



Super-resolution of SDO/HMI Magnetograms Using Novel Deep Learning Methods

Sumiaya Rahman¹, Yong-Jae Moon^{1,2} , Eunsu Park² , Ashraf Siddique³ , Il-Hyun Cho² , and Daye Lim²

¹ School of Space Research, Kyung Hee University, Yongin, 17104, Republic of Korea; moonjy@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 Engineering, Kyung Hee University, Yongin, 17104, Republic of Korea

Received 2020 February 21; revised 2020 May 20; accepted 2020 June 15; published 2020 July 10

Abstract

Image super-resolution is a technique of enhancing the resolution of an image where a high-resolution (HR) image is reconstructed from a low-resolution (LR) image. In this Letter, we apply two novel deep learning models (residual attention model and progressive GAN model) for enhancing Solar Dynamics Observatory (SDO)/Helioseismic and Magnetic Imager (HMI) magnetograms. For this, we consider line-of-sight (LOS) magnetograms taken by SDO/HMI as output and their degraded ones with 4×4 binning as input. Deep learning networks try to find internal relationships between LR and HR images from the given input and the corresponding output image. We consider SDO/HMI magnetograms from 2014 May to August for training, from 2014 October to December for validation, and 2015 January to March for test. We find that the deep learning models generate higher-quality results than the bicubic interpolation in terms of visual aspects and metrics. We apply this model to a full-resolution SDO/HMI magnetogram and then compare the generated magnetogram with the corresponding Hinode/The Solar Optical Telescope Narrowband Filtergrams (NFI) magnetogram. This comparison shows that the generated magnetogram is consistent with the Hinode one with a high correlation (CC: 0.94) and a high similarity (SSIM: 0.93), which are better than the bicubic method.

Unified Astronomy Thesaurus concepts: [Solar magnetic fields \(1503\)](#); [The Sun \(1693\)](#); [Astronomy data analysis \(1858\)](#); [Solar atmosphere \(1477\)](#)

1. Introduction

Solar magnetograms are important tools for studying the distribution of the solar magnetic field, coronal holes, the connection between magnetic fields in active regions, and the interplanetary magnetic field lines (Scherrer et al. 2012). In particular, it is characterized by magnetic field strength, polarity, and its location on the Sun. There are several space missions, such as the Solar Heliospheric Observatory (Domingo et al. 1995), Hinode (Kosugi et al. 2007), and the Solar Dynamics Observatory (SDO; Pesnell et al. 2012), which have observed solar atmosphere and other active phenomena with high spatial and temporal resolution. The Helioseismic and Magnetic Imager (HMI; Schou et al. 2012) is one of SDO's onboard instruments and is designed to provide $0''.504 \text{ pix}^{-1}$ resolution solar image of Doppler shift, continuum intensity, and vector magnetic field. The major activity of the instrument is to investigate the mechanism of solar variability, describe and understand the physical process of the Sun's interior, and observe the magnetic activity of the solar atmosphere. In spite of having such a spatial telescope with polarization selector and image stabilization system, the spatial resolution is not sufficient to observe the small-scale structures of the Sun. A magnetogram with high resolution (HR) can help us to characterize small magnetic features of the solar surface and improve the estimation of the total magnetic flux, which is limited by the resolution of the magnetograms (Baso & Díaz Ramos 2018).

Super-resolution (SR) is a technique of enhancing the resolution of an image from one or more observed low-resolution (LR) images (Park et al. 2003; Dong et al. 2016; Ledig et al. 2017). In astronomy, SR is a process that provides astronomical images with a higher resolution based on observed astronomical images that are captured with the same astronomical scene by adding high-frequency content and eliminating the degradation due to the image processing of the astronomical telescopes (Guo et al. 2019).

In the literature, several approaches have been used to overcome the resolution enhancement problem, such as pixel-reliability-based SR reconstruction algorithm (Li et al. 2018), advanced image restoration techniques (Misra et al. 2018), and image denoising (Yan et al. 2017).

Recently, rapid advancements in deep learning-based image processing have been made using consistent, HR data sets and state-of-the-art algorithms. Deep learning models allow to learn the way humans think and recognize an object using multiple processing layers (Lecun et al. 2015). The depth of the convolutional neural network (CNN; Lecun et al. 1998) is important for SR. However, the deeper networks for image SR are more difficult to train (Zhang et al. 2018) because of the vanishing gradient problem. Because most of the CNN-based SR models are unable to make proper use of the hierarchical features from LR images, relatively low performance was observed (Zhang et al. 2018). To overcome the problem of vanishing gradient, Zhang et al. (2018) considered a residual attention model to get a very deep trainable network that learns important channel-wise features simultaneously. In the literature of image SR, different types of network structure, loss functions, learning principles, and strategies are used. Recently, the residual attention model has proven to be a popular deep learning method that gives a state-of-the-art performance in image SR when compared to the other deep learning models like SRCNN (Dong et al. 2016), EDSR (Lai et al. 2017), and LapSRN (Lim et al. 2017). Furthermore, the generative adversarial network (GAN; Goodfellow et al. 2014) is also a well-known model that enhances the perceptual quality of upsampled images. One of the GAN-based methods is a progressive GAN model (Wang et al. 2018) that follows the progressive multi-scale principle for SR. Direct reconstruction techniques, such as SRCNN and SRGAN (Ledig et al. 2017) can upsample the image in a single step, thus adding blur artifacts. In

contrast, the progressive GAN base model can reduce computational complexity and enlarge the model capability.

