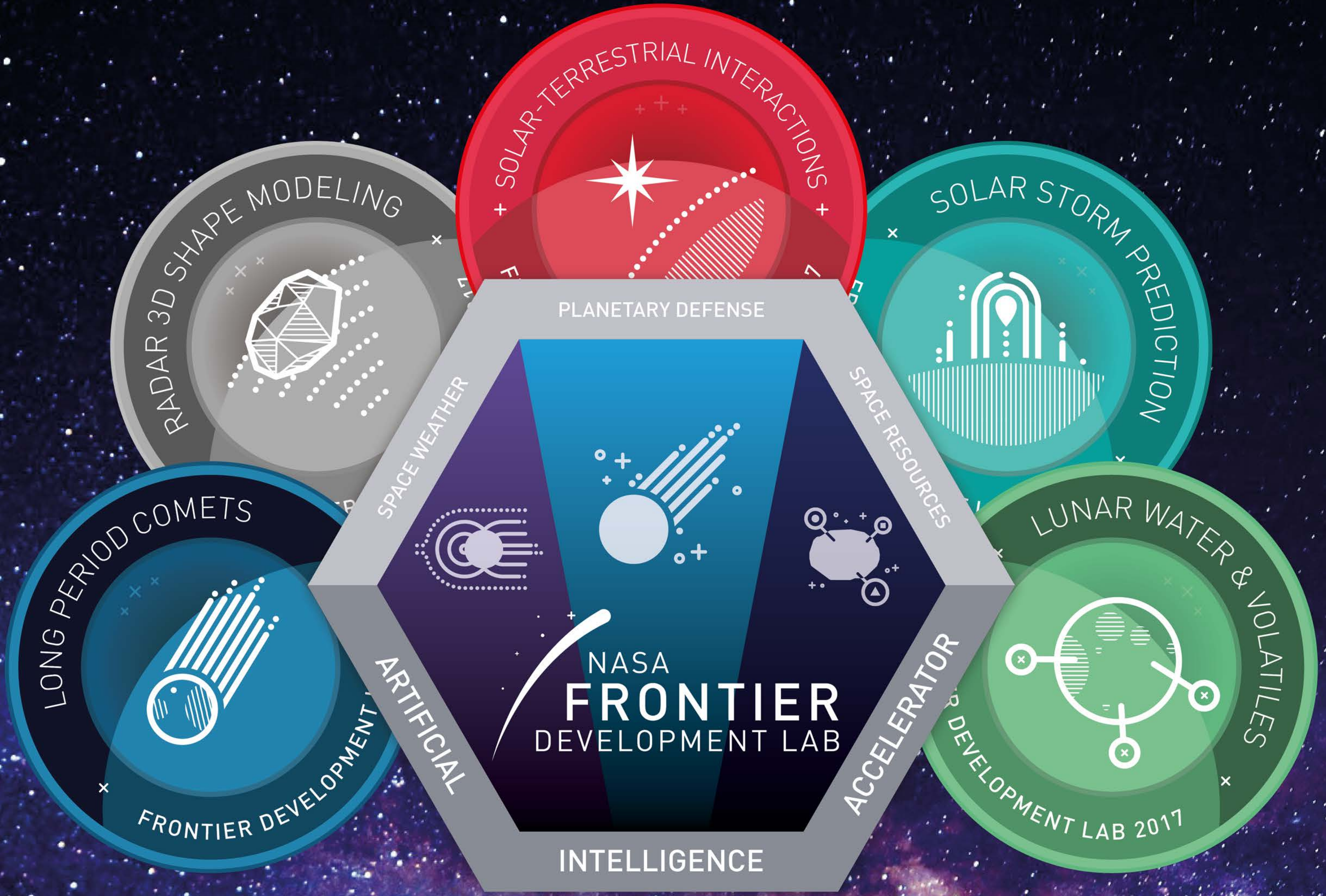


NASA FDL OVERVIEW- DR LIKA GUHATHAKURTA (on behalf of the FDL team.)

« Back to the Blog

What is NASA doing with Big Data today?

October 04, 2012 by Nick Skytland

Open Data | big data | Open Source | open government | Open Innovation | TopCoder

In the time it took you to read this sentence, NASA gathered approximately 1.73 gigabytes of data from our nearly 100 currently active missions! We do this every hour, every day, every year – and the collection rate is growing exponentially. Handling, storing, and managing this data is a massive challenge. Our data is one of our most valuable assets, and its strategic importance in our research and science is huge. We are committed to making our data as accessible as possible, both for the benefit of our work and for the betterment of humankind through the innovation and creativity of the over seven billion other people on this planet who don't work at NASA.

NASA FRONTIER DEVELOPMENT LAB

“Applied artificial intelligence research accelerator that combines the capabilities of NASA, academia, and private sector companies to tackle challenges not only important to NASA, but also to humanity’s future.”

NASA
FRONTIER
DEVELOPMENT LAB

LOCKHEED MARTIN

IBM

SR SPACE RESOURCES.LU

USC

NVIDIA

MISO

SETI INSTITUTE

XPRIZE

kx

intel

SETI enables the public / private partnership

FDL private sector partners provide GPU compute, storage and expertise


FDL POST-DOC TEAMS ARE INTERDISCIPLINARY:
50% DATA SCIENCE / 50% SPACE SCIENCES
NASA FRONTIER DEVELOPMENT LAB

BUT FIRST SOME CONTEXT…

NASA / BIG DATA / AI

WHAT ARE THE OPPORTUNITIES?
HOW CAN FDL HELP NASA MOVE FORWARD?

Artificial Intelligence : A Few Definitions

Artificial Intelligence (AI)

A computer which mimics cognitive functions typically associate with human intelligence.
Examples : goal seeking strategy formulation, complex image recognition, "learning", inference, and creative problem solving.

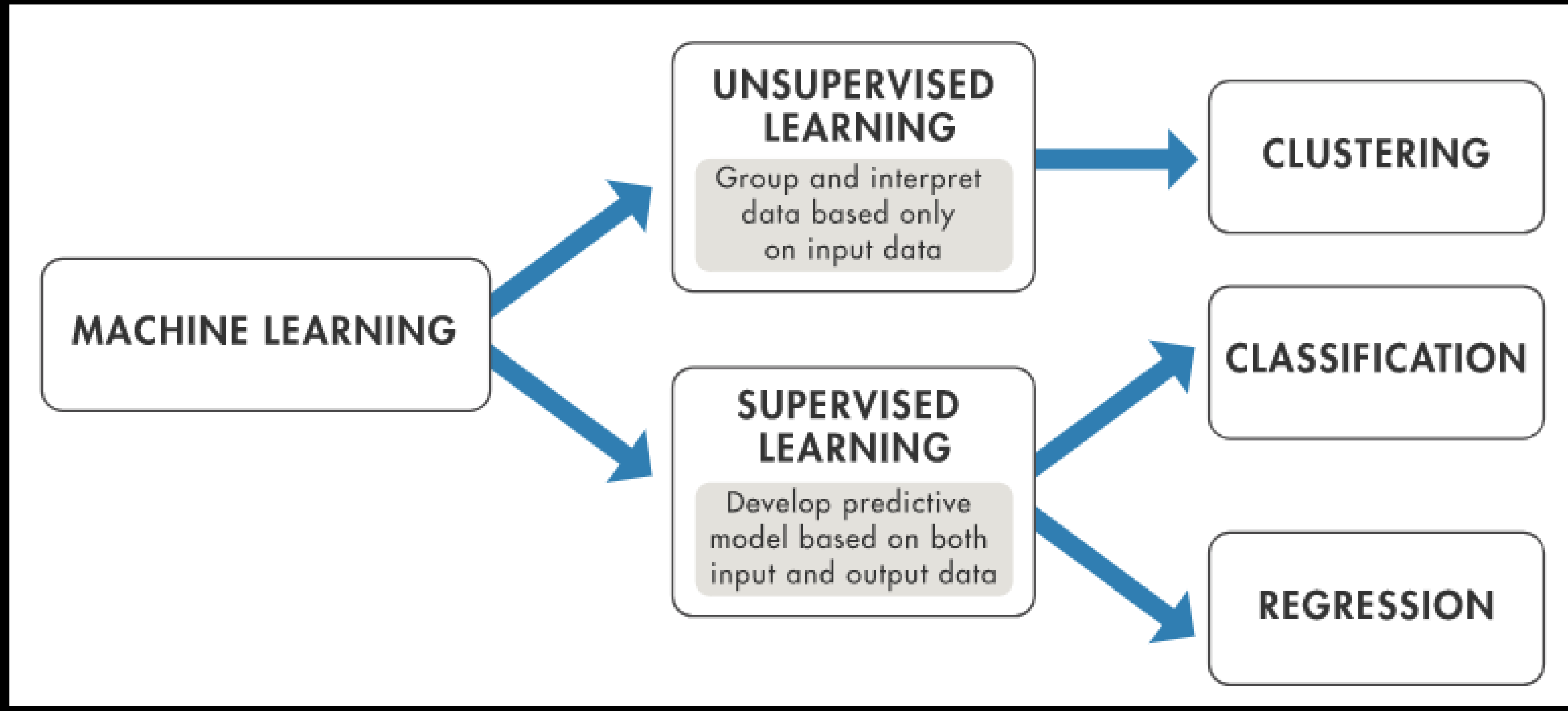
Machines Learning (ML): A branch of artificial intelligence in which a computer progressively improves its performance on a specific task by “learning” from data, without being explicitly programmed.

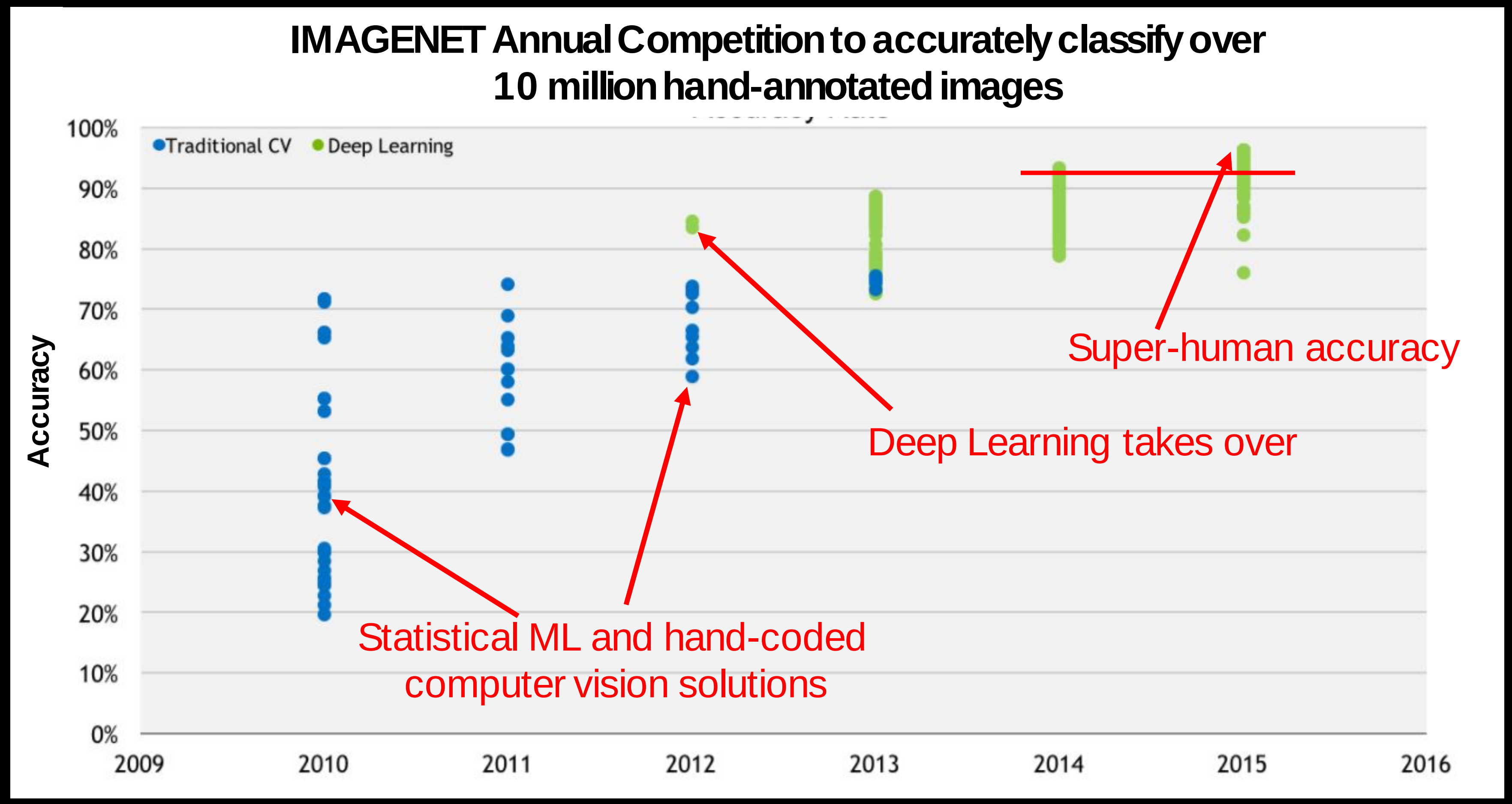
- Closely related to computational statistics, which focuses on prediction and optimization.

Data Mining: Discovering patterns in large data sets using techniques at the intersection of machine learning, statistics, and data management.

Deep Learning (DL): An extension of Machine Learning that uses the mathematical concept of a neural network (NN) to loosely simulate information processing and adaptation patterns seen in biological nervous systems.

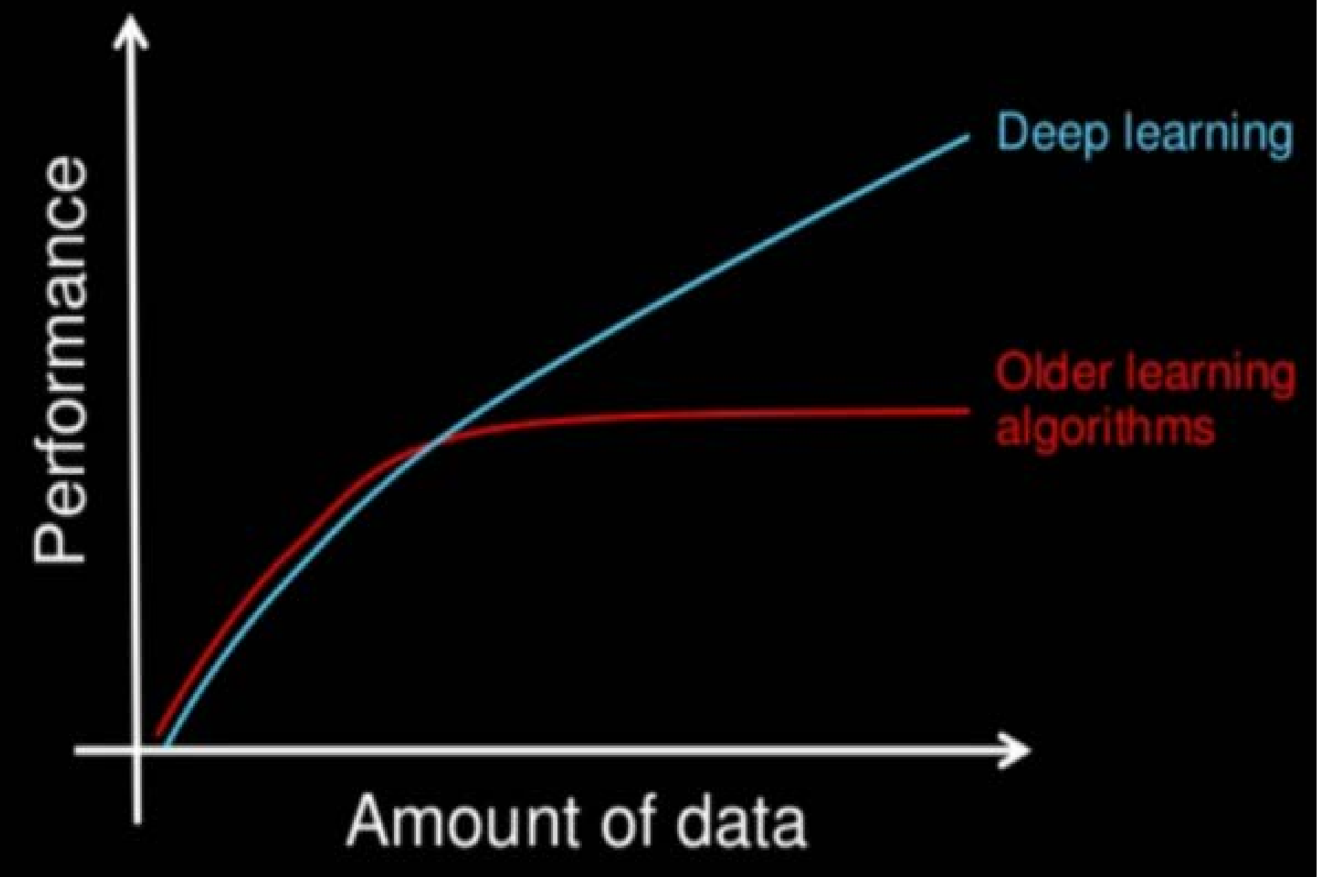
- Many problems which have been traditionally tackled with pensive coding have been overwhelmingly superseded by neural nets that outperform the humans that trained them.
- Exponential investment (patents, publications, funding) has fueled rapid advances in DL capabilities to make predictions, to identify anomalies, and even create new content that mimics what it has previously seen.





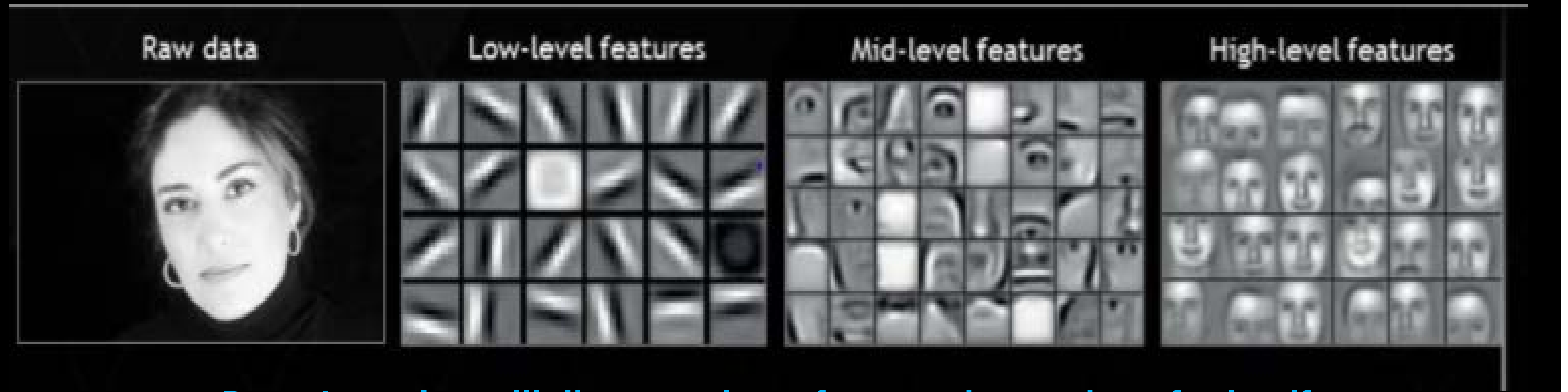
Statistical Machine Learning vs. Deep Learning

Data Scale: When properly architected, the efficacy of DL systems continue to improve with more data, long after statistical models have plateaued.



Interpretation: Machine Learning systems provide "visibility" into their statistical foundations, allowing their results to be interpreted and explained. Deep Learning systems are more of a "black box", although this is improving… and in some cases this is not an impediment (e.g. AI-enhanced science discovery)

Feature Discovery: Machine Learning often requires a human expert to create "feature extractors" that enable the statistical models to learn effectively, but Deep Learning finds these high-level features for itself (often with surprisingly creative results)



Deep Learning will discover these feature abstractions for itself.
Machine Learning needs help to extract features for statistical modeling.

Whole System: Machine Learning typically requires that complex systems be "chunked" into trainable components that are then manually recombined. Deep Learning can often "short circuit" that process and successfully model complex systems from end-to-end

Multiple ML models for each component of the Solar-Terrestrial Environment

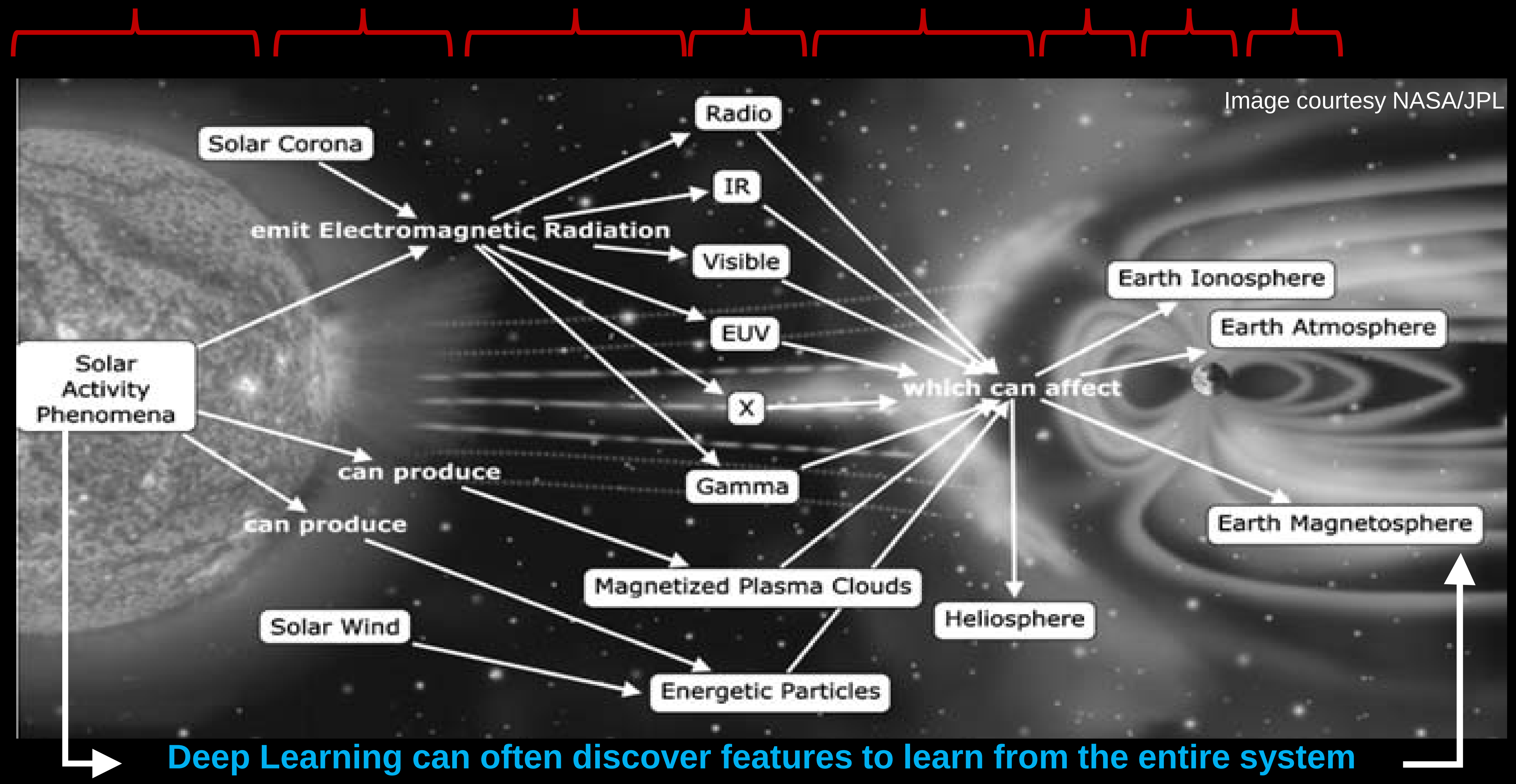
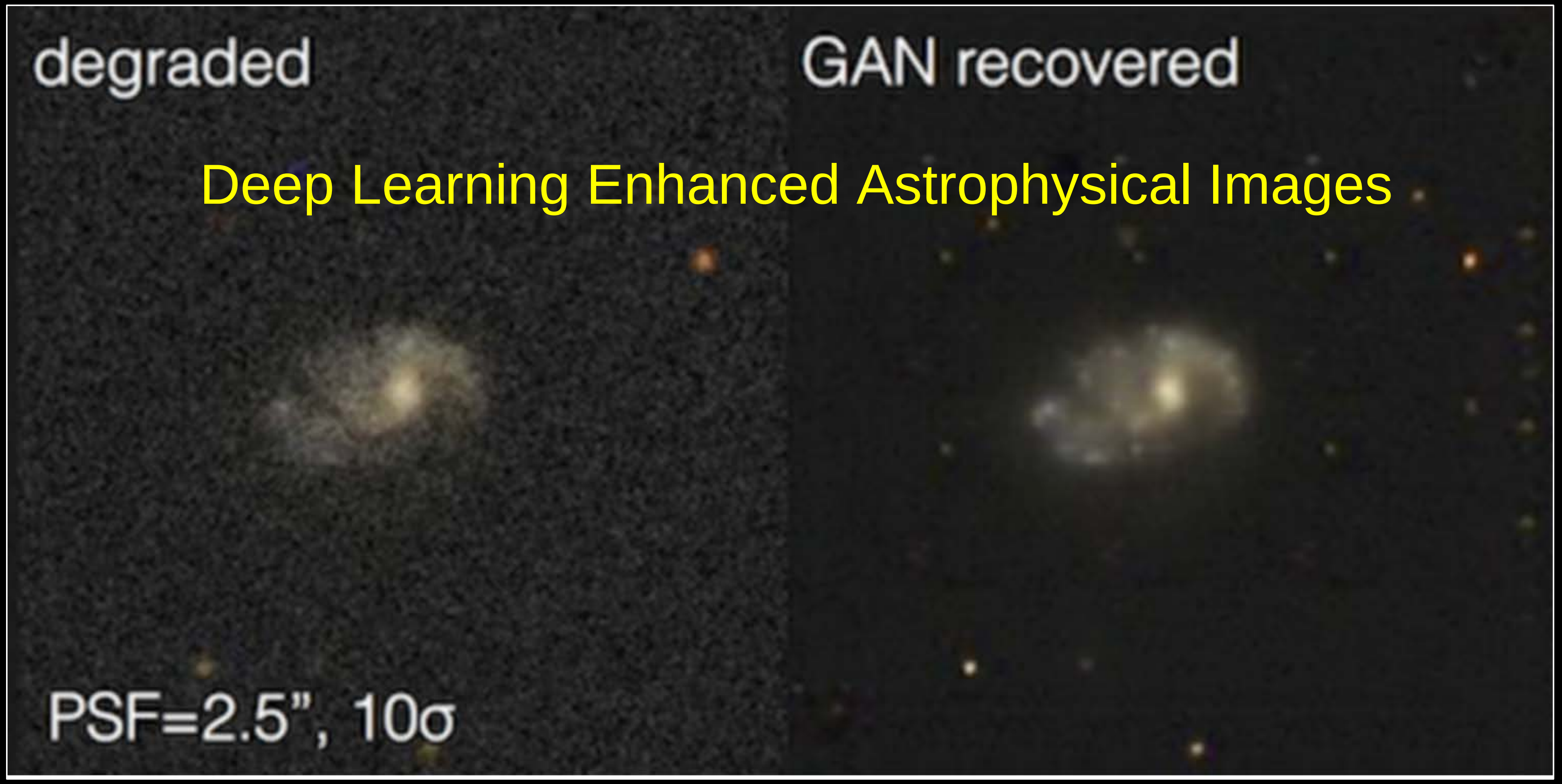


