Single-Frame Super-Resolution of Solar Magnetograms: Investigating Physics-Based Metrics & Losses

Anna Jungbluth
University of Oxford
anna.jungbluth@physics.ox.ac.uk

Xavier Gitiaux
George Mason University
xgitiaux@gmu.edu

Shane A. Maloney
Trinity College Dublin /
Dublin Institute for Advanced Studies
shane.maloney@tcd.ie

Carl Shneider
Center for Mathematics
and Computer Science (CWI)
carl.shneider@cwi.nl

Paul J. Wright
Stanford University
pjwright@stanford.edu

Alfredo Kalaitzis
Element AI
freddie@element.ai

Michel Deudon
Element AI
michel.deudon@elementai.com

Atılım Güneş Baydin
University of Oxford
gunes@robots.ox.ac.uk

Yarin Gal
University of Oxford
yarin@cs.ox.ac.uk

Andrés Muñoz-Jaramillo
Southwest Research Institute
amunozj@boulder.swri.edu

Abstract

Breakthroughs in our understanding of physical phenomena have traditionally followed improvements in instrumentation. Studies of the magnetic field of the Sun, and its influence on the solar dynamo and space weather events, have benefited from improvements in resolution and measurement frequency of new instruments. However, in order to fully understand the solar cycle, high-quality data across time-scales longer than the typical lifespan of a solar instrument are required. At the moment, discrepancies between measurement surveys prevent the combined use of all available data. In this work, we show that machine learning can help bridge the gap between measurement surveys by learning to **super-resolve** low-resolution magnetic field images and **translate** between characteristics of contemporary instruments in orbit. We also introduce the notion of physics-based metrics and losses for super-resolution to preserve underlying physics and constrain the solution space of possible super-resolution outputs.

1   Introduction

For 50 years, a series of instruments have provided images of the Sun's magnetic field (known as *magnetograms*) to study the field's origin and evolution. However, differences in resolution, noise properties, saturation levels, spectral inversion techniques, and other instrument specifics introduce inhomogeneities that prevent direct comparison of magnetograms across instruments. The calibration and homogenization of magnetograms remains an unsolved problem today.

Traditionally, magnetogram calibration has been performed through pixel-to-pixel comparison [1], histogram equalization [2], or harmonic scaling [3]. Recent advances in deep learning highlight the
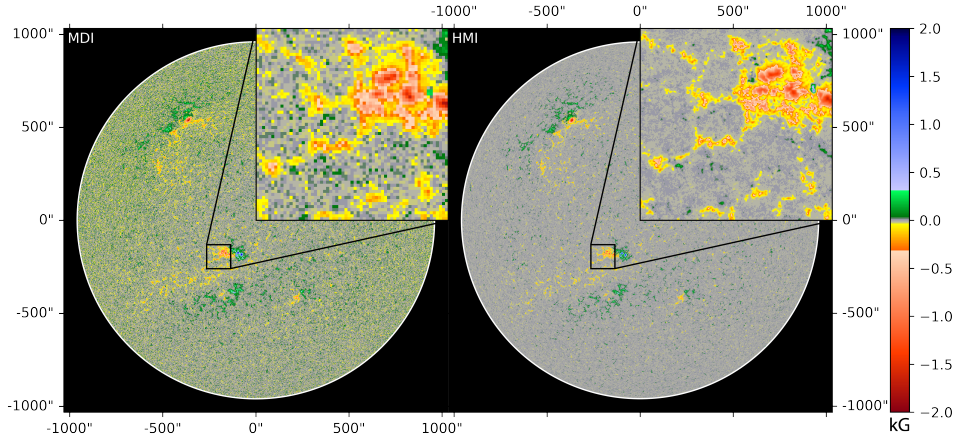
Figure 1: Comparison of co-temporal, co-aligned magnetograms obtained by MDI (*left*), and HMI (*right*). Both images show the full solar disk, re-scaled as if they were observed from 1 Astronomical Unit, and plotted over the range of ± 2000 Gauss. The $128'' \times 128''$ insets were registered using vertical and horizontal shifts to account for the distortion between instruments.

potential to improve and simplify these calibration processes to build a consistent, high-resolution dataset of magnetograms. Deep neural networks can transform and super-resolve natural images [4], and have recently been successfully applied to magnetograms [5, 6, 7, 8].