To our knowledge, there was one attempt to enhance solar images using deep learning. Baso & Díaz Ramos (2018) applied a CNN-based deep learning method to synthetic data, which are simulated images of solar active regions with 2×2 binning, and compared their results with Hinode/The Solar Optical Telescope (SOT; Tsuneta et al. 2008) Broadband Filtergrams continuum images.

In this Letter, we apply two novel deep learning methods (the residual attention model and the progressive GAN model) to SDO/HMI magnetograms. For training, validation, and testing, we consider two data sets: degraded magnetograms with 4×4 binning for input and the corresponding original ones for output. For evaluation, we compare the results of two deep learning models with that of the bicubic method. In addition, we apply the residual attention model to a full-resolution SDO/HMI magnetogram and compare the generated magnetogram with the corresponding Hinode/SOT NFI magnetogram. We evaluate the results of the models in view of metric and visual inspection. This Letter is organized as follows. The data will be described in Section 2. The methods used for SR are discussed in Section 3. The results are given in Section 4. A brief summary and discussion are given in Section 5.

2. Data

2.1. SDO/HMI Magnetogram

SDO is a space mission that was launched on 2010 February 11 (Pesnell et al. 2012). The HMI instrument is part of the SDO mission, which provides continuous coverage of full-disk Doppler velocity, line-of-sight (LOS) magnetic flux, and continuum proxy images (Scherrer et al. 2012). For this study, we consider full-disk (4096×4096) SDO/HMI LOS magnetograms from 2014 and 2015 for every 00:00 UT. First, we make level 1.5 magnetograms by rotating and centering. To make the training computationally feasible, we randomly crop a region from the areas closer to the disk center (about $\pm 500''$). In the limb area, the spatial contrast of the magnetograms is very small so that the deep learning networks cannot reconstruct the structures, thus adding artifacts (Lim et al. 2017). We crop 10,785 pairs of HR SDO/HMI magnetograms with 512×512 pixels and make a 2D Gaussian smoothing with (5, 5) kernel and 2σ . Then we downsample them in size 128×128 pixels with a bilinear interpolation for LR data. For training, validation, and testing, we consider magnetograms with a range of -2000 G for minimum and 2000 G for maximum. We consider 9000 pairs from 2014 May to August for training and 985 pairs from 2014 October to December for validation and 800 pairs from 2015 January to December for testing. To avoid overfitting, we consider 10% of the training data for validation.

2.2. Hinode/SOT NFI Magnetogram

To validate the output of the residual attention model, we compare the output with an NFI magnetogram taken by the Hinode/SOT on 2013 January 7. The NFI magnetogram is defined by the stokes V divided by the stokes I , which are simultaneously observed with the pixel size of $0''.16$ and exposure time of 0.4 s. The NFI magnetogram is aligned to the HMI magnetogram based on visual inspection. The pixel resolution of

the NFI magnetogram is set to be $1/4 \times$ the HMI one by using the cubic interpolation. Then, the NFI magnetogram is calibrated by a cross-comparison (Moon et al. 2007) with the corresponding HMI one. For this, the third-order polynomial fitting of the HMI magnetic field strength is made as a function of NFI one. Finally, all magnetograms are scaled to ± 500 G for display.

3. Method

The network architecture is one of the most important parts of deep learning. In the field of SR, various network design (e.g., residual learning, dense connections, etc.) can be applied. For the SR of the solar magnetogram, we adopt the residual attention model and the progressive GAN model.

3.1. Residual Attention Model

To generate the SR magnetogram, we consider a residual attention model that contains a very deep residual network with channel-wise feature attention learning shown in Figure 1(a). In the residual attention model, a deep network over 400 layers is employed as residual-in-residual structure, where residual groups work as a basic module and long skip connection that allows a coarse level residual learning. Each residual groups contain a stack of residual block with short skip connection that allows recovering the abundant low-frequency information. In order to provide more focus on the informative features, an attention mechanism is applied by exploding the interdependencies among the feature channels. Such a channel attention mechanism enables the network to concentrate the contextual information outside of the local region and enhance the discriminative learning ability.

Such a post-upscaling strategy has been proved to be more efficient in terms of computational complexity, as well as achieve higher performance than pre-upscaling SR methods. We applied several loss functions such as mean square error (MSE) loss, L1 loss, perceptual loss, and adversarial loss to get the optimized results. We find that L1 loss provides better performance for both quantitative and visual evaluation than the others. Here the L1 loss used in this study is given by

$$L_1 = \frac{1}{N} \sum_{i=1}^N |I_i^T - I_i^{SR}|, \quad (1)$$

where I^T and I^{SR} is the target and the generated SR images by the residual attention model, and N is the total number of training data.

For training, we use the ADAM (Kingma & Ba 2014) optimizer with momentum $\beta_1 = 0.9$, $\beta_2 = 0.999$. The initial learning rate is 10^{-4} . In every 2×10^5 iterations of back-propagation, the learning rate is decreased to half from the previous label.

