Project Title: Enhancing Solar Farside Magnetogram Generation through Deep Learning Analysis of Satellite Data

Project Type: Software development and improved performance of algorithms

Background: Previous to the Solar TErrestrial RElationship Observatory (STEREO) mission, astronomers could only see the side of the Sun facing Earth, and had little knowledge of what happened to solar features on the other side of the Sun. Would active regions grow larger, and affect the space weather environment when they rotated back again two weeks later, or would they decay away? What about new active regions forming on the far side of the Sun, waiting to surprise us?

Fortunately, for about 5 years (from 2011-2015) STEREO A and B provided us with data from the Helium II emission line at 304 Angstroms EUVI 304-A of the farside of the Sun. Today (2024) there is no satellite that can measure the magnetic field on the farside of the Sun. STEREO is now near the Earth, the Solar Orbiter and the famous Parker Probe new satellite are also passing in front of the Sun, all will be able to provide any magnetic information of the farside in a few more years. Only the helioseismic monitor that maps of the acoustics of the solar farside is able to tell us that there is a large magnetic region at the back of the Sun. The monitor is located at Stanford University (JSOC server) and it generates 12 h averaged acoustic maps. We have absolute access to the data server directly from the database of the Solar Dynamics Observatory satellite. This is good progress, but this information doesn't contain any signed magnetic polarity at all on the farside.

To be able to understand the solar activity and predict space weather disturbances from the near and farside of the Sun it is essential to estimate the magnetic field of the full Sun. A 3D magnetic mapping of the whole Sun is required by models of the solar wind and coronal mass ejections all the way to Earth. Mathematically, the magnetic field at the photosphere acts as the boundary condition of the mathematical equations needed to extrapolate the global magnetic field of the Sun.

Our AIM is to use an AI cGAN algorithm – which does image to image translation– to create a 3D complete magnetic map of the Sun. For a first task we want to translate EUV STEREO information of the farside into a signed magnetic field detection. This task is at the forefront of solar physics research, with more and more solar data being translated into other type of images, all addressing a critical gap in our understanding and capability in space weather forecasting.

Methodology: Pix2PixHD is one of the popular deep-learning algorithms for image translation of high-resolution images without significant artifacts. In solar physics it was used in a conditional generative adversarial network cGAN to generate solar farside magnetograms. The Pix2PixHD is complex, it consists of two major networks: one is a generative network (generator) and the other is a discriminative network (discriminator). The generator tries to generate realistic output from input, and the discriminator tries to distinguish the more realistic pair between a real pair and a fake pair. The real pair consists of a real input and a real output. The generator tries to generate realistic data that fools the discriminator, thus minimizing the loss.

Kim et al. 2019 generated far side magnetograms from the farside EUV observations. We followed the same steps:

1 They trained/tested a machine learning model that paired the nearside extreme-ultraviolet images (EUV 304- red image in Figure 1) and the near-side SDO/HMI magnetograms (gray image). The model's performance was assessed/validated for the full solar disk including active regions and the quite regions. Focus was in predicting the amount of magnetic field as well as the positive/negative magnetic polarities. We have only worked with the first version of the code.

2 Then the model was applied to the farside images (EUVI STEREO 304-A) to generate farside magnetograms as close as possible to the real configuration.

3 Although the cGAN successfully generated accurate magnetic maps and tracked the temporal evolution of major active regions, numerous uncertainties and incorrect magnetic polarities persist. A second improved version has been published, but we have not yet make that code run. The novel approach was that three EUV passbands were used as input data at 304, 193, and 171 Å, corresponding to the chromosphere, corona, and upper transition region.

4 **Software:** First version of the code (the one we have worked with) was presented by Kim et al (2019): Solar farside magnetograms from deep learning analysis of STEREO/EUVI data were generated. The code is available at: https://github.com/chemron/SolarMagGAN and on Monarch/Monash University. If we want understand this algorithm well, we have to make it faster and then we will be in a better position to test and focus on improving it with elements of ML .

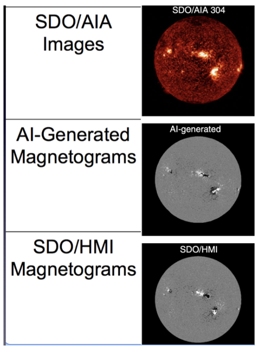

Figure 1: Comparison between HMI magnetograms and AI-generated ones from SDO/AIA 304-Å images: Input data is the SDO satellite data AIA 304 EUV map –red– and the full disk bottom magnetogram –gray– and the AI cGAN generated magnetic map is in the middle. Prediction of the magnetic polarities is good but not perfect.