In this work, we use deep neural networks to super-resolve and cross-calibrate magnetograms taken by the Michelson Doppler Imager (MDI) [9, 10] and the Helioseimic and Magnetic Imager (HMI) [11, 12, 13]. During its tenure from 1995 – 2011, MDI acquired magnetograms every 96 minutes, with a detector of 1024×1024 pixels imaging the Sun at $2''$ (arcseconds) per pixel. In comparison, HMI, which has been operational since 2010, images the Sun every 45 seconds at $0.5''$ per pixel with a detector of 4096×4096 pixels.

MDI and HMI observations overlap from 2010 to 2011, resulting in nearly 9000 co-temporal observations of the same physical structures. By training a deep neural network on a subset of observations, MDI magnetograms are super-resolved and upsampled by a factor of four to the size of HMI observations. We show that penalty terms based on image gradients in the loss function better preserve the physical properties of magnetograms in the super-resolved reconstructions.

2   Data processing

While neural networks can learn systematic differences between instruments, random positional, rotational, and temporal differences between images significantly degrade the model's ability to learn the mapping and have to be removed before training. We therefore apply a series of operations to the MDI and HMI datasets: (1) rotate solar north to the image north to standardize the Sun's orientation in each image, using instrument pointing information contained in the magnetogram headers; (2) scale the image to standardize the detector resolution and the size of the Sun in the detector as if it was observed from 1 Astronomical Unit (AU; the average Sun-Earth distance); (3) register and shift images taken at different times to more closely align features. The effect of parallax, which is less significant than the effect of temporal discrepancies between instruments, is also removed during the above processing steps. Fig 1 shows rotated, distance-corrected, and temporally aligned full-disk magnetograms. The resulting 9000 image pairs are split by randomly selecting one month for validation and testing, and allocating the remaining ten months to training. A monthly split was chosen over a random allocation of images to preserve some of the temporal evolution of the magnetic field during training. In addition, a random allocation of images can lead to many instances where the model is trained and tested on magnetograms spaced only minutes apart. During that time, neither the instrument's perspective nor the magnetic field features have changed enough to avoid a biased evaluation of the model's performance. The full disk magnetograms are broken into 1024 $128'' \times 128''$ patches that are augmented with N-S, E-W reflections and polarity flips during training.

3 Super-resolution

Single-image super-resolution (SISR) approximates the inverse of the blur and down-sampling operators that exist between a high-resolution (HR) image and its low-resolution (LR) counterpart [4]. Super-resolution is an ill-posed problem, as multiple SR outputs can explain a LR input. Compared to classical upsampling methods, such as linear or bicubic upsampling, that do not add new information, learning-based approaches can constrain the non-trivial solution space by exploiting the spatial regularities within a specific distribution of images—in this work, that of magnetograms.

Convolutional Neural Networks (CNN) are the current state-of-the-art in capturing spatial regularities in complex natural images. With a sufficient amount of annotated examples, a CNN can learn salient LR-to-HR correlations that best describe the data [14, 15]. CNNs have successfully been applied to super-resolution tasks in remote sensing, medical imaging and microscopy [16, 17, 18, 19]. Typical loss functions for super-resolution tasks—like either the l_1 or the mean squared error (MSE, l_2)—encode the pixel-level error between the SR reconstruction and the HR ground-truth [20].

In this work, we use HighRes-net [21][1], a state-of-the-art architecture for multi-frame super-resolution (MFSR) that recursively fuses an arbitrary number of LR images of the same view. While our current work focuses on single-frame input images, future work will expand to MFSR to leverage the full capability of the model. Future work will also focus on comparing the performance of HighRes-net to competitive generative models like *Generative Adversarial Networks* (GANs), and *Variational Auto Encoders* (VAEs). For now, the model's output is compared to a bicubic upsampling baseline. Since bicubic upsampling is a trivial operation, it serves as an ideal "no-effort" baseline to compare the model's output to.

