Generation of Modern Satellite Data from Galileo Sunspot Drawings in 1612 by Deep Learning

Harim Lee[1] , Eunsu Park[1] , and Yong-Jae Moon[1,2]

[1] Department of Astronomy and Space Science, College of Applied Science, Kyung Hee University, Yongin, 17104, Republic of Korea; moonyj@khu.ac.kr
[2] School of Space Research, Kyung Hee University, Yongin, 17104, Republic of Korea

Abstract

Historical sunspot drawings are very important resources for understanding past solar activity. We generate solar magnetograms and EUV images from Galileo sunspot drawings using a deep learning model based on conditional generative adversarial networks. We train the model using pairs of sunspot drawings from the Mount Wilson Observatory and their corresponding magnetograms (or UV/EUV images) from 2011 to 2015 except for every June and December by the Solar Dynamic Observatory satellite. We evaluate the model by comparing pairs of actual magnetograms (or UV/EUV images) and the corresponding AI-generated ones in June and December. Our results show that bipolar structures of the AI-generated magnetograms are consistent with those of the original ones and their unsigned magnetic fluxes (or intensities) are consistent with those of the original ones. Applying this model to the Galileo sunspot drawings in 1612, we generate Helioseismic and Magnetic Imager-like magnetograms and Atmospheric Imaging Assembly-like EUV images of the sunspots. We hope that the EUV intensities can be used for estimating solar EUV irradiance at long-term historical times.

Unified Astronomy Thesaurus concepts: The Sun (1693); Sunspots (1653); Sunspot cycle (1650); Solar active regions (1974); Solar magnetic fields (1503)

1. Introduction

Sunspot drawing is an important resource for understanding the long-term variation of solar activity (Vaquero 2007; Arlt & Vaquero 2020). In particular, a record of sunspot numbers and their locations played an important role in discovering the 11 yr solar cycle and solar rotation (Hoyt & Schatten 1998a, 1998b; Usoskin et al. 2003a; Casas et al. 2006). Furthermore, the relationship between solar activity and the Earth's climate has been studied by estimating magnetic field strength and solar irradiance through simple physical models using historical sunspot number and area information (Lockwood et al. 1999; Usoskin et al. 2003b; Solanki & Krivova 2004; Zharkov & Zharkova 2006; Krivova et al. 2007; Nagovitsyn et al. 2016; Pevtsov et al. 2019).

Nowadays, it is possible to conduct extensive research on solar activity using high-quality solar satellite observations (Solanki 2003; Munoz-Jaramillo & Vaquero 2019). We propose a new way to conduct solar long-term variation studies using a deep learning model that transforms historical sunspot drawings into modern satellite images and their physical parameters. This study is expected to serve as a bridge between historical sunspot drawing data and modern satellite data.

For this we adopt a deep learning model, which is a general purposed solution for the image-to-image translation, named "pix2pix" (Isola et al. 2016), to sunspot drawings, in order to generate solar magnetograms and UV/EUV images. For this we make eight deep learning models for image translations. For training and evaluation data sets, we consider pairs of the Mount Wilson Observatory (MWO) sunspot drawings and Solar Dynamics Observatory (SDO; Pesnell et al. 2012)/Helioseismic and Magnetic Imager (HMI; Schou et al. 2012) line-of-sight magnetograms from 2011 to 2015. We also consider seven sets of data: pairs of the sunspot drawings and Atmospheric Imaging Assembly (AIA; Lemen et al. 2012) seven wavelength images.

This study is organized as follows. The data will be described in Section 2, and the method in Section 3. Results

and a discussion are present in Section 4 and our conclusion is given in Section 5.

2. Data

We use the MWO sunspot drawings from 2011 to 2015 for input data. First, we align the sunspot drawings with the corresponding SDO magnetograms and EUV/UV images. Then, we manually remove all letters and lines except for sunspots. In order to use relatively clear sunspot drawings, we make 8 bit scale images as follows (Figure 1(a)): 255 for solar disk and 0 for umbra, penumbra, and outside of solar disk.