Image courtesy NASA/JPL

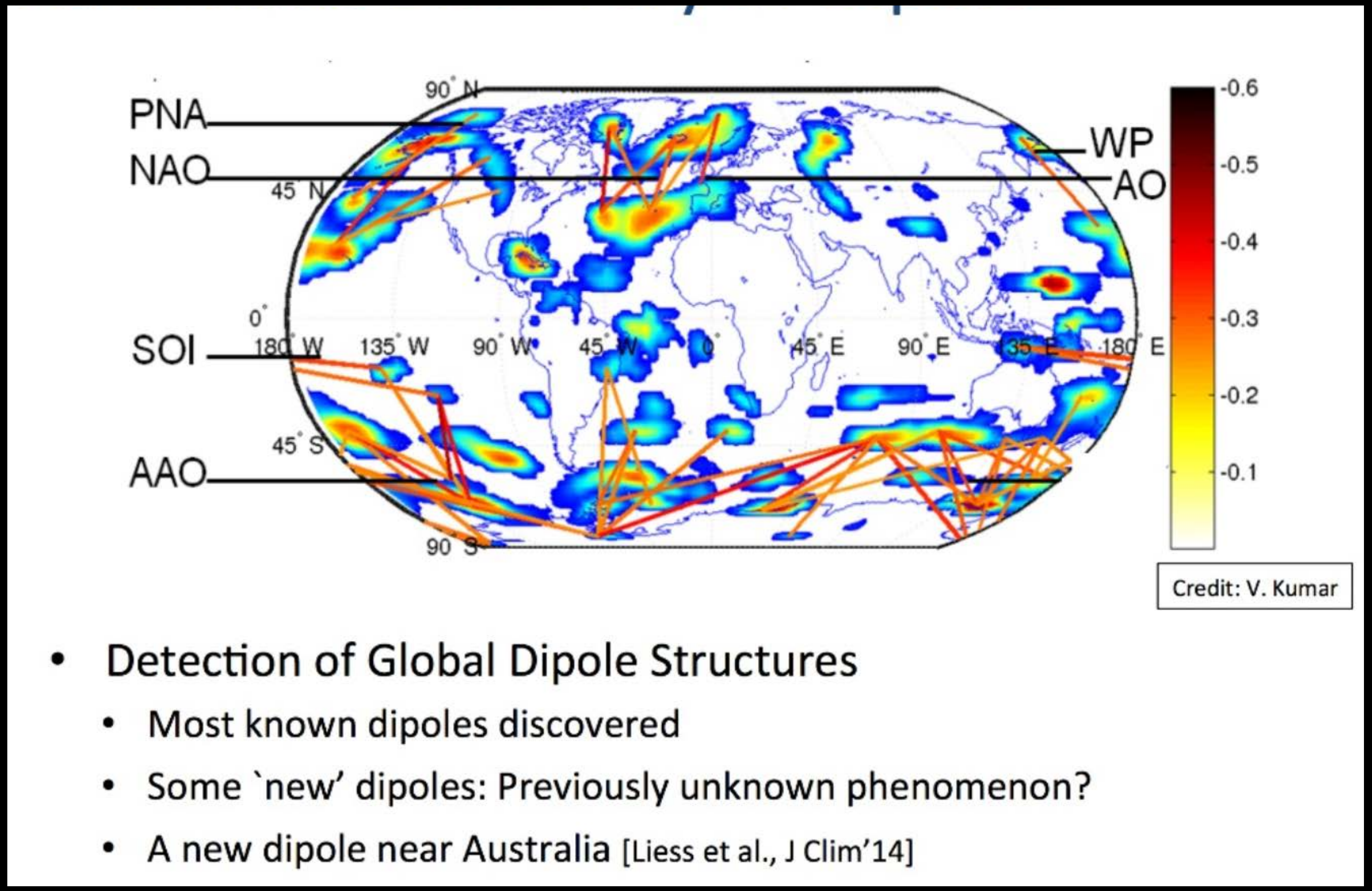
Deep Learning can often discover features to learn from the entire system

Examples of Deep Learning in Space Science

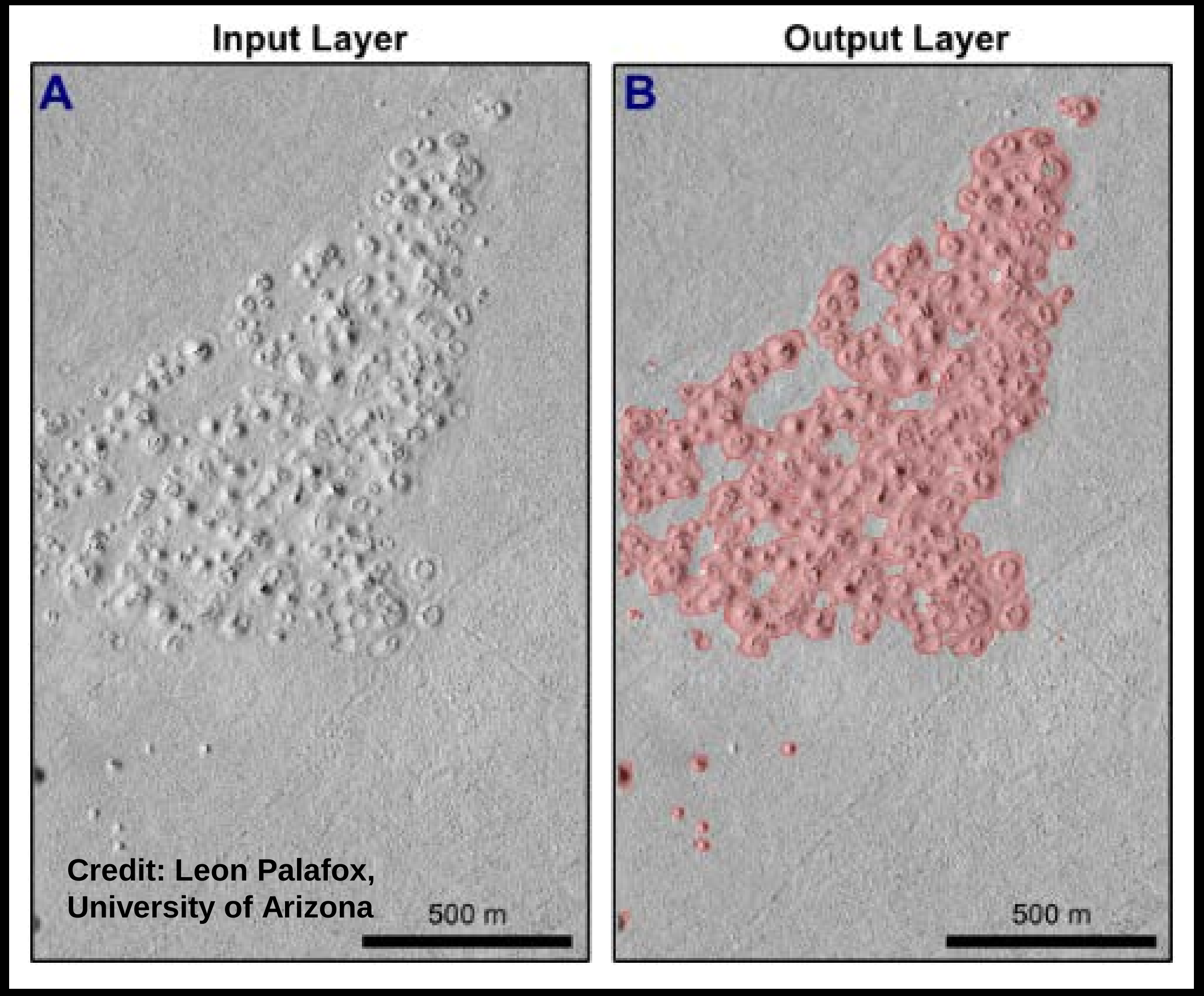


Kevin Schawinski et al, Generative Adversarial Networks recover features in astrophysical images of galaxies beyond the deconvolution limit, Royal Astronomical Society, 2017

Discovery of Dipoles using Neural Networks

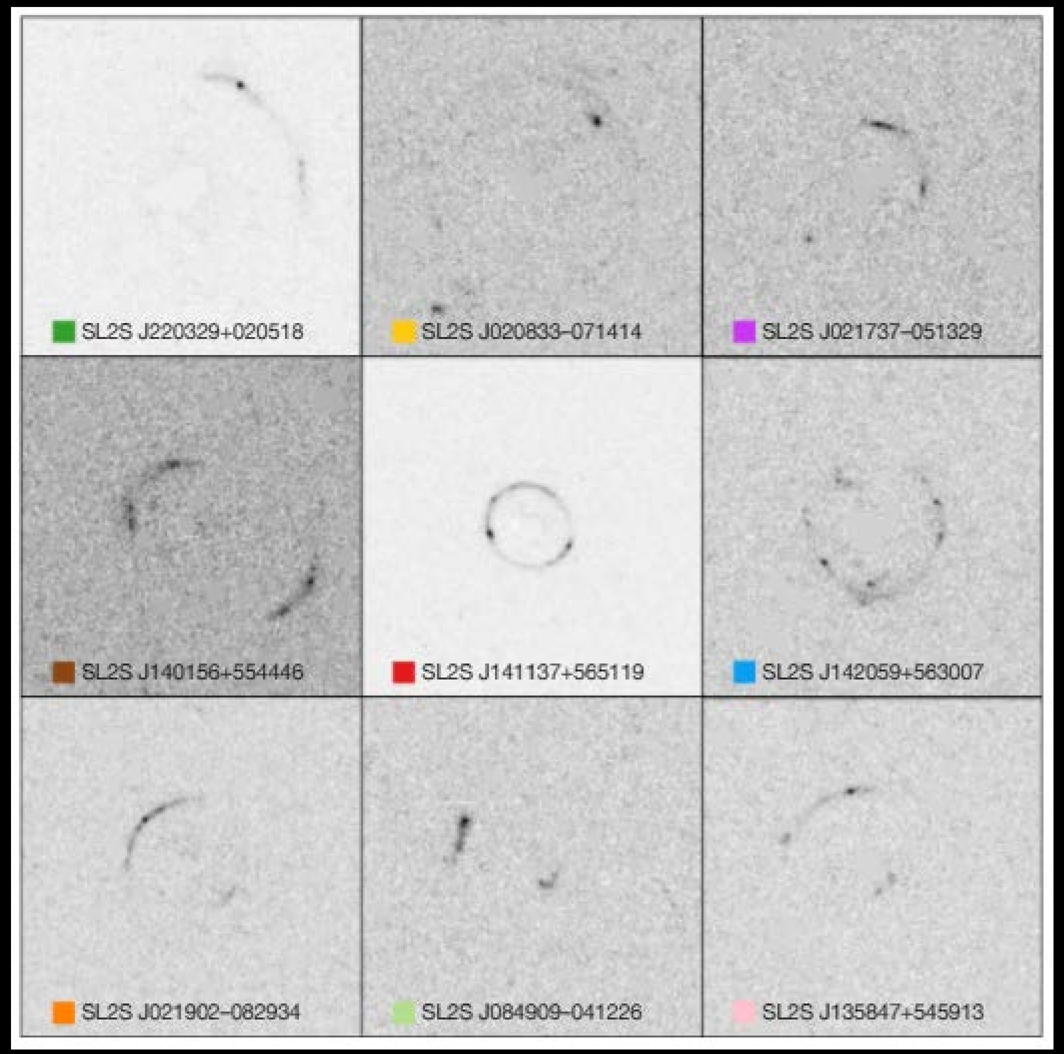


Neural Net Analysis of Mars HiRISE Images



Identification of Martian volcanic rootless cones within HiRISE images (96% classification accuracy)

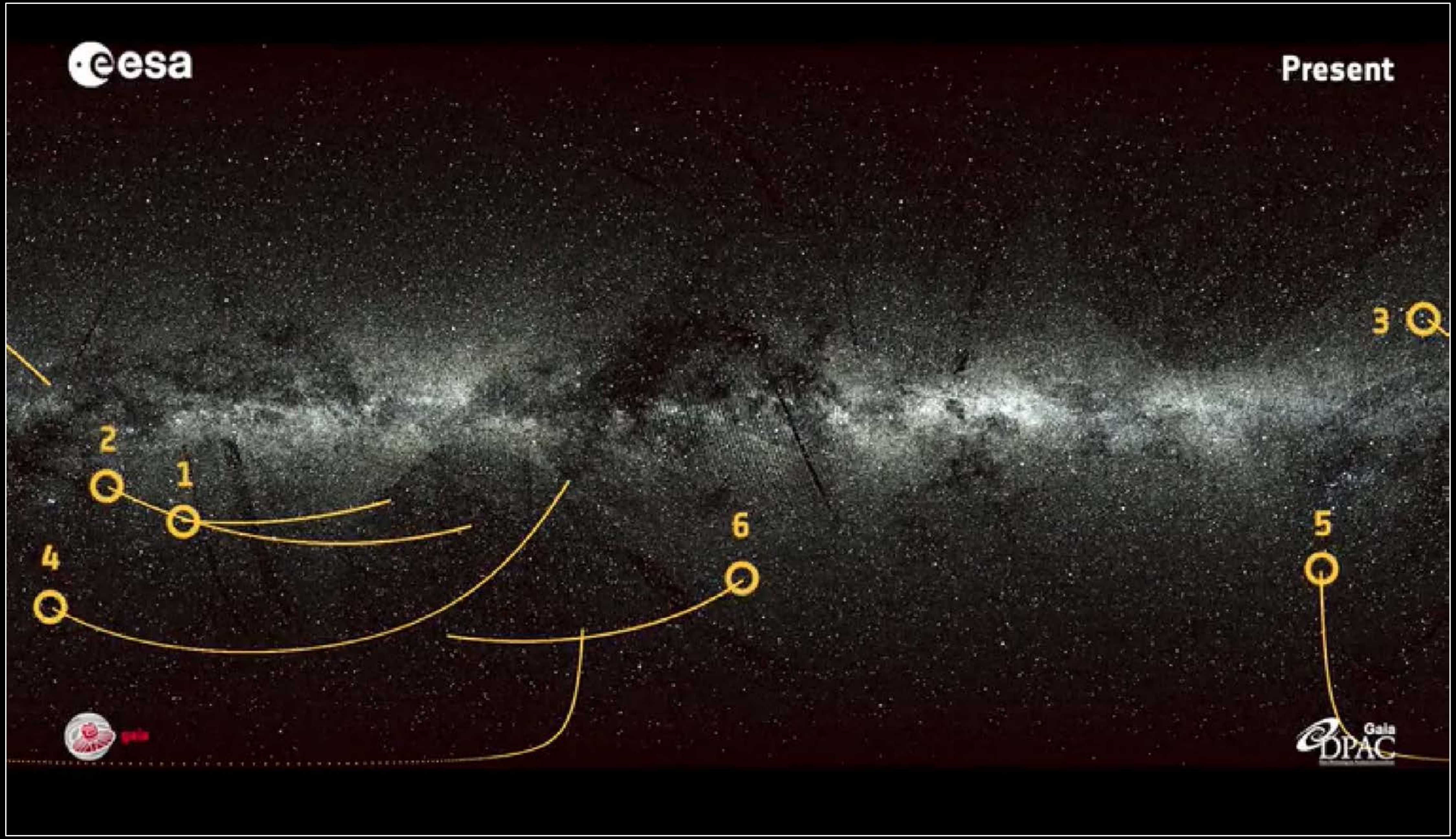
Neural Network discovery and analysis of gravitational lenses



Yashar D. Hezaveh et al. "Fast automated analysis of strong gravitational lenses with convolutional neural networks", *Nature,* Aug 2017

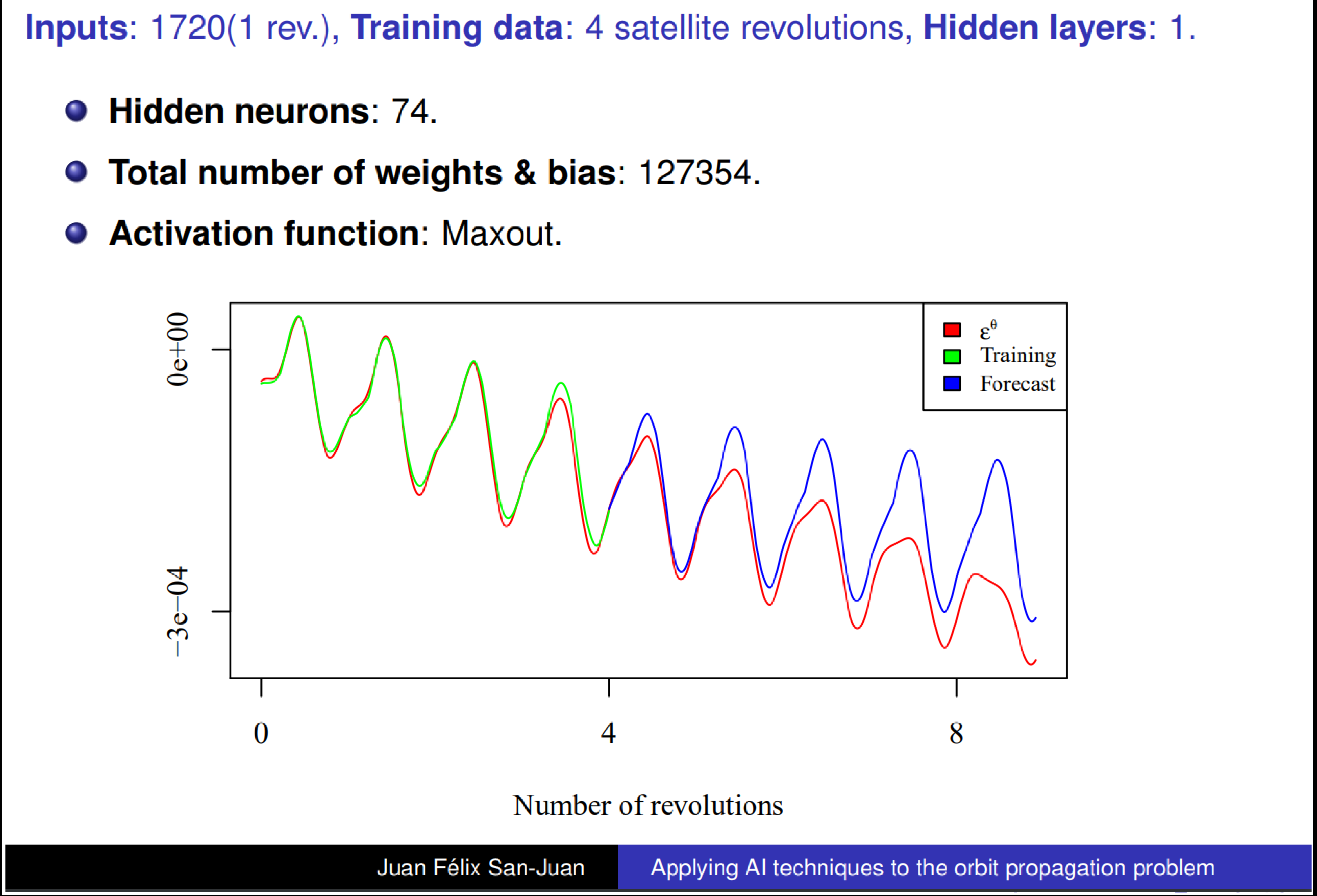
Examples of Deep Learning in Space Science

Deep Learning Discovery of Hypervelocity Stars



Elena Rossi, et al. Discovery of hypervelocity stars using an artificial neural network with ESA Gaia data, European week of Astronomy and Space Science, 2017

Applying Deep Learning AI techniques to the Orbit Propagation Problem



Juan Felix San-Juan, International Round Table on Intelligent Control for Space Missions November 24, 2017

AI & Deep Learning at NASA

- Some Deep Learning exploratory projects are underway at NASA. Examples…
  - NASA DeepSAT: A Deep Learning Approach to Tree-Cover Delineation in 1-m NAIP Imagery. (S. Ganguly, AGU 2016)
  - Anomaly detection in aviation data using extreme learning machines. (V. Manikandan, et al. International Joint Conference on Neural Networks, 2016)
  - Multi-Objective Reinforcement Learning-Based Deep Neural Networks for Cognitive Space Communications. (P. Ferreria, et al. NASA/TM–2017)

… but more experience is needed in order to establish an overarching strategy.

- FDL provides a low-risk / low-cost mechanism for NASA to move forward:
  - Program is managed by the SETI Institute, but with NASA guidance on the problem definitions
  - Private sector partnerships provide infrastructure, resources and much of the funding
  - NASA experts participate, learn, and observe best practice: allows NASA's strategy for AI to move forward in a more informed manner

"Frontier Development Lab is proving its value at training early career professionals/students to apply modern data science techniques to sticky analysis problems confronting NASA science and exploration programs. […] The BDTF finds that this type of program aligns with its recommendations to NASA that there needs to be more formal, long term education as well as more short-form workshops dedicated to introducing modern data science methodologies as approaches for improving the discoveries in its vast science data archives."