4 Loss functions & performance metrics

Any application of super-resolution within the physical sciences would be amiss if during training, the model had no incentive to preserve the underlying physics of the task at hand[2]. The need for physical accuracy allows us to further constrain the SR solution space. While typical *reconstruction*-based losses (e.g. MSE) suffice to obtain visually convincing SR reconstructions, quantitative methods are required to evaluate the scientific quality of the model's output. This motivates our investigation of (1) differentiable loss functions for physics-driven training, and (2) post-training performance metrics to assess the quality of the SR outputs.

4.1 Loss functions

Compared to natural images, the range of pixel values in magnetograms is not constrained, even though most pixels show close to zero magnetic field. Active regions, with large magnetic fields, occur outside the disk center and are orders of magnitude more rare. This is especially problematic for MDI to HMI conversions, as the year of overlap (2010-2011) occurred at a minimum of the solar cycle with few active regions.

As a consequence of this uneven representation, the model likely learns to preserve small fields well but struggles to reconstruct active regions. While a simple mean squared error can reconstruct broad features in magnetograms, it fails to reconstruct strong magnetic fields (Fig 4). This is also due to the MSE sensitivity to large errors, which leads to compensating by over-smoothing the output. This over-smoothing can be offset by using an l_1 or *Huber loss* [22] instead of an l_2 loss. Future work will focus on expanding our investigations to alternative loss functions.

To improve the reconstruction of large pixel values, our current work focuses on investigating loss functions composed of an MSE loss with additional penalty terms. So far, we have mostly focused on an image gradient-based penalty term, but are investigating adding the *Kullback-Leibler divergence* [23].

Image gradient-based loss With a Sobel kernel [24] we approximate the x and y gradients of the SR output (I_{SR}) and the HR ground truth (I_{HR}). By computing the mean squared difference between

---

[1] `https://github.com/ElementAI/HighRes-net`.
[2] For instance, zero divergence of the magnetic field needs to be preserved.
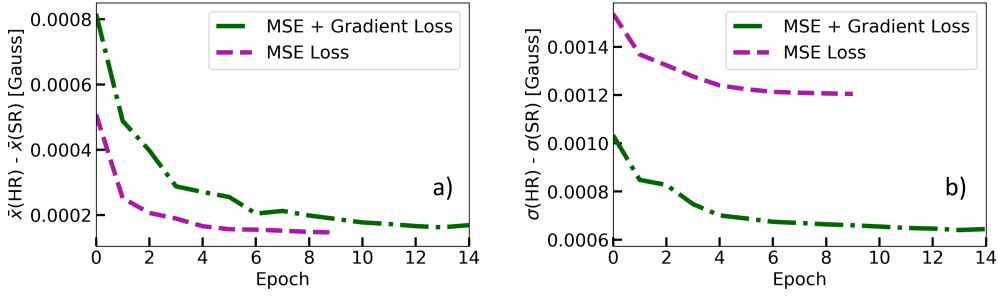
Figure 2: Difference between the a) mean and b) standard deviation of the pixel values of the HR HMI target and SR output as a function of training epochs.

the pixel gradients (G) of the images, we incentivize the model to preserve both magnetic field values and spatial variations:

$$Loss = MSE(I_{SR}, I_{HR}) + n \cdot MSE(G_{SR}, G_{HR}) \tag{1}$$

In this case, the weighting n of the gradient term is chosen to be 100 to allow for equal magnitude contributions of both terms to the overall loss function.

In computer vision, image gradients are often used for edge detection and texture matching [25]. For magnetograms, edge detection can help define sharper edges around non-zero magnetic field regions and texture matching can aid in recovering small-scale details that make up a high-resolution image.

