

---

# Probabilistic Super-Resolution of Solar Magnetograms: Generating Many Explanations and Measuring Uncertainties

---

**Xavier Gitiaux**  
George Mason University  
xgitiaux@gmu.edu

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

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

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

**Paul J. Wright**  
Stanford University  
pjwright@stanford.edu

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

**Michel Deudon**  
Element AI  
michel.deudon@elementai.com

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

**Alfredo Kalaitzis**  
Element AI  
freddie@element.ai

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

## Abstract

Machine learning techniques have been successfully applied to super-resolution tasks on natural images where visually pleasing results are sufficient. However in many scientific domains this is not adequate and estimations of errors and uncertainties are crucial. To address this issue we propose a Bayesian framework that decomposes uncertainties into epistemic and aleatoric uncertainties. We test the validity of our approach by super-resolving images of the Sun’s magnetic field and by generating maps measuring the range of possible high resolution explanations compatible with a given low resolution magnetogram.

## 1 Introduction

Deep learning has been successful at super-resolution (SR) of natural images, i.e. reconstructing higher resolution (HR) images from low resolution (LR) inputs [1]. However, to the extent of our knowledge, there is no existing work measuring uncertainty in super-resolution tasks. For scientific applications, estimating the uncertainty of a SR output is as important as the prediction itself, all the more as super-resolution is an ill-posed problem with many super-resolved images being consistent with the same low-resolution input.

To obtain robust uncertainties, we propose a Bayesian framework as in Kendall & Gal [2] that decomposes uncertainty into *epistemic* and *aleatoric* uncertainty. *Epistemic* uncertainty relates to our ignorance of the true data generating process, and *aleatoric* uncertainty captures the inherent noise in the data. We apply our framework to images of the Sun’s derived magnetic field (magnetograms), which are used to study the solar corona [3] and to predict space-weather events [4]. These applications often require magnetograms spanning time-ranges longer than the lifetime of any single instrument. SR can compensate for inhomogeneities and discontinuities between instruments by converting multiple surveys to a single common resolution (e.g. [5]).

In practice, we convert a state of the art super-resolution encoder-decoder architecture, HighRes-net [6]<sup>1</sup>, into a Bayesian deep learning framework by adding dropout at each convolutional layer, and tracking both the mean and variance of magnetic field values. We test the effectiveness of our framework by super-resolving magnetograms from the Helioseismic and Magnetic Imager [HMI; 7].

We demonstrate that modelling *epistemic* uncertainty allows us to measure the range of high resolution explanations consistent with a low-resolution input. Figure 1 illustrates how our super-resolution architecture can generate two different extrapolations of the same input and how the difference is more attenuated once downsampled<sup>2</sup>. Moreover, high resolution explanations have larger variance in regions of the Sun with a large magnetic field (so called active regions). We show that this larger variance cannot be properly accounted for unless disentangled from the larger amount of noise in regions with larger magnetic fields. In fact, *aleatoric* uncertainty is a full order of magnitude larger in active regions.

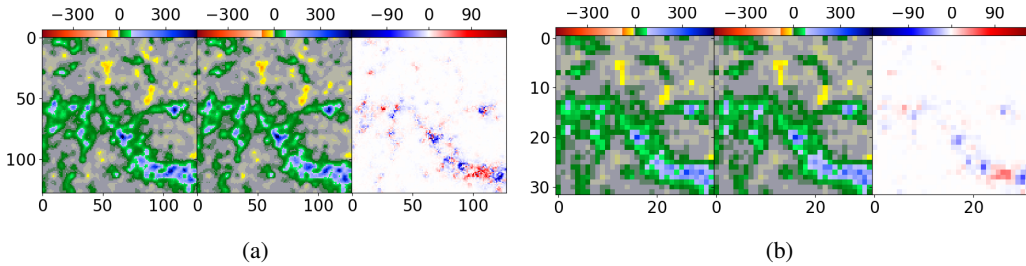


Figure 1: Comparison of two realisations from HighRes-net, trained with dropout and heteroskedastic loss (see section 2), obtained from the same input using dropout during inference. HR models outputs and differences (a) and HR outputs degraded to match LR input (b). The differences are clearly visible in the HR (a) but reduced in the LR, which illustrates how many model’s explanations might be consistent with the same low-resolution input.