Existing progress and problems

We are familiar with standard coding and debugging and compilation in C, UNIX scripts and python. We understand the mathematics of a cGAN architecture. The algorithm works now on local small computers, it is dormant on the Monarch cluster at Monash University using tensorflow and keras and pix2pix algorithms. Latency between running tests makes the understanding of the algorithm and its output difficult, slow. This requires updating the code with a new version of python or other pytorch or keras dependencies, and add a selection of GPUs and CPUs to speed up the code. An honours student at Monash in 2020-2021 helped me with this code (github link provided, 2019). We also wrote python codes for the download and preparation of the input pairs of data. A significant challenge for us was also lacking a productive method to assess the quality of the generated data. The algorithm's performance is hampered by its slow execution speed and inefficient configuration on GPU/CPU setups. Also, we were not able to make forward scientific planning for modelling large scale magnetic maps of the Sun because, we are stalling in generating precise magnetic maps with polarities included and assess its scientific significance.

Request for Expertise We recognize the need to expand our scientific aims by acquiring expertise and skills currently beyond our capacity.

1. Operational: Our primary objective is to make the deep learning algorithm cGANs operational for supercomputer platforms, utilizing the latest Python environments, TensorFlow packages, and other dependencies on GPU/CPU configurations that we understand. This will enable us to perform detailed performance testing of the algorithm and its architecture for various solar data pairs.

2. Desired Improvements on Speed: We require expertise: to profile the code, to identify performance problems, to find solutions to avoid bottle necks within the large flow of data that have to pass though a generator and a discriminator (two networks), therefore to enhance the computational efficiency of our model, we are aiming to reduce the processing time from 12 hours to 3 hours. This fourfold increase in speed is crucial for deeper insights into the algorithm's functioning and how to explore GPU/CPU configurations.

3. Visualise intermediate steps inside the code to assess the progress of generating images, that clearly shows the intrinsic competition between the generator and discriminator.

4. Data: Once a profile of the code is performed, perhaps the code can be improved by assessing its data flow and prepare for a larger influx of data at. better cadence (data as png or as fits files? normalisation or not of data?).

Impact:

The improvements to the code will enable more effective training of the algorithm, which is vital for advancing the science that is currently out of our reach. Additionally, training PhD students will cultivate essential skills within our group.

A primary goal is to enhance international collaboration with Dr. Kiran Jain's team at the National Solar Observatory, renowned for their continuous pipeline of farside helioseismic images sourced from a network of six ground-based GONG observatories. They are keen to apply this algorithm to their data, aiming to establish a permanent monitoring system for the Sun and a robust protocol to exchange data results and h unlimited access to data.

We will also create quantifiable impacts, such as accelerated research output and strengthened international/national collaborations. Once the code is optimized, we will lay the groundwork for advanced research in this field. Our goal is to tackle a complex challenge and pioneer the use of this algorithm to convert helioseismic images into magnetic maps (see Fig 2). Unfortunately the farside helioseismic maps do not have the resolution of the training set, and image manipulation such as applying another cutting-edge machine learning technique called 'super-resolution' and more software expertise may be required.

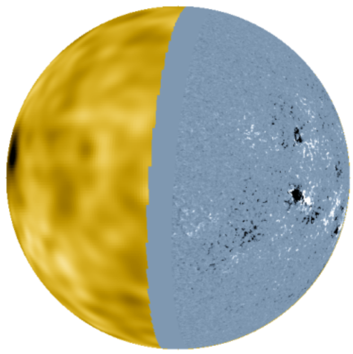

Figure 2: The ultimatum future challenge: Ensemble image of the active regions on the Sun: magnetograms are on the near side, helioseismic maps –low resolution– are on the farside of the Sun.

Improving the code by reaching up to understand a second version of it, may lead to new complexities in the performance of the algorithm (which we will tackle once a desired learning experience and skills is accumulated in the first phase), which may lead to request further ADACS support.

We expect that the improved code will lead to create tailored solutions applied for other applications with image to image protocols. In the fields of astronomy and geophysics, many multi-wavelength observations are available, so the model can be used to extend these kinds of data.

As we can see there are exiting further steps than can be done and we are looking forward to accelerate our learning. I would estimate that 1-1.5 full year, with constant interaction between members of the project may lead to a superb code that offers a solution for space weather forecasting done, for the first time in Australia.

References

1. Kim, T., Park, E., Lee, H. et al. Solar farside magnetograms from deep learning analysis of STEREO/EUVI data. Nat Astron 3, 397–400 (2019). The code is available at: https://github.com/chemron/SolarMagGAN.

2. Jeong et al. (2020) "Solar coronal magnetic field extrapolation from synchronic data with AI-generated farside", The Astrophysical Journal Letters, Volume 903, Issue 2, id.L25, 9,https://github.com/JeongHyunJin/