting that the noise level of HMI magnetograms is about 10 G (Liu et al. 2012). We compute the metrics for 436 full disk
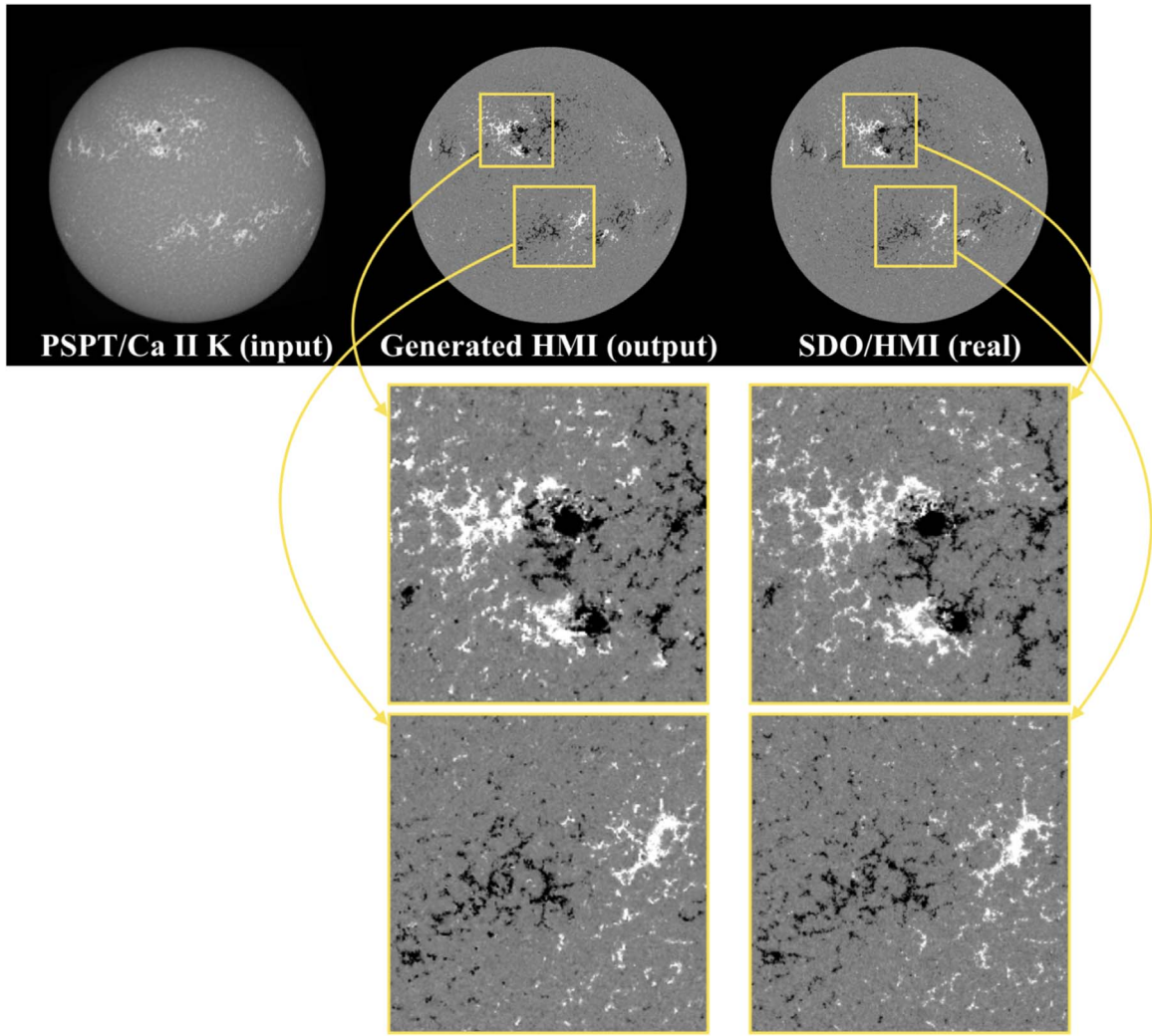
Figure 2. (Upper left) a Ca II K input intensity image, (upper middle) a generated HMI-like magnetogram, and (upper right) the corresponding real SDO/HMI. Ca II K and real SDO/HMI are taken at 18:48 UT 2012 January 31. (Lower left column) enlarged views of active regions in the generated magnetogram. (Lower right column) those in the real magnetogram. The white area denotes positive polarities and the black area denotes negative ones. For visual purpose, magnetic flux densities are expressed within ±100 G.

magnetograms, 510 active regions with 128×128 pixel size, and 436 quiet regions. In addition, we compute the pixel-to-pixel CC and the relative error of absolute magnetic flux density maps. When computing CC, we calculated CC from each pair of a generated and a real magnetogram and then average them. The results are shown in Table 1, which also includes the results of Kim et al. (2019) for comparison. In their work, they translated 304 Å images taken by the Atmospheric Imaging Assembly (AIA; Lemen et al. 2012) at SDO to the corresponding SDO/HMI magnetograms based on the pix2pix model.