4.2 Metrics

Apart from tracking the mean and standard deviation of magnetic field values (Fig 2), we evaluate the network's performance using three main metrics: (1) correlation plots between the target and output pixel values (Fig 3); (2) distribution histograms of pixel values to assess the recovery of small and large magnetic fields (Fig 5-a); and (3) the *information entropy* of an image at multiple scales of spatial binning (Fig 5-b). The information entropy of a random variable X is defined as $H[X] = -\sum_x p(x) \ln p(x)$. By comparing the information entropy of the output and target image, we assess whether the reconstruction contains less detail (over-smoothing) or more detail (hallucination) relative to the ground-truth complexity. At each scale of 2^k pixels, for $k \in \{0, 1, 2, ...\}$, we convolve the image with an averaging kernel of size $2^k \times 2^k$ and compute its entropy.
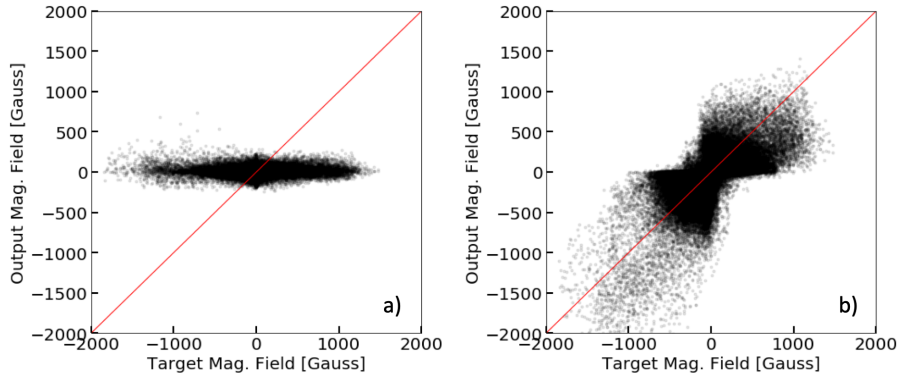

Figure 3: Correlation plots of pixel values of the HR HMI target and SR output, using a) an MSE loss and b) an MSE + gradient-based loss function.
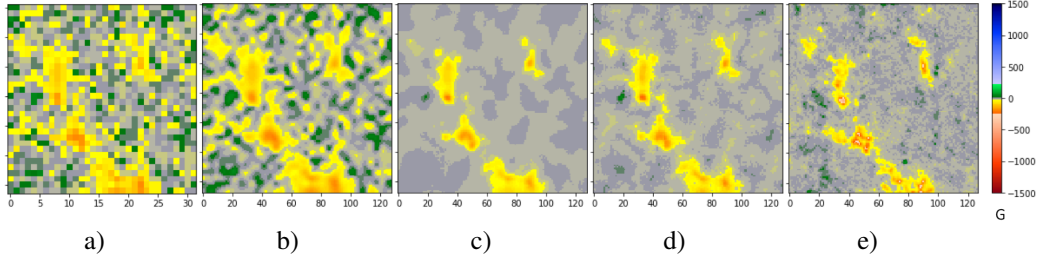
4

Figure 4: Comparison of magnetogram patches using different loss functions. Left to right: a) LR MDI input, b) bicubic upsampling, c) MSE loss, d) MSE + gradient-based loss, e) HR HMI target.