We use SDO/HMI line-of-sight magnetograms and SDO/AIA seven wavelength images (94, 131, 171, 193, 211, 304, and 335 Å) for target data. We first make level 1.5 images by calibrating, rotating, and centering the images. We divide the data numbers of all AIA images by exposure time to make a homogeneous exposure condition (DN s^{-1}). To compensate for the instrument degradation over time for seven AIA EUV passbands (Boerner et al. 2012), we apply their degrading factors to the images using a SolarSoft routine (aia_get_response.pro) with a reference date of 2011 January 1. Also, we coalign the AIA and HMI images by fixing the solar disk size, and downsample the images to 512 by 512. Magnetic flux densities are considered within ±1000 Gauss. Since sunspot drawings only represent strong magnetic fields, we calculate the total unsigned magnetic fluxes (TUMFs) of magnetograms for only strong field areas whose absolute field strengths are larger than 50 Gauss, which approximately correspond to 5 times of noise levels (Liu et al. 2012).

We use an AIA image with the range of 0 DN s^{-1} for minimum and $2^6 - 1 \sim 2^{13} - 1$ (DN s^{-1}) for maximum. The maximum value of each passband is determined by the brightness in active regions without flares (see Table 5 of Boerner et al. 2012).

We exclude image pairs with poor quality: e.g., images that are too noisy because of solar flares, incorrect header information,
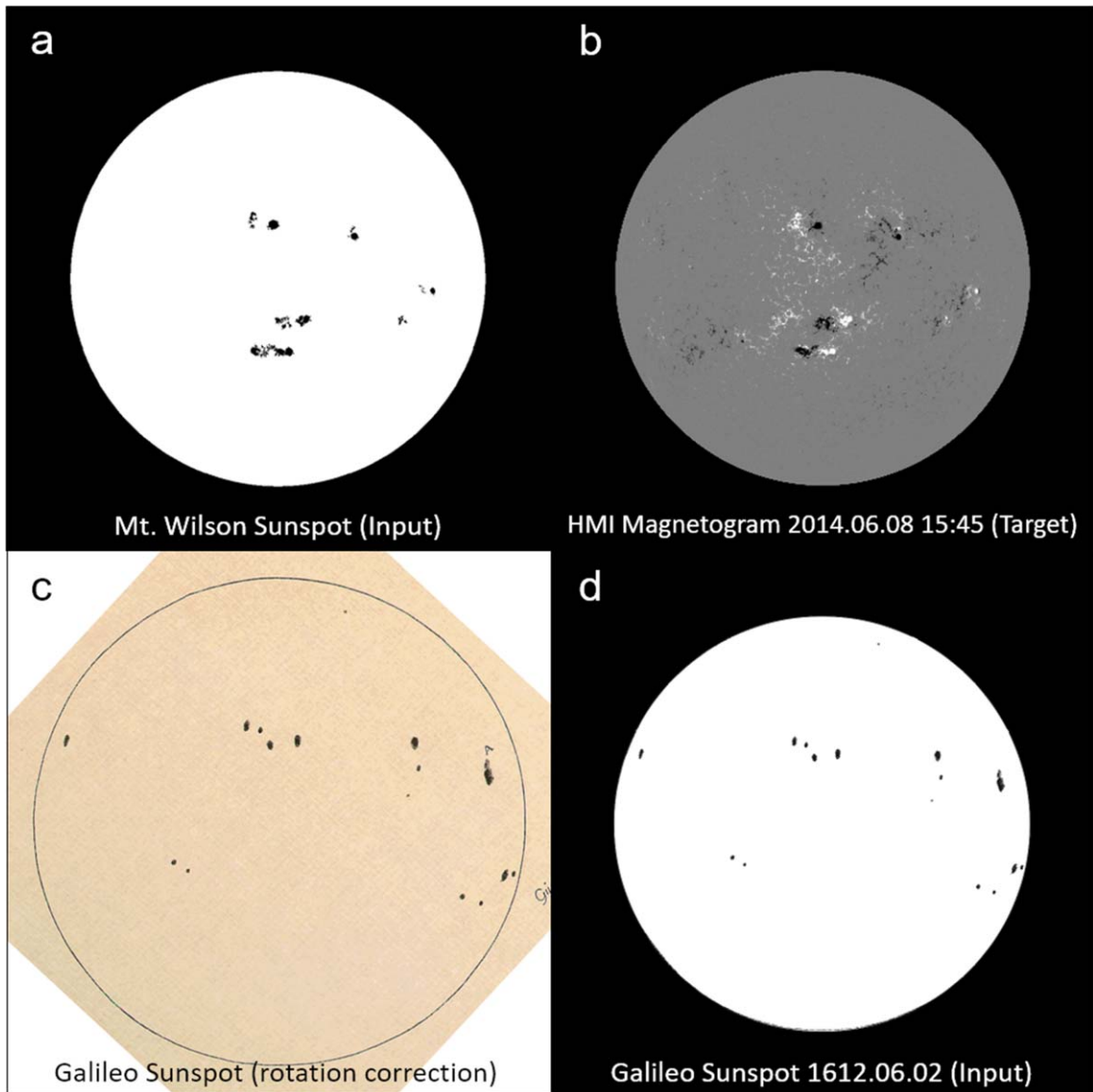
Figure 1. An example of input and target images for test and application. (a) Mount Wilson sunspot drawing, which is an input image on 2014 June 8. (b) The corresponding SDO/HMI image as the target, which is byte-scaled with ±500 Gauss for only display. (c) Galileo sunspot drawing with rotation correction on 1612 June 2. (d) Its modified image for input data.