As can be seen in Table 1, CC values of TUMF are 0.99 for both full disk and active regions, and 0.95 for quiet regions. These CC values show that our model is able to produce reliable TUMF values.

Pixel-to-pixel CC values after 8×8 binning show 0.74, 0.81, and 0.24 for full disk, active, and quiet regions, respectively. The values with other binning size are given in Table 1. These values demonstrate that our model cannot only well produce shapes of magnetic features but also assign polarities, at least for active regions. In fact, the magnetic flux densities with polarities for quite regions are not successful. For

more discussions of the results for active regions and quite regions, please refer to Kim et al. (2019). In the case of absolute flux density maps, CC values show 0.93, 0.94, and 0.81, respectively. These values strongly support that the generated absolute magnetic flux densities are highly consistent with SDO/HMI ones, even to the magnetic features in weak field regions. As for the relative error, its mean values are -0.085, -0.040, and -0.054 for full disk, active regions, and quiet regions, and its standard deviations are 0.032, 0.064, and 0.069, respectively. This implies that the generator slightly underestimates TUMF values more than the real ones but not significantly. For SSIM, its average values are 0.98, 0.80, and 0.97 for full disk, active, and quiet regions. This shows that the model is able to reconstruct magnetic structures in accordance with ones in SDO/HMI.

Our results of TUMF are better than those of baseline and Kim et al. (2019) for full disk and active regions, and much better for quiet regions. For pixel-to-pixel CC and the relative error, our results show better or comparable results with Kim et al. (2019). For example, the average pixel-to-pixel correlation with 8 by 8 binning (0.81) for active regions is better than that of Kim et al. (2019). These results demonstrate that it is
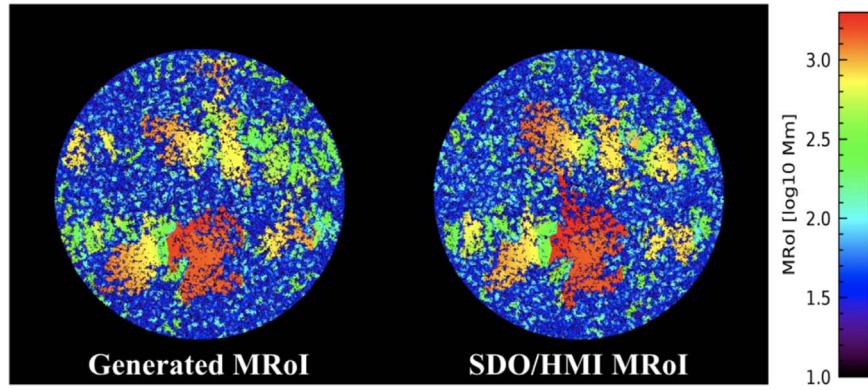
Figure 3. (Left) MRoI from the generated HMI and (right) MRoI from SDO/HMI at 18:48 UT on 2012 January 31. Small values present closed magnetic fields whereas large values present open magnetic fields.

useful to generate magnetograms from Ca II K images using pix2pixHD model. One impressive thing is that our model has been successfully applied to the translation from ground-based data to satellite data.

5. Conclusion

In this paper we have developed a deep learning model to translate from full disk Ca II K 393.3 nm intensity images to the corresponding high-resolution (1024 × 1024) HMI-like magnetograms. This model is called "pix2pixHD" which is specially devised for a high-resolution image translation task. We train the model with 2465 pairs of Ca II K images taken from PSPT in Rome and corresponding full disk line-of-sight magnetograms from SDO/HMI from 2011 January to 2015 June except for January and July. We evaluate the model with the remaining 436 pairs.

