# Data Augmentation of Magnetograms for Solar Flare Prediction using Generative Adversarial Networks

Allison Liu[1] and Wendy Carande[1]

[1]Laboratory for Atmospheric and Space Physics

November 23, 2022

**Abstract**

Space weather forecasting remains a national priority in the United States due to the impacts of events like solar flares to life on Earth. High energy bursts of radiation originating from solar flares have the potential to disrupt critical infrastructure systems, including the power grid and GPS and radio communications. The rise of machine learning and the development of higher-quality instruments has greatly improved solar flare prediction models over the past decade. However, the magnetogram data used for solar flare forecasting—taken by the Solar and Heliospheric Observatory/Michelson Doppler Interferometer (SOHO/MDI) and the NASA Solar Dynamic Observatory/Helioseismic and Magnetic Imager (SDO/HMI) instruments—are incompatible due to differences in the cadence, resolution, and size of the data. Furthermore, many studies only focus on data from a single instrument which disregards decades worth of potential training data that is necessary to understand solar cycles. In this work, we show Generative Adversarial Networks (GANs) can be used to super-resolve the historic lower-quality SOHO/MDI data set to match SDO/HMI quality to create a standardized magnetogram data set. The implementation of a Pix2Pix GAN produced some undesirable artifacts in the synthetic image while image translation methods CycleGAN and CUT preserved solar features present in the data more accurately, even in the absence of paired data. The resulting combined, higher-quality data set will be used to improve the predictive power of current solar flare forecasting models.

# Data Augmentation of Magnetograms for Solar Flare Prediction using Generative Adversarial Networks (GANs)

Allison Liu[1] and Wendy Carande[1]
[1]Laboratory for Atmospheric and Space Physics (LASP)

Many solar flare prediction models disregard decades worth of training data because the magnetogram datasets used are incompatible.

**We produced a higher-quality magnetogram dataset for solar flare prediction using generative adversarial networks (GANs).**

## Motivation

**Solar flares – high energy bursts of radiation from the sun – have the potential to disrupt critical infrastructure systems on Earth, including the power grid and GPS and radio communications.** Understanding solar flares also helps us protect astronauts in space from harmful radiation. The two magnetogram datasets used for solar flare forecasting are incompatible due to differences in the cadence, resolution, and size of the data.

## Data and Preprocessing

We use line-of-sight, full-disk magnetograms from

1. the NASA Solar Dynamic Observatory/Helioseismic and Magnetic Imager (SDO/HMI), 720s cadence
2. the Solar and Heliospheric Observatory/Michelson Doppler Interferometer (SOHO/MDI), 96m cadence