## 2 Quantifying Uncertainties in Super-resolution

If an image  $y$  is downsampled by an unknown transformation into a LR image  $x$ , SR consists of learning the inverse transformation  $f$  from  $x$  to  $y$ . This is an ill-posed problem as many SR outputs map to the same LR input. Despite this difficulty, deep learning architectures can achieve SR, obtaining  $y$  by parameterizing  $f$  with an encoder-decoder neural network  $f_w(x)$ . To model *epistemic* uncertainty we use a Bayesian framework from Kendall & Gal [2] and impose a prior distribution on the model parameters  $w$ . Given a sample  $\mathcal{D}_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$  of magnetograms, we evaluate the posterior distribution  $p(w|\mathcal{D}_n)$  by Bayesian inference. In practice, this inference is approximated by dropout variational inference [8], i.e. adding dropout to each layer during training and inference to sample a series of SR realizations.

*Aleatoric* uncertainty is modelled by assigning a distribution to the outputs of the model  $p(y|f_w(x))$ . We assume that for each pixel  $j$ , the noise is Gaussian with a variance  $\sigma_j(x)^2$  that depends on the input image and pixel position  $j$  in the detector. We chose to model this heteroskedascity, i.e. the variation of *aleatoric* uncertainty across pixels, because projection distortions are a function of pixel

<sup>1</sup><https://github.com/ElementAI/HighRes-net>.

<sup>2</sup>Downsampling consists in two operations: (i) smoothing that convolves the high resolution magnetogram with a Gaussian kernel; (ii) downsampling that averages magnetic fields on each block of the size of the downscaling factor.

position and conditional on a low-resolution magnetic field, the distribution of magnetic field is noisier in active regions of the Sun (Figure 4).

Assuming that pixel noise is independent across images and pixels, the negative log-likelihood of a sample  $\mathcal{D}_n$  is

$$\sum_{i=1}^n \sum_{j=1}^m \frac{1}{2\sigma_j(x_i)^2} (y_{ij} - f_{w,j}(x))^2 + \frac{1}{2} \log(\sigma_j(x_i)^2), \quad (1)$$

where  $m$  is the number of pixels in each high resolution image. To estimate both *epistemic* and *aleatoric* uncertainties, we use a Bayesian neural network that outputs both  $f_w(\cdot)$  and  $\sigma(\cdot)$  and is trained to minimize the heteroskedastic loss (Equation 1). For an input image  $x$ , the magnetic field uncertainty is obtained by sampling  $T$  times the network weights, getting  $f_t(x)$  and  $\sigma_t(x)^2$  for each sample  $t$  and computing predictive uncertainty as

$$\underbrace{\frac{1}{T} \sum_{t=1}^T f_t(x)^2 - \left( \frac{1}{T} \sum_{t=1}^T f_t(x) \right)^2}_{\text{Epistemic}} + \underbrace{\frac{1}{T} \sum_{t=1}^T \sigma_t(x)^2}_{\text{Aleatoric}}. \quad (2)$$

### 3 Experiments

#### 3.1 Data and Architecture

We use 400,000 ( $4096 \times 4096$  pixel) magnetograms collected by HMI between 2010 and 2019. Our experiment consists of super-resolving HMI magnetograms that have been artificially reduced by a factor of 4 in resolution by smoothing the HR image with a Gaussian kernel before down-sampling by averaging the magnetic fields of each 2 by 2 block of the HR magnetogram. The data is split into training, validation, and test sets by allocating one month randomly drawn every year to each of the test and validation sets and the remaining months to the training set.

We conduct this experiment with a state-of-the-art super-resolution architecture, HighRes-net [6]<sup>3</sup> which we adapted to output both mean and variance of each pixel value in high resolution. Dropout is modelled at each layer by a Bernoulli distribution with parameter  $p = 0.2$ . We run a set of experiments on  $128 \times 128$  pixel patches cropped from the center of HR images. Variational inference is done by sampling 500 realizations of the model.