atypical images due to reasons such as the eclipse of a planet. We adopt SDO data, which were observed within ±36 minutes of the observation time of each sunspot drawing, in order to minimize the effect of solar rotation. This time corresponds to 1.5 pixels by solar rotation. As a result, we make eight data sets, which include 1250 pairs of each data set: MWO sunspot drawings and SDO images. For training we use 1046 pairs from 2011 to 2015 except for every June and December. For evaluating our model, the remaining 204 pairs are used. Figure 1 shows an example of input image, target one, and Galileo sunspot drawing. Figure 1(a) is an MWO sunspot drawing (input) on 2014 June 8. Figure 1(b) is the corresponding SDO/HMI image (target) at 15:45 UT on 2014 June 8.

For application of our model, we use the 35 Galileo sunspot drawings (Galilei et al. 1613) processed by Al Van Helden and Owen Gingerich from 1612 June 2 to July 8 (images are available at galileo.rice.edu). First, we align the Galileo sunspot drawings with the MWO sunspot drawings, after correcting the rotational axis so north is up. We use the rotational degree from the electronic supplementary material of Vokhmyanin & Zolotova (2018). Then,

we remove the letters and make 8 bit images as we did for our training and test data. Figure 1(c) is the Galileo sunspot drawing with rotation correction on 1612 June 2 and Figure 1(d) is an input image for our model.

