

# X-ray, Optical, and Radio Solar Event Identification and Prediction using Neural Networks

Ted Grosson, Cody Meng, Preston Tracy, Jackson White, Yiwen Zhu

Advisor: Chris Tunnell — Rice University

Submitted February 22, 2021

## Project Pitch

Solar events, such as flares and prominences, can lead to detrimental effects on Earth systems by disrupting electronics and communication systems, particularly in satellites with limited protection from Earth’s magnetic field. Due to the potential harm of these events, it would be remarkably useful to identify and predict these events when or before they occur to allow rapid preparatory actions to mitigate potential damages as much as possible. We propose applying machine learning techniques to historical solar data from NASA’s Solar Dynamics Observatory in order to train a set of neural networks to identify and predict solar events from live multi-waveband imaging of the Sun. The Solar Dynamics Observatory (SDO) satellite was launched in 2010 and continuously observes the Sun on roughly 10-minute intervals with multiple imaging instruments. In particular, the Atmospheric Imaging Assembly (AIA) aboard SDO images the sun in wavelengths ranging from the UV to optical, providing a constant stream of live, high-resolution solar information [1]. Using the images obtained from AIA, we should be able to easily identify x-ray and optical events and prominences emanating from the Sun. By analyzing historical AIA data prior to and during documented solar events, we may be able to construct a tool that predicts solar events before and as they actually occur.

## Core Scientific Question

In this project, we plan to identify and predict specific solar events from multi-waveband observations of the Sun obtained from the Solar Dynamics Observatory Atmospheric Imaging Assembly. The cause of solar flares and prominences are most often attributed to magnetic fields, but we lack a concrete mechanism by which such events take place [2]. Determining whether visible precursors to solar events exist, how early they take place, and how prominently they appear in each waveband, may provide valuable constraints on the theoretical mechanisms by which these solar events are created. Even if we fail to produce a neural network capable of predicting solar events before they occur, this failure may indicate that solar events lack temporal precursors large or coherent enough to appear on the Sun’s surface in x-ray, ultraviolet, and optical wavelength images. Because we are applying similar methods to both identification and prediction of solar events, we can perform at least a cursory analysis comparing the efficacy of our tools applied to both problems, which could provide a point of comparison between the visibility of solar events precursors to the events themselves.

We also hope to determine whether we can precisely identify solar radio events from optical or ultraviolet images of the Sun, as it is unclear as to the degree of precision by which we can identify

solar events outside of their principal wavelengths. Should some solar events be recognizable in shorter wavelength images, we may be able to use this information to constrain or classify radio events. For example, if one type of radio event produces an identifiable signature in the optical while another does not, we may be able to use our tools to aid in classifying these events, providing additional clues as to the mechanisms by which either type of event can take place. In addition, if we are able to predict radio events in the optical or ultraviolet, this would provide similar benefits to those described in the previous paragraph: namely, determining how early, prominently, and in what wavelengths precursors to these events can occur.

## **Project Objectives**

Our project objectives fall under three overarching steps, which we will discuss in detail within this section.

First, we will attempt to train a neural network to identify solar events from solar images, using historical SDO AIA solar images in 5-6 wavebands between 90 and 300 Angstroms alongside solar event reports obtained from NOAA's Space Weather Prediction Center. The images and solar event reports will be aligned in time, so that the neural network will learn to associate features within the images with specific types of solar events. We will initially focus on two types of events, x-ray events and optical flares.

Next, we will apply this same methodology to radio bursts. We will attempt to use the longest wavelengths available to us on the AIA, but radio emission is well outside of the low-frequency limit of the accessible bandpasses. However, we will still train a neural network on similar sets of AIA images alongside simultaneous radio burst reports. If radio bursts carry signature features present in higher-frequency wavebands, our neural network may be able to identify them.

Finally, should the previous steps be successful, we will transition to prediction of solar events rather than identification. Our overall methodology is the same, but we will train our neural networks on solar event data alongside images taken at various time intervals before the solar events occurred. We will experiment with various time lags, training several neural networks on different lags simultaneously, and evaluating what time lags are optimal or effective for certain kinds of events. With an understanding of predictive accuracy as a function of time lag, we can determine whether and when notable surface signatures arise as precursors to x-ray, optical, and radio solar events. The previous two steps will both serve as points of comparison between the efficacy of our identification and prediction networks as well as provide a proof of concept demonstration of how our predictive network may function under ideal conditions, with a time lag of zero.