The main results of this study are summarized as follows. Our model can generate reliable HMI-like magnetograms from Ca II K images even to fine magnetic structures. The average pixel-to-pixel CC for MRoI after 8 by 8 binning is 0.82. For TUMF CC, the generated magnetograms show higher than 0.95 for full disk, active, and quiet regions. The pixel-to-pixel CCs after 8 × 8 binning are 0.74, 0.81, and 0.24 for full disk, active, and quiet regions, respectively. The CCs for absolute magnetic flux density maps are 0.93, 0.94, and 0.81, respectively. The mean relative errors are -0.085, -0.040, and -0.054 and their standard deviations are 0.032, 0.064, and 0.069, respectively. The mean SSIM are 0.98, 0.80, and 0.97, respectively. In comparison with results from Kim et al. (2019) which translated SDO/AIA 304 Å images to HMI-like magnetograms using a pix2pix model, our results are similar to or even better than their results. Our preliminary experiments show that a pix2pixHD model is much better than a pix2pix model, especially for ground-based data whose quality strongly depends on atmospheric and/or instrumental conditions. For further improvements of the model in terms of reconstructing complex magnetic structures, we believe that introducing an additional loss from physical parameters such as MRoI to the discriminator can be beneficial. We leave this point to following studies.

Lastly, our study shows a sufficient feasibility to reconstruct reliable past magnetograms from historic Ca II K data. Routine observations of Ca II K started in 1904 (Kodaikanal Observatory in India, Foukal et al. 2009) and its data have been digitized (Chatzistergos et al. 2019b). By applying our model

to these data we hope that our model is able to produce pseudo-magnetograms for this period. For this, cross-calibrations between past and modern Ca II data sets are essential (Chatzistergos et al. 2018). The modern magnetic field data such as Kitt Peak magnetograms started in 1974 (Livingston et al. 1976) can be used for evaluating our model. This kind of study is expected to offer more information on the long-term evolution of solar magnetic fields and their related studies such as long-term solar irradiance.

Appendix A
Objective

The objective (sometimes called loss function or simply loss) of cGAN is given as

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{(x,y)}[\log D(x, y)] \\ + \mathbb{E}_x[\log(1 - D(x, G(x)))], \quad (A1)$$

where G, D, x, and y denote a generator, a discriminator, input data, and target (real) data, respectively. $G(x)$ means an output from the generator for a given input image. While G tries to minimize the objective, D tries to maximize it. For the generator to get a lower value from the objective, it should generate an output in a way that $D(x, G(x))$ has a value close to 1. For the discriminator to have a higher value, it should output a value close to 1 when it takes a real pair (x, y) and a
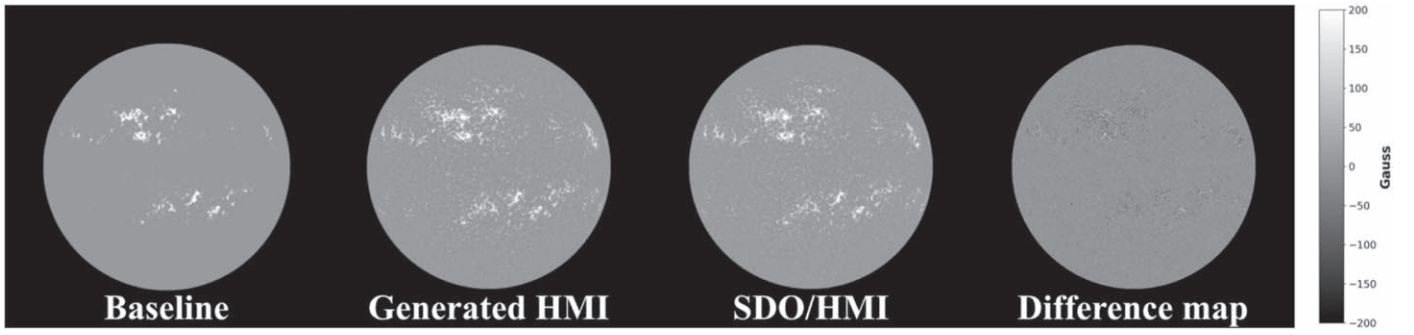
Figure 4. (Left) An estimated absolute flux density map based on a simple pixel to pixel calibration curve, (mid-left) an absolute flux density map from a generated HMI-like magnetogram, (mid-right) an absolute flux density map from the real SDO/HMI taken at 18:48 UT on 2012 January 31, and (right) the difference map between the generated and SDO/HMI magnetogram. Absolute flux densities are clipped to be under 100 G for the purpose of presentation. The black area denotes where the generator predicts higher than the real value and white area for the opposite.

value close to 0 when it takes a fake pair $(x, G(x))$. In practice, we use a mean squared loss function instead of log function following Wang et al. (2017). It is a well-known technique that brings stability to training GAN (Mao et al. 2016; Zhu et al. 2017).