Source: Final Report of the Big Data Task Force, NASA Advisory Council Science Committee, 2017. https://science.nasa.gov/science-committee/subcommittees/big-data-task-force

NASA FRONTIER DEVELOPMENT LAB

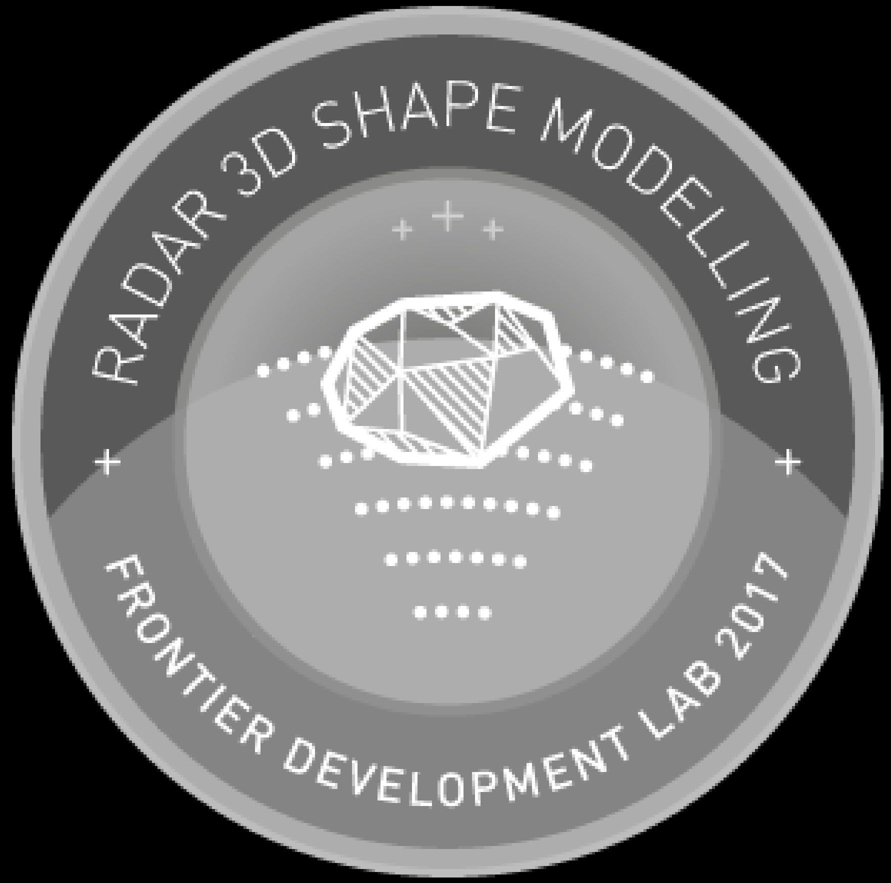
- PROGRAM STRUCTURE
- RESULTS & PROGRESS
- FUTURE PLANS

3 projects in 2016

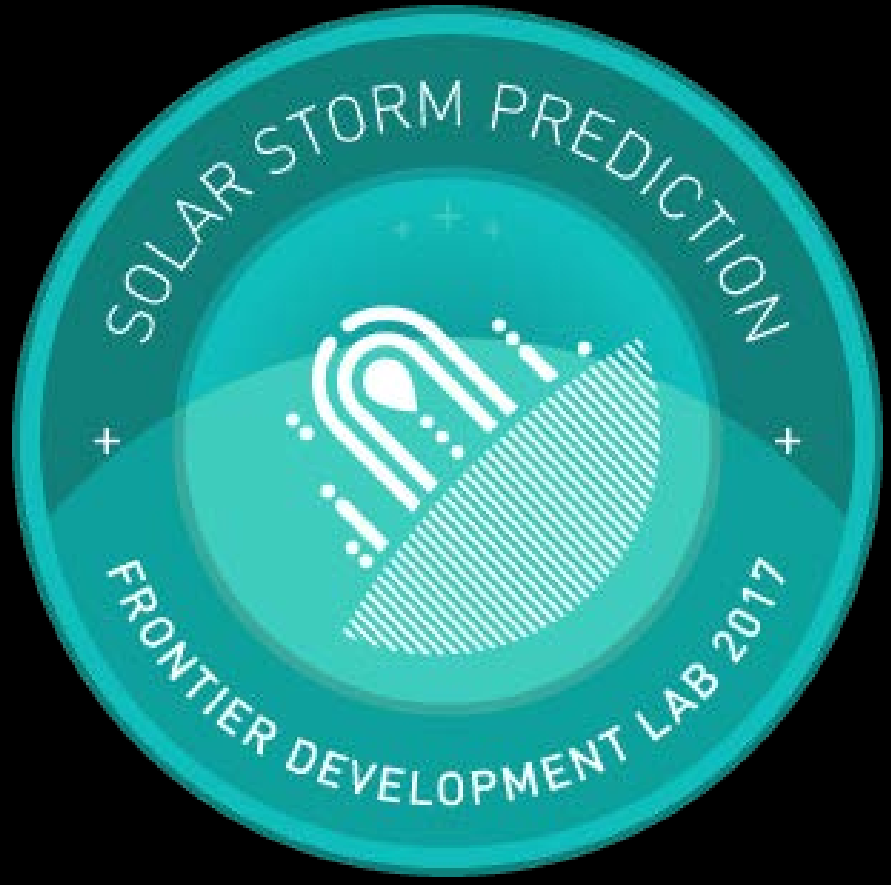
6 projects in 2017

12 projects being assessed for 2018

PLANETARY DEFENSE

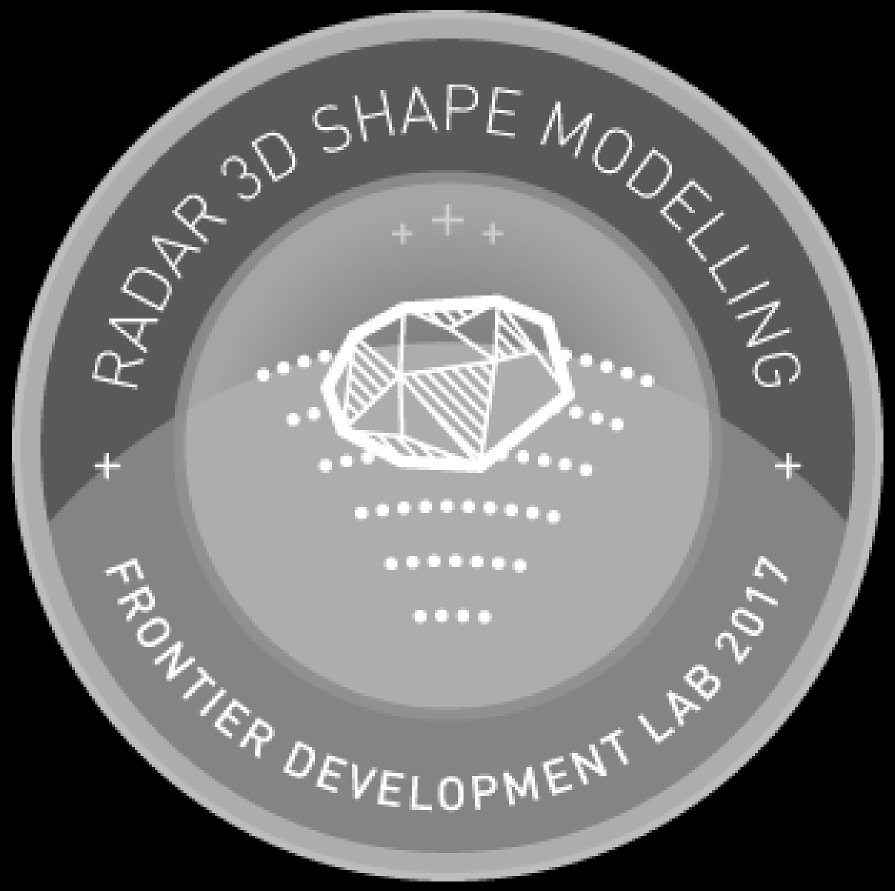


HELIOPHYSICS

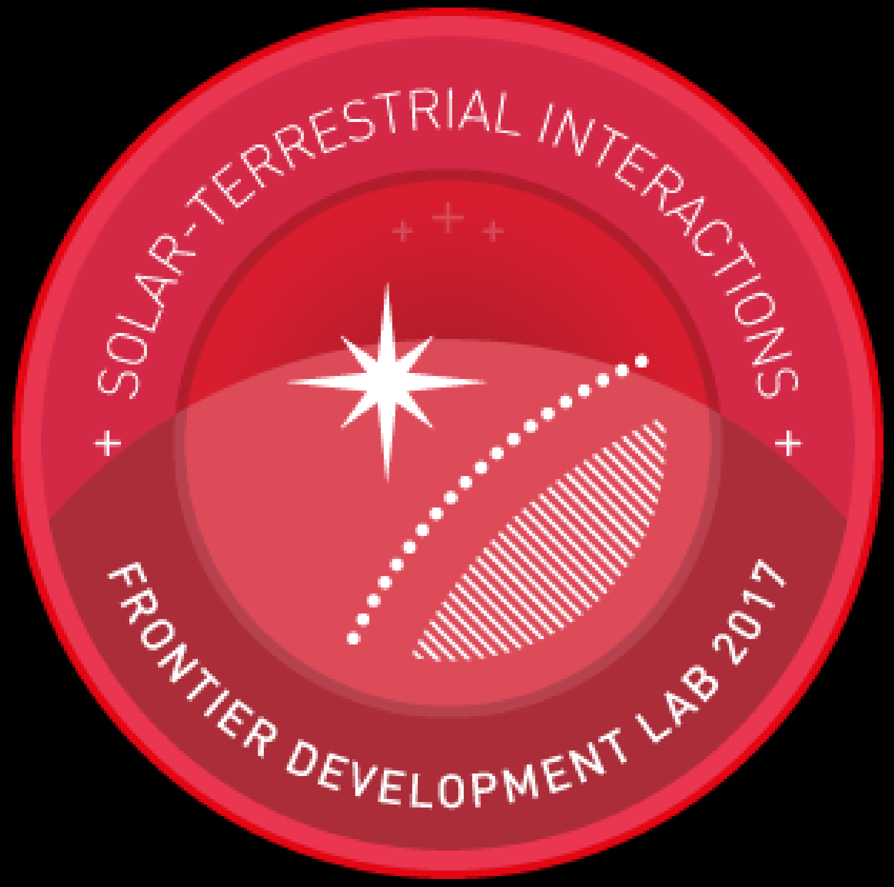


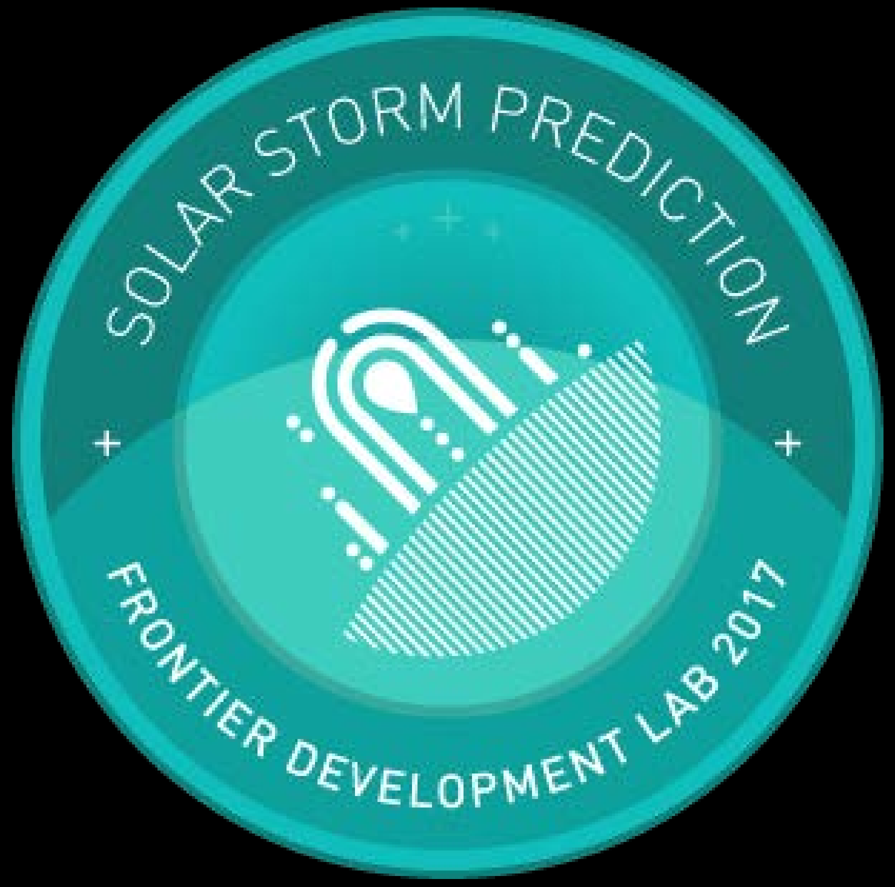
ADJANCENT BUT RELATED PROBLEM DOMAINS

PLANETARY DEFENSE



HELIOPHYSICS





ALLOWS USEFUL OVERLAP OF EXPERTISE AND TALENT


LONG-PERIOD
COMETS
dkolyer@seti.org
SOLAR
July 5, 2017
Irradiance
Radio burst
Cosmic ray flux
solar wind
Mean magnetic field (Solar)
Already timeseries
Terrestrial
Cloud
Earthquakes
Hurricanes

IBM's Executive Project Manager briefs the FDL team on the compute resource available for each team.

Google’s Francois Chollet - inventor of the Keras.io framework briefs the FDL team. (Python for machine learning.)

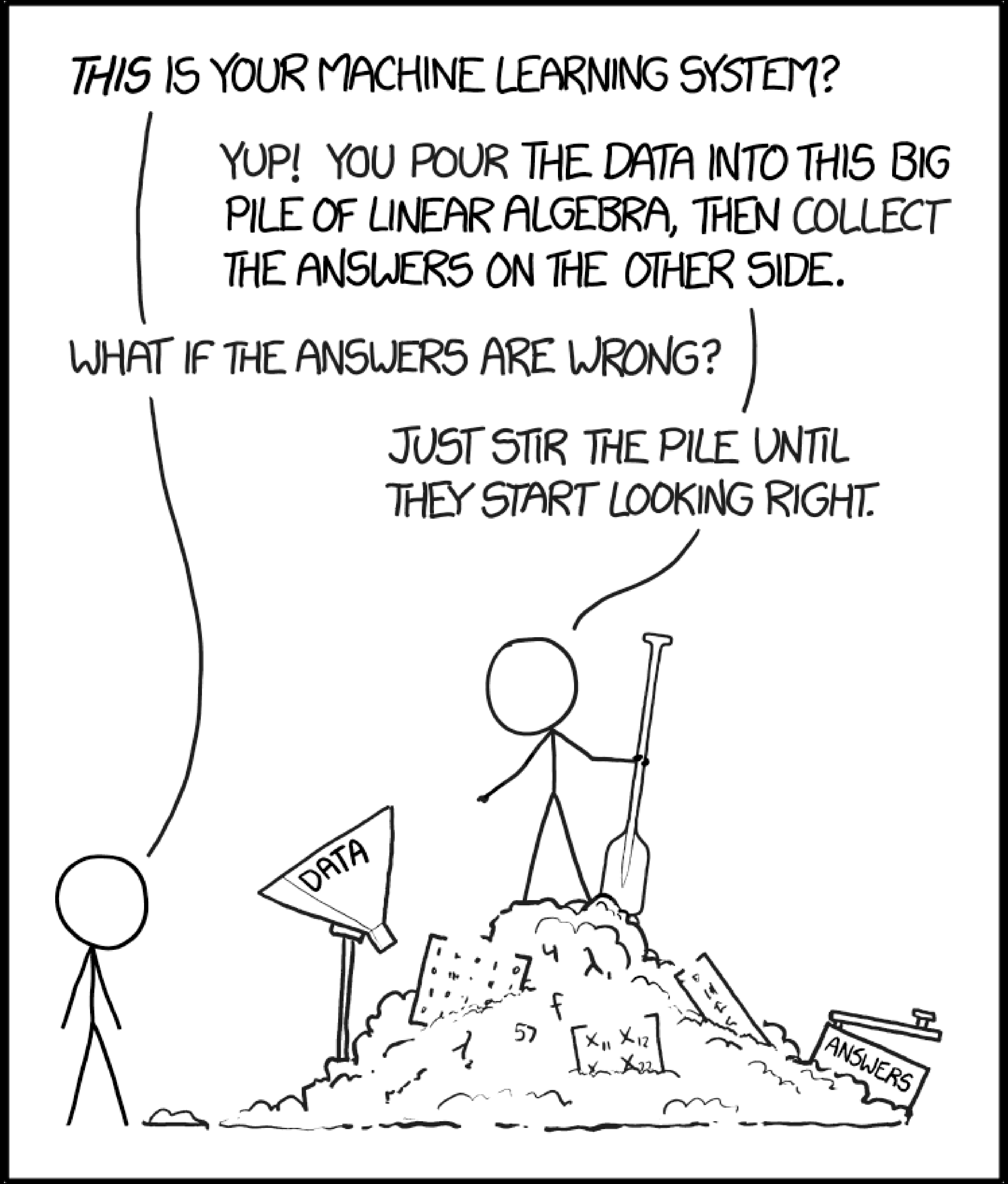
- **Current operational flare forecasting relies on human morphological analysis of active regions and the persistence of solar flare activity.**

- **The FDL team performed analyses of solar magnetic complexity and deployed convolutional neural networks to connect solar UV images taken by SDO/AIA into forecasts of maximum x-ray emissions.**

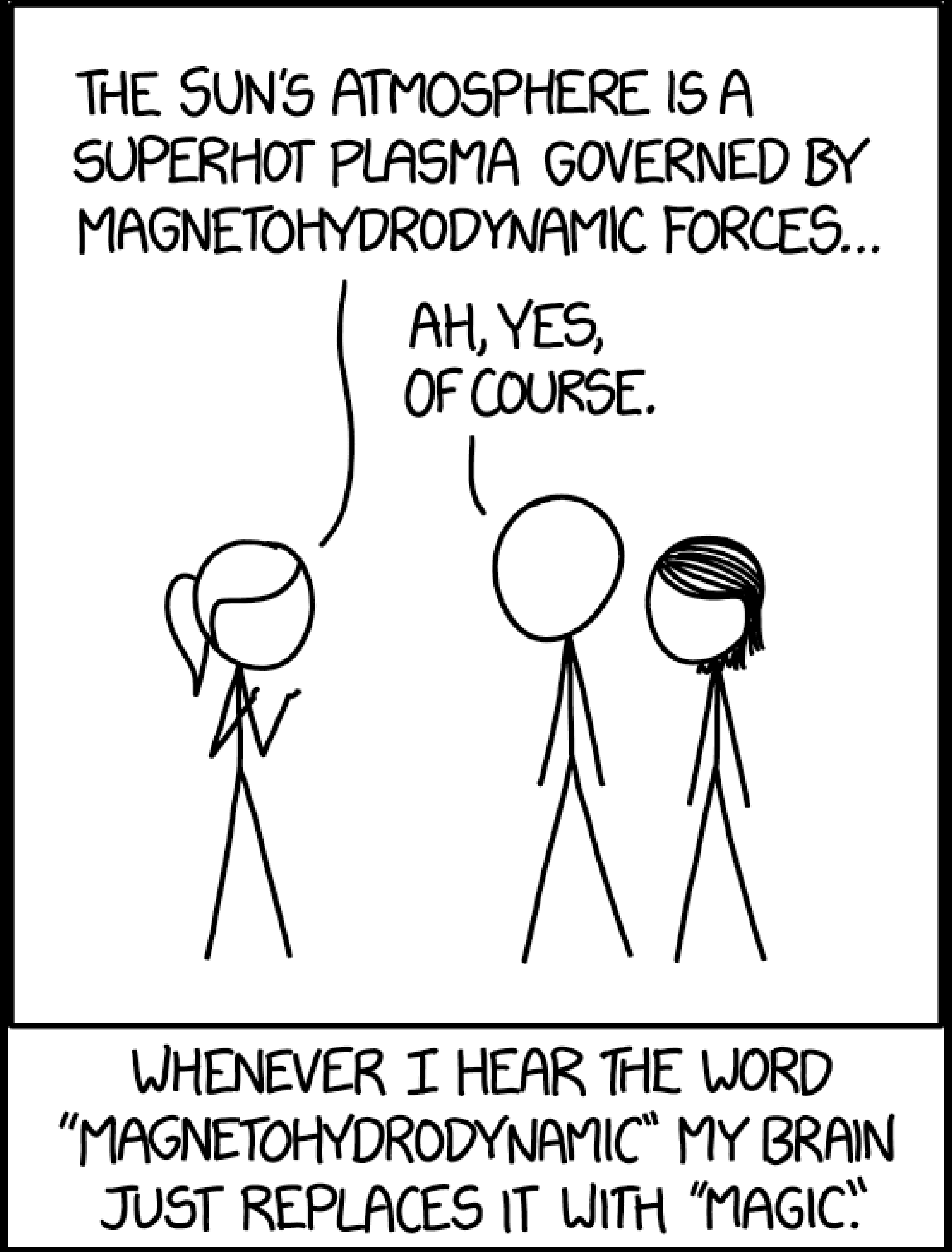
- **The technique has the potential to improve both the reliability and accuracy of solar flare predictions.**