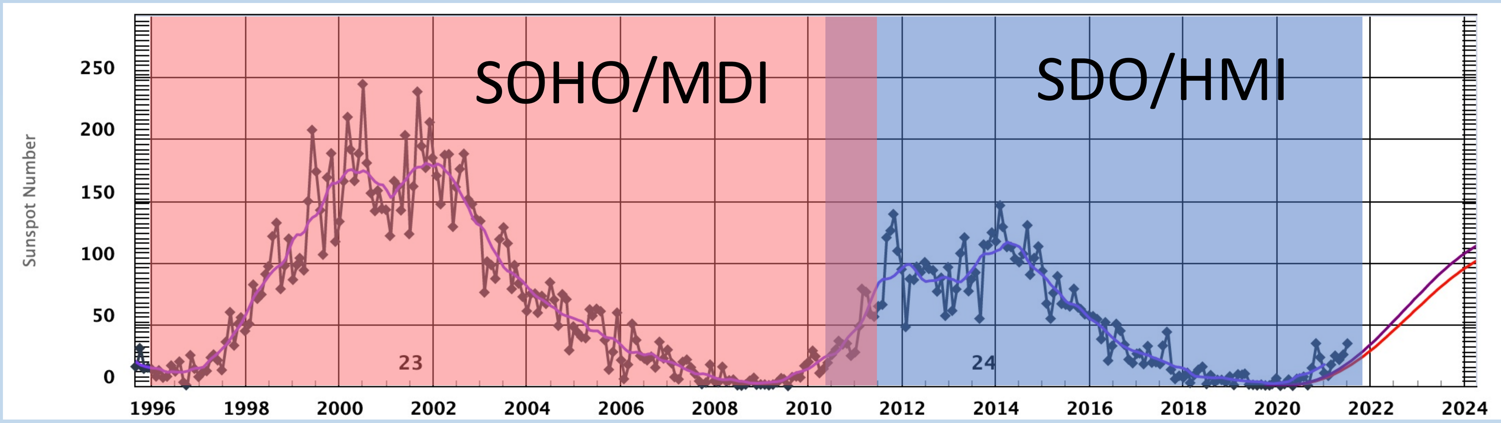


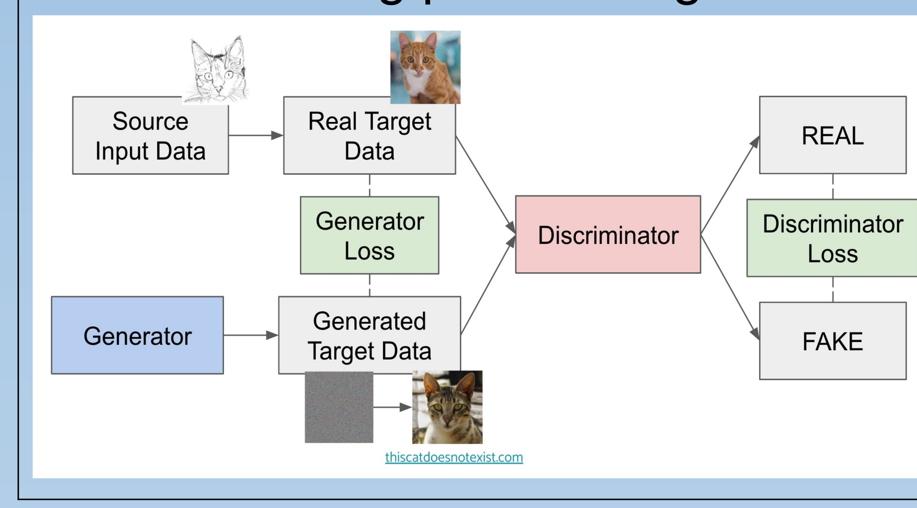

**Figure 1.** The data overlap of SOHO/MDI and SDO/HMI over solar cycles 23 and 24
Source: SWPC

Training and testing images were paired based on approximate time, then checked for missing pixels. Images were then rotated and centered.

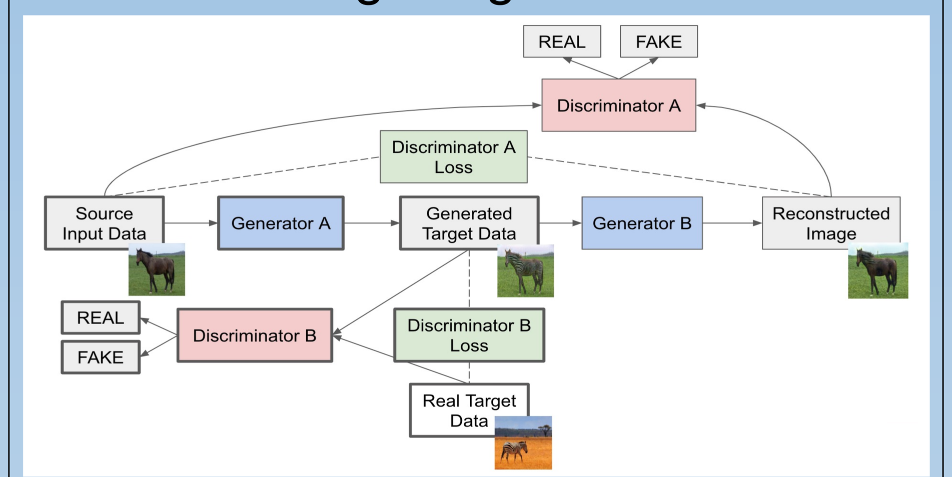
## Methods

Goal: Image-translation. We want to learn a mapping between input and output images
We compare the performance of three GAN models:

