Auto-Calibration of Remote Sensing Solar Telescopes with Deep Learning

Brad Neuberg
NASA Frontier Development Lab

Souvik Bose
Rosseland Center for Solar Physics,
University of Oslo

Valentina Salvatelli
NASA Frontier Development Lab & IQVIA

Luiz F.G. dos Santos
CUA-IACS, NASA-GSFC

Mark Cheung
Lockheed Martin Solar and Astrophysics Laboratory &
Stanford University

Miho Janvier
Institut d'Astrophysique Spatiale, Université Paris-Sud

Atilim Gunes Baydin
University of Oxford

Yarin Gal
OATML, University of Oxford

Meng Jin
Lockheed Martin Solar and Astrophysics Laboratory & SETI

Abstract

As a part of NASA's Heliophysics System Observatory (HSO) fleet of satellites, the Solar Dynamics Observatory (SDO) has continuously monitored the Sun since 2010. Ultraviolet (UV) and Extreme UV (EUV) instruments in orbit, such as SDO's Atmospheric Imaging Assembly (AIA) instrument, suffer time-dependent degradation which reduces instrument sensitivity. Accurate calibration for (E)UV instruments currently depends on periodic sounding rockets, which are infrequent and not practical for heliophysics missions in deep space. In the present work, we develop a Convolutional Neural Network (CNN) that auto-calibrates SDO/AIA channels and corrects sensitivity degradation by exploiting spatial patterns in multi-wavelength observations to arrive at a self-calibration of (E)UV imaging instruments. Our results remove a major impediment to developing future HSO missions of the same scientific caliber as SDO but in deep space, able to observe the Sun from more vantage points than just SDO's current geosynchronous orbit. This approach can be adopted to perform autocalibration of other imaging systems exhibiting similar forms of degradation.

1 Introduction

Solar activity plays a major role in influencing the interplanetary medium and space-weather around us. Understanding the complex mechanisms that govern such a dynamic phenomenon is important and challenging. Remote-sensing instruments on board heliophysics missions can provide a wealth of information on the Sun's activity, especially via the measurement of magnetic fields and the emission of light from the multi-layered solar atmosphere. NASA currently operates the Heliophysics System Observatory (HSO) that consists of a fleet of satellites constantly monitoring the Sun, its extended atmosphere and space environments around the Earth and other planets of the solar system [1].

One of the flagship missions of the HSO is NASA's Solar Dynamics Observatory (SDO) [2]. Launched in 2010, it consists of three instruments: the Atmospheric Imaging Assembly (AIA) [3], the Helioseismic & Magnetic Imager (HMI) [4], and the EUV Variability Experiment (EVE) [5]. The SDO has been generating terabytes of observational data every day and has been constantly monitoring the Sun with the highest temporal and spatial resolution for full-disk observations.

Unfortunately, the (E)UV instruments in orbit suffer time-dependent degradation which reduces instrument sensitivity [6]. Accurate calibration for EUV instruments currently depends on sounding rockets (e.g., for SDO/EVE, and SDO/AIA), that are infrequent [7]. Since SDO is in a geosynchronous orbit, sounding rockets can be used for calibration, but calibration experiments may not be practical for missions in deep space (e.g., STEREO satellites) [8].

In the present work, we develop a neural network that auto-calibrates the SDO/AIA channels, correcting sensitivity degradation, by exploiting spatial patterns in multi-wavelength observations to arrive at a self-calibration of (E)UV imaging instruments. This removes a major impediment to developing future HSO missions that can deliver solar observations from different vantage points beyond Earth-orbit.

2 Methodology

2.1 Dataset

The processed SDO AIA dataset [9, hereafter SDOML] was used in order to provide a machine-learning ready dataset for the auto-calibration problem. More specifically, the dataset, which consists of a subset of the original SDO data dating from 2010 to 2018, is comprised of 7 EUV channels, 2 UV channels, and HMI vector magnetograms (three components). Both AIA and HMI data are spatially co-registered, have identical spatial sampling, and all instruments are synchronous. The temporal cadence is 6 min for AIA, 12 min for HMI, while the spatial resolution is ~4.8 arcsec per pixel (each full disk image has 512 x 512 pixels). Crucially, the images are corrected for instrumental degradation over time and have exposure corrections (determined from sounding rocket experiments). The end result is that in this dataset, changes in pixel brightness are from changes in solar conditions and not due to instrument performance.
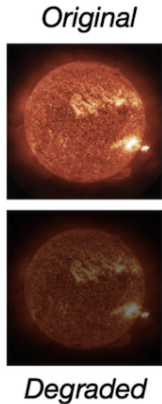