In order to facilitate better learning for the generator, we introduce an additional objective function $\mathcal{L}_{FM}$, which is called "feature matching loss" to the generator (Wang et al. 2017). This loss function is defined as

$$
\mathcal{L}_{FM}(G, D) = \mathbb{E}_{(x,y)} \sum_{i=1}^{T} \frac{1}{N_i}
$$
$$
\times [\|D^{(i)}(x, y) - D^{(i)}(x, G(x))\|], \qquad (A2)
$$

where T, i, N_i denote the total number of layers in the discriminator, the serial number of the layers, and the number of pixels in output feature maps of each layer, respectively. The role of this objective is to regularize the fake pair to have more similar statistics with the real pair throughout the discriminator's intermediate layers. Our final objective is as follows:

$$
\min_G \left(\left(\max_{D_1, D_2} \sum_{k=1,2} \mathcal{L}_{cGAN}(G, D_k) \right) \right.
$$
$$
\left. + \lambda \sum_{k=1,2} \mathcal{L}_{FM}(G, D_k) \right), \qquad (A3)
$$

where k and λ denote the serial number of two discriminators and a hyperparameter which controls the importance of $\mathcal{L}_{cGAN}$ and $\mathcal{L}_{FM}$. We use 10 for λ as in Wang et al. (2017).

Appendix B
Data Augmentation

In Kim et al. (2019), it was noticed that the generator struggles to properly reconstruct bipolar structures where an unusual tilt angle appears between a preceding sunspot and the following one. We see that this is due to a lack of such examples and augment data by allowing random rotation within $\pm 30°$ in Ca II K/magnetogram pairs in order to compensate for an insufficient number of data for the case. We find that the model trained with the rotational augmentation shows better results overall, especially in terms of pixel-to-pixel CC after 8×8 binning where the performance gain is around 0.04. This shows that a better result can be drawn with more number of data for rare cases and/or a proper data augmentation. Throughout this Letter, we use results from the model trained with the rotational augmentation method.

Appendix C
Implementation Details

C.1. Training Process

The learning process of our model is made as follows (see Figure 1).

1. A Ca II K image x is given to a generator and it produces an HMI-like magnetogram $G(x)$.
2. The same Ca II K image x and the generated magnetogram $G(x)$ are combined and given to the discriminators A and B. The discriminators yield matrices $D(x, G(x))$ whose values are between 0 (fake) and 1 (real). Then we give a matrix, whose every element is 0 and whose size is the same with each of the discriminator's output, so that the discriminators get a loss from differences between their output $D(x, G(x))$ and the given matrix of 0 values. With the loss, the discriminators learn a relation between Ca II images and generated magnetograms.
3. In order to make the discriminators learn a relation between Ca II images and the corresponding SDO/HMI magnetograms, a Ca II K image x and the target magnetogram y are combined and given to the discriminators A and B. Then we give matrices whose elements are all 1 and the discriminators are learned up to the difference between their output $D(x, y)$ and the given matrix of 1 values.
4. The generator gets updated with losses from $D(x, G(x))$ and $D(x, y)$.
5. We train the model by iterating from steps 1 to 4 about 490,000 times (about 200 epochs).

C.2. Hyperparameters

We initialize weights in convolutional and transposed convolutional layers with a normal distribution whose mean is 0 and standard deviation is 0.02. For optimization, we use the Adam solver (Kingma & Ba 2014) with an initial learning rate of 2×10^{-4}, and set coefficients β_1 and β_2 as 0.5 and 0.999 for computing running averages of gradients and their squares.

ORCID iDs

Gyungin Shin https://orcid.org/0000-0003-1793-665X

Table 1
Model Evaluation with Total Unsigned Magnetic Flux CC, Pixel-to-pixel CC, and Relative Error

		Full Disk 436 Images (1024 × 1024)			Active Region 510 Images (128 × 128)			Quiet Region 436 Images (128 × 128)		
		Baseline	Ours	Kim et al. (2019)	Baseline	Ours	Kim et al. (2019)	Baseline	Ours	Kim et al. (2019)
Total unsigned magnetic flux CC		0.36[a]	0.99	0.97	0.58[a]	0.99	0.95	−0.047[a]	0.95	0.74
pixel-to-pixel CC	1 × 1 binning	(0.39)	0.42 (0.58)	⋯	(0.40)	0.57 (0.66)	⋯	(0.34)	0.09 (0.31)	⋯
	2 × 2 binning	(0.47)	0.55 (0.75)	⋯	(0.45)	0.66 (0.79)	⋯	(0.49)	0.15 (0.53)	⋯
	4 × 4 binning	(0.54)	0.67 (0.87)	⋯	(0.51)	0.75 (0.88)	⋯	(0.64)	0.20 (0.72)	⋯
	8 × 8 binning	(0.62)	0.74 (0.93)	0.77	(0.59)	0.81 (0.94)	0.66	(0.74)	0.24 (0.81)	0.21
	16 × 16 binning	(0.73)	0.78 (0.97)	⋯	(0.71)	0.85 (0.97)	⋯	(0.78)	0.28 (0.84)	⋯
Relative error	Mean	−0.79[a]	−0.085	0.067	−0.72[a]	−0.040	0.072	−0.78[a]	−0.054	0.091
	Standard deviation	0.23[a]	0.032	0.036	0.13[a]	0.064	0.019	1.10[a]	0.069	0.075
Structural similarity index		⋯	0.98	⋯	⋯	0.80	⋯	⋯	0.97	⋯

