Space Weather Workshop (University Corporation for Atmospheric Research (UCAR)) April 17-21, 2023 Boulder, CO

A Machine Learning Ecosystem for Filament Analysis – Phase I: A Manually Annotated Dataset of Filaments

Samuel McDonald *, Rohan Adhyapak *, Kartik Chaurasiya *, Laxmi Alekhya Nagubandi *, Apaar Bawa *, Patrick Copeland *, Aya Abdelkarem *, Aparna Venkataramanasastry †, Alexei Pevtsov ‡, Luca Bortello ‡, Niles A. Oien ‡, Petrus C. Martens *, Azim Ahmadzadeh *

* Georgia State University † Bay Area Environmental Research Institute ‡ National Science Foundation's National Solar Observatory

Abstract

Considering the challenges associated with processing and fully utilizing the high-volume solar image data streaming in from the National Science Foundation's Global Oscillation Network Group and the lack of robust solar filament datasets, we proposed the Machine Learning Ecosystem for Filament Detection (ML Eco Fi), an NSF-funded, multiyear project that will produce an open-source collection of high-fidelity solar filament data and computer vision software for space weather research. In cooperation with NSF's National Solar Observatory and the NSO Integrated Synoptic Program, ML Eco Fi will assist in automatically detecting, classifying, localizing, and segmenting solar filaments in full-disk H-alpha images.

ML Eco Fi

https://www.mlecofi.net

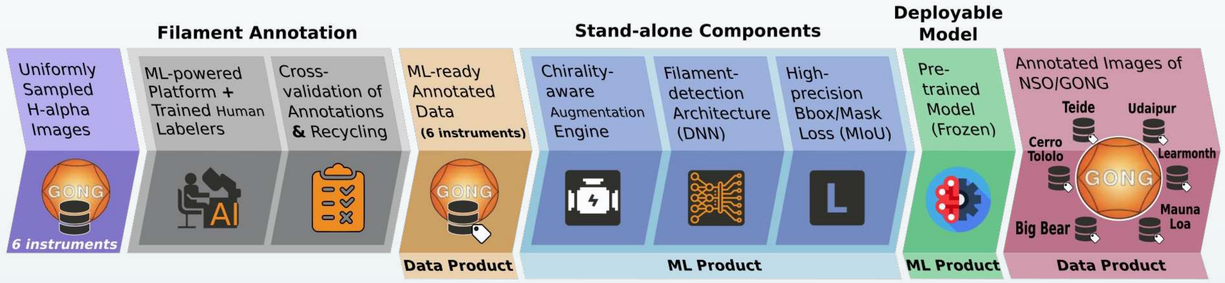
Background and Motivation

Timely detection and classification of solar filaments is critical in forecasting Earth-affecting transient solar events, including large solar flares and coronal mass ejections. Undetected, these space weather events can cause catastrophic geomagnetic storms resulting in substantial economic damage and death. NSO deployed a network of six ground-based solar observatories, GONG, in 1995 to provide continuous observation of solar activity and aid in space weather research efforts [1]. In 2010, GONG began archiving open-source full-disk H-alpha FITS images for the research community, collected year-round at an uninterrupted one-minute cadence. Based on previous investigations of filaments [2][3], and identifying that existing automatic segmentation models fail to capture fine structure information [4], the ML Eco Fi team set out to develop a new automatic segmentation model to aid in NSO's mission.

The GONG solar observatories. Image credit: NSO

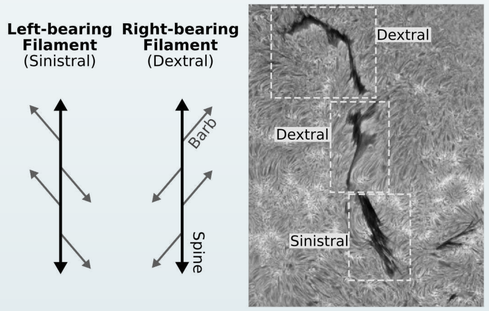
Project Overview

The present phase of research aims to produce a dataset of thousands of GONG H-alpha images, with all solar filaments' magnetic chirality, bounding box, and segmentation mask manually annotated through morphological analysis following strong quality assurance and data validation standards. This high-fidelity dataset will help advance research of filaments and filament-related topics. This dataset, as the first ML Eco Fi product, will facilitate the development of upcoming ML Eco Fi products, including a browser-based tool for easily viewing and analyzing GONG H-alpha images in cadence, chirality-aware filament data augmentation engine, high-precision image segmentation machine learning loss function, deep neural network segmentation and classification model, and filament detection module that is planned for deployment into NSO's live infrastructure for the research community.