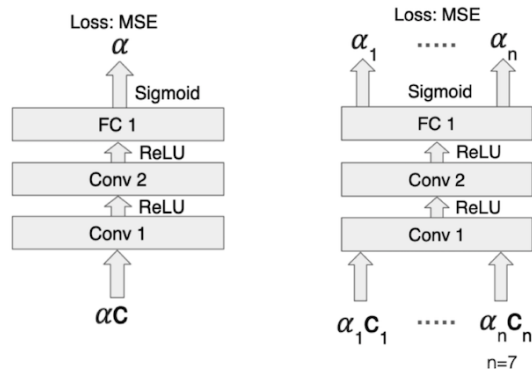Figure 1: A schematic of the auto-calibration problem. α is the brightness degradation factor.

Figure 2: CNN taking in one channel (left) and multiple channels (right).

2.2 Statement of Problem

Consider a set $\mathbf{C} = \{C_i, i \in [1, ..., n]\}$ of multi-channel, contemporaneous set of SDO/AIA images sampled from the SDOML dataset. C_i denotes the i-th channel image in the set. Consider a corresponding set of dimmed images $\alpha\mathbf{C} = \{\alpha_i C_i, i \in [1, ..., n]\}$, where each dimming factor α_i is individually sampled from a uniform distribution $[0.01, 1.0]$. Note the dimming factors α_i

apply uniformly per channel and is not spatially dependent. The spatial dependence portion of the degradation is assumed to be accounted for by periodic flatfields applied to AIA images. The goal is to find a model $M : \alpha\mathbf{C} \rightarrow \alpha$, i.e. find a model that retrieves the set of multi-channel dimming factors α_i (see Fig. 1).

In actuality, the realized dimming factors $\alpha_i(t)$ (where t is time since beginning of the SDO mission) trace a single trajectory in n-dimensional space (starting with $\alpha_i(t=0) = 1 \forall\, i \in [1, ..., n]$ at the beginning of the mission). During model training, we intentionally hide this time-dependence from the model in order to make it more general and robust. This is done by (1) using the SDOML dataset which has already been corrected for degradation effects, and (2) not assuming any relation between t and α and not using t as an input feature.

The training set comprises multi-channel images $\mathbf{C}$ taken during the months January to July during the years 2012 to 2013 (2 years). From one training epoch to another, the same multi-channel image set $\mathbf{C}$ can be dimmed by a completely different set of α dimming factors. This is a data augmentation strategy that allows the model to generalize and be effective in retrieving dimming factors over a range of solar conditions.

The test set comprises images taken during the months August to October during the years 2012 to 2013. The outcome of the learning algorithm is tested via a metric, the "binary frequency". This metric represents the percentage of predictions matching ground truth with an error less than a given tolerance across all the channels. We chose a binary frequency tolerance of 0.05 to gauge model success, which is below the uncertainty of the AIA degradation calibration [6].

2.3 Baseline Model

In order to compare the results of our deep learning approach, we created the following baseline model, which is motivated by the assumption that the EUV brightness outside magnetically active regions is invariant in time.

From the SDOML dataset, we picked a set of reference images per channel C_{ref} at time $t = t_{\mathrm{ref}}$. We computed the histograms (probability density functions) of the pixel values per channel. However, since the level of solar activity (in terms of active regions covering the solar disk) continuously evolves in time, so does the computed histograms. We propose that by including only pixels where the HMI line-of-sight magnetic flux density is $B_{\mathrm{los}} < 5\,\mathrm{Mx\,cm^{-2}}$, where $\mathrm{Mx\,cm^{-2}}$ is a unit of magnetic flux density, the histograms would capture only information about the Sun's ambient background state (the so-called "quiet Sun"). For each AIA channel, the reference histogram has a maximum $I_{i,\mathrm{ref}}^{\mathrm{mp}}$ corresponding to the most probable brightness value.