#### 3.2 Results

We test how including epistemic and aleatoric uncertainty affects the accuracy of the SR task. Table 1 compares the mean squared error (MSE) of our baseline, HighRes-net trained with a MSE loss to (i) HighRes-net trained with dropout and MSE loss; (ii) HighRes-net trained with a heteroskedastic loss (Equation 1); and, (iii) HighRes-net trained with both dropout and heteroskedastic loss. Although accounting separately for *epistemic* and *aleatoric* uncertainty does not seem to degrade the performance of the neural network, the MSE is 8% higher with a combination of aleatoric and epistemic uncertainty. We found that the model with combined uncertainty is sensitive to the initialization of heteroskedastic variances. In our current implementation, we compute pixel  $j$ 's heteroskedastic variance as  $\sigma_j^2(x) = \tilde{\sigma}_j^2 + \bar{\sigma}^2$ , where  $\bar{\sigma}^2$  is the mean squared error obtained with homoskedastic variances. However, more work is needed to improve the model performance

Modelling *epistemic* uncertainty allows us to capture that the super-resolution of magnetograms is a more ill-posed problem in active regions of the Sun with large magnetic fields. Figure 1 shows that predictions consistent with the same low-resolution input are more likely to disagree for large magnetic fields. A possible explanation is the lack of examples of active regions, particularly around the solar Equator.

However, this lack of data is confounded by the presence of more noise in active regions: Figure 4 shows that the distribution of magnetic fields in high resolution has higher variance conditional on its counterpart value in low-resolution. *Aleatoric* uncertainty effectively measures the large amount of

<sup>3</sup><https://github.com/ElementAI/HighRes-net>.

Models	MSE
HighRes-net	88.45
+ Epistemic	90.00
+ Aleatoric	90.47
+ Epistemic & Aleatoric	98.10

Table 1: Super-resolution performance: Mean Squared Error of HighRes-net with and without dropout and heteroskedastic loss. Lower is better. This shows that accounting for aleatoric or epistemic uncertainty separately while training HighRes-net does not degrade the model accuracy. However, the MSE is 8% higher with a combination of aleatoric and epistemic uncertainty

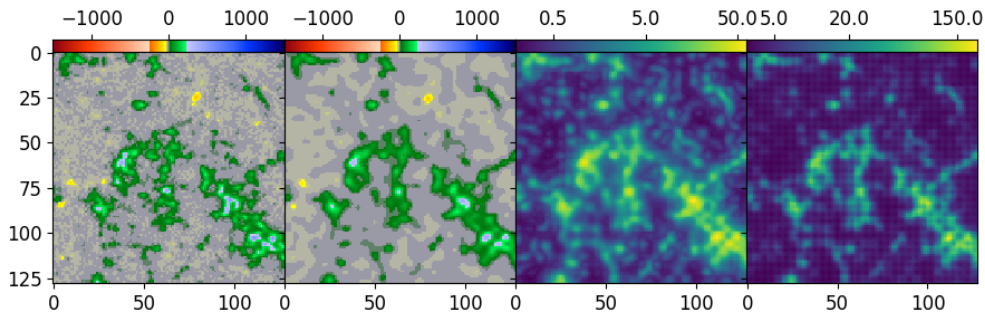


Figure 2: (left to right) An example of a HR target image ( $128'' \times 128''$ ) plotted over  $\pm 1500$  Gauss; the corresponding mean of MC-dropout samples; model uncertainty; and estimated noise.

noise in regions with a large magnetic field (Figure 2 and 3). By accounting for both *aleatoric* and *epistemic* uncertainty, we can properly disentangle the ill-posed nature of the super-resolution task from the heteroskedasticity of the noise (Figure 2 and 3). We observe that in active regions, *aleatoric* uncertainty is larger than *epistemic* uncertainty.