5 Results

The results of our model, trained with either an MSE or an MSE + image gradient-based loss, are compared to a simple bicubic upsampling of the input magnetograms. Fig 4 shows the same patch of a magnetogram taken on 2011-04-01 by both MDI or HMI, compared to the bicubic baseline and the model's outputs, and Fig 6 shows the full disk reconstruction. Fig 5-a shows the distribution histograms of pixel values and Fig 5-b the multi-scale entropy for all patches of the full disk image. While bicubic upsampling is able to preserve the pixel distributions well, the quality of the output image is poor (Fig 4-b). In addition, the multi-scale entropy of the bicubic upsample exceeds that of the target. This may indicate that low-resolution MDI magnetograms contain noise that is not in the high-resolution HMI target. Bicubic usampling extrapolates noise to the high-resolution output. This highlights that bicubic upsampling is not good at instrument conversion because it cannot account for detector specificities. Using HighRes-net with an MSE loss does well at removing MDI's noise floor (Fig. 4-c), but fails at reconstructing strong magnetic fields (Fig 5-a). The correlation plot in Fig 3-a confirms the strong clamping of magnetic field values between -500 and 500 Gauss. In comparison, the composite of the MSE and the image gradient-based loss leads to a substantial improvement in the model's performance. Not only are the pixel distributions of the output and target almost identical, the multi-scale entropies are also more closely matched. Visually, the features in Fig 4-d have sharper edges and resemble those of the target image (Fig 4-e) more closely. Fig 4-d also shows finer small-scale features, compared to the large, blocky patches of low-magnetic field regions in Fig 4-c. Since image gradients are often used for improved edge detection and texture matching, a visual improvement of the magnetograms does not necessarily come as a surprise. Crucially, visuals are not enough to provide evidence of the scientific value of super-resolution for magnetograms. Taking image gradients into account during training improves the model's performance on all presented metrics. Compared to natural images, the image gradients of a magnetogram do not only hold information on texture and feature shapes, but are a direct measure of the magnetic field gradients on the surface of the Sun. Consequently, incentivizing the model to preserve both the magnitude and the gradient of the magnetic field leads to qualitatively and quantitatively promising results for the application of super-resolution to magnetograms.
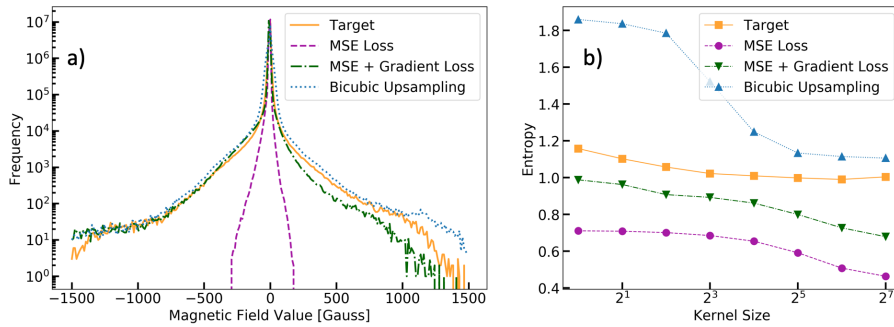

Figure 5: a) Pixel histograms and b) multi-scale entropy of the HR HMI target, bicubic upsampling, and the SR output using an MSE, or MSE + gradient-based loss function.
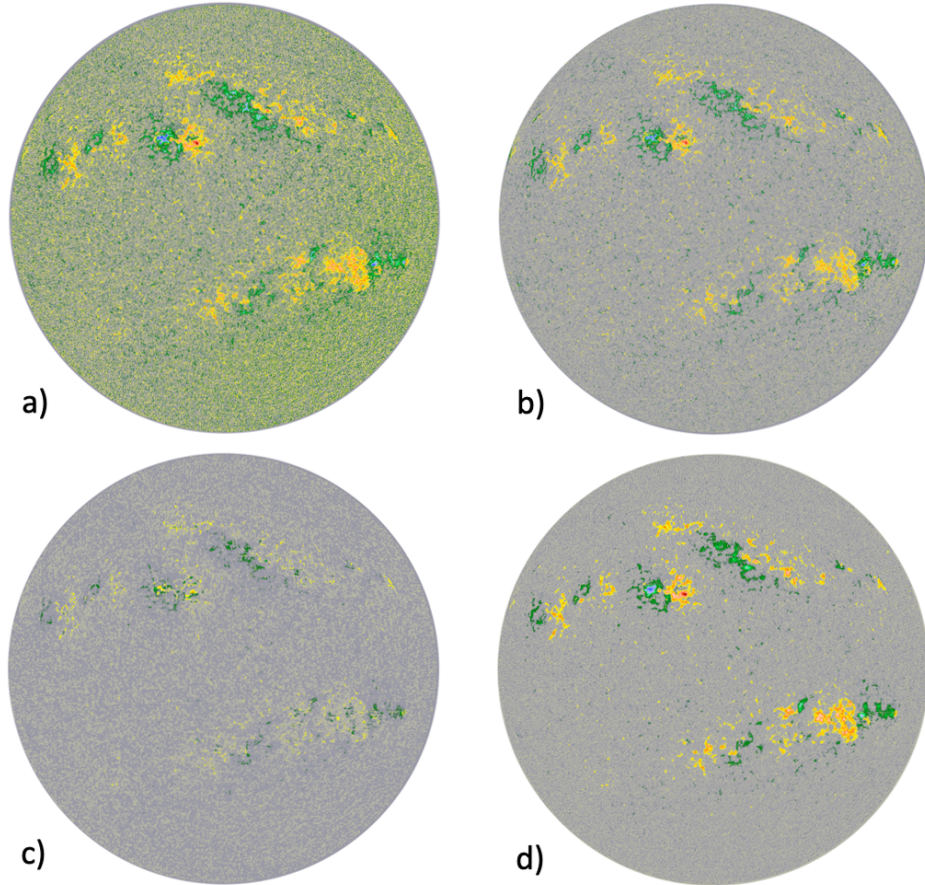
5