The baseline model works as follows. For any multi-channel image $\mathbf{C}$, we extract the channel-wise most probable value I_i^{mp}. The baseline model estimates dimming factors according to:

$$\alpha_i := I_i^{\mathrm{mp}} / I_{i,\mathrm{ref}}^{\mathrm{mp}}. \tag{1}$$

In terms of the binary frequency metric, the best results for the dimming factor retrieved from the ratio of the most probable values are 90% (304 Å) and 76% (131 Å), and we find ~54% mean success across all channels, for a tolerance level (the absolute difference between the ground truth degradation factor and the predicted one) of 0.05.

3 Convolutional Neural Network Model

As a first step, we tried a Convolutional Neural Network (CNN) [10] taking in one AIA wavelength channel, as represented in the schematic on the left of Fig. 2.

The architecture of the CNN consists of two convolutional layers and one fully connected (FC) layer, as well as two ReLUs (Rectified Linear Units) [11] and one final sigmoid as the activation functions. We use a L2 loss function, measuring the mean squared error of the distance between the predicted degradation factor and the ground truth value. The software library PyTorch was used for both training and inference of the CNN [12].

The results are given in Table 1, showing that for a tolerance level of 0.05, using a single channel increases the metric from 54% for the baseline to 93% for the CNN (taking the mean of the metric

Model	Binary Frequency Correct (tolerance=0.05)
Mean baseline	54%
Mean single channel CNN	93%
Worst single channel CNN	75%
Multi-channel CNN	**100%**

Table 1: Comparison between baseline, single channel Convolutional Neural Network (CNN) model (both mean and the worst wavelength), and multi-channel CNN model over 693 hold out test images.
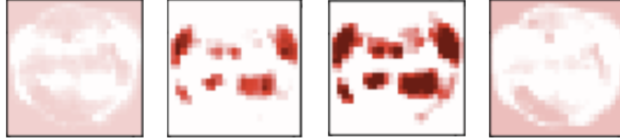

Figure 3: Subset of activations from the final convolutional layer, indicating that the convolutional layers have learned to distinguish between different morphological structures on the Sun.

across all wavelength channels). Furthermore, the metric for the worst wavelength degradation recovery (for 211 Å) is at 75%, which is still significantly better than the baseline.

We also test the same CNN taking as input degraded images from multiple channels, as indicated in the schematic on the right of Fig. 2. The hypothesis tested here is that the multi-channel CNN can learn more effectively than a single channel CNN by taking into account structures appearing in the different channels. A visualization of what structures are learned by the final convolutional layer, produced via [13], can be seen in Fig. 3. Such a mapping allows to see that the network learns both to distinguish the structures of the Sun and to separate them from the background. The results are presented in Table 1: we can see that the multi-channel CNN recovers the degradation factor, for all channels and for a tolerance level of 0.05, to 100%, over 693 test images. This is better than both the baseline as well as the CNN architecture that uses only one channel.

3.1 Dependence on Inter-Channel and Morphology/Brightness Relationships

In setting out to solve this problem, our premise was that the relationships between the brightness and morphology of solar features across multiple channels would facilitate the model to accurately estimate dimming factors.

To test this premise, we generated synthetic, idealized fake "Sun" images, in which the Sun is replaced by a 2D Gaussian profile of height A and characteristic width σ, viz.

$$C_i[x, y] = A_i \exp\left(-[x^2 + y^2]\sigma^{-2}\right) \qquad (2)$$

where σ is drawn from a uniform distribution between 0 and 1. We tested how (a) the presence/absence of a relationship between brightness A_i and size σ, and (b) the presence/absence of a relationship between A_i for different channels affects performance. When $A_0 \propto \sigma$ (linear relationship between size and brightness) and $A_i = A_0^i$ (i.e. relationship across channels), the CNN solution performed the best. When the inter-channel relationship was removed (i.e. each A_i was randomly sampled), the performance suffered. Finally, when both A_i and σ_i were both independently randomly sampled for all channels, the model achieved a performance that is equivalent to randomly guessing. This is expected behaviour since A_i, A_j, σ_i, and σ_j are uncorrelated for $i \neq j$. These experiments suggest the CNN solution on SDOML images works (partly) because of inherent relationships between the morphology of solar structures and their brightness, and because of relationships between different wavelength channels.