3.2. Progressive GAN Model

We also apply a progressive GAN model shown in Figure 1(b) to SDO/HMI magnetograms. In this model, we considered a progressive multi-scale GAN for enhancement of the perceptual quality of the upscaled magnetogram. This network is designed to reconstruct an HR image in intermediate steps by progressively performing a $2 \times$ upsampling of the input from the previous level. For each level, dense compression units are applied after being adapted from the corresponding dense blocks. This network provides residual $R_s(X)$ as output that is added to bicubic

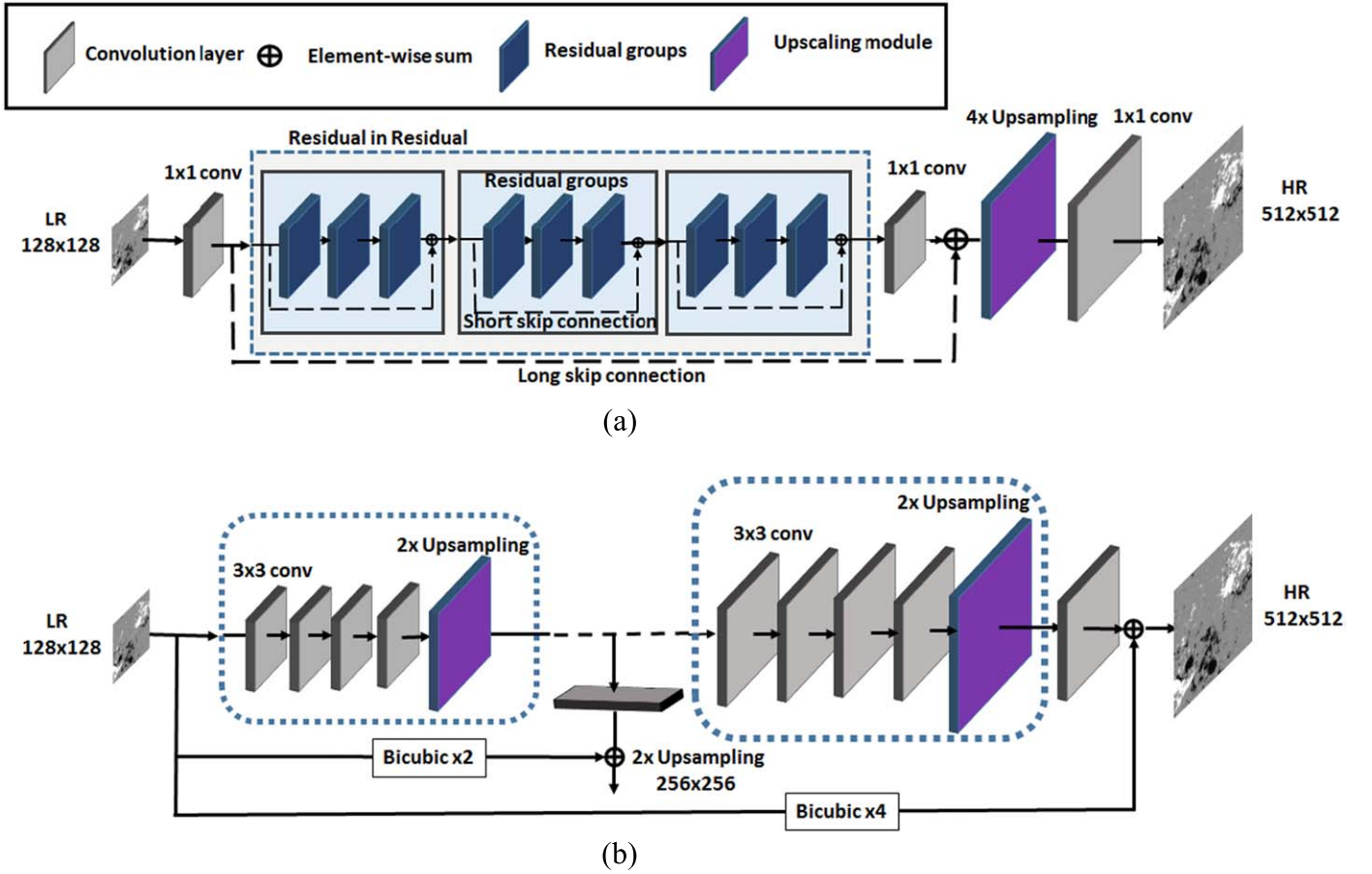


Figure 1. Two flow charts represent two novel deep learning models: (a) the residual attention model and (b) the progressive GAN model. LR is a degraded low-resolution image, HR is a high-resolution image.

upsampling I^{Bic} of input image X to generate final output I^{SR} given by

$$I_s^{\text{SR}} = I_s^{\text{Bic}} + R_s(X). \quad (2)$$

In order to enable multi-scale GAN-enhanced SR, a progressive discriminator network similar to the generator network is used. For training, the least-square loss is applied from VGG16 (Simonyan & Zisserman 2014), which is added to the generator. Expressing the predicted residual and real residual as $R(I^{\text{SR}})$ and $R(I^T)$, the discriminator and generator loss can be expressed as

$$L(D_s) = (D(R(I_s^{\text{SR}})))^2 + (D(R(I_s^T)) - 1)^2, \quad (3)$$

$$L(G_s) = (D(R(I_s^{\text{SR}})) - 1)^2 + \sum_{k \in \{2,4\}} \|\phi_k(I_s^{\text{SR}}) - \phi_k(I_s^T)\|^2, \quad (4)$$

where $L(G_s)$ is the generator loss, $L(D_s)$ is the discriminator loss, s is scaling factor, and ϕ_k denotes the k th pooling layer input in VGG16.