## **Data Description**

### **AIA Observations**

In this project we will be using two primary datasets: Atmospheric Imaging Assembly (AIA) instrument observations of the sun, from the Solar Dynamics Observatory (SDO), and space weather reports from the National Oceanic and Atmospheric Administration (NOAA). The AIA takes observations at ten different wavelengths, which are summarized in Table 1, and each one probes

## AIA Filter Summary

Wavelength (Å)	Primary Emission Source	Region of Atmosphere	log(T)
4500	Continuum	Photosphere	3.7
1700	Continuum	Photosphere	3.7
1600	C IV + Continuum	Upper Photosphere / Transition Region	4.7
335	Fe XVI	Flaring Regions	6.8
304	He II	Chromosphere / Transition Region	5.8
211	Fe XIV	Active-Region Corona	6.3
193	Fe XII, XXIV	Corona and Hot Flare Plasma	6.1, 7.3
171	Fe IX	Quiet Corona / Upper Transition Region	5.0
131	Fe VIII, XX, XXIII	Flaring Regions	5.6, 7.0, 7.2
94	Fe XVIII	Flaring Regions	6.8

Table 1: The wavelengths observed by the AIA, along with the primary source of emission at each wavelength, and the approximate location and temperature of the portion of the Sun that each wavelength probes. [5]

different layers of the sun. The AIA has been observing the sun since 2010, and observes at 4500 Å once every hour, at 1600 and 1700 Å once every 24 seconds, and at the remaining wavelengths once 12 seconds. The images are stored as 4096x4096 pixel .fits files, where each pixel represents the flux captured by the corresponding pixel on the AIA CCD. The headers of the .fits files also include relevant information such as the exposure time and time of the observation. These files can be queried and downloaded from within python, using the sunpy package. Using sunpy we can select images based on their observation time and filter, which will allow us to easily select and download observations which correspond to events in the space weather reports.

The different wavelengths observed by AIA were chosen to examine specific portions of the Sun’s surface or atmosphere. Each of these wavelengths are centered on a different emission line to examine the features visible at the temperature of the emission line. The EUV wavelengths observe extremely hot temperatures, and are centered on different iron ions, ranging from Fe IX (171 Å) to Fe XXIII (131 Å), and observing temperatures from 60,000 K (304 Å) to 20,000,000 K (193 Å) [3]. The 1700 Å and 4500 Å wavelengths show continuum images of the Sun in the UV and optical, respectively. The 1600 Å wavelength examines the transition region between the chromosphere and the corona. The 304 Å wavelength also examines the transition region, as well as the chromosphere. The 171 Å wavelength shows the “quiet” (low magnetic activity) corona and coronal loops, and the 193 Å examines a hotter region of the corona, as well as the hotter material in solar flares. The 211 Å and 335 Å wavelengths both examine the hot, magnetically active regions of the corona. Finally, the 94 Å and 131 Å wavelengths both examine flaring regions, with wavelengths centered at different temperatures [4]. The temperature coverage provided by these wavelengths allows for a more complete reconstruction of thermal structure than previous missions [5].

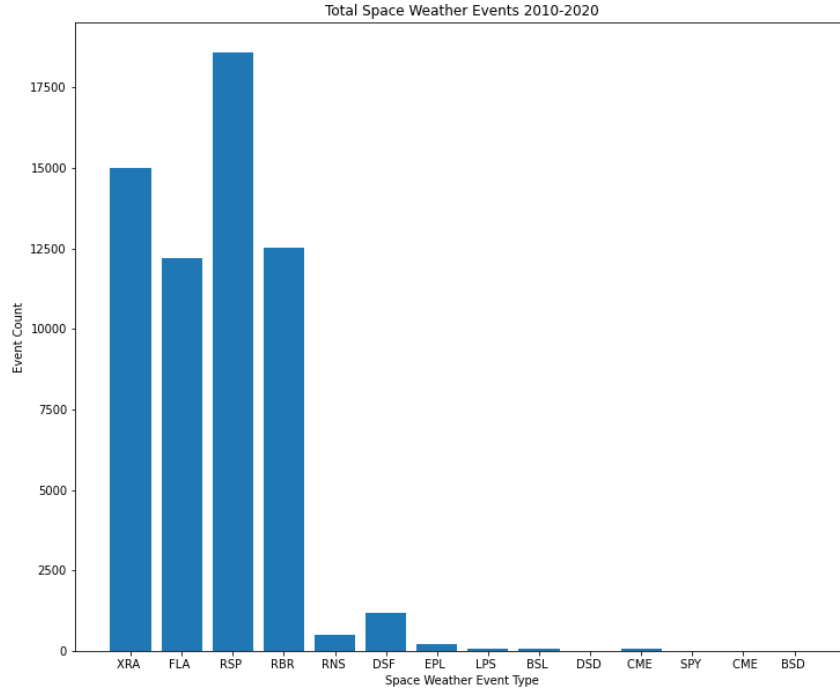


Figure 1: Total number of space weather events observed since 2010, grouped by event type (Event types are listed in table 2)