4 Summary & Future Work

In this paper we have shown a CNN that takes multiple wavelengths as input and which auto-calibrates SDO/AIA data, that performs significantly above a non-CNN baseline (100% vs. 54%) for a tolerance level of 0.05.

This opens the doors for future HSO missions to auto-calibrate their E(UV) instruments without using sounding rockets, enabling deep space Sun observatories from many vantage points.

References

[1] Steven Clarke. NASA's Heliophysics System Observatory. In *EGU General Assembly Conference Abstracts*, volume 18, pages EPSC2016–18529, Apr 2016.

[2] William Pesnell, Barbara Thompson, and Phillip Chamberlin. The Solar Dynamics Observatory. *solphys*, 275:3–15, 11 2012.

[3] J. R. Lemen, A. M. Title, D. J. Akin, P. F. Boerner, C. Chou, J. F. Drake, D. W. Duncan, C. G. Edwards, F. M. Friedlaender, G. F. Heyman, N. E. Hurlburt, N. L. Katz, G. D. Kushner, M. Levay, R. W. Lindgren, D. P. Mathur, E. L. McFeaters, S. Mitchell, R. A. Rehse, C. J. Schrijver, L. A. Springer, R. A. Stern, T. D. Tarbell, J.-P. Wuelser, C. J. Wolfson, C. Yanari, J. A. Bookbinder, P. N. Cheimets, D. Caldwell, E. E. Deluca, R. Gates, L. Golub, S. Park, W. A. Podgorski, R. I. Bush, P. H. Scherrer, M. A. Gummin, P. Smith, G. Auker, P. Jerram, P. Pool, R. Soufli, D. L. Windt, S. Beardsley, M. Clapp, J. Lang, and N. Waltham. The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Solar Physics*, 275:17–40, January 2012.

[4] 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*, 275:229–259, January 2012.

[5] T. N. Woods, F. G. Eparvier, R. Hock, A. R. Jones, D. Woodraska, D. Judge, L. Didkovsky, J. Lean, J. Mariska, H. Warren, D. McMullin, P. Chamberlin, G. Berthiaume, S. Bailey, T. Fuller-Rowell, J. Sojka, W. K. Tobiska, and R. Viereck. Extreme Ultraviolet Variability Experiment (EVE) on the Solar Dynamics Observatory (SDO): Overview of Science Objectives, Instrument Design, Data Products, and Model Developments. *Solar Physics*, 275:115–143, January 2012.

[6] Paul Boerner, Christopher Edwards, James Lemen, Adam Rausch, Carolus Schrijver, Richard Shine, Lawrence Shing, Robert Stern, Theodore Tarbell, Alan Title, C. Jacob Wolfson, Regina Soufli, Eberhard Spiller, Eric Gullikson, David McKenzie, David Windt, Leon Golub, William Podgorski, Paola Testa, and Mark Weber. Initial Calibration of the Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Solar Physics*, 275(1-2):41–66, Jan 2012.

[7] S. Wieman, L. Didkovsky, T. Woods, A. Jones, and C. Moore. Sounding Rocket Observations of Active Region Soft X-Ray Spectra Between 0.5 and 2.5 nm Using a Modified SDO/EVE Instrument. *Solar Physics*, 291:3567–3582, December 2016.

[8] M. L. Kaiser, T. A. Kucera, J. M. Davila, O. C. St. Cyr, M. Guhathakurta, and E. Christian. The STEREO Mission: An Introduction. *Space Science Reviews*, 136:5–16, April 2008.

[9] Richard Galvez, David F. Fouhey, Meng Jin, Alexandre Szenicer, Andrés Muñoz-Jaramillo, Mark C. M. Cheung, Paul J. Wright, Monica G. Bobra, Yang Liu, James Mason, and Rajat Thomas. A Machine-learning Data Set Prepared from the NASA Solar Dynamics Observatory Mission. *The Astrophysical Journal Supplement Series*, 242(1):7, May 2019.

[10] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, 1(4):541–551, Dec 1989.

[11] Vinod Nair and Geoffrey E. Hinton. Rectified Linear Units Improve Restricted Boltzmann Machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pages 807–814, USA, 2010. Omnipress.

[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.

[13] Utku Ozbulak. PyTorch CNN Visualizations. `https://github.com/utkuozbulak/pytorch-cnn-visualizations`, 2019.