4. Results and Discussion

Figure 2 represents a comparison among the residual attention model, the progressive GAN model, and the bicubic method outputs. The deep learning models show much better results in $4\times$ SR of SDO/HMI magnetograms in both weak and strong fields than the bicubic method. It is very impressive that the magnetograms generated by the deep learning models are consistent with the corresponding target magnetograms.

The results show that deep learning models can reconstruct small-scale magnetic structures much better than the bicubic method. Also, the polarity inversion line is sharper in the deep learning models results than in the bicubic method. It is interesting to note that the deep learning models can successfully alleviate the blurring artifacts when compared to the interpolation methods.

For quantitative evaluation of the results, we calculate four types of metrics for the test data sets. First, we calculate the peak signal-to-noise ratio (S/N) that is commonly applied to measure the reconstruction quality of the model,

$$\text{Peak S/N} = 10 \log \left(\frac{\text{MAX}_I^2}{\text{MSE}_i} \right), \quad (5)$$

where i is the number of testing samples, MAX_I is equal to 4000 maximum value of data, and MSE is the mean square error between the target and the generated magnetogram. Typical values for the peak S/N vary from 30 to 50, where higher is better. The second metric is a pixel-to-pixel Pearson's correlation coefficient (CC) between the target and the generated magnetograms. The third metric is the root mean square error (RMSE) given by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_i^T - I_i^{\text{SR}})^2}, \quad (6)$$

where N , I^T , and I^{SR} are the total number of testing data, target, and generated SR magnetograms, respectively. The last metric

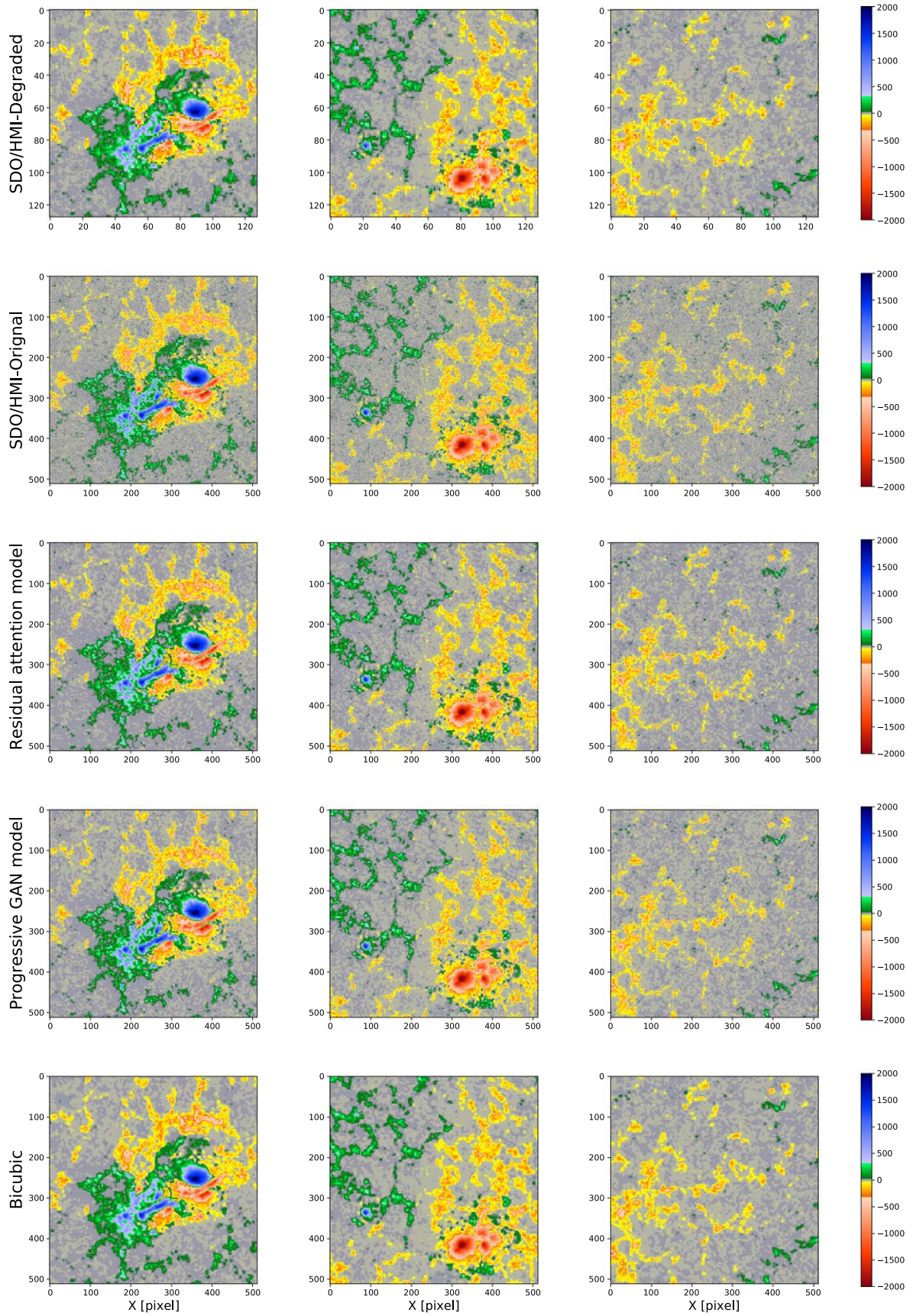


Figure 2. Comparison among input, target, and generated magnetograms by the residual attention model, the progressive GAN model, and the bicubic method. From top to bottom, each row: degraded LR SDO/HMI magnetograms as input data, SDO/HMI magnetograms as target data, generated magnetograms by the residual attention model, the progressive GAN model, and the bicubic method.