Figure 6: Reconstruction of full disk magnetograms; a) LR MDI input, b) HR HMI target, c) SR output using an MSE loss, d) SR output using an MSE + gradient-based loss.

6 Conclusions and future work

We demonstrate the combination of a reconstruction (MSE) and an image gradient-based loss as a way of super-resolving magnetic field images taken by MDI to the resolution of HMI. By adding an image gradient-based penalty term to the loss function, we are able to visually improve the model's output, as well as improve the reconstruction of large magnetic field regions. These results show the potential of physics-based metrics as regularization terms to constrain ill-defined super-resolution tasks. Moving forward, we focus on expanding this work to multi-frame super-resolution of all 50 year's worth of magnetogram instruments, paving the way to generate the uniform, high-resolution dataset of magnetograms needed for space weather studies.

References

[1] Y. Liu, J. T. Hoeksema, P. H. Scherrer, J. Schou, S. Couvidat, R. I. Bush, T. L. Duvall, K. Hayashi, X. Sun, and X. Zhao. Comparison of Line-of-Sight Magnetograms Taken by the Solar Dynamics Observatory/Helioseismic and Magnetic Imager and Solar and Heliospheric Observatory/Michelson Doppler Imager. *Solar Physics*, 279(1):295–316, Jul 2012.

[2] P. Riley, M. Ben-Nun, J. A. Linker, Z. Mikic, L. Svalgaard, J. Harvey, L. Bertello, T. Hoeksema, Y. Liu, and R. Ulrich. A Multi-Observatory Inter-Comparison of Line-of-Sight Synoptic Solar Magnetograms. *Solar Physics*, 289(3):769–792, Mar 2014.

[3] Ilpo Virtanen and Kalevi Mursula. Photospheric and coronal magnetic fields in six magnetographs. III. Photospheric and coronal magnetic fields in 1974-2017. *Astronomy and Astrophysics*, 626:A67, Jun 2019.

[4] W. Yang, X. Zhang, Y. Tian, W. Wang, J. Xue, and Q. Liao. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, pages 1–1, 2019.

[5] Sean McGregor, Dattaraj Dhuri, Anamaria Berea, and Andrés Muñoz-Jaramillo. FlareNet: A Deep Learning Framework for Solar Phenomena Prediction. In *Neural Information Processing Systems (NIPS) 2017 workshop on Deep Learning for Physical Sciences (DLPS), Long Beach, CA, US, December 8, 2017*, 2017.

[6] A. Asensio Ramos, J. de la Cruz Rodríguez, and A. Pastor Yabar. Real-time, multiframe, blind deconvolution of solar images. *Astronomy & Astrophysics*, 620:A73, Dec 2018.

[7] C. J. Díaz Baso and A. Asensio Ramos. Enhancing SDO/HMI images using deep learning. *Astronomy & Astrophysics*, 614:A5, Jun 2018.

[8] Taeyoung Kim, Eunsu Park, Harim Lee, Yong-Jae Moon, Sung-Ho Bae, Daye Lim, Soojeong Jang, Lokwon Kim, Il-Hyun Cho, Myungjin Choi, and Kyung-Suk Cho. Solar farside magnetograms from deep learning analysis of STEREO/EUVI data. *Nature Astronomy*, 3:397–400, Mar 2019.

[9] P. H. Scherrer, R. S. Bogart, R. I. Bush, J. T. Hoeksema, A. G. Kosovichev, J. Schou, W. Rosenberg, L. Springer, T. D. Tarbell, A. Title, C. J. Wolfson, I. Zayer, and MDI Engineering Team. The Solar Oscillations Investigation - Michelson Doppler Imager. *Solar Physics*, 162(1-2):129–188, Dec 1995.