3. Method

We adopt a deep learning model based on pix2pix. The pix2pix is based on the generative adversarial network (GAN; Goodfellow et al. 2014), which is an novel deep learning algorithm for the generation tasks. The pix2pix is a combination of the conditional generative adversarial network (cGAN; Mirza & Osindero 2014) and the deep convolutional generative adversarial network (DCGAN; Radford et al. 2015). Our model consists of two networks: one is a generator and the other is a discriminator. The rule of the generator is to generate a target-like image from an input image by minimizing the difference between the target image and the generated one. The rule of the discriminator is to distinguish the real pair from the generated pair. The real pair
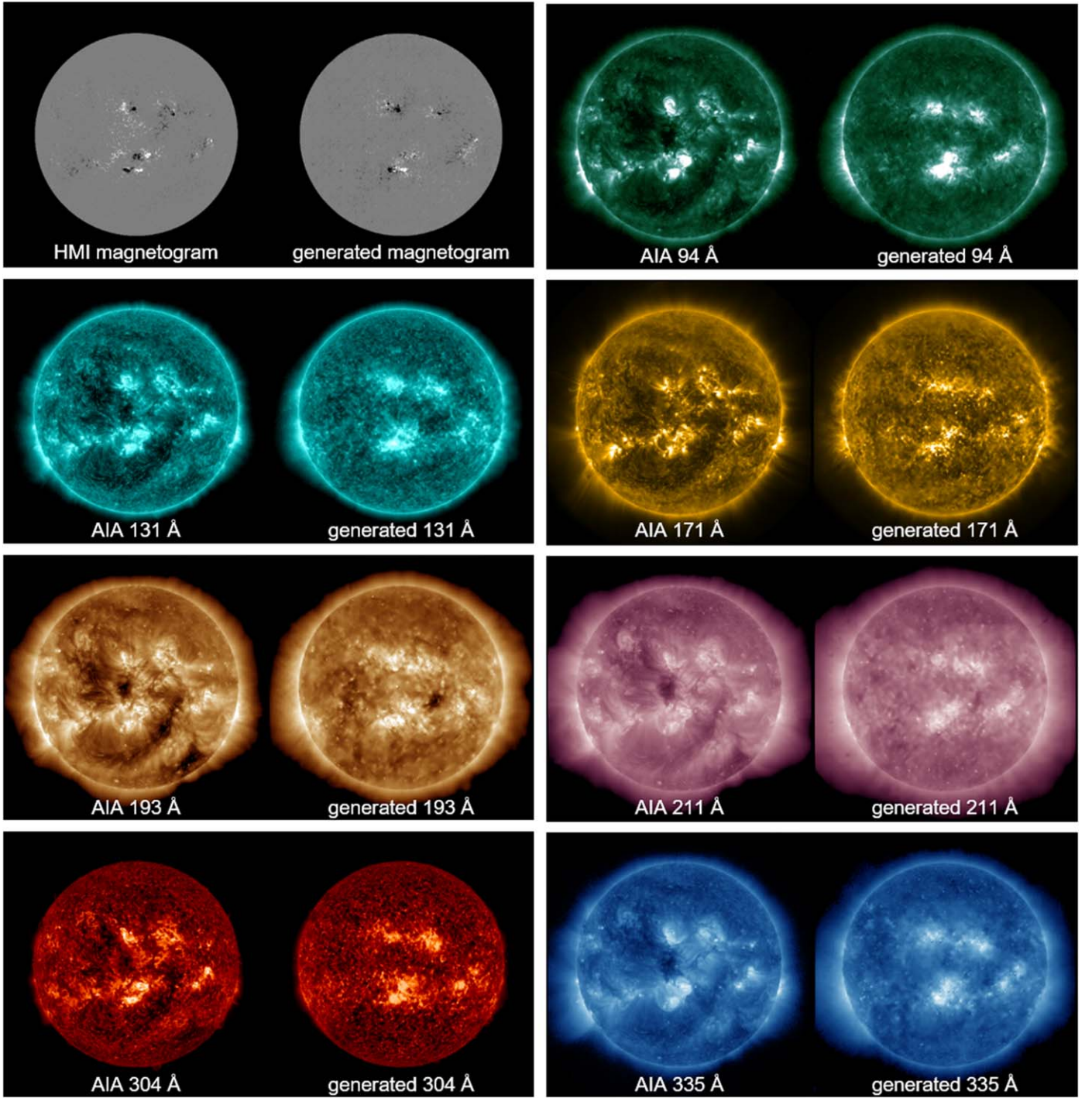
Figure 2. Comparison between real SDO images and the generated ones at 15:45 UT on 2014 June 8. The first columns represent SDO/HMI and SDO/AIA images, the second columns are AI-generated images. Here two magnetograms are byte-scaled by ±500 Gauss for display only.

consists of the input image and the target one. The generated pair consists of the input image and the one generated by our model.