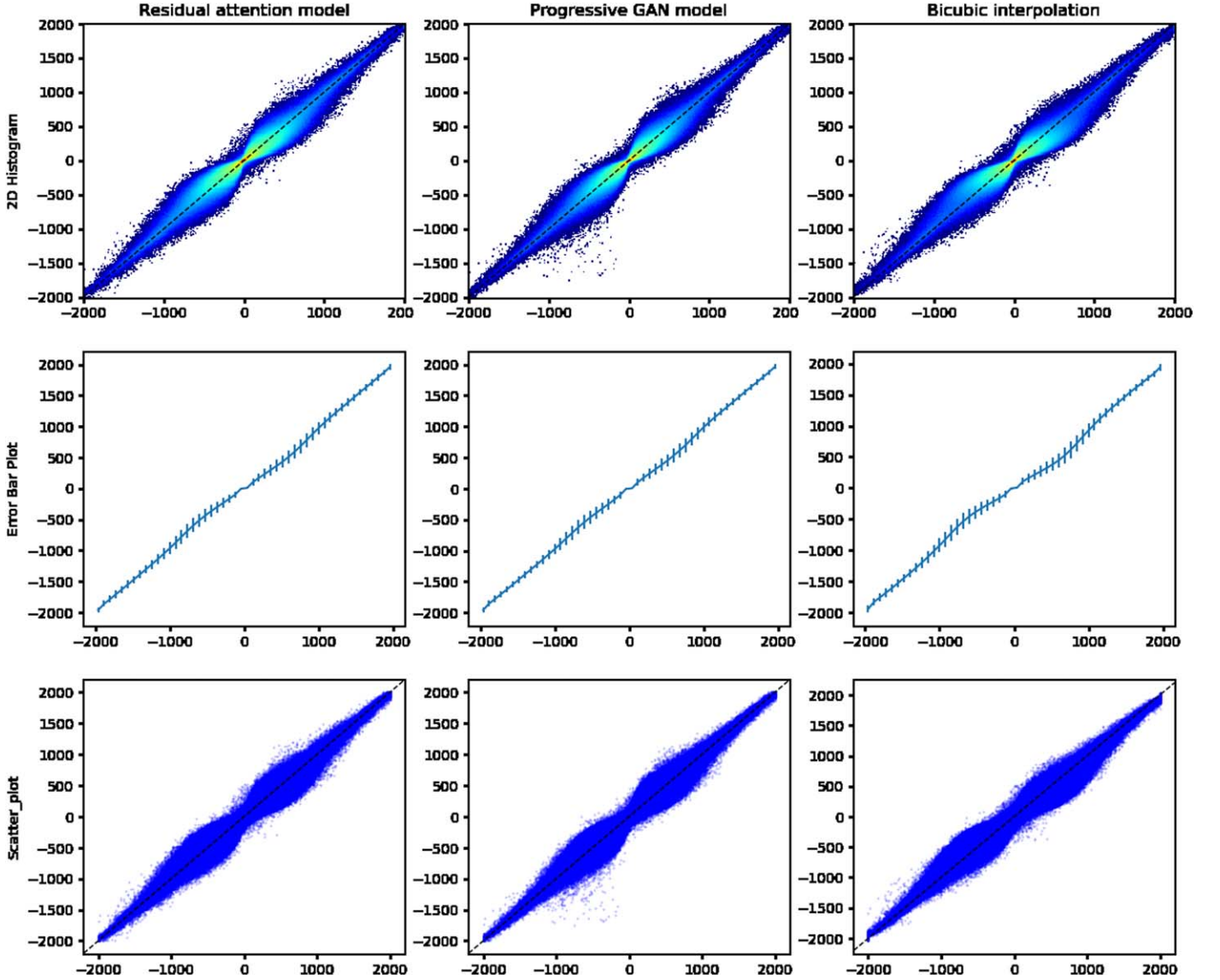


Figure 3. The top row shows 2D histograms, the middle row shows error bar plots, and the bottom row shows scatter plots of the target and generated magnetograms of the testing data set with range of -2000 to 2000 G. From left to right, each column: the residual attention model, the progressive GAN model, and the bicubic ones. The error bar plot shows the mean pixel value, and the error bars indicate the standard deviation of the pixel values. The 2D histogram illustrates the density of the pixel distribution, and scatter plots illustrate the relationship between targets and outputs.

is a structural similarity (SSIM) index, which measures the perceptual difference between the target and the generated image. Here SSIM is given by

$$\text{SSIM} = \frac{(2\mu_T\mu_{SR} + c_1)(2\sigma_{T,SR} + c_2)}{(\mu_T^2 + \mu_{SR}^2 + c_1)(\sigma_T^2 + \sigma_{SR}^2 + c_2)}, \quad (7)$$

where μ_T , μ_{SR} , σ_T , σ_{SR} , $\sigma_{T,SR}$ are the mean, variance, and co-variance of target and generated magnetograms. Here c_1 and c_2 are constant value. SSIM value varies from 0 to 1, where higher is better.