SPACE WEATHER: SOLAR STORM PREDICTION

Interdisciplinary Collaboration



Heliophysicist's view of ML



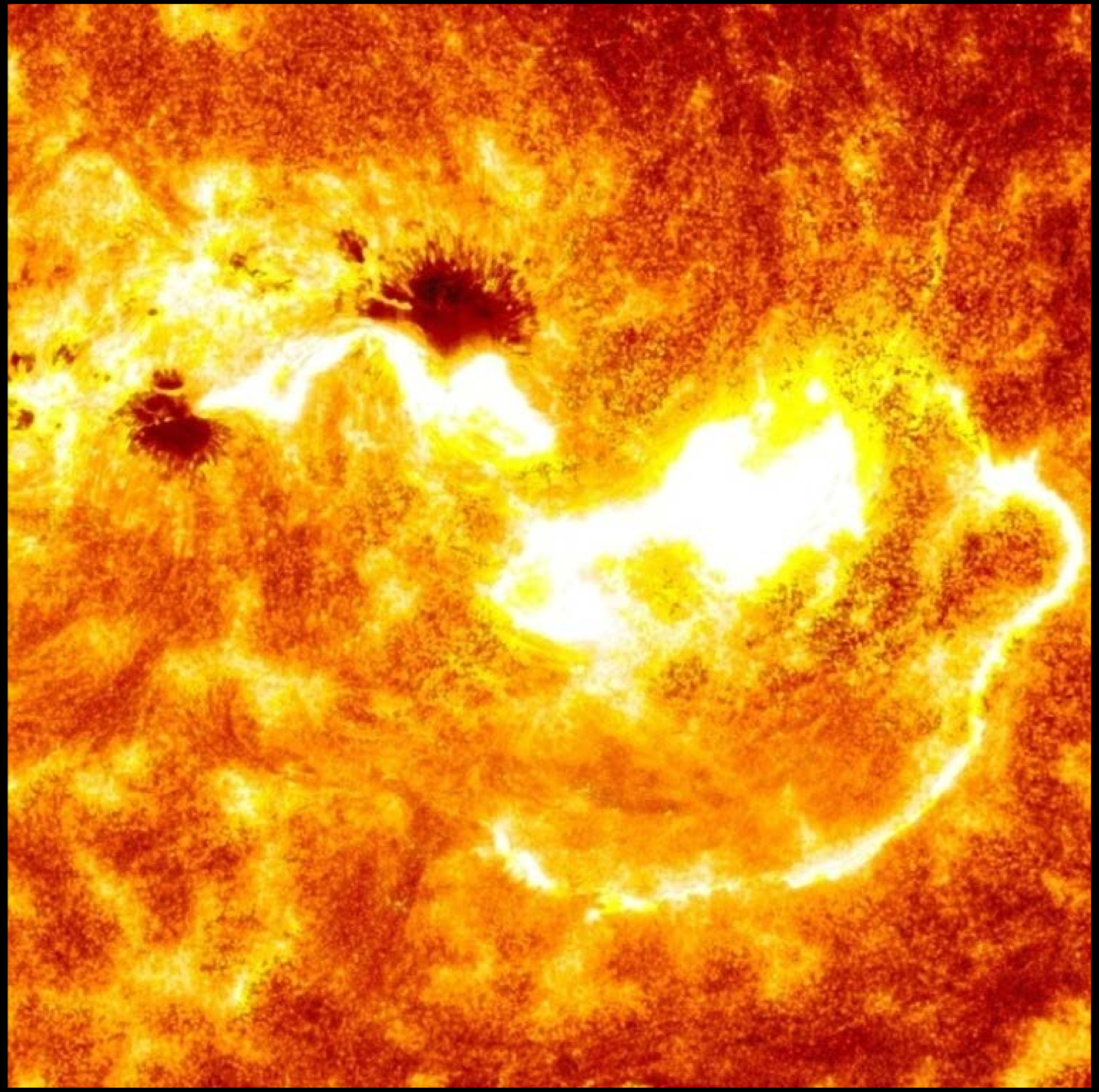
Data scientist's view of HP

SOLAR STORM PREDICTION
FRONTIER DEVELOPMENT LAB 2017

SPACE WEATHER: SOLAR STORM PREDICTION

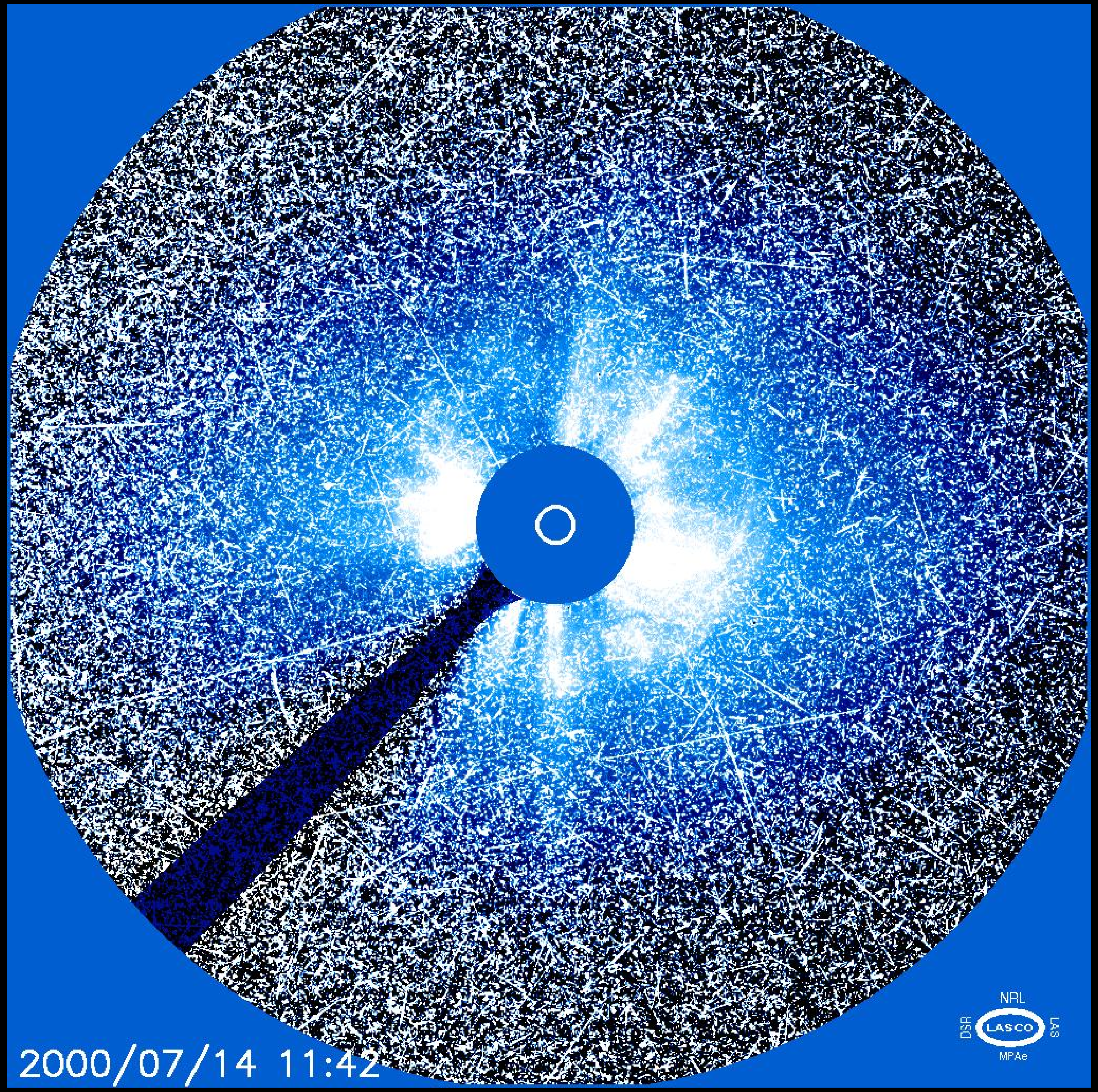
Types of Space Weather

FLARES



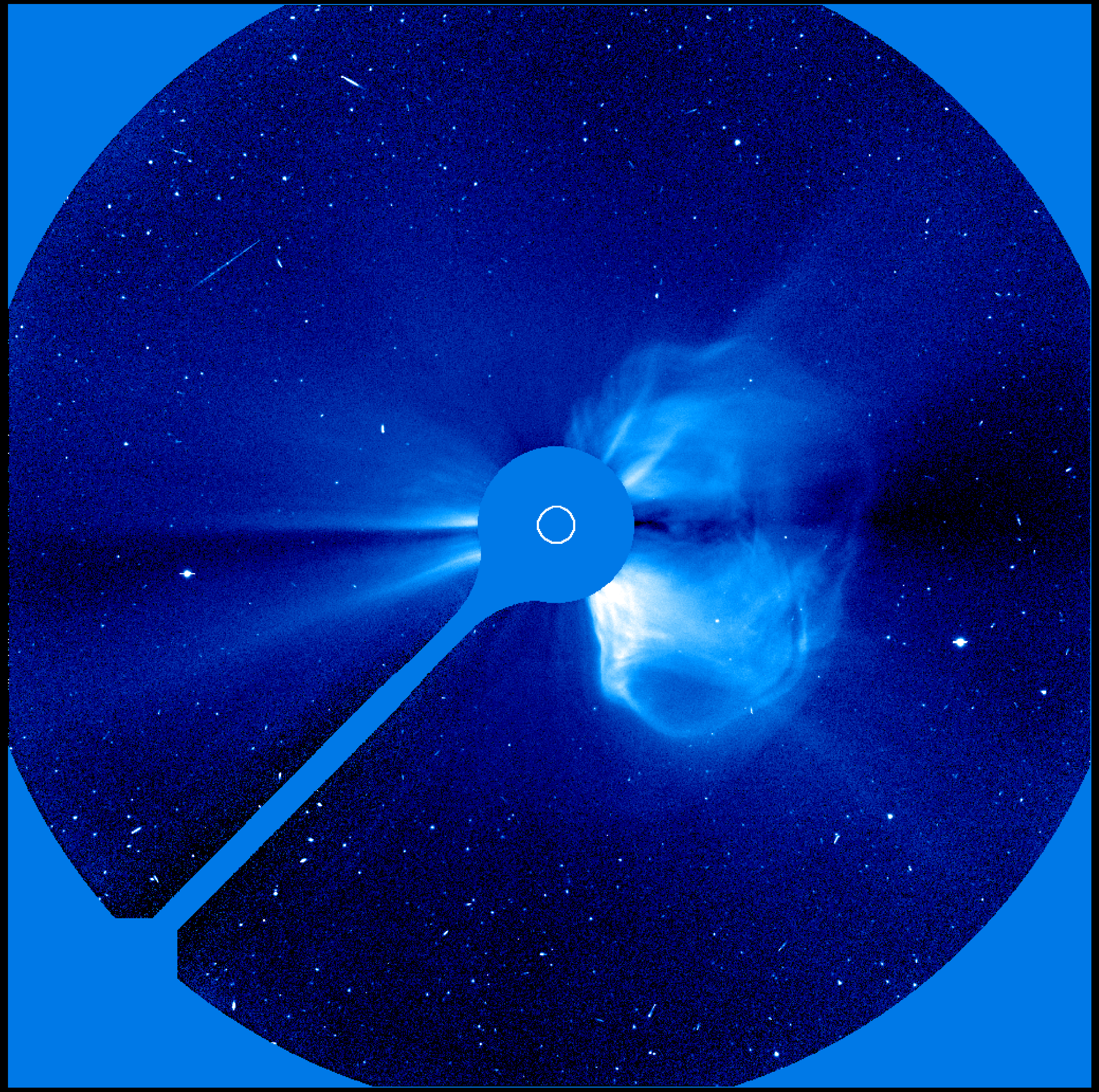
Electromagnetic Radiation

ENERGETIC PARTICLES



Particle Radiation

MASS EJECTIONS



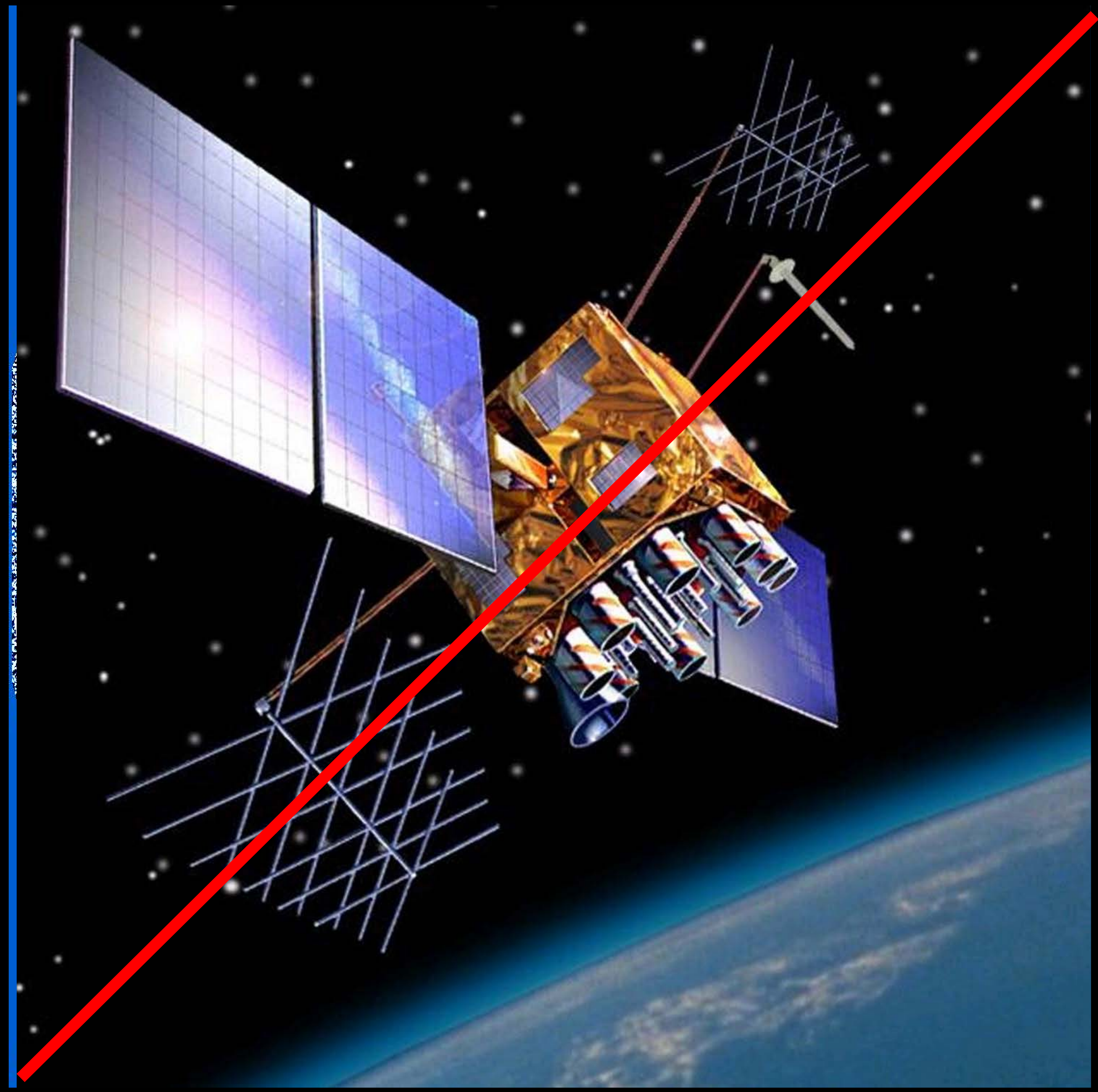
Massive Magnetic Ropes

Types of Space Weather

FLARES

Disruption of Communications

ENERGETIC PARTICLES



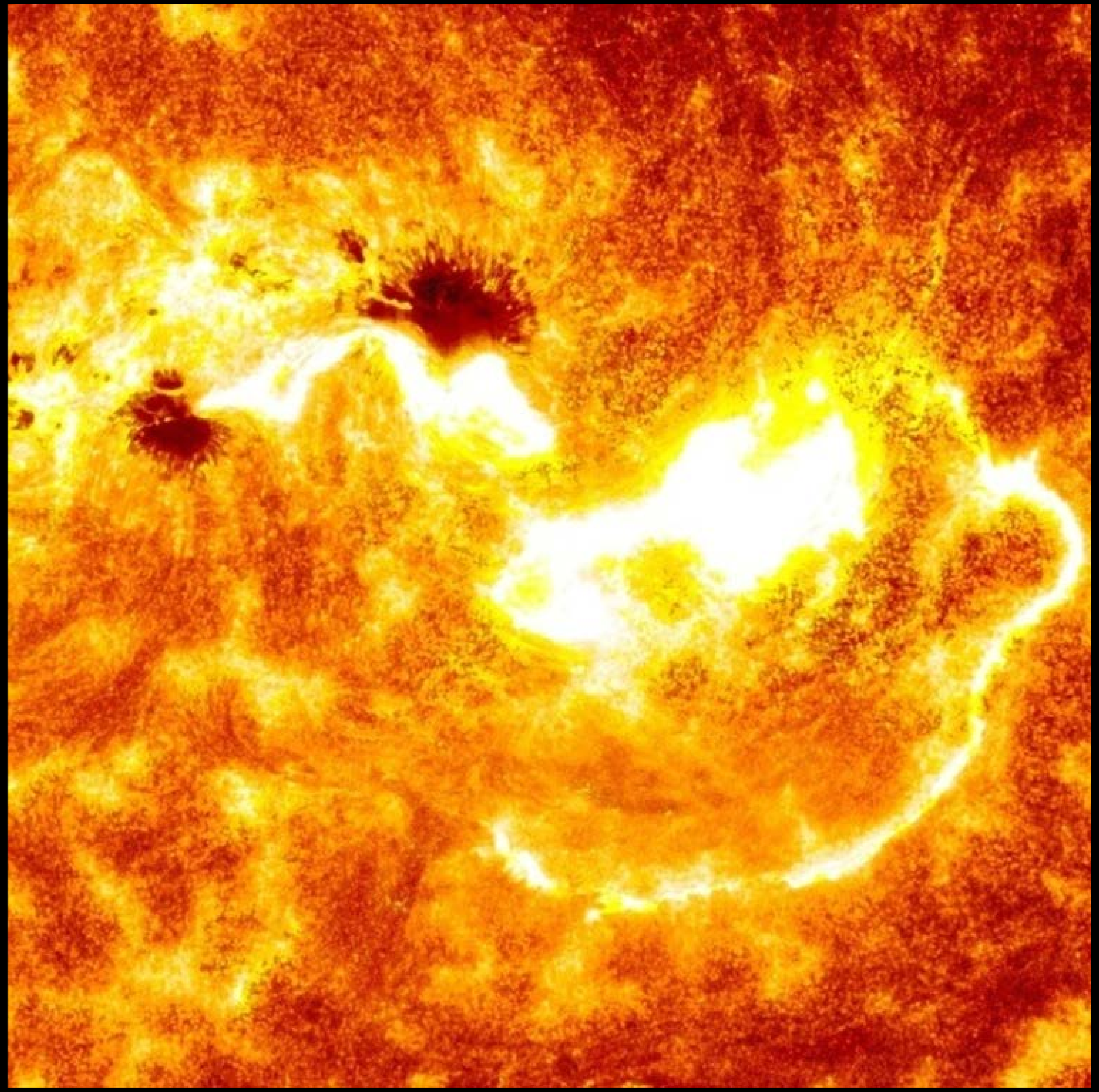
Satellite Damage

MASS EJECTIONS

Power grid Disruption

Why Solar Flare prediction is important?

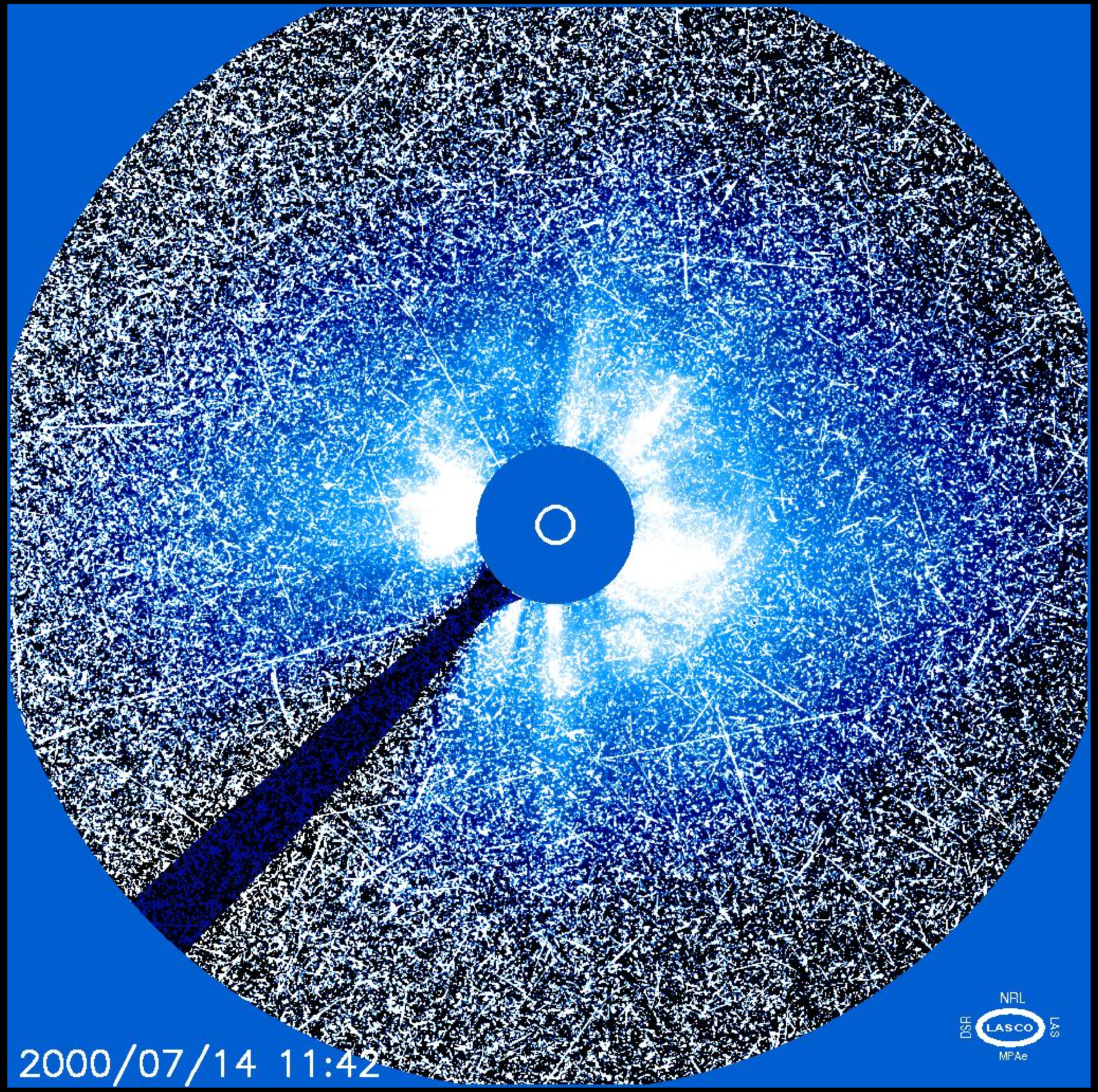
FLARES



Speed of Light

No warning

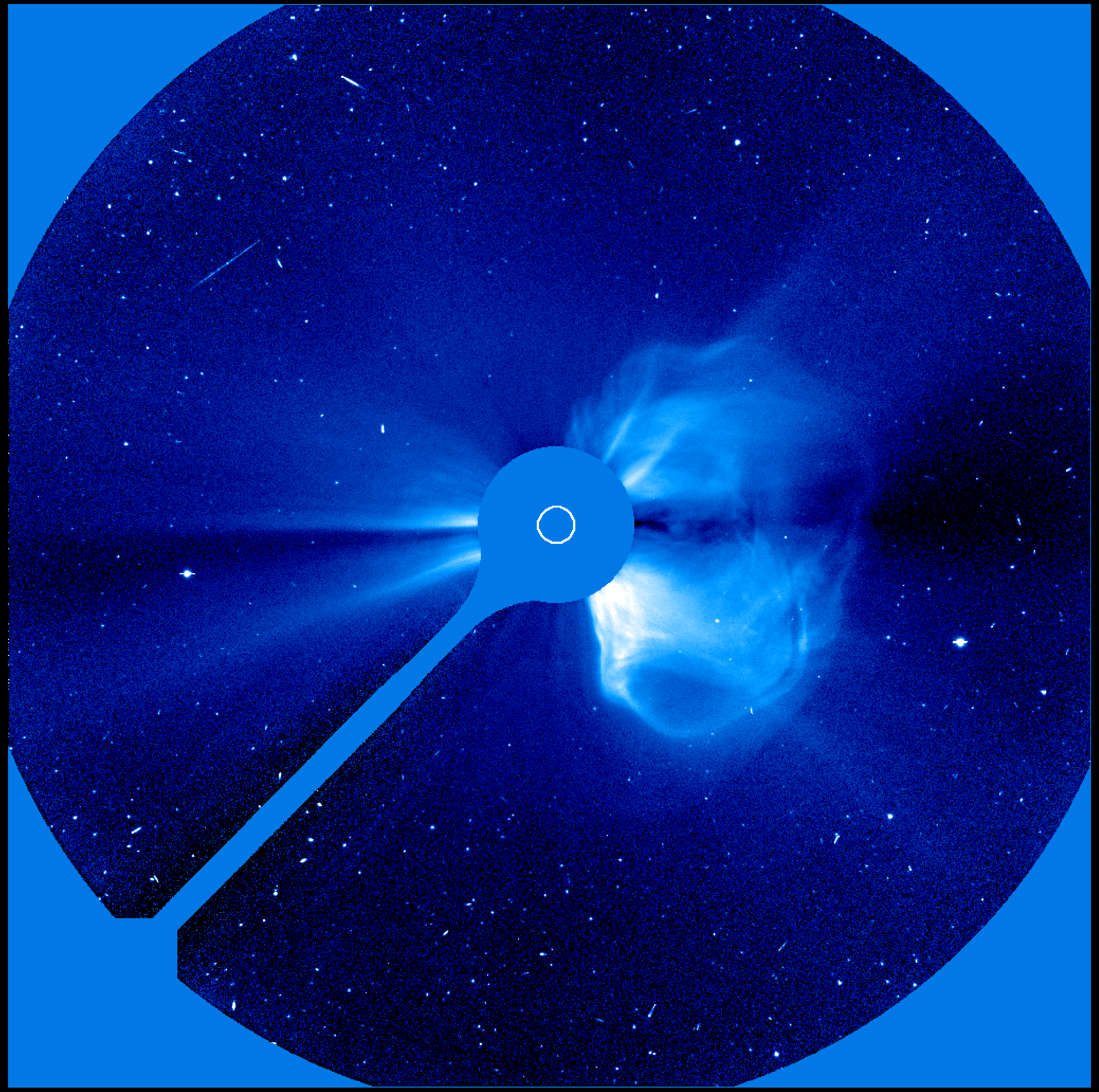
ENERGETIC PARTICLES



Relativistic speeds

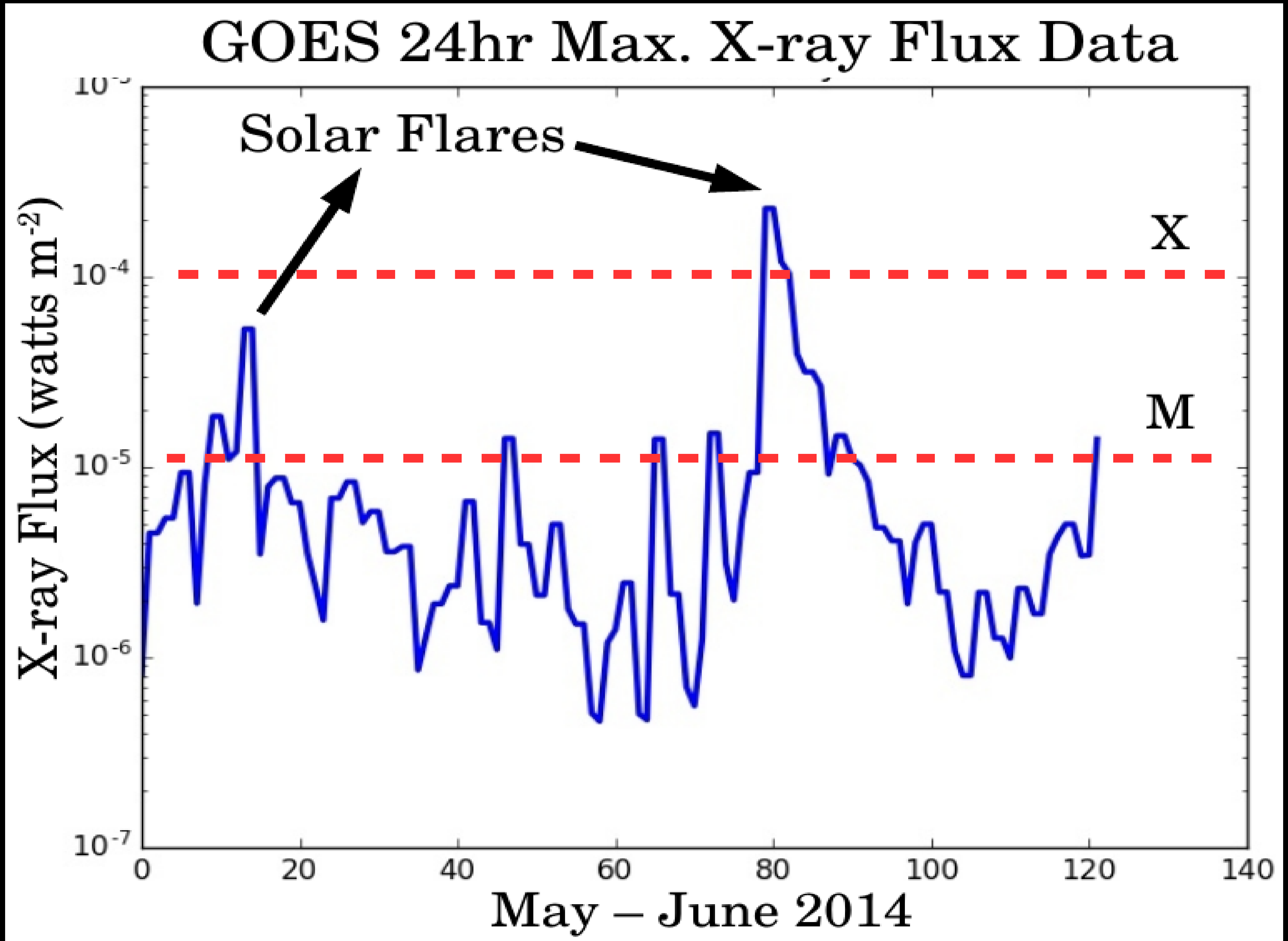
20 minute warning

MASS EJECTIONS



20 hour warning

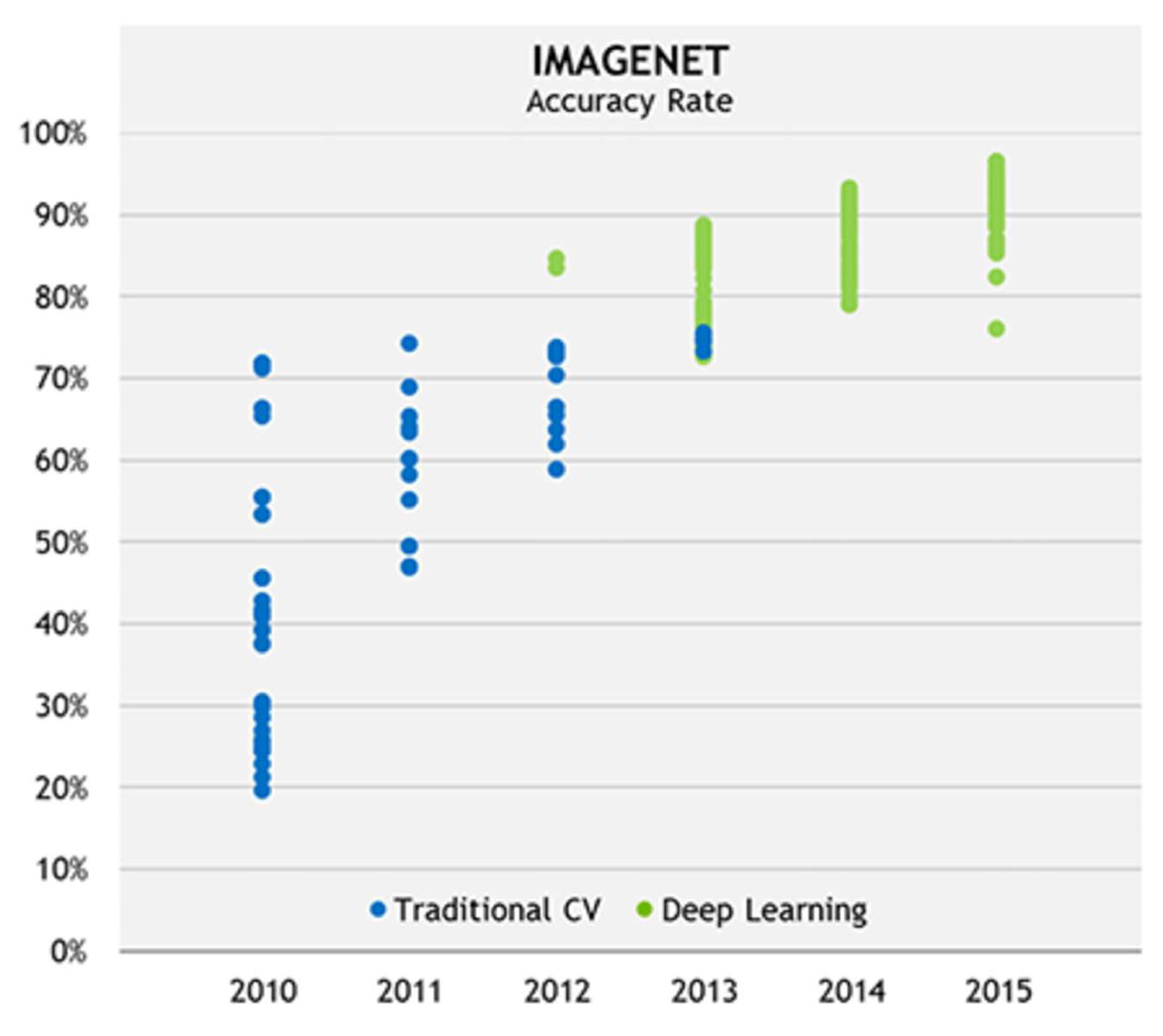
How is a flare defined?