The pix2pix, proposed by Isola et al. (2016), uses 256 × 256 size images for training. We modify the data pipeline and depth of the generator network because the size of our data is 512 × 512. The loss function and other hyper parameters are the same as those of Isola et al. (2016). For each translation from sunspot drawings to a specific type of solar image, we make one deep learning model.

We save the generator (and the discriminator) every 10,000 iterations to check the training process, to avoid over-training, and

to find the best model. The best model is taken when it gives the highest mean correlation coefficient (CC) value for the evaluation data set. We empirically find that the models are sufficiently trained before 210,000 iterations (∼200 epochs). Here one iteration is when one pair of images is trained in our model, and one epoch is when an entire training data set of 1046 pairs is done in our model.

4. Results and Discussion

Figure 2 shows eight pairs of target images and their corresponding AI-generated ones at 15:45 UT on 2014 June 8.
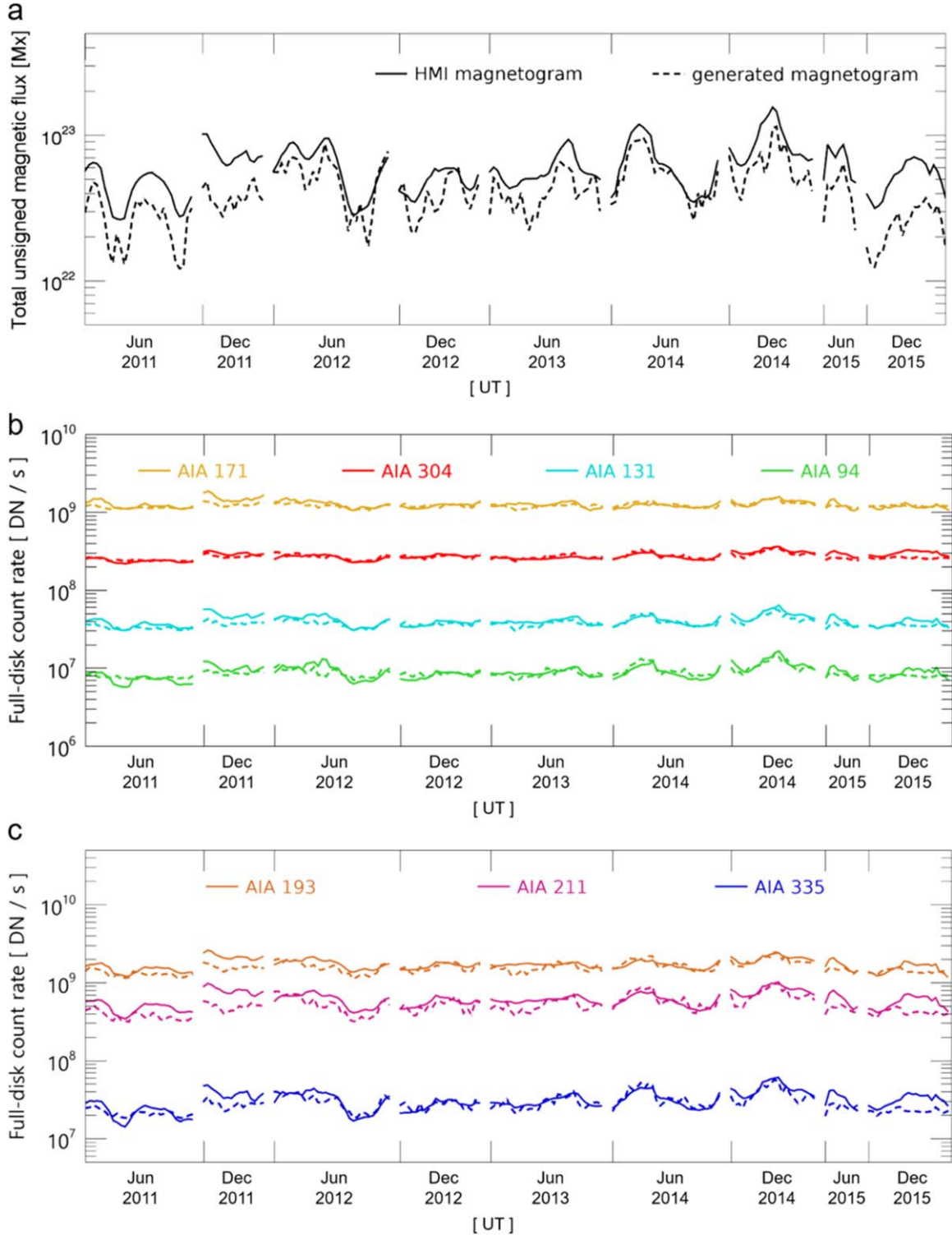
Figure 3. Temporal variations of the total unsigned magnetic flux (a) and full-disk count rates ((b) and (c)) from 2011 to 2015. The solid lines show variations from the real images and dashed lines correspond to those from the generated ones. Small vertical space is given the period of training.