Table 1 represents the average results of peak S/N, pixel-to-pixel Pearson's CC, RMSE, and SSIM value between SDO/HMI magnetograms and generated magnetograms. As shown in Table 1, the peak S/N value of the residual attention model is 48.35 dB, which is significantly higher than those of the other two models. In view of peak S/N, the residual attention

Table 1
Average Values of Pearson's CC, Peak S/N, RMSE, and SSIM of the Testing Data Set

Method	Peak S/N (dB)	Pixel to Pixel (CC)	RMSE (G)	SSIM
Bicubic	46.87	0.87	19.1	0.94
Progressive GAN model	48.32	0.92	16.1	0.98
Residual attention model	48.35	0.93	15.65	0.98

model can reconstruct and recover small-scale features more efficiently than the other models. The pixel-to-pixel correlation value of the residual attention model is 0.93, which is higher than the progressive GAN model and the bicubic method. The average RMSE of the residual attention model is 0.037 G,

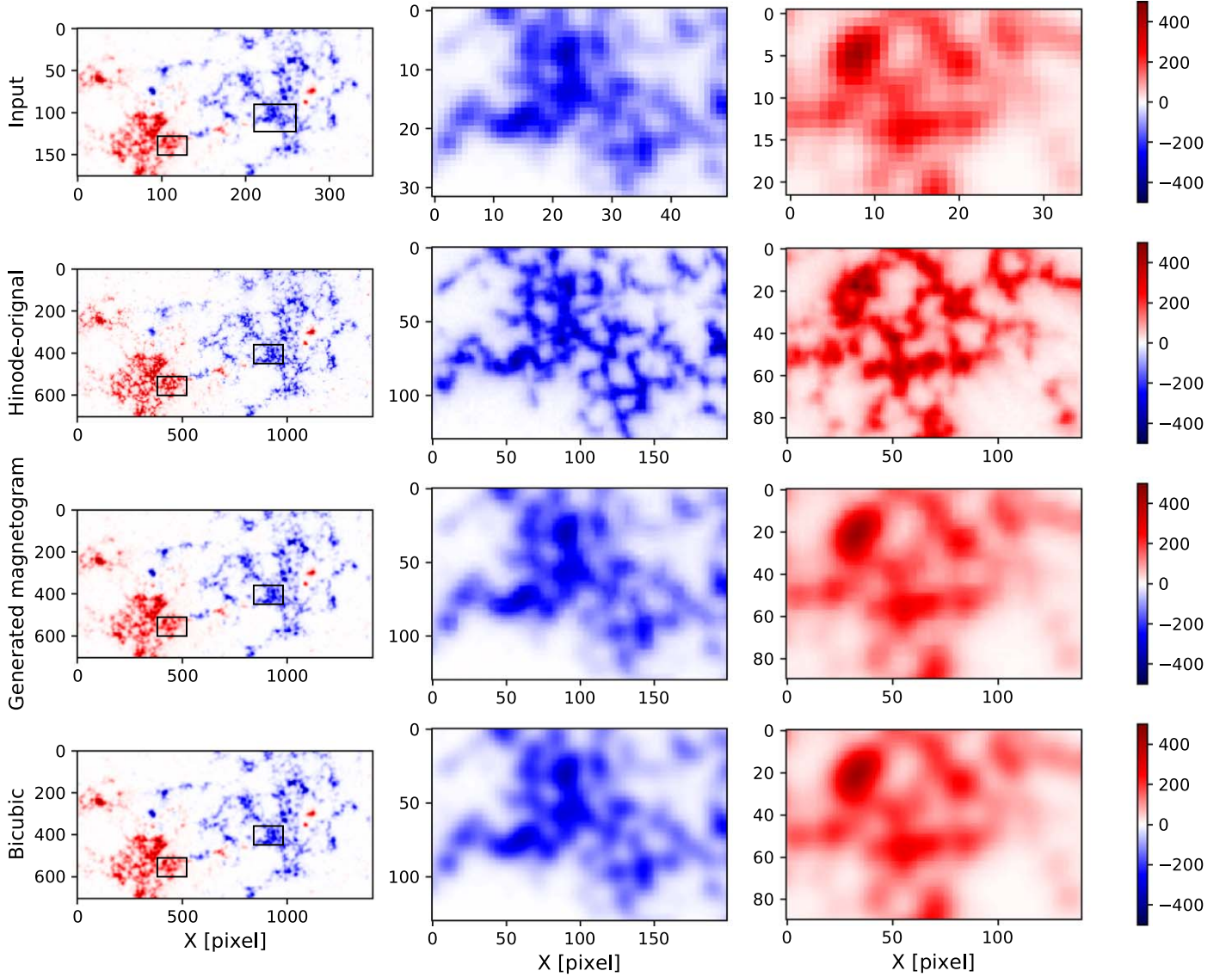


Figure 4. Comparison of the SR magnetogram generated by the residual attention model with the corresponding Hinode/SOT NFI magnetogram observed on 2013 January 7. From top to the bottom: SDO/HMI magnetogram as input, Hinode/SOT NFI magnetogram, SR magnetogram by the residual attention model, and one by the bicubic method.

which is very small and much better than those of the other methods. The average SSIM indices of three models show that the magnetograms generated by the deep learning models (0.98) are more similar to the target ones than those of the other method.