Using X-ray flux as measured by the GOES satellite

SPACE WEATHER: SOLAR STORM PREDICTION

Deep Learning

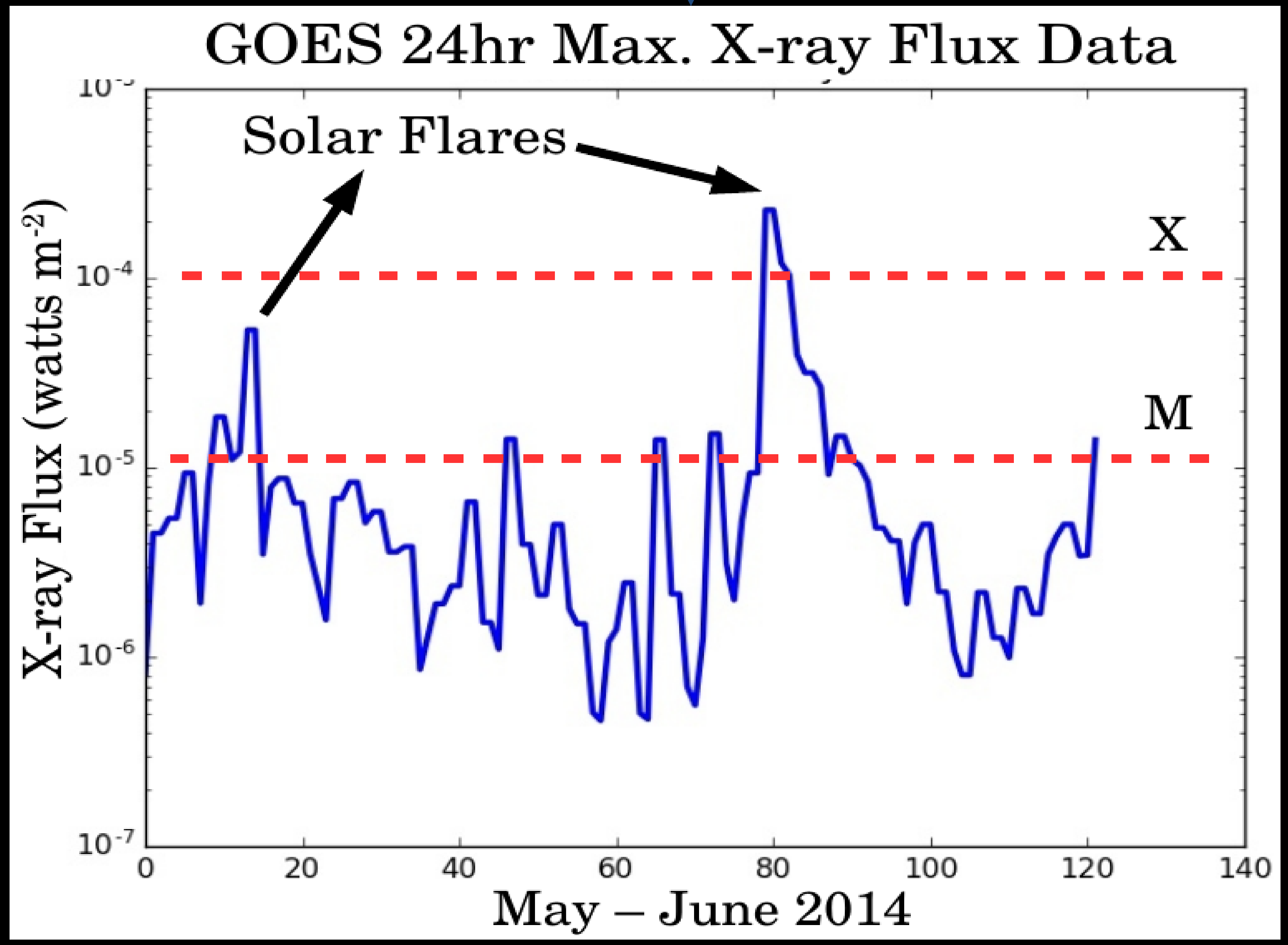


Deep learning has revolutionized the way we do image classification.

Target Breakthroughs

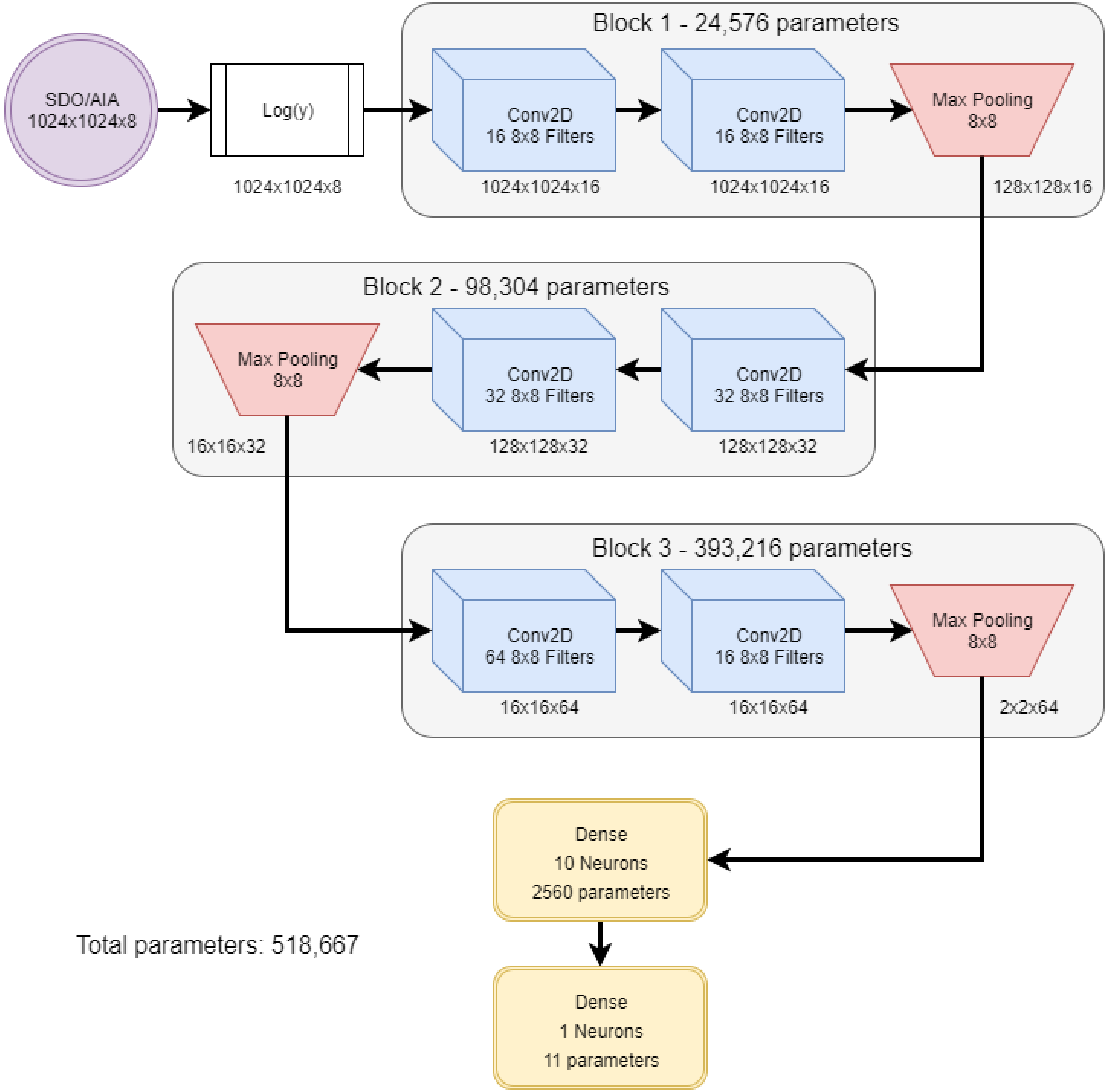
1. Dataset Preparation: *Take advantage of big data*
2. Software: *Build scientific process*
3. Prediction: *Enable Flare Forecasting*
4. Science: *Visualize Results*
    - Discover Flare Precursors
- Providing new physical insight
- New Physics?

SDO/AIA Image Channels



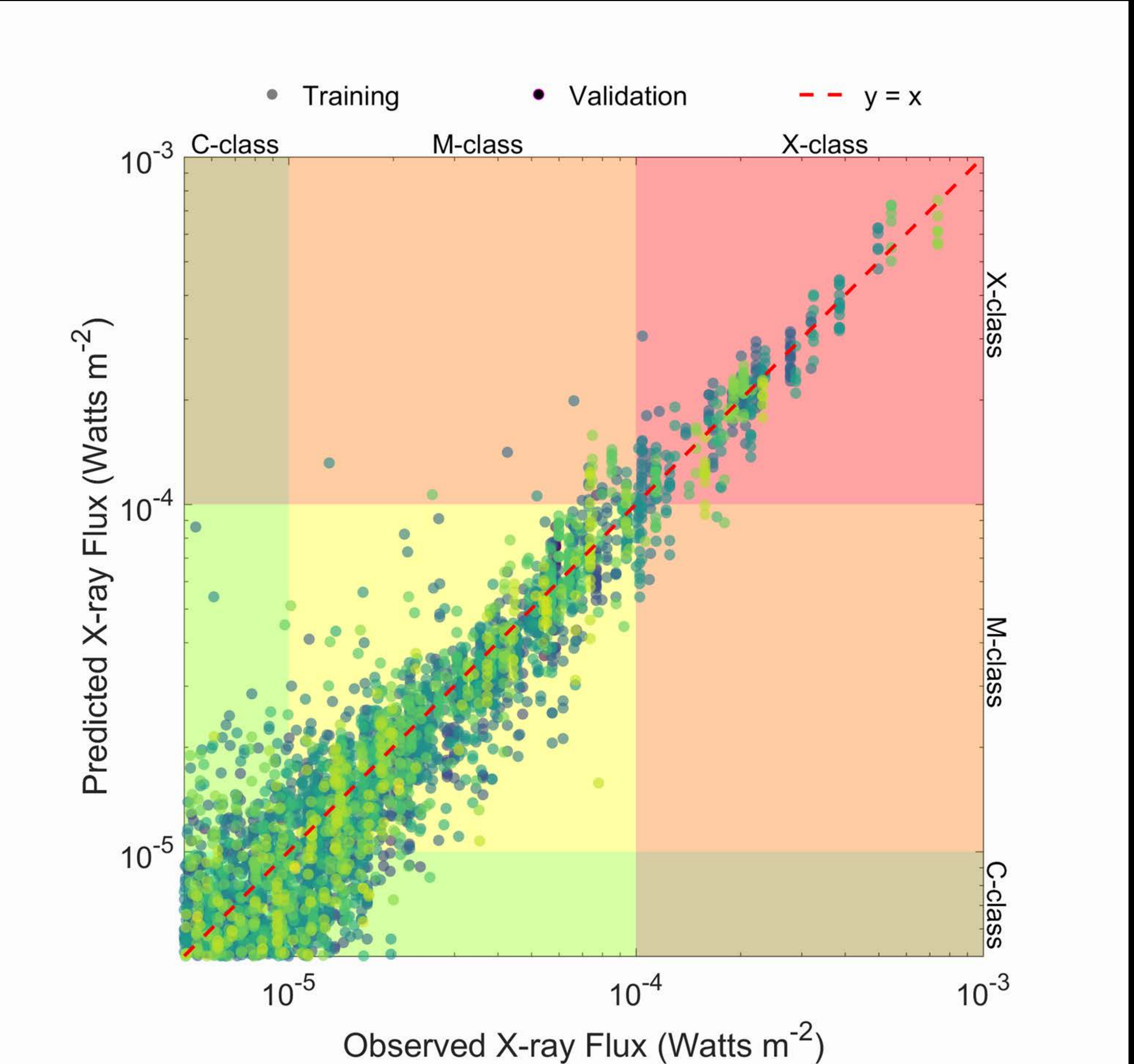
Can we use deep learning to connect AIA images with flare strength?

FlareNet


SDO/AIA
1024x1024x8
Log(y)
1024x1024x8
Block 1 - 24,576 parameters
Conv2D
16 8x8 Filters
1024x1024x16
Conv2D
16 8x8 Filters
1024x1024x16
Max Pooling
8x8
128x128x16
Block 2 - 98,304 parameters
Conv2D
32 8x8 Filters
128x128x32
Conv2D
32 8x8 Filters
128x128x32
Max Pooling
8x8
16x16x32
Block 3 - 393,216 parameters
Conv2D
64 8x8 Filters
16x16x64
Conv2D
16 8x8 Filters
16x16x64
Max Pooling
8x8
2x2x64
Dense
10 Neurons
2560 parameters
Dense
1 Neurons
11 parameters
Total parameters: 518,667

Memorization vs. Generalization

All flares used for training

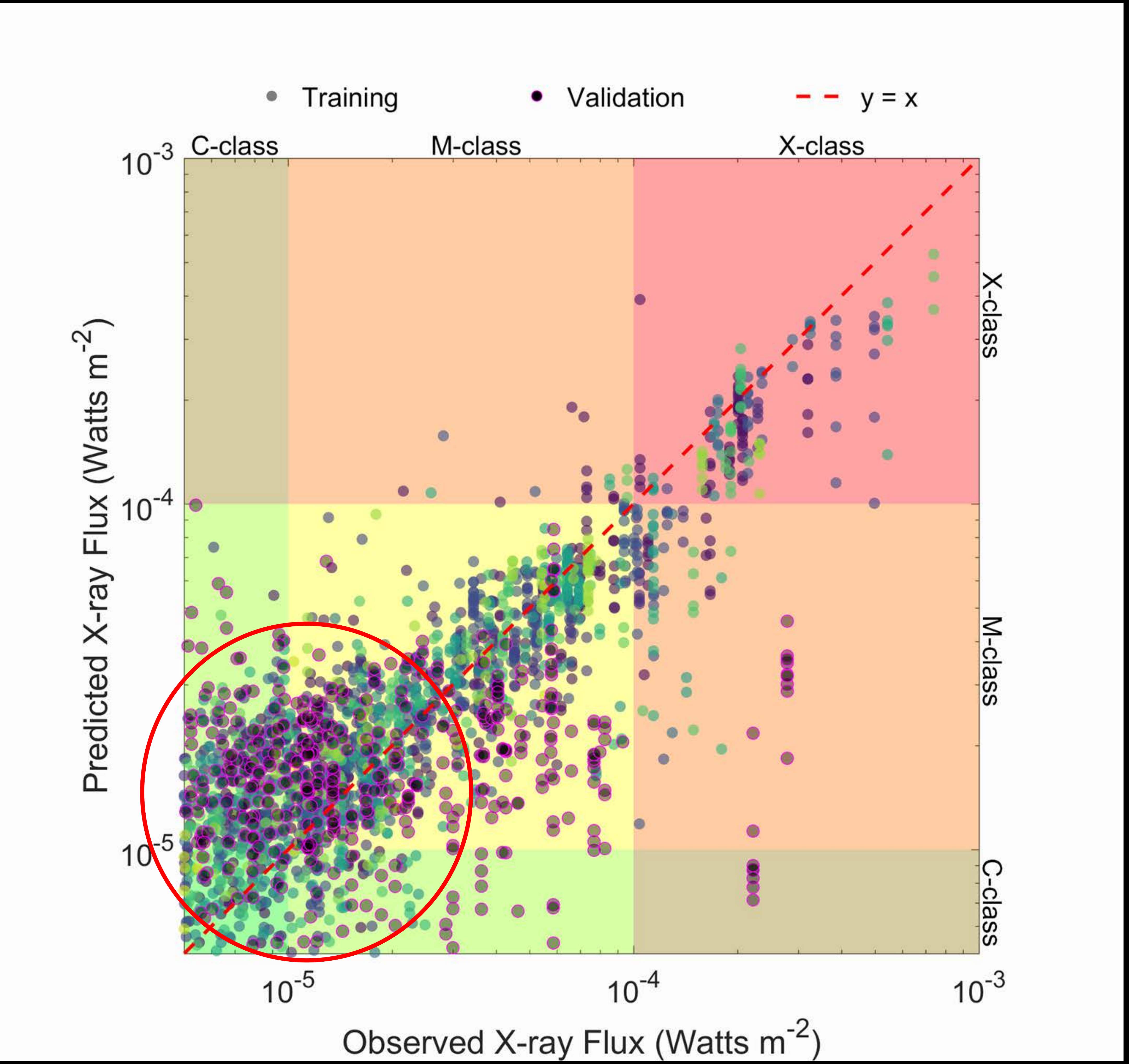


Our first goal was to see if the neural network could connect AIA images with flare X-ray amplitude.

The concern is whether the neural network is simply memorizing the images.

SPACE WEATHER: SOLAR STORM PREDICTION

Memorization vs. Generalization

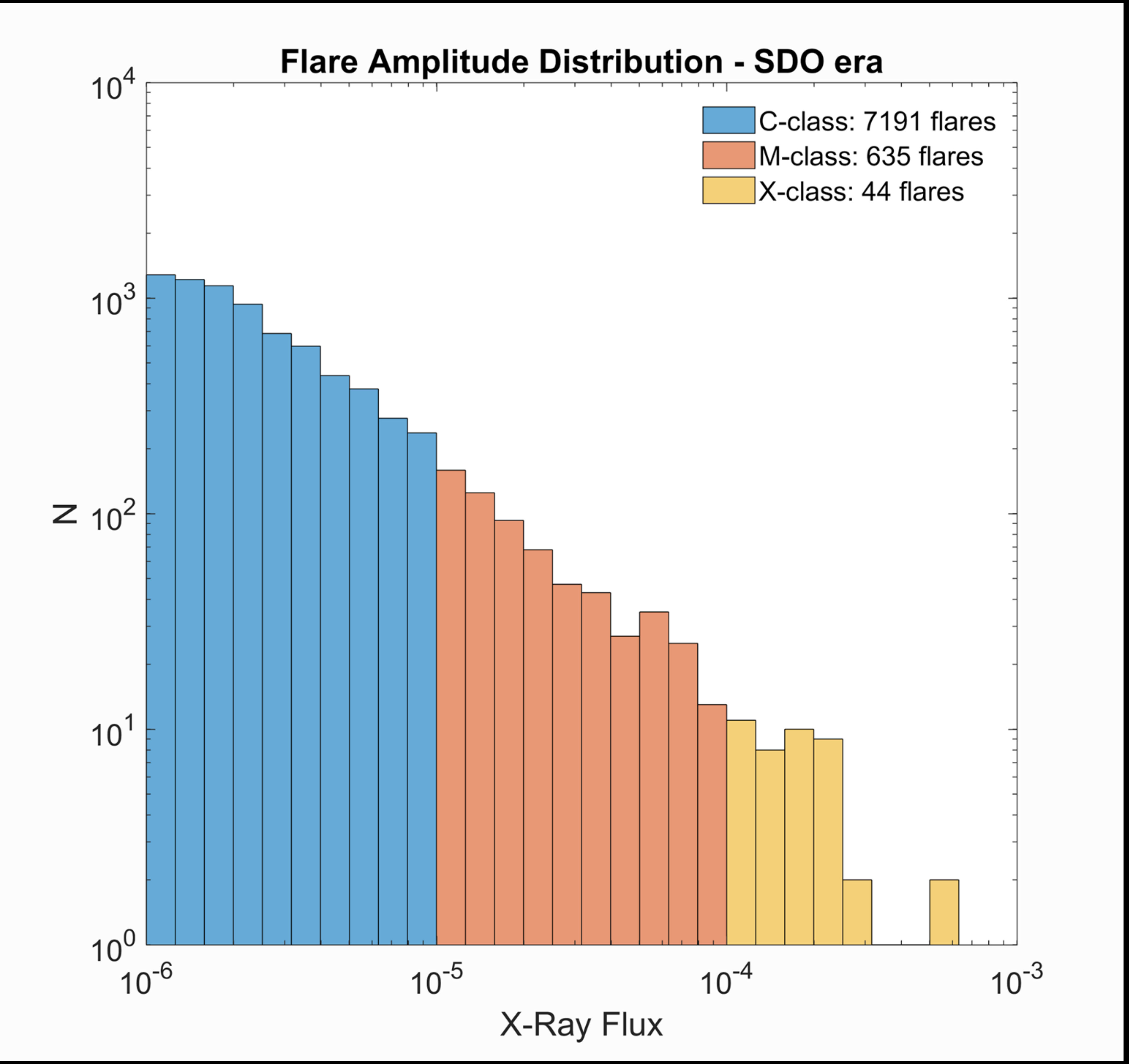


Only flares observed
prior to 2015
used for training

Our current neural network seems to be able to generalize for weak flares (C-class), but not yet for stronger flares .

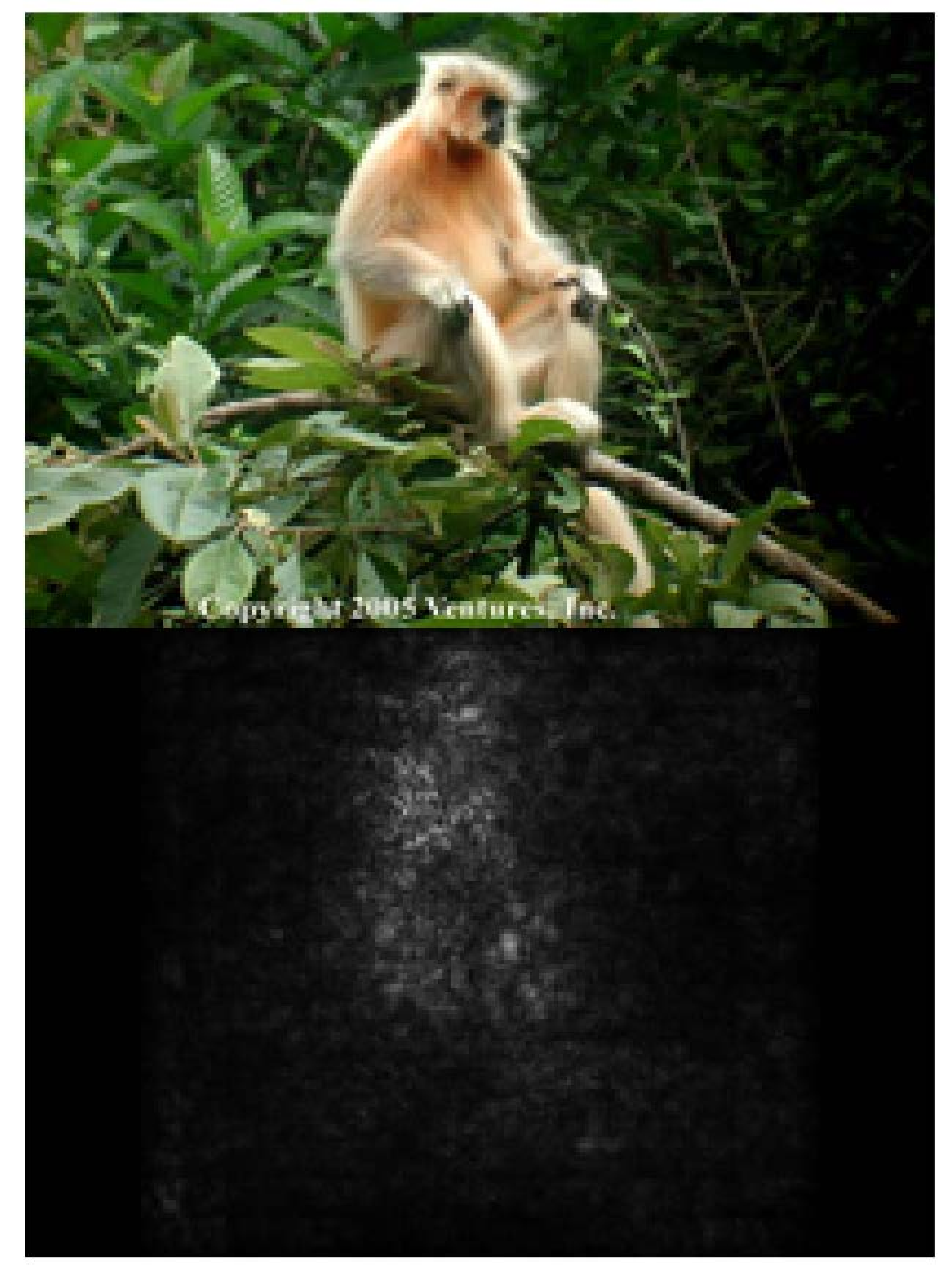
SPACE WEATHER: SOLAR STORM PREDICTION

Memorization vs. Generalization



Our current biggest challenge is class imbalance!