A comparison between target and AI-generated magnetograms shows that the bipolar structures of the HMI magnetograms are approximately restored. Even though we do not have any prior conditions such as preceding or following sunspots for sunspot drawings, bipolar structures in AI-generated magnetograms mostly follow Hale's law (Hale & Nicholson 1925). Our model learns the polarity pattern in the training step, then reproduces

such a pattern in the evaluation and the generation step. It makes sense in that most of the active regions follow Hale's law. However, our model does not successfully generate active regions that do not follow Hale's law. Note that the polarity of the solar magnetic field is reversed cycle by cycle. Since all data are from the twenty-fourth solar cycle, there is no problem producing the Hale's law pattern in this cycle. Hence, our model would be
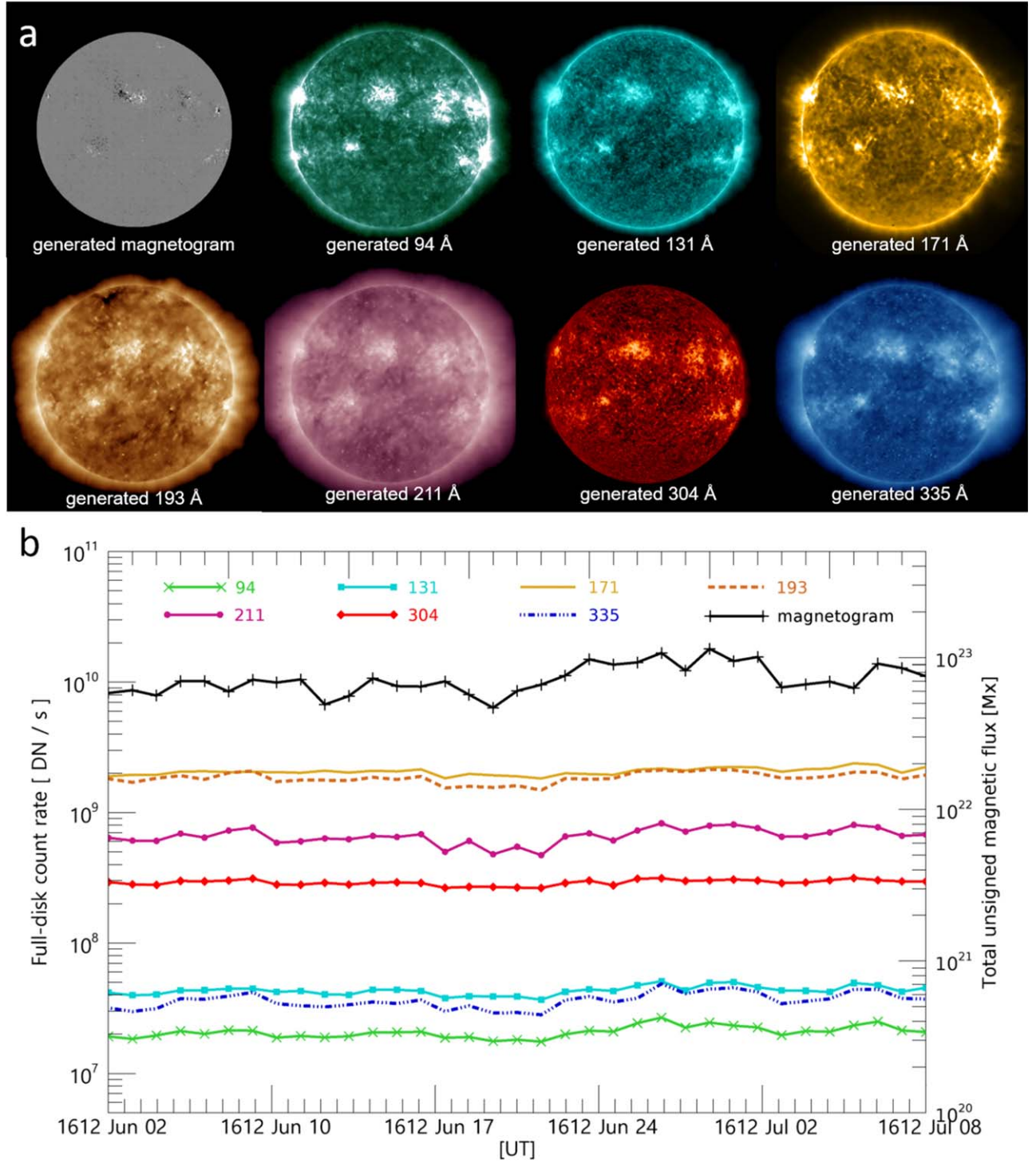
Figure 4. Results of our model when the Galileo sunspot drawings are used for input data. (a) AI-generated magnetogram and UV/EUV images from the Galileo sunspot drawing on 1612 June 2. (b) Total unsigned magnetic flux and full-disk count rates estimated from the AI-generated Galileo magnetogram and UV/EUV images from 1612 June 2 to July 8.