[10] V. Domingo, B. Fleck, and A. I. Poland. The SOHO Mission: an Overview. *Solar Physics*, 162(1-2):1–37, Dec 1995.

[11] P. H. Scherrer, J. Schou, R. I. Bush, A. G. Kosovichev, R. S. Bogart, J. T. Hoeksema, Y. Liu, T. L. Duvall, J. Zhao, A. M. Title, C. J. Schrijver, T. D. Tarbell, and S. Tomczyk. The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO). *Solar Physics*, 275(1-2):207–227, Jan 2012.

[12] J. Schou, P. H. Scherrer, R. I. Bush, R. Wachter, S. Couvidat, M. C. Rabello-Soares, R. S. Bogart, J. T. Hoeksema, Y. Liu, T. L. Duvall, D. J. Akin, B. A. Allard, J. W. Miles, R. Rairden, R. A. Shine, T. D. Tarbell, A. M. Title, C. J. Wolfson, D. F. Elmore, A. A. Norton, and S. Tomczyk. Design and Ground Calibration of the Helioseismic and Magnetic Imager (HMI) Instrument on the Solar Dynamics Observatory (SDO). *Solar Physics*, 275(1-2):229–259, Jan 2012.

[13] W. Dean Pesnell, B. J. Thompson, and P. C. Chamberlin. The Solar Dynamics Observatory (SDO). *Solar Physics*, 275(1-2):3–15, Jan 2012.

[14] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *European conference on computer vision*, pages 184–199. Springer, 2014.

[15] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[16] Marcus Märtens, Dario Izzo, Andrej Krzic, and Daniël Cox. Super-resolution of proba-v images using convolutional neural networks. *Astrodynamics*, Aug 2019.

[17] Yu Tao and Jan-Peter Muller. Super-resolution restoration of misr images using the ucl magigan system. *Remote Sensing*, 11(1):52, 2019.

[18] Chi-Hieu Pham, Aurélien Ducournau, Ronan Fablet, and François Rousseau. Brain mri super-resolution using deep 3d convolutional networks. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 197–200. IEEE, 2017.

[19] Yair Rivenson, Zoltán Göröcs, Harun Günaydin, Yibo Zhang, Hongda Wang, and Aydogan Ozcan. Deep learning microscopy. *Optica*, 4(11):1437–1443, 2017.

[20] Sina Farsiu, M Dirk Robinson, Michael Elad, and Peyman Milanfar. Fast and robust multiframe super resolution. *IEEE transactions on image processing*, 13(10):1327–1344, 2004.

[21] Anonymous. Highres-net: Multi-frame super-resolution by recursive fusion. In *Submitted to International Conference on Learning Representations*, 2020. under review.

[22] Peter J. Huber. Robust estimation of a location parameter. *Ann. Math. Statist.*, 35(1):73–101, 03 1964.

[23] David JC MacKay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[24] Irwin Sobel. An isotropic 3x3 image gradient operator. *Presentation at Stanford A.I. Project*, 1968.

[25] David A Forsyth and Jean Ponce. *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference, 2002.

[26] The SunPy Community, Stuart J. Mumford, Steven Christe, David Pérez-Suárez, Jack Ireland, Albert Y. Shih, Andrew R. Inglis, Simon Liedtke, Russell J. Hewett, Florian Mayer, Keith Hughitt, Nabil Freij, Tomas Meszaros, Samuel M. Bennett, Michael Malocha, John Evans, Ankit Agrawal, Andrew J. Leonard, Thomas P. Robitaille, Benjamin Mampaey, Jose Iván Campos-Rozo, and Michael S. Kirk. SunPy—Python for solar physics. *Computational Science and Discovery*, 8(1):014009, Jan 2015.

[27] Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science and Engineering*, 13(2):22–30, Mar 2011.

[28] Wes McKinney. Data structures for statistical computing in python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.

[29] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in PyTorch. In *NIPS Autodiff Workshop*, 2017.