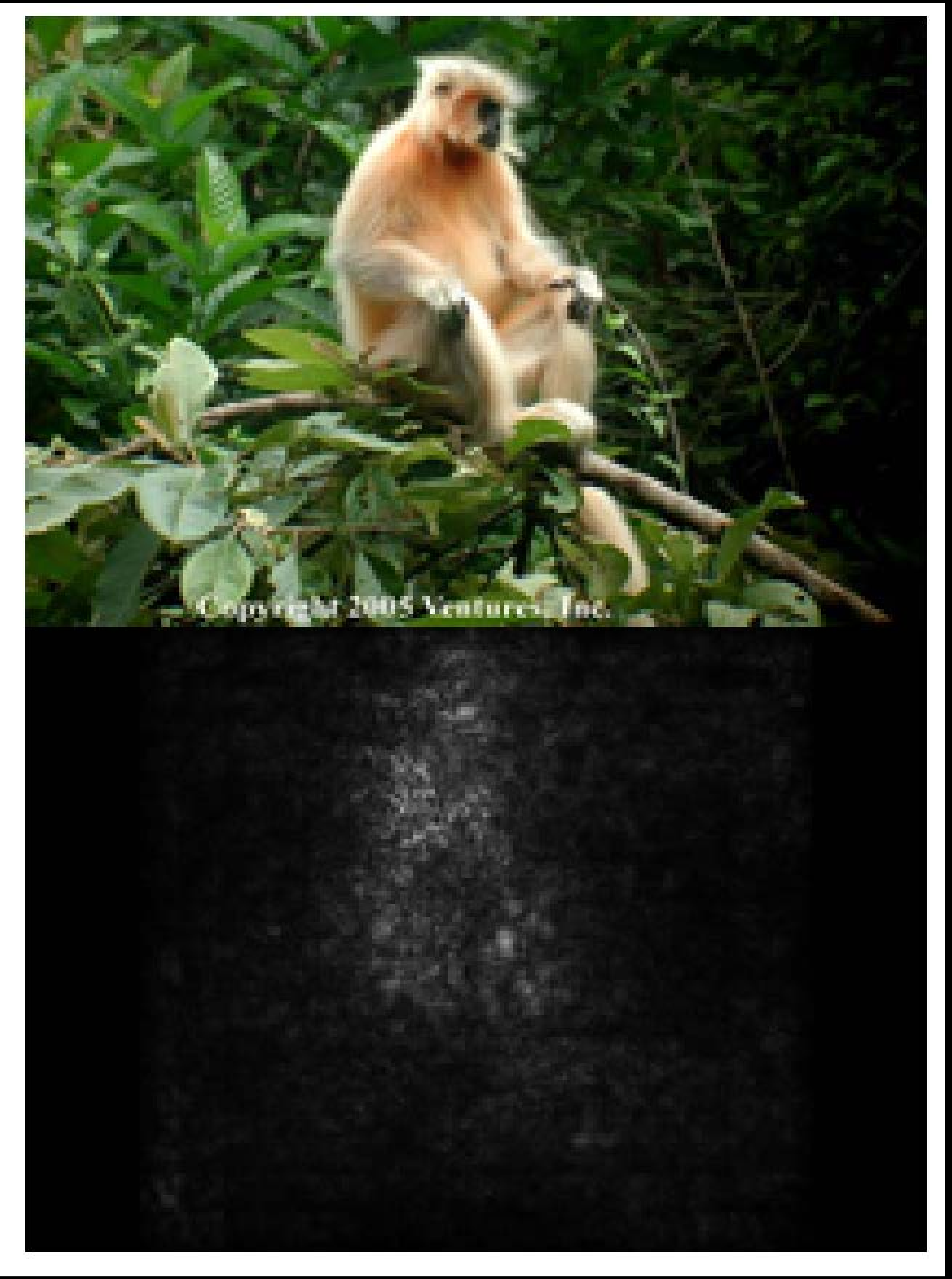
Analysis Scripts: Saliency



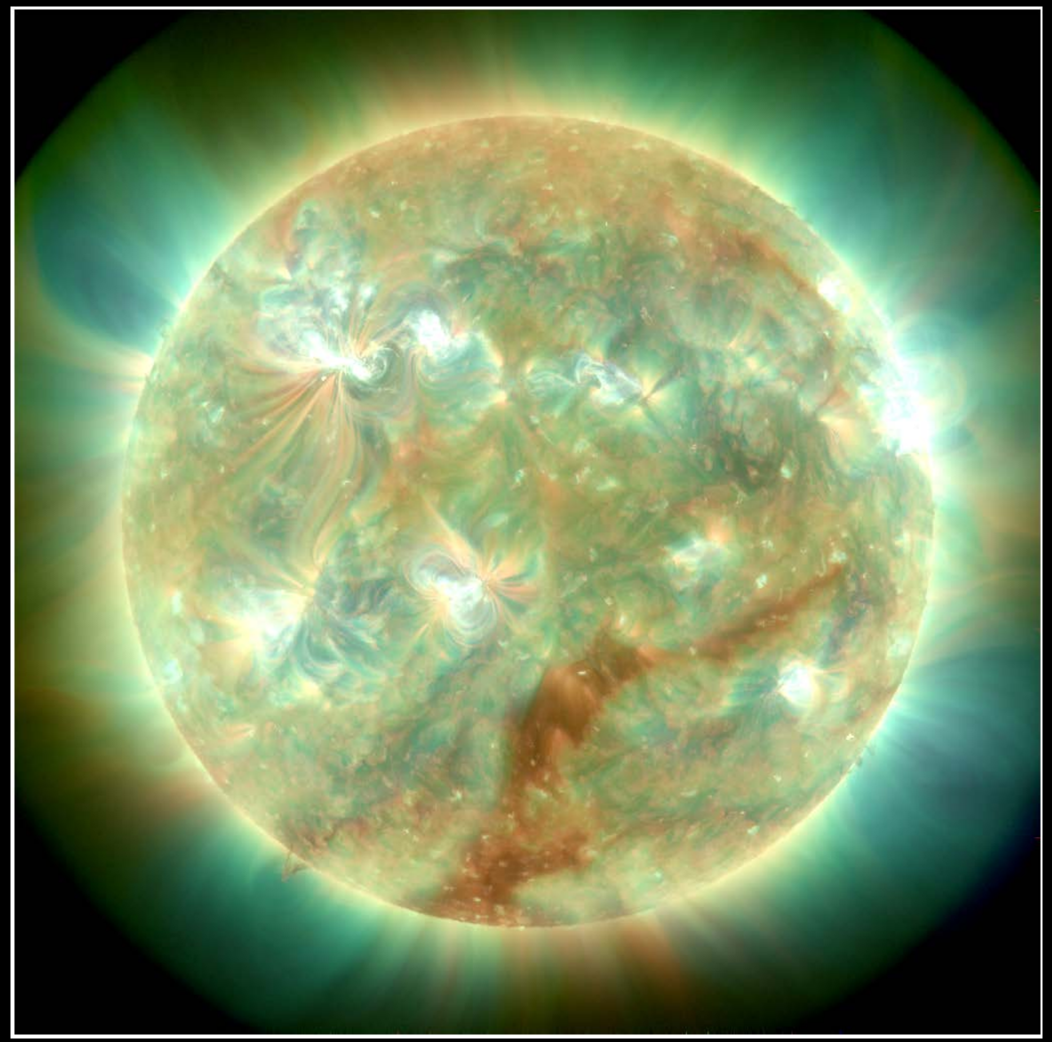
What does a convolutional neural network pay attention to?

Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.

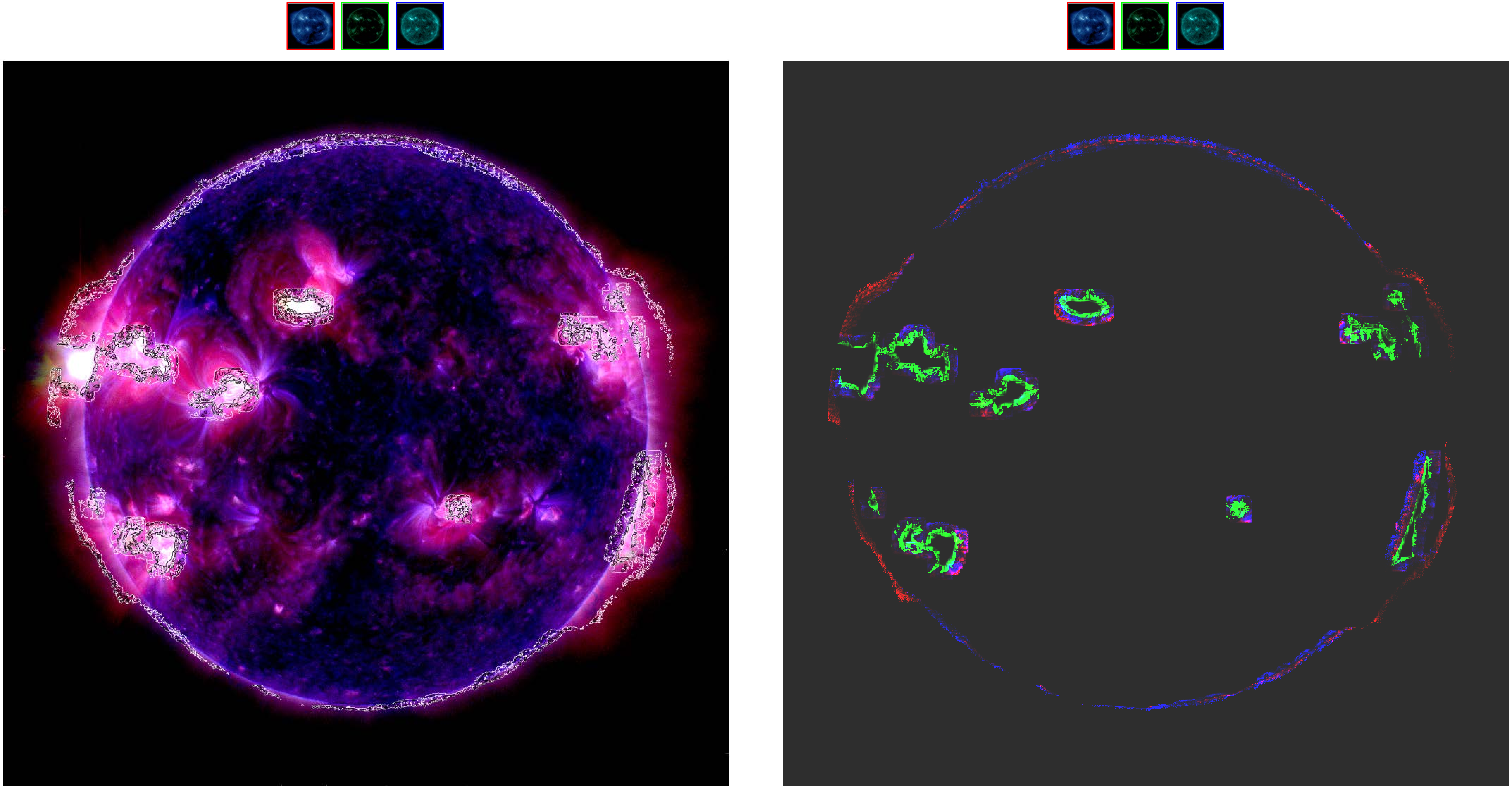
Analysis Scripts: Saliency



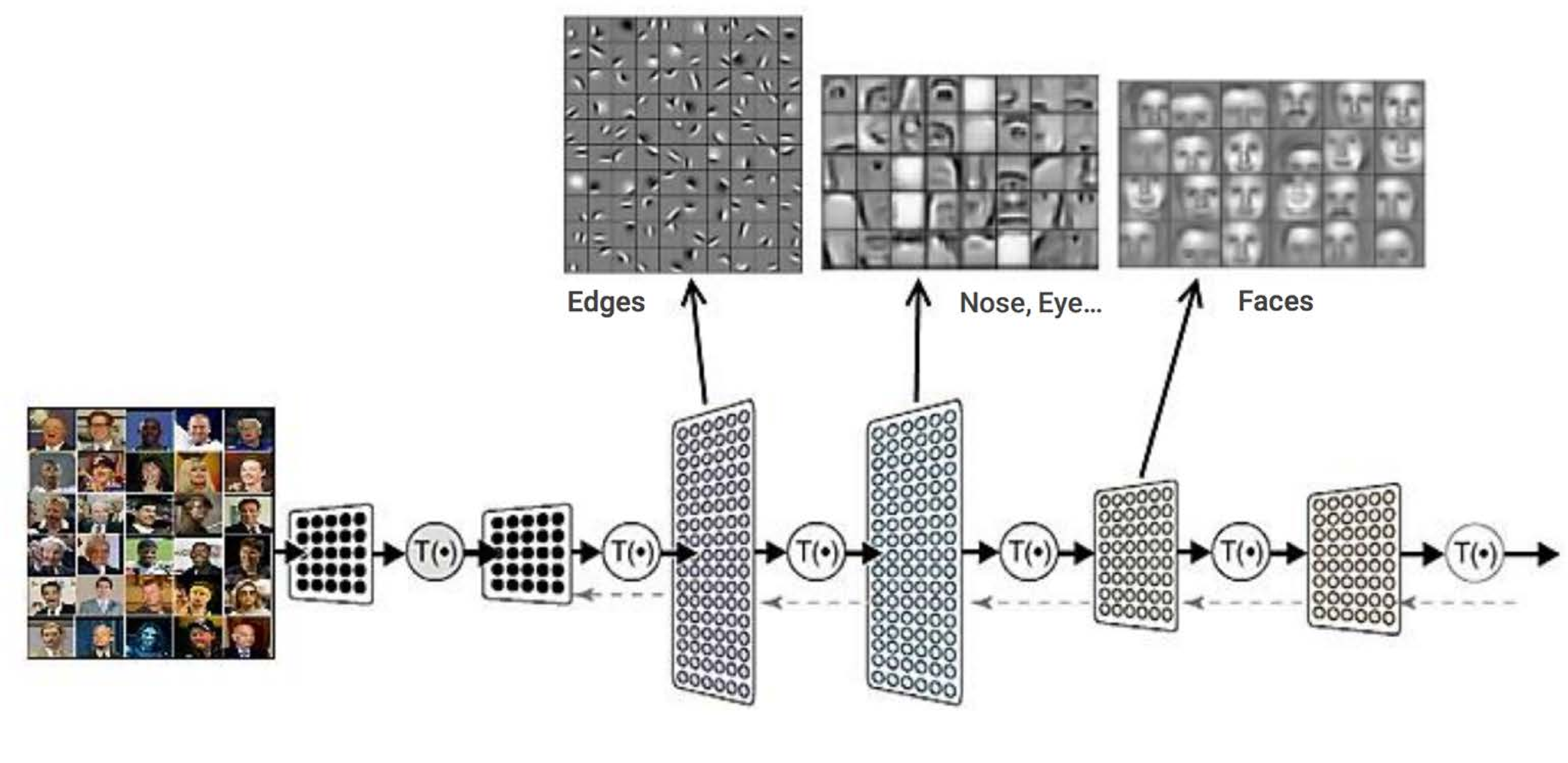
Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.



?



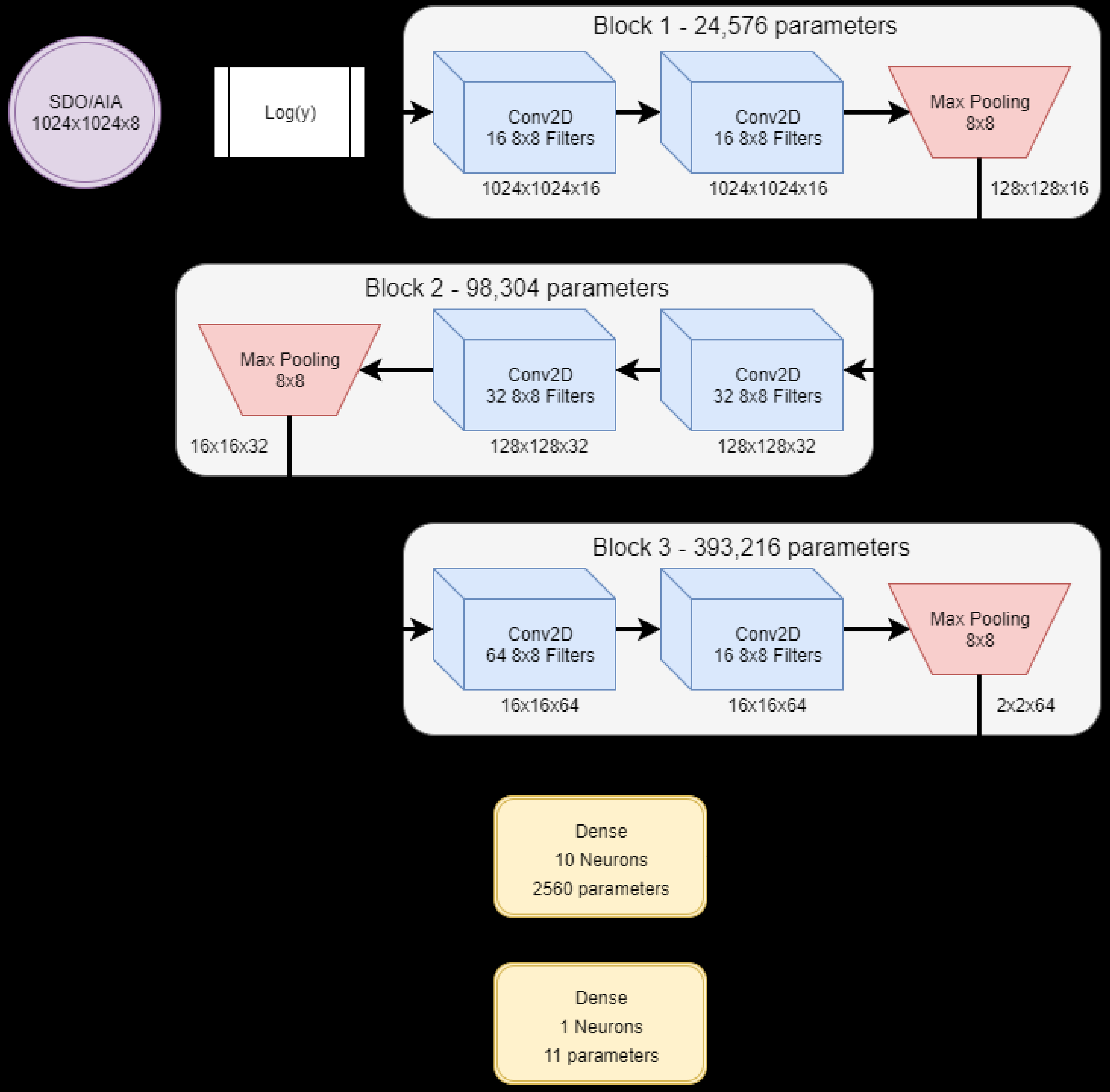
FlareNet’s filter activations

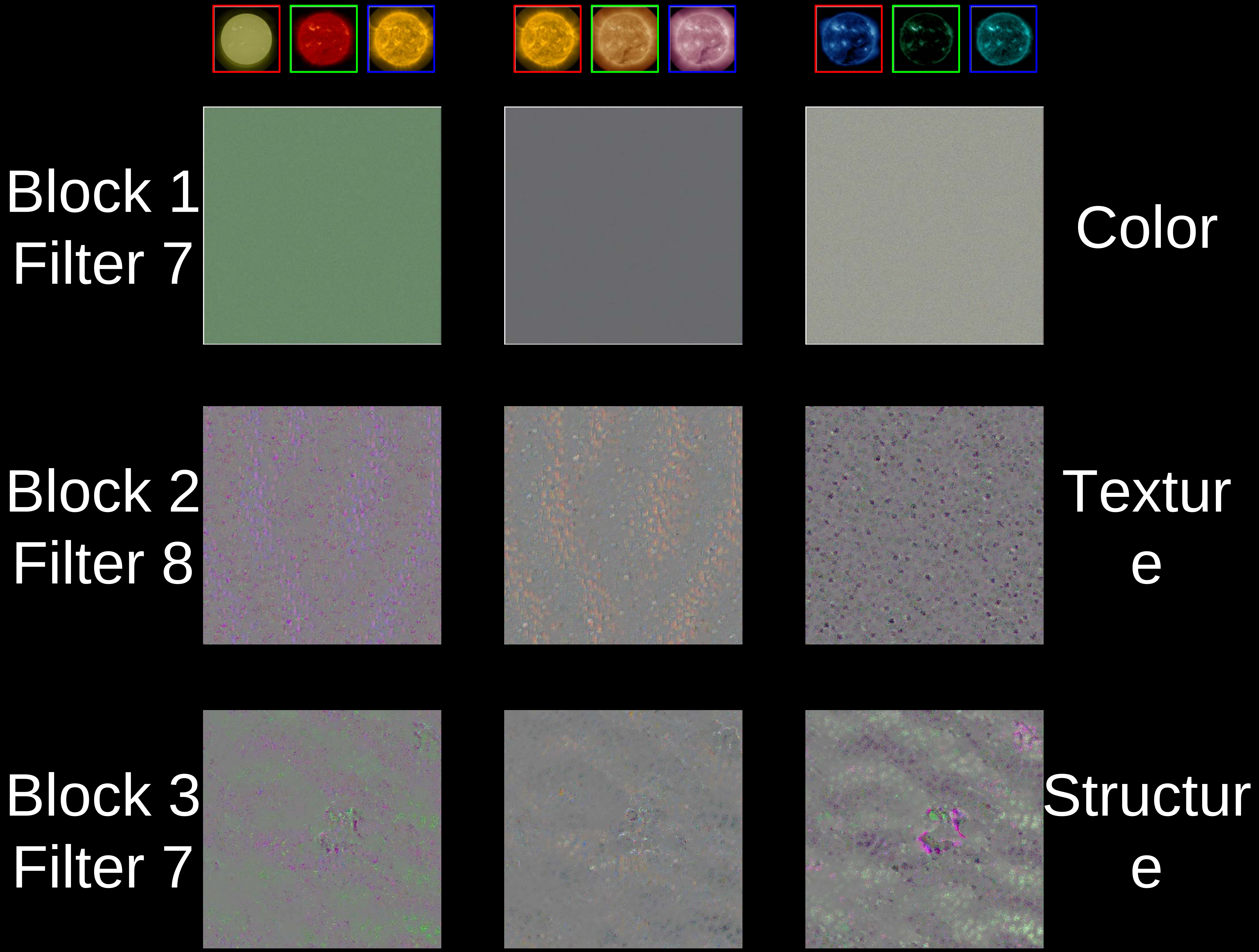


Several convolutional layers allow the neural network to recognize features of increased complexity