effective for even solar cycles, but should be tested or the reversed polarity for odd cycles. For further discussion on this issue, please refer to Kim et al. (2019). As seen in the figure, the UV/EUV brightness at active regions of the AIA images is mostly restored. Other detailed structures are not reproduced well, because sunspot drawings do not have information on other structures such as filaments and coronal holes. These results mean that our model can only reproduce active regions.

We estimate TUMF and full-disk count rates (CR) of the evaluation data for both generated and real images and their temporal variations are given in Figure 3. We only consider pixels within 0.98 solar radius for all images to avoid uncertainties near the limb. The average CC of TUMF and CR between generated and real ones are 0.82 for magnetograms, 0.74 for 94 Å, 0.75 for 131 Å, 0.56 for 171 Å, 0.65 for 193 Å, 0.67 for 211 Å, 0.70 for 304 Å, and 0.76 for 335 Å,

respectively. The average CC is the highest for magnetograms because the sunspot drawings trace photospheric magnetic field distributions. It is the lowest for 171 Å, thus it seems that sunspot drawings cannot figure out the detailed configuration of EUV coronal loops, which are most evident for this passband image. The normalized rms error of TUMF and CR ranges from 0.08 to 0.37. The TUMF and CR from generated images are comparable or slightly underestimated to those from real ones. This may be caused by the fact that the model generates active regions well, but not for quiet regions and/or EUV coronal structures. Nevertheless, their overall trends from both data sets are approximately consistent with each other.