Notes. For pixel-to-pixel CC and relative error, we also evaluate them from absolute magnetic flux density maps (written in parenthesis). Every right column in each area subsection are taken from Kim et al. (2019) for comparison.
[a] We count pixels over +30 G for evaluating baseline as we found the simple pixel to pixel calibration curve does not properly present weak fields.

Yong-Jae Moon ⓘ https://orcid.org/0000-0001-6216-6944
Eunsu Park ⓘ https://orcid.org/0000-0003-0969-286X
Hyunjin Jeong ⓘ https://orcid.org/0000-0003-4616-947X
Harim Lee ⓘ https://orcid.org/0000-0002-9300-8073

References

Arjovsky, M., Chintala, S., & Bottou, L. 2017, arXiv:1701.07875
Babcock, H. W., & Babcock, H. D. 1955, ApJ, 121, 349
Balasubramaniam, K. S., & Pevtsov, A. 2011, Proc. SPIE, 8148, 814809
Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019a, A&A, 626, A114
Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019b, SoPh, 294, 145
Chatzistergos, T., Ermolli, I., Solanki, S. K., & Krivova, N. A. 2018, A&A, 609, A92
Domingo, V., Fleck, B., & Poland, A. I. 1995, SoPh, 162, 1
Ermolli, I., Criscuoli, S., Centrone, M., Giorgi, F., & Penza, V. 2007, A&A, 465, 305
Ermolli, I., Fofi, M., Bernacchia, C., et al. 1998, SoPh, 177, 1
Foukal, P., Bertello, L., Livingston, W. C., et al. 2009, SoPh, 255, 229
Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv:1406.2661
Hale, G. E., Ellerman, F., Nicholson, S. B., & Joy, A. H. 1919, ApJ, 49, 153
Isola, P., Zhu, J., Zhou, T., & Efros, A. A. 2016, CoRR, arXiv:1611.07004
Kim, T., Park, E., Lee, H., et al. 2019, NatAs, 3, 397
Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
Leamon, R. J., & McIntosh, S. W. 2009, ApJL, 697, L28
LeCun, Y., & Bengio, Y. 1998, The Handbook of Brain Theory and Neural Networks (Cambridge, MA: MIT Press)

Leighton, R. B. 1959, ApJ, 130, 366
Lemen, J. R., Title, A. M., Akin, D. J., et al. 2012, SoPh, 275, 17
Liu, Y., Hoeksema, J. T., Scherrer, P. H., et al. 2012, SoPh, 279, 295
Livingston, W. C., Harvey, J., Slaughter, C., & Trumbo, D. 1976, ApOpt, 15, 40
Mao, X., Li, Q., Xie, H., Lau, R. Y. K., & Wang, Z. 2016, CoRR, arXiv:1611.04076
McIntosh, S. W., Davey, A. R., & Hassler, D. M. 2006, ApJL, 644, L87
McIntosh, S. W., Leamon, R. J., Davey, A. R., & Wills-Davey, M. J. 2007, ApJ, 660, 1653
Mirza, M., & Osindero, S. 2014, CoRR, arXiv:1411.1784
Park, E., Moon, Y.-J., Lee, J.-Y., et al. 2019, ApJL, 884, L23
Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in Neural Information Processing Systems 32, ed. H. Wallach et al., 32 (Red Hook, NY: Curran Associates, Inc.), 8026
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Pevtsov, A. A., Virtanen, I., Mursula, K., Tlatov, A., & Bertello, L. 2016, A&A, 585, A40
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, SoPh, 275, 229
van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, PeerJ, 2, e453
Wang, T., Liu, M., Zhu, J., et al. 2017, CoRR, arXiv:1711.11585
Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E. 2004, ITIP, 13, 600
Zhu, J., Park, T., Isola, P., & Efros, A. A. 2017, CoRR, arXiv:1703.10593