The first row of Figure 3 shows 2D histograms of target and SR magnetograms generated by the residual attention model, the progressive GAN model, and the bicubic method for the testing data set. The histograms show that the deep learning models can preserve the pixel distributions much better than the bicubic methods. Error bar plots in the second row of Figure 3 show that the standard deviations of the deep learning models are much lower than those of the bicubic method, and the mean pixel values are better correlated with the target ones. The standard deviation indicates the dispersion of generated magnetogram pixel values from the mean pixel value for a given data range. The scatter plot of the deep learning models show a linear relationship between the target and generated magnetograms. The error bar plots show slightly non-linear

features for smaller values, especially for the bicubic method. Because these patterns mostly occur at magnetic fields weaker than about 30 Gauss, we think that SR by deep learning models cannot significantly change scientific results.

Figure 4 presents a comparison between the Hinode/SOT NFI magnetogram observed on 2013 January 7 and the SR magnetogram generated by the residual attention model. Visually, the generated magnetogram cannot produce small-scale structures like NFI. The second and third columns show that the features of the generated magnetogram have sharper edges of magnetic flux and show a close resemblance to the target magnetogram in both the positive and the negative magnetic regions. The bicubic one is much more blurred than the generated magnetogram and the magnetic flux is consistently well distributed in the generated one. Other metrics of the residual attention model are as follows: CC = 0.94, peak S/N = 39.11 dB, and SSIM = 0.93. These values are all better than those of the bicubic method that gives: CC = 0.93, peak S/N = 38.92 dB, and SSIM = 0.91.

5. Conclusion and Summary

In this Letter, we apply two novel deep learning models for SR to solar magnetograms. For this work, we consider the residual attention and the progressive GAN model. We select 10,785 pairs of degraded LR SDO/HMI magnetograms as input data and corresponding HR magnetograms as target data. We trained our models using 9000 pairs from 2014 May to August, validated our models using 985 pairs from 2015 October to December, and tested using 800 pairs from 2015 January to March.

The major results of this study are as follows. First, all considered models can successfully generate HR solar magnetograms from LR magnetograms. The generated magnetograms are consistent with the target magnetograms. Moreover, we observe that perceptually the deep learning models provide more small-scale and texture-wise information than the bicubic method. From the correlation coefficient values, we show that the residual attention model (0.93) results are highly correlated with the target magnetograms. The higher peak S/N and SSIM values represent that the residual attention model can reconstruct the magnetograms quite efficiently. Furthermore, the metric values (CC, peak S/N, RMSE, and SSIM) of the residual attention model are much better than the other models. In addition, we compare the generated magnetogram with the Hinode/SOT NFI magnetogram and the bicubic method. The generated magnetogram is well correlated with the Hinode one with higher peak S/N and SSIM values. Finally, 2D histograms, error bar plots, and scatter plots show better pixel distribution, mean pixel values, and linear relationship in the deep learning models than the other interpolation methods.

SR is an ill-defined problem for solar magnetograms (Yang et al. 2019). Our study shows that deep learning models can extract additional information and features for generating a higher resolution magnetogram. The generated magnetogram will allow us to study small-scale details, magnetic features, sunspot structure, and other physical features of the Sun. We have also tried a larger scale SR like 8×8 binning; however, the results are not satisfactory. In the future, we look forward to

overcoming this problem using other available techniques. More feasible loss function, optimization method, and model structure may improve the results. In the future, we have a plan to train our model using other types of data such as HR solar continuum images showing detailed granule structures from the Goode Solar Telescope (Cao et al. 2010) and the Daniel K. Inouye Solar Telescope (Tritschler et al. 2016). This methodology can be applied to not only solar images, but also to other astronomical images as well.

We appreciate the referee’s constructive comments. We thank the numerous team members who have contributed to the success of the SDO and Hinode mission. This work was supported by the BK21 plus program through the National Research Foundation (NRF) funded by the Ministry of Education of Korea, the Basic Science Research Program through the NRF funded by the Ministry of Education (NRF-2016R1A2B4013131, NRF-2019R1A2C1002634, NRF-2020R1C1C1003892), the Korea Astronomy and Space Science Institute (KASI) under the R&D program “Development of a Solar Coronagraph on International Space Station” (project No. 2019-1-850-02)” supervised by the Ministry of Science and ICT, and the Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (2018-0-01422, Study on analysis and prediction technique of solar flares). We acknowledge the community effort devoted to the development of the following open-source packages that were used in this work: SunPy (sunpy.org), pytorch (pytorch.org), TensorFlow (tensorflow.org), and SolarSoft.

Appendix

Figure A1 shows 2D histogram plots and scatter plots of Figure 4. As shown in Figure A1, the residual attention model and bicubic interpolation method are similar to one another. Visually and the matrices results of the residual attention model are little better than those of the bicubic method.