Now we apply our model to Galileo sunspot drawings in 1612, in order to generate modern satellite solar images. Figure 4(a) shows the SDO/HMI-like magnetogram and SDO/AIA-like images on 1612 June 2. Noting that 1755 is the first year of the first solar cycle and the solar cycle period is 11 years, we assume that 1612 belongs to an even solar cycle. Therefore our model generates the polarity pattern of the active regions in the AI-generated magnetogram like twenty-fourth solar cycle. We can see the EUV brightness of active regions in the AI-generated UV/EUV images. Since other structures such as coronal loops, coronal hole, and filaments are not well reproduced, we admit that the present model cannot produce the detailed morphology of solar corona. However, it is possible to estimate magnetic flux and EUV intensity of solar active regions. Figure 4(b) shows the temporal variation of TUMF and CR from 1612 June 2 to July 8. From this result, we can see the temporal variation of magnetic flux and EUV intensity over a month in 1612.

5. Conclusion

In this study, we proposed a new attempt to generate modern satellite data and their related physical parameters from historical sunspot drawings. We demonstrated the validity of this attempt using modern data sets: Mount Wilson sunspot drawings and SDO data. Finally, our model produces modern satellite images and related physical values from Galileo sunspot drawings. This study is expected to offer more information on the long-term evolution of solar magnetic fields and their related studies such as long-term variation of solar irradiance.

ORCID iDs

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

References

Arlt, R., & Vaquero, J. M. 2020, LRSP, 17, 1
Boerner, P., Edwards, C., Lemen, J., et al. 2012, SoPh, 275, 41
Casas, R., Vaquero, J. M., & Vazquez, M. 2006, SoPh, 234, 379
Galilei, G., Welser, M., & de Filiis, A. 1613, Istoria E dimostrazioni intorno alle macchie solari E loro accidenti comprese in tre lettere scritte all'illvstrissimo signor Marco Velseri (Rome: G. Mascadi)
Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv:1406.2661
Hale, G. E., & Nicholson, S. B. 1925, ApJ, 62, 270
Hoyt, D. V., & Schatten, K. H. 1998a, SoPh, 171, 189
Hoyt, D. V., & Schatten, K. H. 1998b, SoPh, 171, 491
Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. 2016, arXiv:1611.07004
Kim, T., Park, E., Lee, H., et al. 2019, NatAs, 3, 397
Krivova, N. A., Balmaceda, L., & Solanki, S. K. 2007, A&A, 467, 335
Lemen, J. R., Title, A. M., & Waltham, N. 2012, SoPh, 275, 17
Liu, Y., Hoeksema, J. T., Scherrer, P. H., et al. 2012, SoPh, 279, 295
Lockwood, M., Stamper, R., & Wild, M. N. 1999, Natur, 399, 437
Mirza, M., & Osindero, S. 2014, arXiv:1411.1784
Munoz-Jaramillo, A., & Vaquero, J. M. 2019, NatAs, 3, 205
Nagovitsyn, Y. A., Tlatov, A. G., & Nagovitsyna, E. Y. 2016, ARep, 60, 831
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Pevtsov, A. A., Tlatova, K. A., Pevtsov, A. A., et al. 2019, A&A, 628, A103
Radford, A., Metz, L., & Chintala, S. 2015, arXiv:1511.06434
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, SoPh, 275, 229
Solanki, S. K. 2003, A&ARv, 11, 153
Solanki, S. K., & Krivova, N. A. 2004, SoPh, 224, 197
Usoskin, I. G., Mursula, K., & Kovaltsov, G. A. 2003a, SoPh, 218, 295
Usoskin, I. G., Solanki, S. K., Schussler, M., Mursula, K., & Alanki, K. 2003b, PhRvL, 91, 211101
Vaquero, J. M. 2007, AdSpR, 40, 929
Vokhmyanin, M. V., & Zolotova, N. V. 2018, SoPh, 293, 31
Zharkov, S. I., & Zharkova, V. V. 2006, AdSpR, 38, 868