## 4 Conclusions and Future Work

We employ a Bayesian deep learning framework as a way to measure how uncertain predictions of super-resolved magnetograms are, particularly in active regions of the Sun. The task is uniquely challenged by the confounding effect of larger noise in active regions and relative scarcity of regions with large magnetic fields. We model *aleatoric* and *epistemic* uncertainty to properly measure the range of high resolution explanations compatible with a given low resolution input. This work is a first step toward generating super-resolved magnetograms useful to the heliophysics community. Future avenues for research include (i) expanding our results to the full solar disk; and, (ii) generating an initialisation procedure that improves the accuracy of our Bayesian model.

### Acknowledgments

This work was conducted at the NASA Frontier Development Laboratory (FDL) 2019. NASA FDL is a public-private partnership between NASA, the SETI Institute and private sector partners including Google Cloud, Intel, IBM, Lockheed Martin, NVIDIA, and Element AI. These partners provide data, expertise, training, and compute resources necessary for rapid experimentation and iteration in data-intensive areas. P. J. Wright acknowledges support from NASA Contract NAS5-02139 (HMI) to Stanford University. This research has made use of the following open-source Python packages SunPy [9], NumPy [10], Pandas [11], and PyTorch [12]. We thank Santiago Miret and Sairam Sundaresan (Intel) for their advice on this project.

## References

- [1] Wenming Yang, Xuechen Zhang, Yapeng Tian, Wei Wang, and Jing-Hao Xue. Deep Learning for Single Image Super-Resolution: A Brief Review. *arXiv e-prints*, page arXiv:1808.03344, Aug 2018.
- [2] Alex Kendall and Yarin Gal. What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems*, pages 5574–5584, 2017.
- [3] J. A. Linker, Z. Mikić, D. A. Biesecker, R. J. Forsyth, S. E. Gibson, A. J. Lazarus, A. Lecinski, P. Riley, A. Szabo, and B. J. Thompson. Magnetohydrodynamic modeling of the solar corona during whole sun month. *Journal of Geophysical Research: Space Physics (1978–2012)*, 104(A5):9809–9830, 5 1999.
- [4] Gábor Tóth, Igor V. Sokolov, Tamas I. Gombosi, David R. Chesney, C. Robert Clauer, Darren L. De Zeeuw, Kenneth C. Hansen, Kevin J. Kane, Ward B. Manchester, Robert C. Oehmke, Kenneth G. Powell, Aaron J. Ridley, Ilia I. Roussev, Quentin F. Stout, Ovsei Volberg, Richard A. Wolf, Stanislav Sazykin, Anthony Chan, Bin Yu, and József Kóta. Space weather modeling framework: A new tool for the space science community. *Journal of Geophysical Research: Space Physics (1978–2012)*, 110(A12), 12 2005.
- [5] C. J. Díaz Baso and A. Asensio Ramos. Enhancing SDO/HMI images using deep learning. *Astronomy & Astrophysics*, 614:A5, Jun 2018.
- [6] Anonymous. Highres-net: Multi-frame super-resolution by recursive fusion. In *Submitted to International Conference on Learning Representations*, 2020. under review.
- [7] 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.
- [8] Yarin Gal and Zoubin Ghahramani. Bayesian convolutional neural networks with Bernoulli approximate variational inference. *arXiv preprint arXiv:1506.02158*, 2015.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.