## NOAA Space Weather Reports

The second dataset, the space weather reports, are created once per day by the NOAA and can contain 13 different types of space weather events, which are listed in Table 2. These reports exist back until 1997, so we have corresponding reports for all AIA data. A sample report from 2015 is depicted in Figure 2. The entire collection of reports amounts to only 18 MB in storage size, so we are able to store all the reports on google drive. 79.6% of the reports since 2010 have at least one observed space weather event, and there are on average 15.0 events per day. There are significant frequency discrepancies between the different types of space weather events. As shown in Figure 1, the most common events are X-Ray events (XRA), optical flares in H-Alpha (FLA), and Sweep-Frequency Radio Bursts (RSP), and Fixed-Frequency Radio Bursts (RBR). Each of these events comprise almost a quarter of the dataset.

## Data Storage Concerns

Every event has a listed start time and end time, among other parameters. We will use these start time and end time values to identify when different events should be present, and use those time frames to select the images that will make up our neural net training and testing data sets. Each single .fits file is about 10MB in size, so we will need to be selective about how many images we

Acronym	Full Name
BSL	Bright Surge on Limb
DSF	Filament Disappearance
EPL	Eruptive Prominence on Limb
FIL	Filament
FLA	Optical Flare in H-Alpha
FOR	Forbush Decrease (CR Decrease)
GLE	Ground-Level Event (CR Increase)
LPS	Loop Prominence System
PCA	Polar Cap Absorption
RBR	Fixed-Frequency Radio Burst
RNS	Radio Noise Storm
RSP	Sweep-Frequency Radio Burst
SPY	Spray
XRA	X-ray Event

Table 2: List of Event Types tracked in space weather reports

store. The mean duration for the four most common types of events (XRA, FLA, RBR, and RSP) since 2010 are 20, 18.5, 3.75, and 48 minutes respectively. Considering the average duration of each event, the number of total number events since the AIA began observations, and the frequency of AIA observations, we would need almost a petabyte of storage to save every image in every wavelength for which a space weather event is occurring. Focusing instead on only 5-6 filters, and using one image per event, we should drop our storage needs down to about one terabyte per event type, which should be a more reasonable target.

## Background Information

Solar flares are highly energetic events of localized increased brightness on the sun over a time period ranging from milliseconds to over an hour. This brightness can be seen across many wavelengths, including x-ray, optical, and radio, and can be seen in Figure 3. They tend to be associated with groups of sunspots — localized regions of cooler material and strong magnetic fields — and are often accompanied by the ejection of charged particles. These particles, as well as high-energy electromagnetic radiation, can affect electrical systems and the Earth’s ionosphere, and have the potential to cause major disruptions [2]. The main goal of the Solar Dynamics Observatory is to understand the mechanisms of these events which have the potential of affecting life on Earth [1]. The precise cause of solar flares remains uncertain, but a leading theory is that energy is suddenly released from the strong magnetic field in sunspots through a process called reconnection. This occurs when a magnetic field loop “breaks” and “reconnects” in a more stable path, releasing most of the energy stored in the loop [2]. An additional process which appears in our images is granulation, which occurs as a result of convection in the solar atmosphere. Granules tend to span around 700 km and have lifetimes of five to ten minutes [2]. Since we are analyzing how the Sun

is changing over time, the granulation pattern could affect our identifications, as seen in Figure 4.

There are some successful prior efforts to predict the solar flares based on changes in the Sun’s magnetic field [6]; however, because solar flare events are associated with changes in the Sun’s magnetic field, it may not be possible to directly predict them through purely visual means. Should our prediction method turn out to be satisfactory, it would provide an unprecedented view about the nature of solar flares that the solar flares are correlated to the dynamics of the sun’s surface in a way that we can use the emission from the sun’s surface to detect and predict a solar flare.