SPACE WEATHER: SOLAR STORM PREDICTION

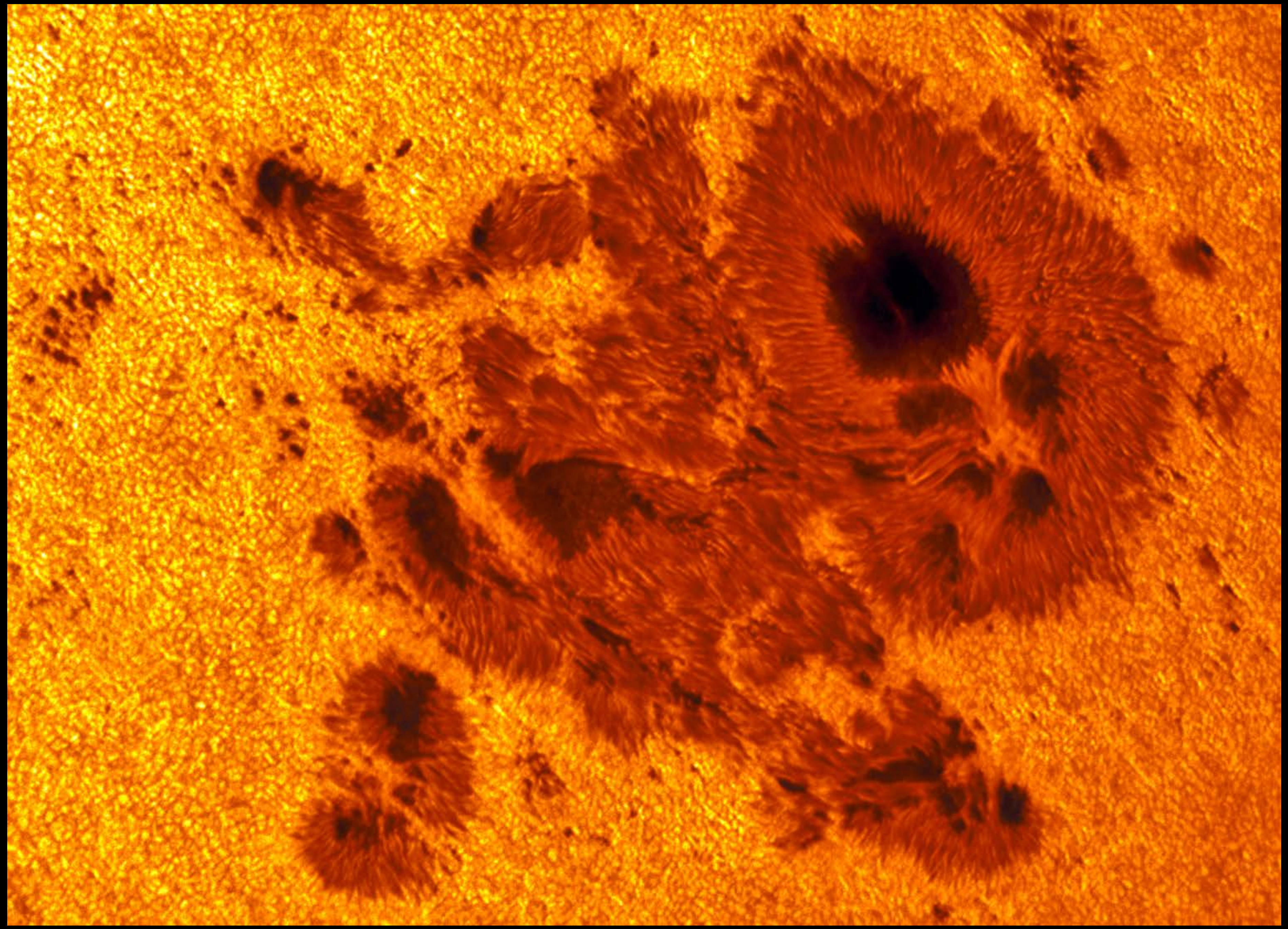
FlareNet's filter activations





SPACE WEATHER: SOLAR STORM PREDICTION

FlareNet’s filter activations

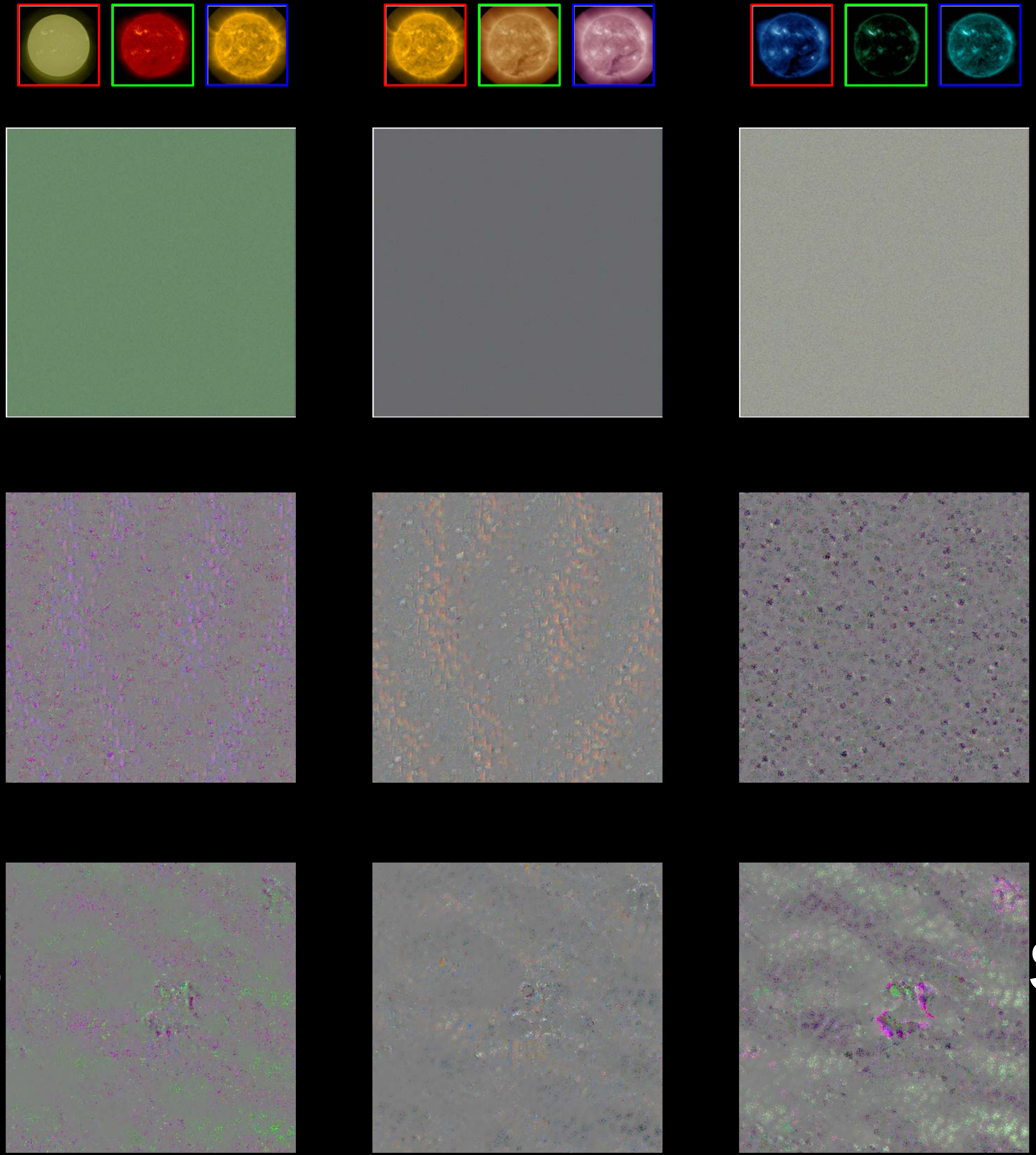


FlareNet learned the importance of active regions

Block 1 Filter 7

Block 2 Filter 8

Block 3 Filter 7



Color

Textur
e

Structur
e

Achievements

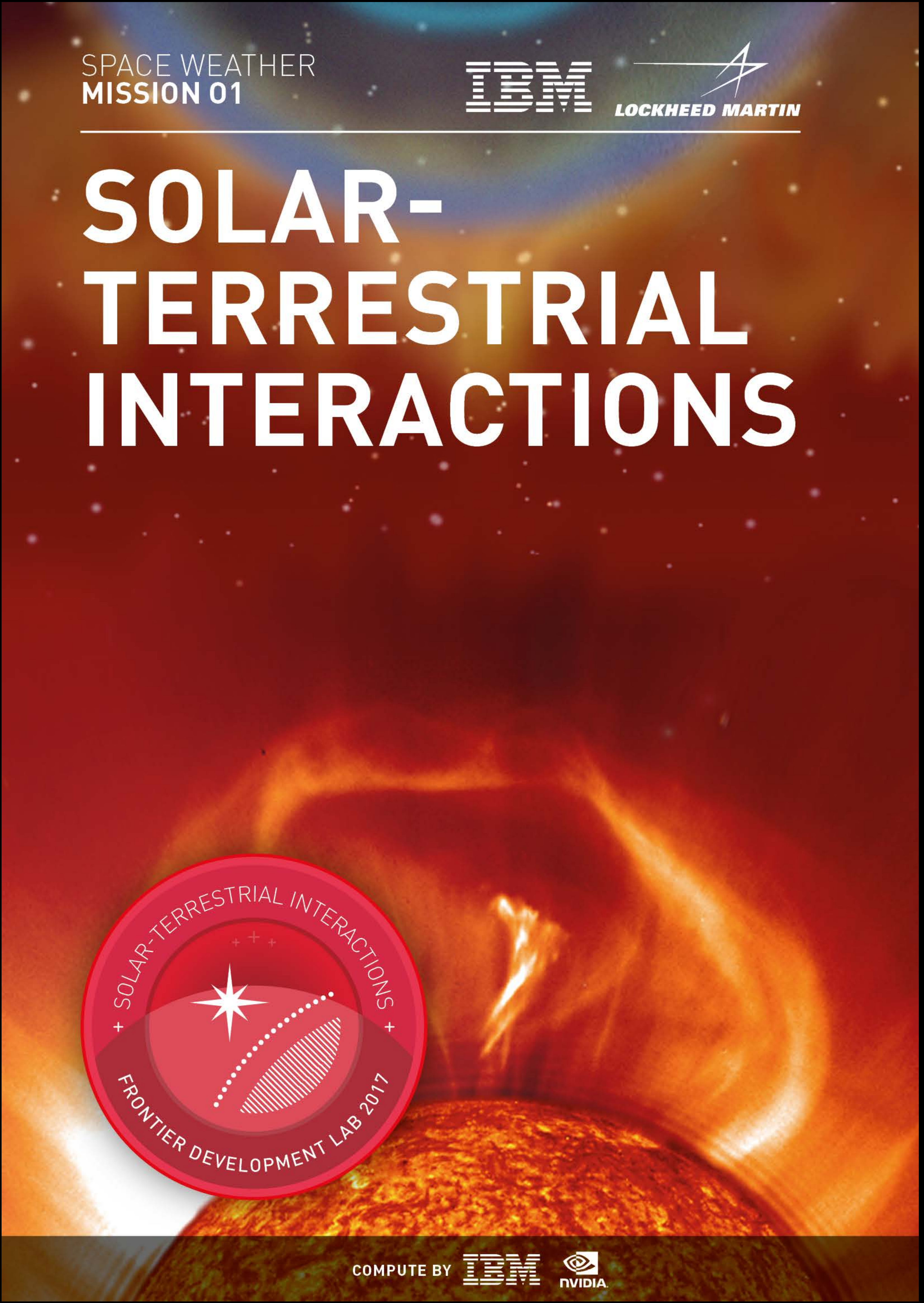
- Developed a framework to apply CNNs to heliophysics problems.
- Developed a CNN visualization framework to mine trained networks for physical insight.
- Demonstrated the capability of CNNs to identify structures of flaring relevance.

Future Work

1. Expand our data enhancement capabilities.

1. Explore the possibility of adding other instruments to increase our flare pool (Stereo, SOHO, GOES.)

1. Try alternative problem definitions besides regression (distribution, classification.).



- The vast amounts of data collected by satellites and observatories operated by government agencies such as NASA, NOAA and the US Geological Survey remains a largely untapped resource for discovering how the Sun interacts with Earth.

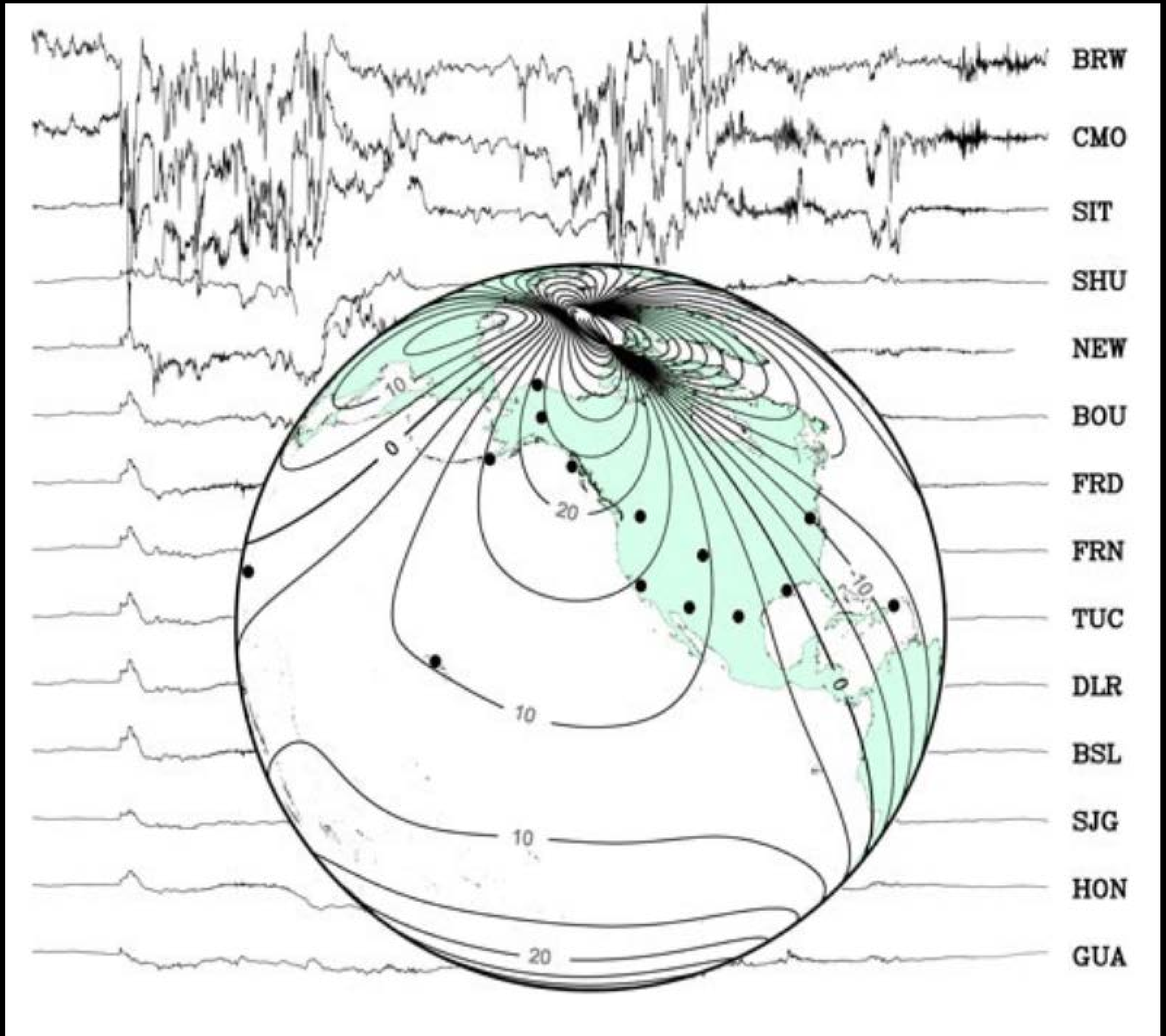
- The FDL team built a knowledge discovery module named STING (Solar Terrestrial Interactions Neural Network Generator) on top of industry-standard, open source machine learning frameworks to allow researchers to further explore these complex datasets.

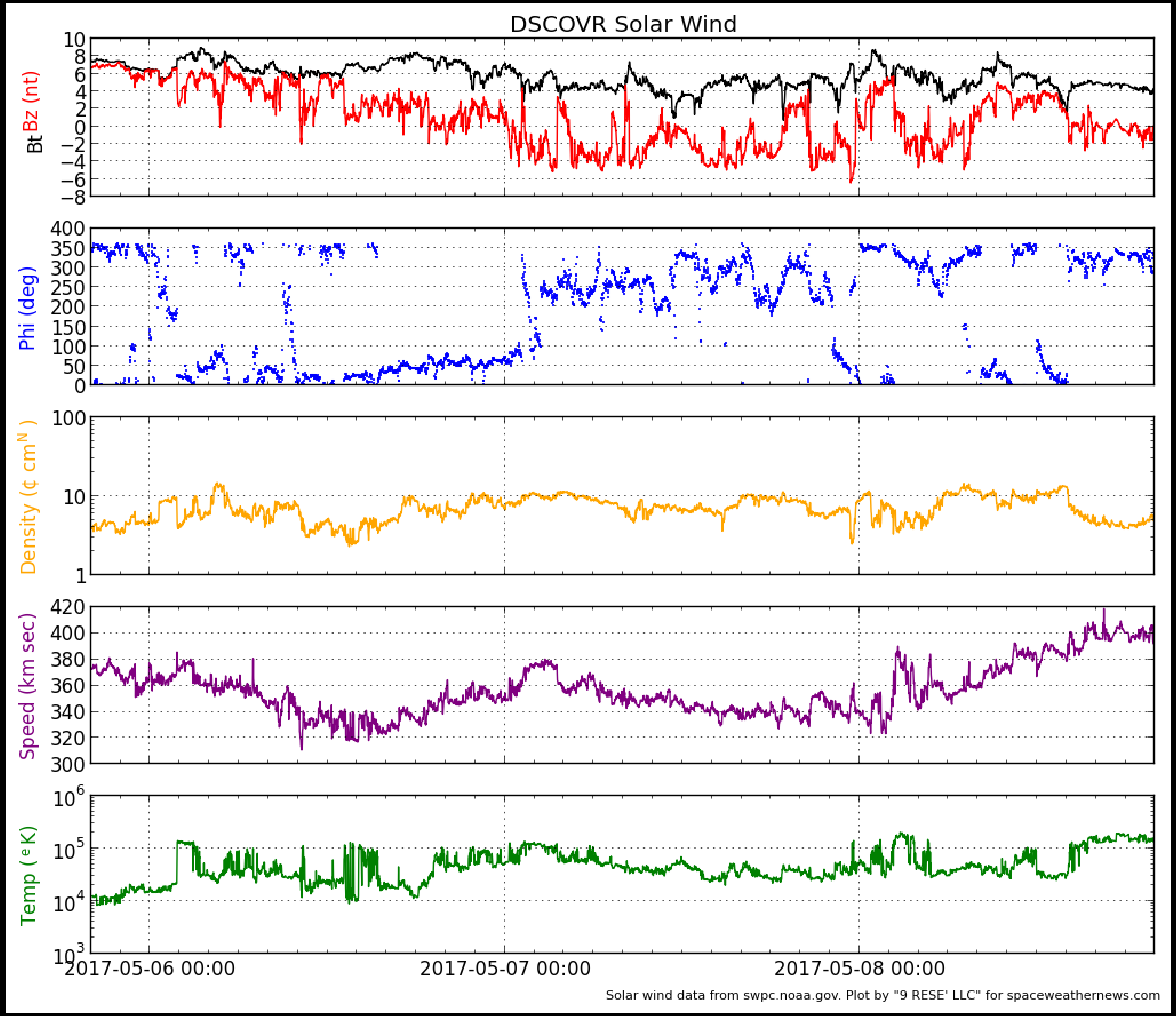
- STING showed the ability to accurately predict the variability of Earth's geomagnetic fields in response to solar driving - specifically the KP index.

- In the process the tool discovered the imprint of the magnetospheric ring current in precursors of geomagnetic storms - an example of an AI derived discovery.

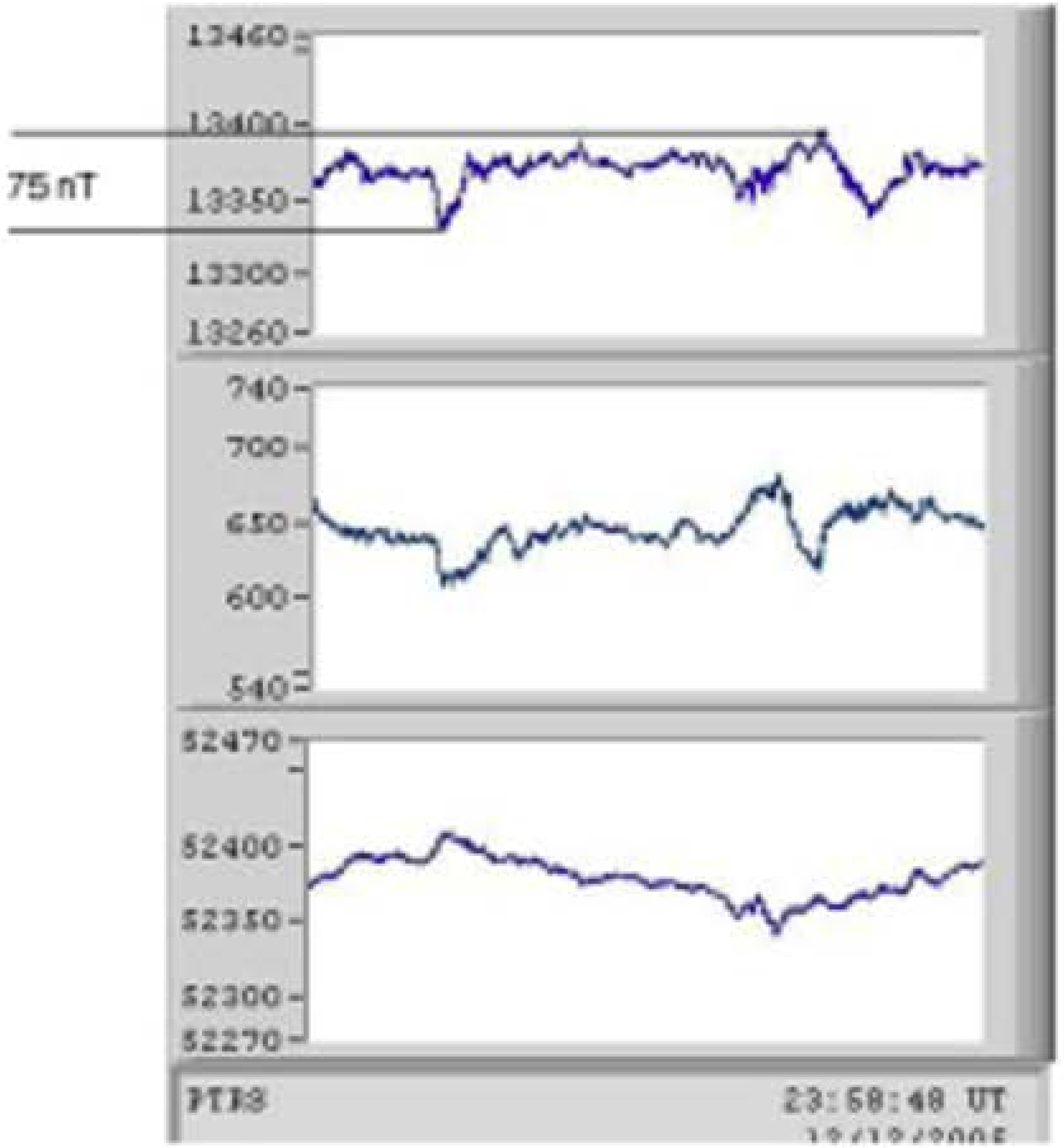
SPACE WEATHER: SOLAR TERRESTRIAL INTERACTIONS

DATA SOURCES





Kp INDEX



Petersburg, AK magnetometer data with a 75 nT change in the X-direction (Magnetic North)

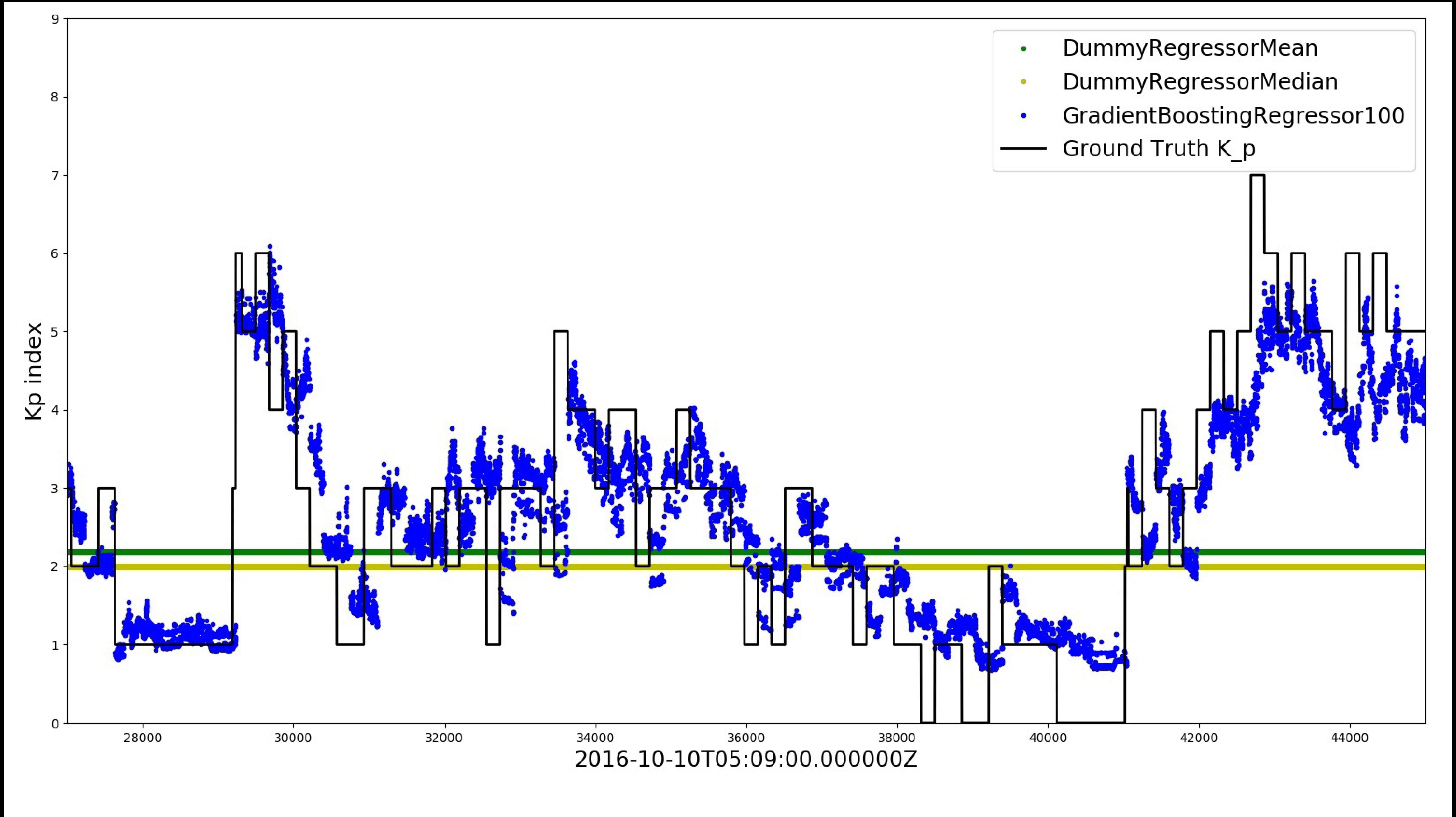
Use this table on the right to convert the difference in the maximum and minimum x-values for today to a K index. The larger the K index, the stormier it is in Earth's magnetic field.