## A Appendix

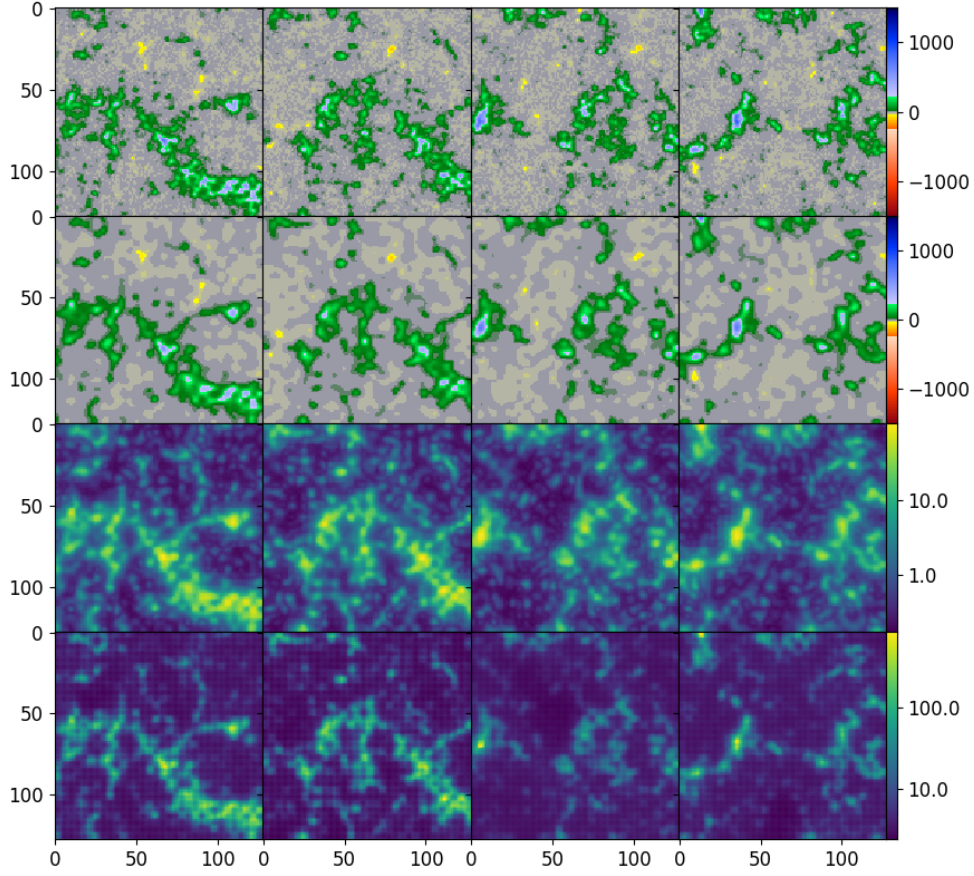


Figure 3: Sample of four HR target images ( $128'' \times 128''$ ) plotted over  $\pm 1500$  Gauss (1<sup>st</sup> row) the corresponding mean of MC-dropout samples (2<sup>nd</sup> row) model uncertainty (3<sup>rd</sup> row) and estimated noise (4<sup>th</sup> row).

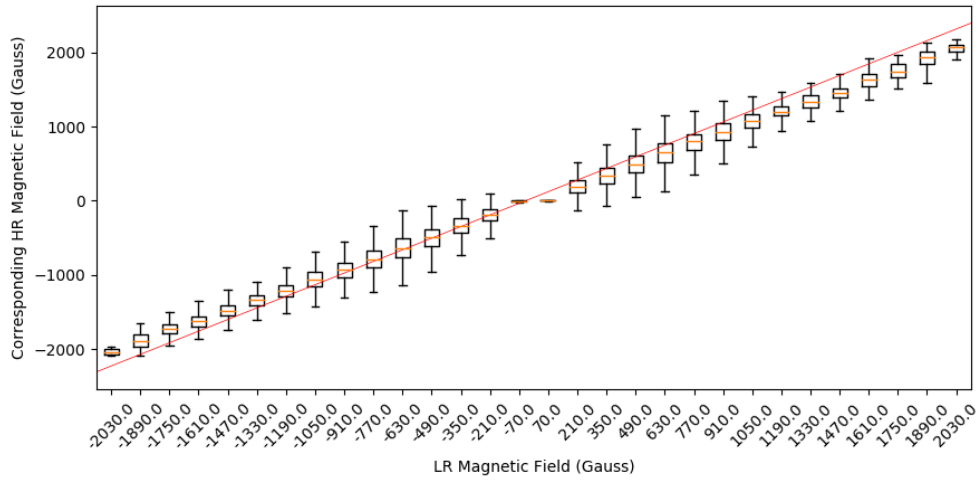


Figure 4: (left) Empirical mapping between single full disk LR and HR magnetograms. There is a clear increase in the difference of the mean and variance as a function of the LR field strength.