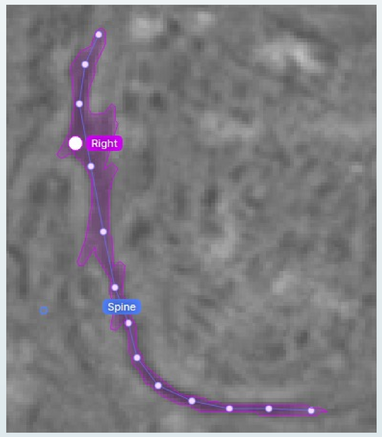


Results

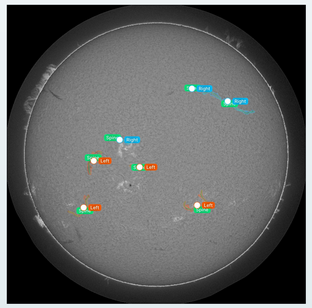
The data annotation phase is ongoing, but sampled here are examples of annotations in the final dataset. The ML Eco Fi annotations include instance segmentation polygons and masks, bounding boxes, spine polygons, and magnetic chirality labels, for all filaments in focus in each full-disk H-alpha image. As shown in the second figure, care is taken to ensure the fine structural details of each filaments such as the size and shape of its barbs, are accurately captured in the segmentation mask. Existing research datasets for the morphological analysis of filaments often lack this information. It is the ML Eco Fi team's objective that this additional effort will yield a dataset that meets and exceeds the quality of current data sources.



Morphological analysis of filaments.
Image credit: Martin, S.F. "Formation of Filaments"
(https://doi.org/10.1023/A:1005026814076)



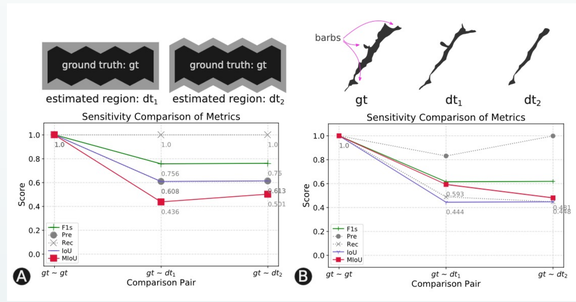
Annotated solar filament, showing pixel-precise masking of barbs and polygonal trace of spine



Annotation of filaments on full-disk H-alpha GONG image

Conclusion and Future Work

The ML Eco Fi team's ongoing annotation and validation efforts are proceeding well. When the dataset is completed, it will be made available publicly, alongside detailed documentation of its production and an open-source repository of related software. A novel machine learning segmentation loss function based on a fine-structure-sensitive object similarity metric called Multiscale Intersection over Union (MIoU) is currently being developed by the ML Eco Fi team. The MIoU segmentation loss function will be integrated into a deep neural network filament-detection architecture, trained using the ML Eco Fi filament dataset augmented by our chirality-aware data augmentation engine. Once implemented, this model will be able to automatically classify and segment filaments in new H-alpha images from GONG.



Comparison of MIoU segmentation loss metric to IoU
Image credit: Ahmadzadeh, A. et al, "Multiscale IoU"
(https://doi.org/10.1109/ICIP42928.2021.9506337)

Follow our Work!

Feedback Form

References and Acknowledgements

[1]Hill, F., "The Global Oscillation Network Group Site Survey - Part One", *Solar Physics*, vol. 152, no. 2, pp. 321–349, 1994. https://doi.org/10.1007/BF00680443.

[2]Bernasconi, P.N. "Advanced Automated Solar Filament Detection And Characterization Code: Description, Performance, And Results", *Solar Physics*, vol. 228, pp. 97–117, 2005. https://doi.org/10.1007/s11207-005-2766-y.

[3]Hazra, S., "Hemispheric Preference and Cyclic Variation of Solar Filament Chirality from 2000 to 2016", *The Astrophysical Journal*, vol. 865, 108, 2018. https://doi.org/10.3847/1538-4357/aadb42.

[4]Ahmadzadeh, A., "Toward Filament Segmentation Using Deep Neural Networks," *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, 2019, pp. 4932-4941. https://doi.org/10.1109/BigData47090.2019.9006340.

This project has been supported by funding from CSSI, Directorate for Computer and Information Science and Engineering (CSE), Office of Advanced Cyberinfrastructure (OAC), under NSF award #2209912.