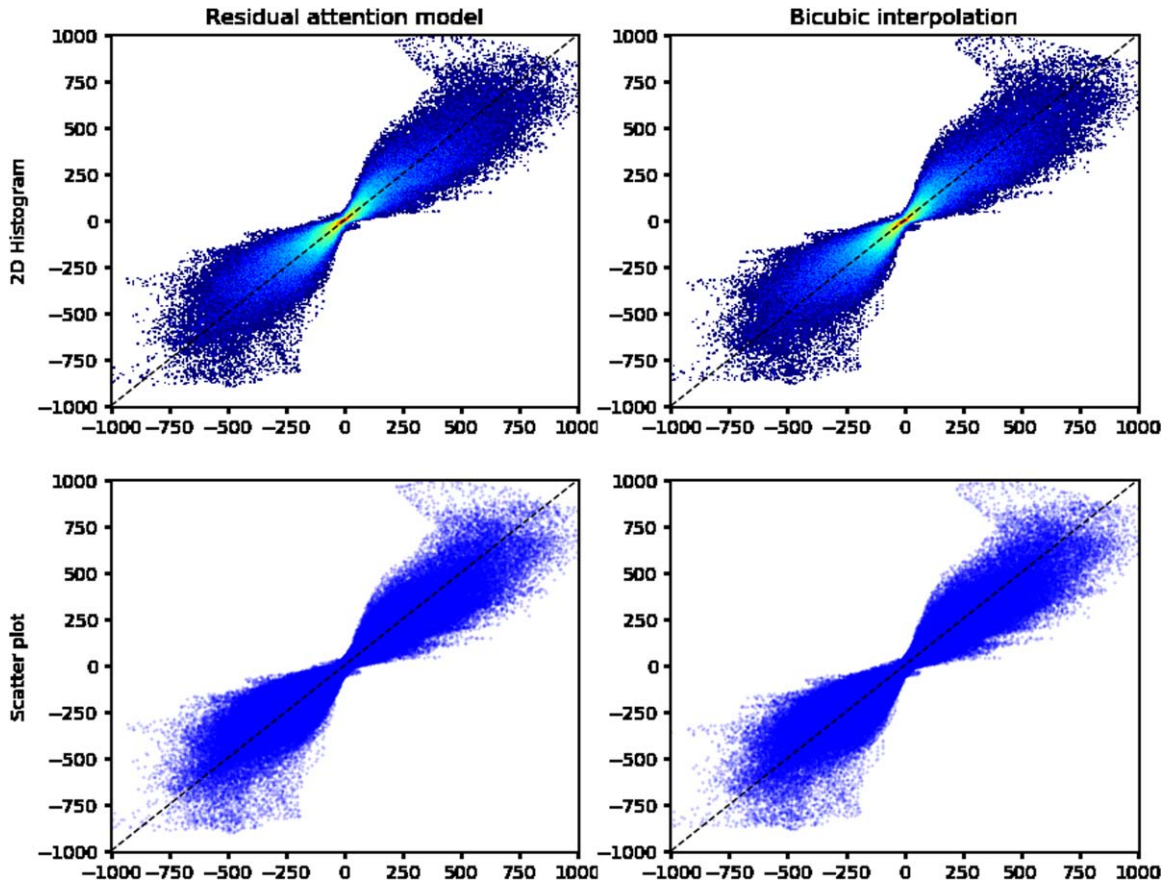


Figure A1. The top row shows 2D histograms and the bottom row shows scatter plot of Hinode/SOT NFI magnetogram and generated magnetogram. From left to right, each column: the residual attention model and the bicubic ones.

ORCID iDs

Yong-Jae Moon <https://orcid.org/0000-0001-6216-6944>
 Eunsu Park <https://orcid.org/0000-0003-0969-286X>
 Ashraf Siddique <https://orcid.org/0000-0003-2186-5735>
 Il-Hyun Cho <https://orcid.org/0000-0001-7514-8171>
 Daye Lim <https://orcid.org/0000-0001-9914-9080>

References

- Baso, C. J., & Díaz Ramos, A. A. 2018, *A&A*, **614**, A5
 Cao, W., Gorceix, N., Coulter, R., et al. 2010, *AN*, **331**, 636
 Domingo, V., Fleck, B., & Poland, A. I. 1995, *SoPh*, **162**, 1
 Dong, C., Loy, C. c., He, K., et al. 2016, *ITPAM*, **38**, 295
 Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv:1406.2661
 Guo, R., Shi, X., & Wang, Z. 2019, *JEL*, **28**, 023032
 Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
 Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, *SoPh*, **243**, 3
 Lai, W.-S., Huang, J.-B., Ahuja, N., et al. 2017, in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (New York: IEEE), 624
 Lecun, Y., Bengio, Y., & Hinton, G. 2015, *Natur*, **521**, 436
 Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, *IEEEP*, **86**, 2278
 Ledig, C., Theis, L., Huszár, F., et al. 2017, in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (New York: IEEE), 4681
 Li, Z., Peng, Q., Bhanu, B., et al. 2018, *Ap&SS*, **363**, 92
 Lim, B., Son, S., Kim, H., et al. 2017, in Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops (New York: IEEE), 136
 Misra, D., Mishra, S., & Appasani, B. 2018, arXiv:1812.09702
 Moon, Y.-J., Kim, Y.-H., Park, Y.-D., et al. 2007, *PASJ*, **59**, S625
 Park, S. C., Park, M. k., & Kang, M. G. 2003, *ISPM*, **20**, 21
 Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, *SoPh*, **275**, 3
 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
 Simonyan, K., & Zisserman, A. 2014, arXiv:1409.1556
 Tritschler, A., Rimmele, T., Berukoff, S., et al. 2016, *AN*, **337**, 1064
 Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. 2008, *SoPh*, **249**, 167
 Wang, Y., Perazzi, F., McWilliams, B., et al. 2018, arXiv:1804.02900
 Yan, L., Liao, W., Chang, Y., et al. 2017, in IEEE International Geoscience and Remote Sensing Symp. (IGARSS) (New York: IEEE), 3425
 Yang, W., Zhang, X., Tian, Y., et al. 2019, *IEEE Trans. Multimedia*, **21**, 3106
 Zhang, Y., Li, K., Li, K., et al. 2018, in Proc. European Conference on Computer Vision (ECCV), 286