## Potential Risks

Although the data for this project is easy to access, the biggest risk is the failure to identify events from the images we have. If our produced neural networks are not capable of detecting events, then there is little hope for them to be able to predict such events beforehand. There are curated datasets available which could improve the quality of the images in some ways; however, these are only available a week after the images are taken at the earliest, offering little use in trying to predict events in real time [7]. In the event that the lower-quality images are unable to detect events, the curated datasets could still be useful in determining whether flare events are associated with visual changes in the Sun, which is valuable information regardless of whether it can be used to predict flares in real time. Additionally, it is possible that we will need to train our neural networks on more images than expected, in order to successfully detect different objects. This could pose a significant challenge to overcome if we encounter difficulties securing additional storage space or computational power.

## References

- [1] W. Dean Pesnell, B. J. Thompson, and P. C. Chamberlin. The Solar Dynamics Observatory (SDO). *Solar Physics*, 275(1-2):3–15, January 2012.
- [2] Bradley W. Carroll and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. Cambridge University Press, 2 edition, 2017.
- [3] Title A.M. Akin D.J. et al. Lemen, J.R. The atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo). *Solar Physics*, 275:17–40, January 2012.
- [4] Holly Zell. How sdo sees the sun, Feb 2015.
- [5] AIA Science Team. Atmospheric Imaging Assembly for the Solar Dynamics Observatory, Concept Study Report, Appendix A (AIA Science Plan) , Feb. 2004.
- [6] Abbas Raboonik, Hossein Safari, Nasibe Alipour, and Michael Wheatland. Prediction of solar flares using unique signatures of magnetic field images. *The Astrophysical Journal*, 834, 10 2016.
- [7] 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*, 242(1):7, May 2019.

## Appendix

```

:Product: 20150617events.txt
:Created: 2015 Jun 18 2127 UT
:Date: 2015 06 17
# Prepared by the U.S. Dept. of Commerce, NOAA, Space Weather Prediction Center
# Please send comments and suggestions to SWPC.Webmaster@noaa.gov
#
# Missing data: ////
# Updated every 5 minutes.
#
# Edited Events for 2015 Jun 17
#
#Event      Begin      Max      End  Obs  Q  Type  Loc/Frq  Particulars      Reg#
#-----
3910        0040      ////      0041  PAL  C   RSP  025-075  III/1
4060        A0116      ////      A1226  HOL  3   DSF  N39W23   7
3930        0120      0120      0120  LEA  G   RBR  245      110
3940        0205      ////      0205  PAL  C   RSP  075-180  III/1
3950        0207      0207      0207  LEA  U   RBR  245      240
3960        0546      0606      0635  G15  5   XRA  1-8A     C1.6    4.5E-03  2367
3960 +      B0548      U0558      A0653  SVI  2   FLA  S19E04   1F      ERU      2367
4280        0654      ////      0739  LEA  U   RSP  025-037  VI/1
3970        0706      0706      0706  LEA  G   RBR  245      120

```

Figure 2: Sample Space Weather Report text file From June 6, 2015. Along with event types and times, the reports also contain information about the event strengths and locations.