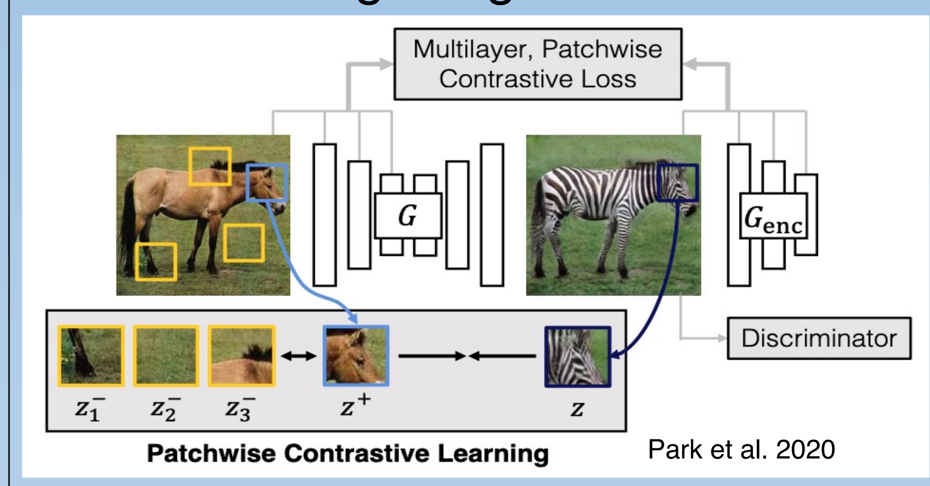
**Pix2Pix**
Relating paired images




**CycleGAN**
Relating image domains




**CUT**
Relating image domains




## Results

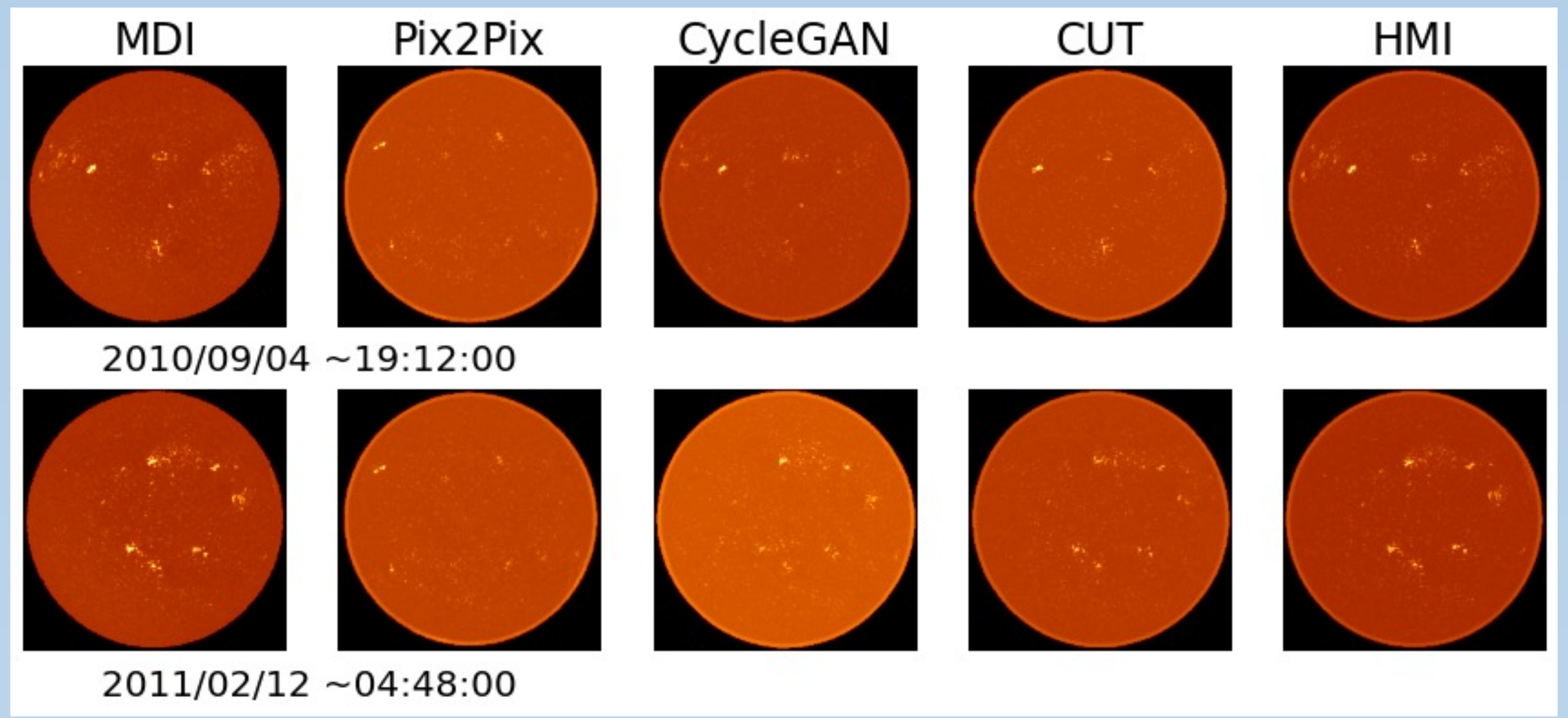



**Figure 2.** Synthetic images generated by different GAN models.


**Figure 3.** Plot of absolute error (a) and mean squared error (b) per image.

| Model | Mean Abs Error | MSE |
|---|---|---|
| Pix2Pix | 728.1717 | 0.001027 |
| CycleGAN | 608.7130 | 0.00745 |
| CUT | 623.8944 | 0.00074 |
| Baseline | 1007.9503 | 0.001288 |

**Table 1.** Average values of absolute errors and mean squared errors for each model

## Takeaways

We have demonstrated that GANs can be used to upconvert the historic SOHO/MDI dataset to SDO/HMI quality. **We obtained best results from unpaired models CycleGAN and CUT, which generalized the best and allow for a training dataset that extends beyond the one-year overlap of SOHO/MDI and SDO/HMI data.**

Laboratory for Atmospheric and Space Physics
University of Colorado Boulder

Contact: allison.liu@colorado.edu or wendy.carande@lasp.colorado.edu

This work was supported by the SWx TREC Deep Learning Laboratory and the National Science Foundation REU Program, award #1659878

For references and more information:
SCAN ME