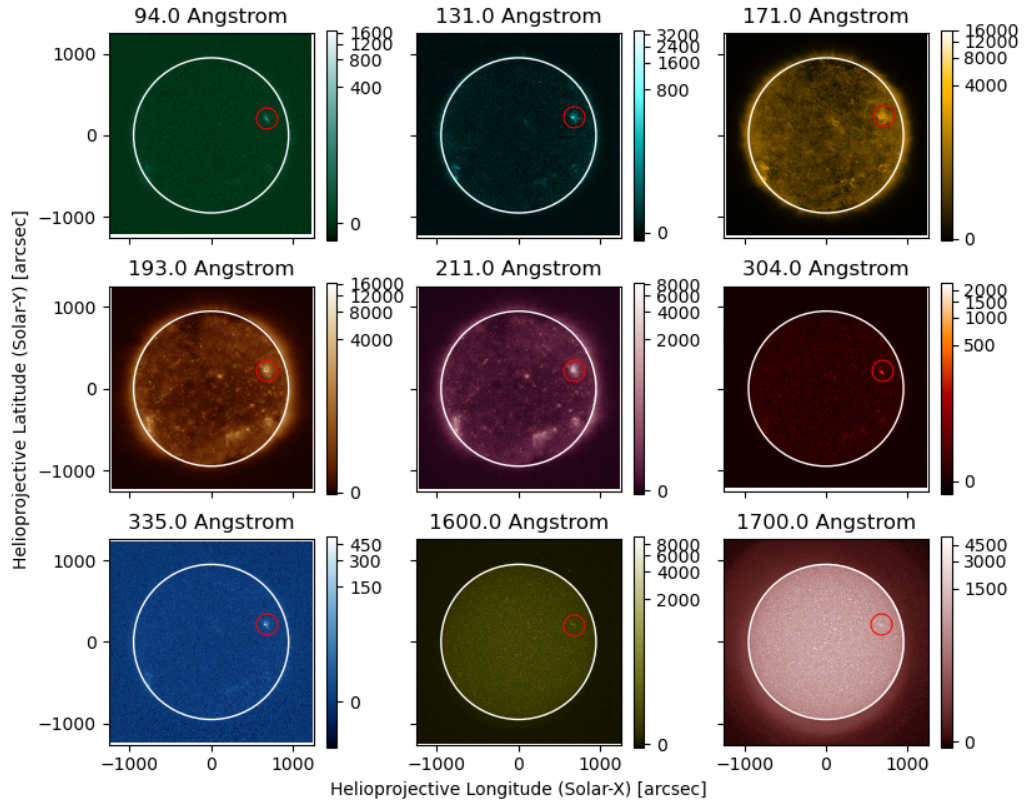


Figure 3: A small x-ray flare which occurred August 19, 2020 in each AIA wavelength apart from 4500 Å. The flare, circled in red on each image, shows up in all wavelengths to varying degrees.

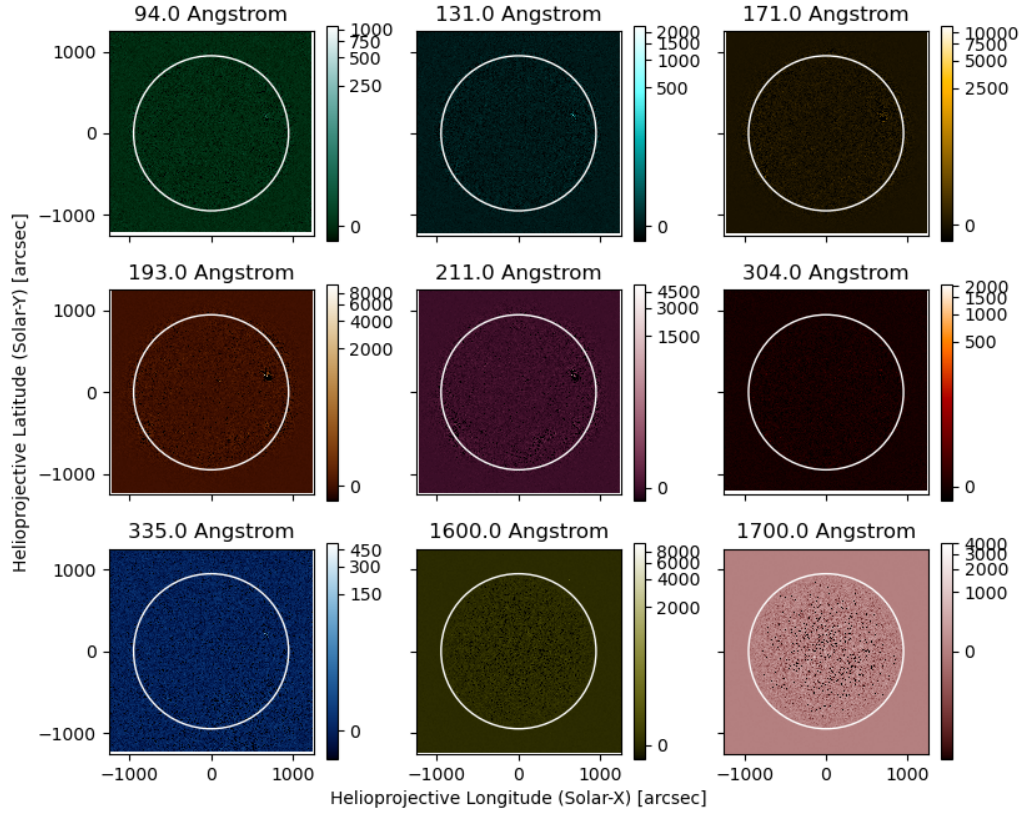


Figure 4: The same flare as above, after subtracting a previous image from the current one. This process is sometimes called difference imaging, and is useful in seeing how an object changes over time, such as the appearance of a flare. These difference images were taken over a span of two minutes, but varying this time could optimize our detection process. The changing granulation pattern is especially apparent in the 1700 Å image, where the flare is completely drowned out.