K index	nT diff.
0	0-5
1	5-10
2	10-20
3	20-40
4	40-70
5	70-120
6	120-200
7	200-330
8	330-500
9	>500

Planetary Kp Index
(Bartels, 1938)

SPACE WEATHER: SOLAR TERRESTRIAL INTERACTIONS

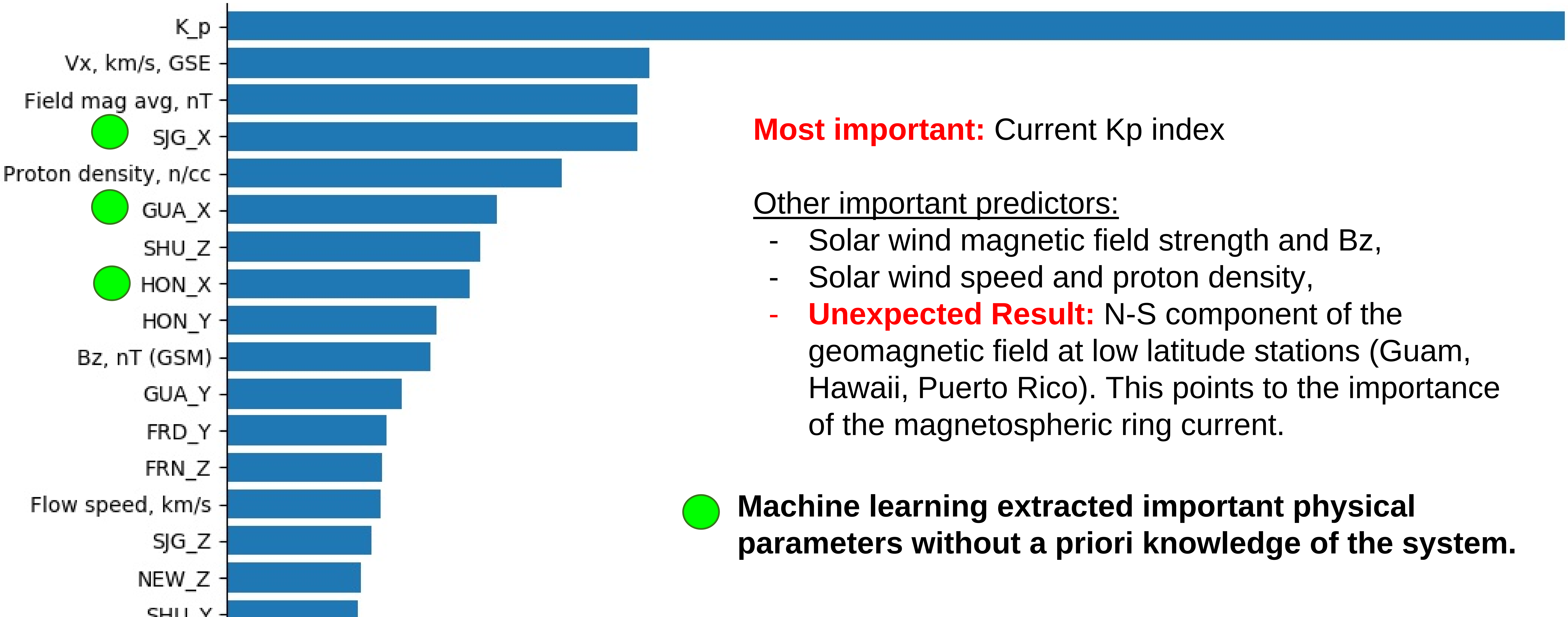
GRADIENT BOOSTING RESULTS



SPACE WEATHER: SOLAR TERRESTRIAL INTERACTIONS

FEATURE DISCOVERY

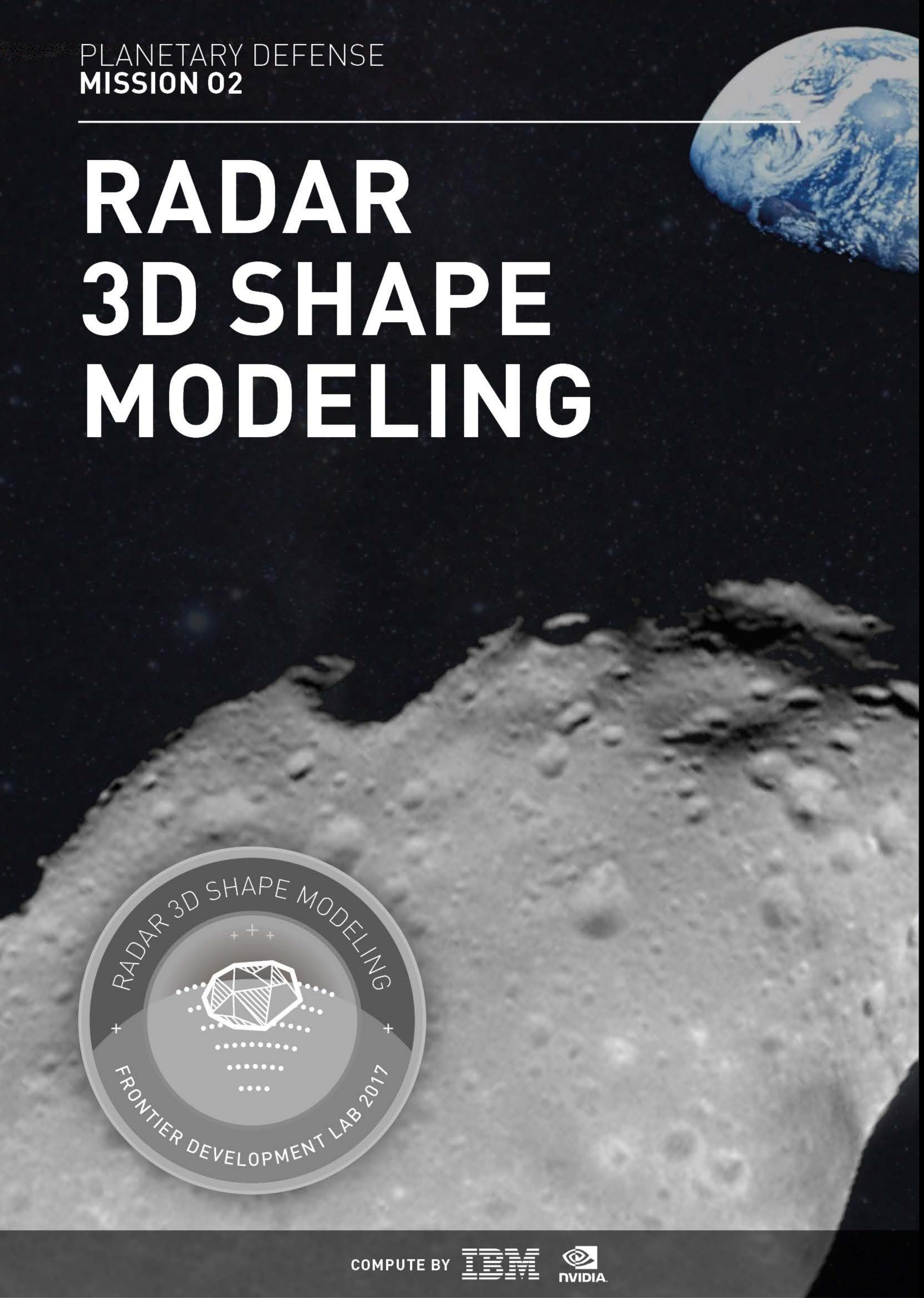
This plot shows the relative importance of the physical parameters for Kp prediction.



Most important: Current Kp index

Other important predictors:
- Solar wind magnetic field strength and Bz,
- Solar wind speed and proton density,
- **Unexpected Result:** N-S component of the geomagnetic field at low latitude stations (Guam, Hawaii, Puerto Rico). This points to the importance of the magnetospheric ring current.

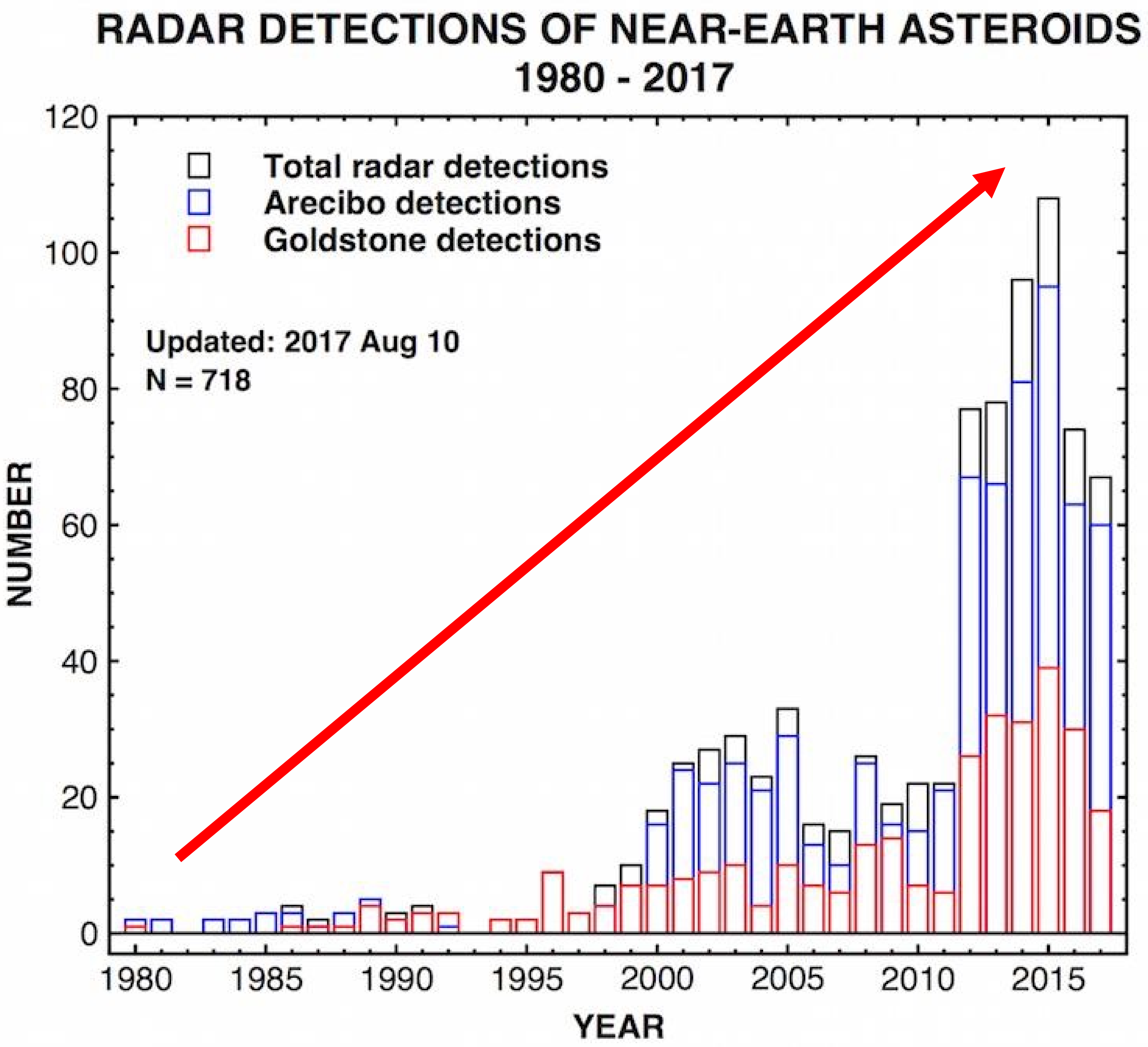
Machine learning extracted important physical parameters without a priori knowledge of the system.



- The FDL team tackled the task of automating task of creating 3D shape models of NEOs from sparse radar data
- The process currently takes up to four weeks of manual interventions by experts using established software.
- The team demonstrated a pipeline for automation that allows NEOs to be modelled in several hours.
- This result will hopefully support researchers render 3D models of the current backlog of radar imaged asteroids.

PLANETARY DEFENSE: RADAR 3D SHAPE MODELING

RADAR OBSERVATIONS

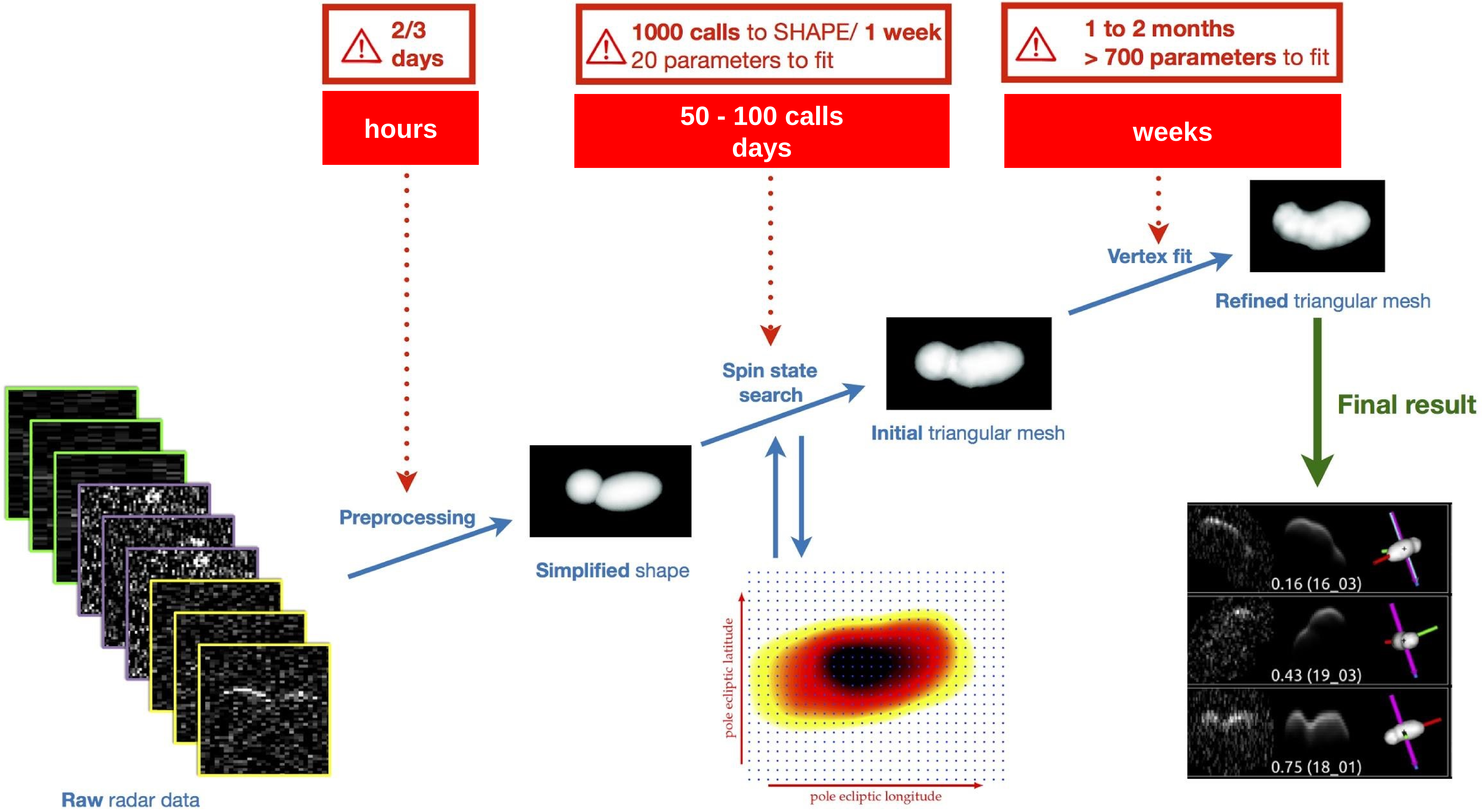


We are observing NEAs faster than we can analyze them!

https://echo.jpl.nasa.gov/~lance/Radar_detected_neas.html

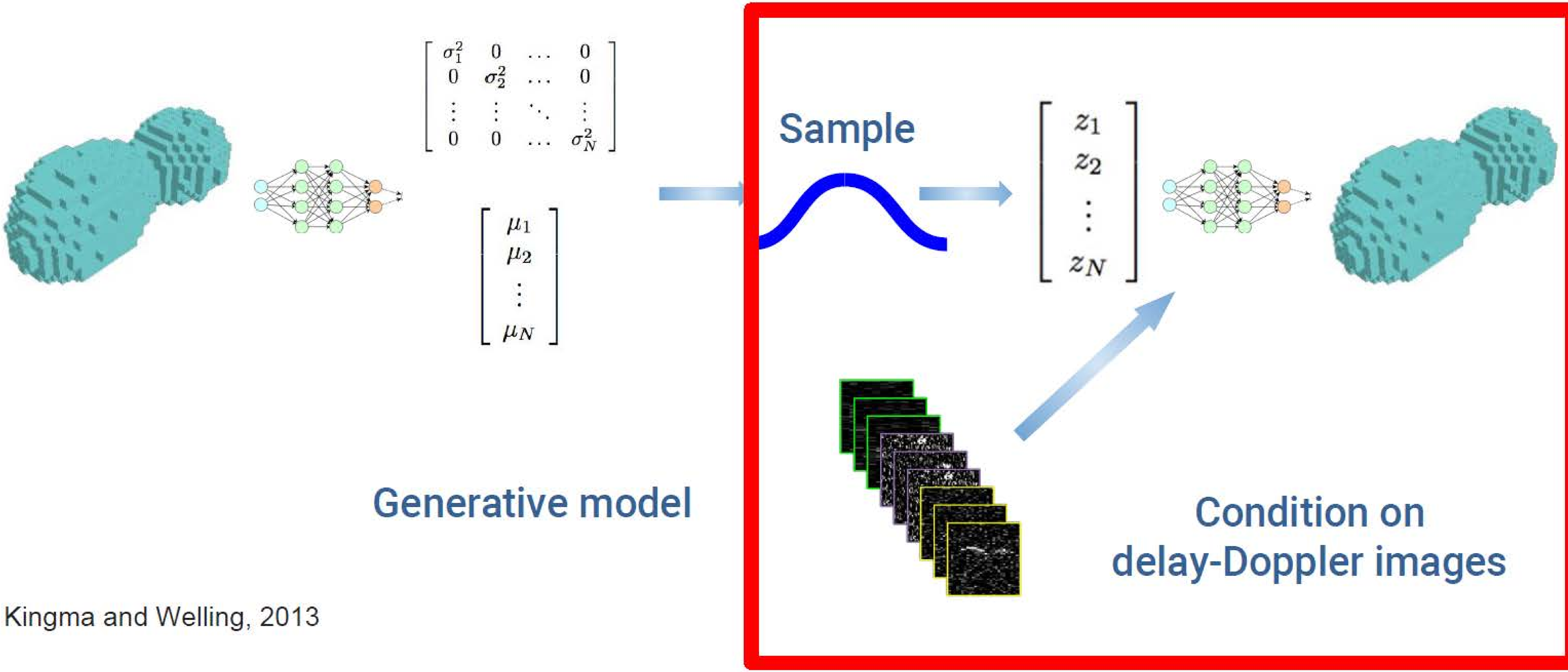
PLANETARY DEFENSE: RADAR 3D SHAPE MODELING

SHAPE MODELING PIPELINE



PLANETARY DEFENSE: RADAR 3D SHAPE MODELING

VARIATIONAL AUTOENCODER



Kingma and Welling, 2013

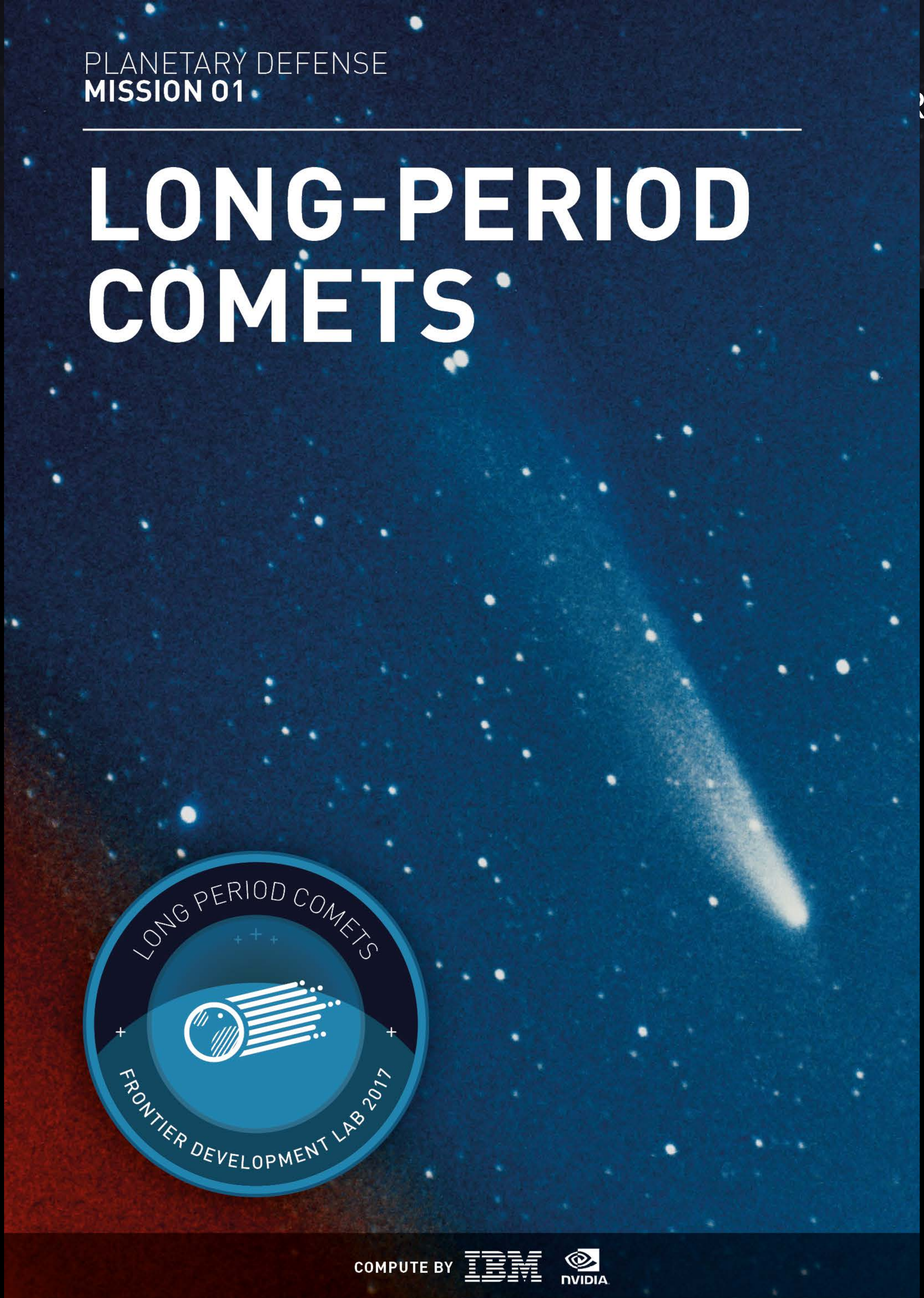
OUR SOLUTIONS

a) Pre-processing is *faster*

b) Spin axis determination is *faster*

c) Training data generation is *improved*

d) Neural network is *improved*



RADAR 3D SHAPE MODELING

- **Meteor showers caused by the previous-return ejecta of long period comets can guide deep searches, and improve warning time, for potentially hazardous long period comets that passed near Earth's orbit in the past ten millennia.**

- **The FDL team showed how the data reduction of the 'CAMS' meteor shower survey program could be successfully automated by using deep learning approaches.**

- **By using dimensionality reduction (t-SNEs) the team were able to identify yet uncatalogued meteor shower clusters - a promising direction for further investigation.**

LONG PERIOD COMETS
FRONTIER DEVELOPMENT LAB 2017

PLANETARY DEFENSE: LONG-PERIOD COMETS

Meteor shower surveys

Convolutional Neural Network (CNN)

Results: Precision: 88.6% Recall: 90.3%

PLANETARY DEFENSE: LONG-PERIOD COMETS

Mapping meteors in the sky

cams.seti.org/FDL/

[NO REAL DATA YET: SITE IS UNDER CONSTRUCTION]

CAMS

All Networks:

BeNeLux

California

EXOSS

Florida

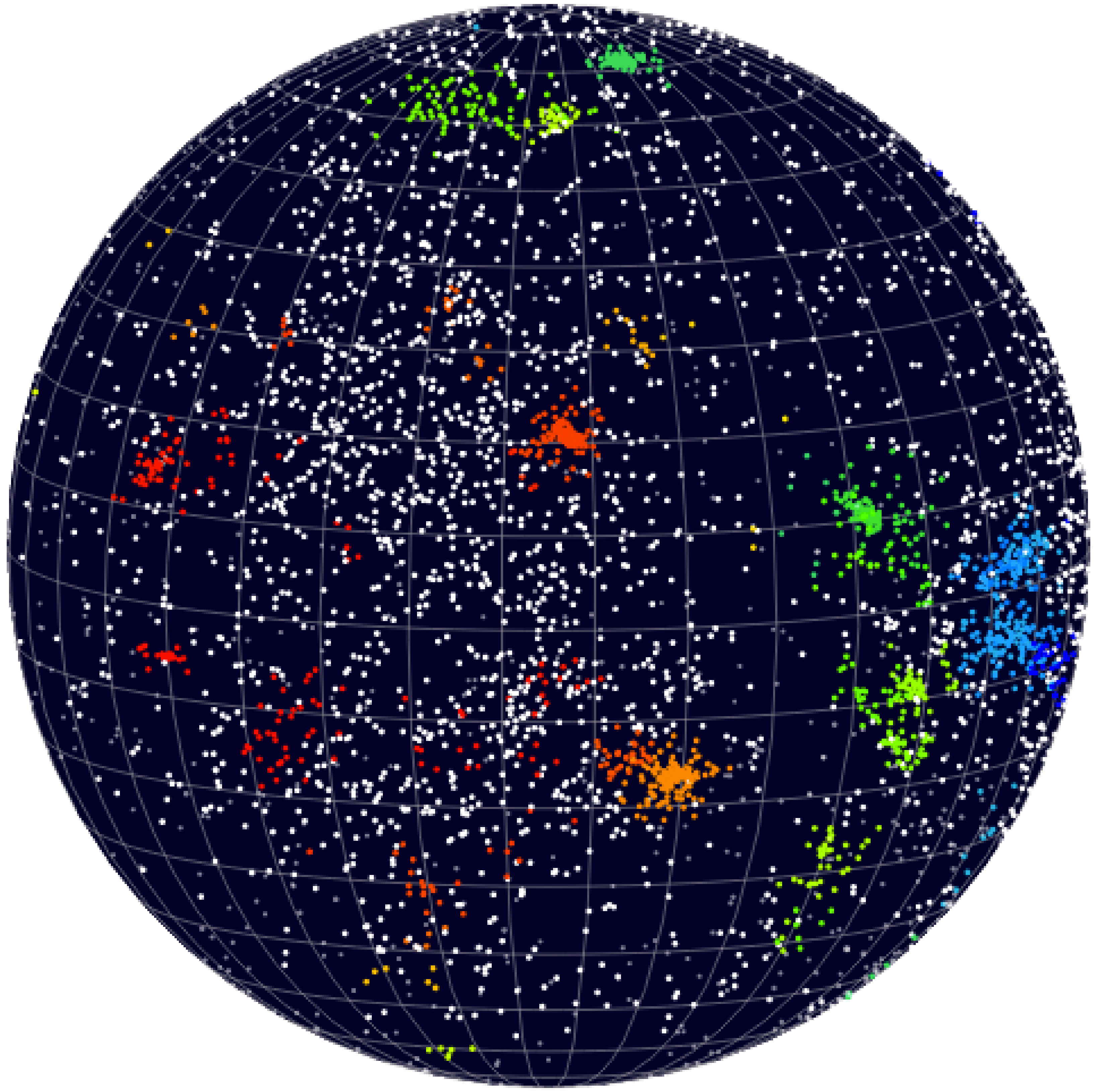
LOCAMS

Mid-Atlantic

New Zealand

South Africa

UACN

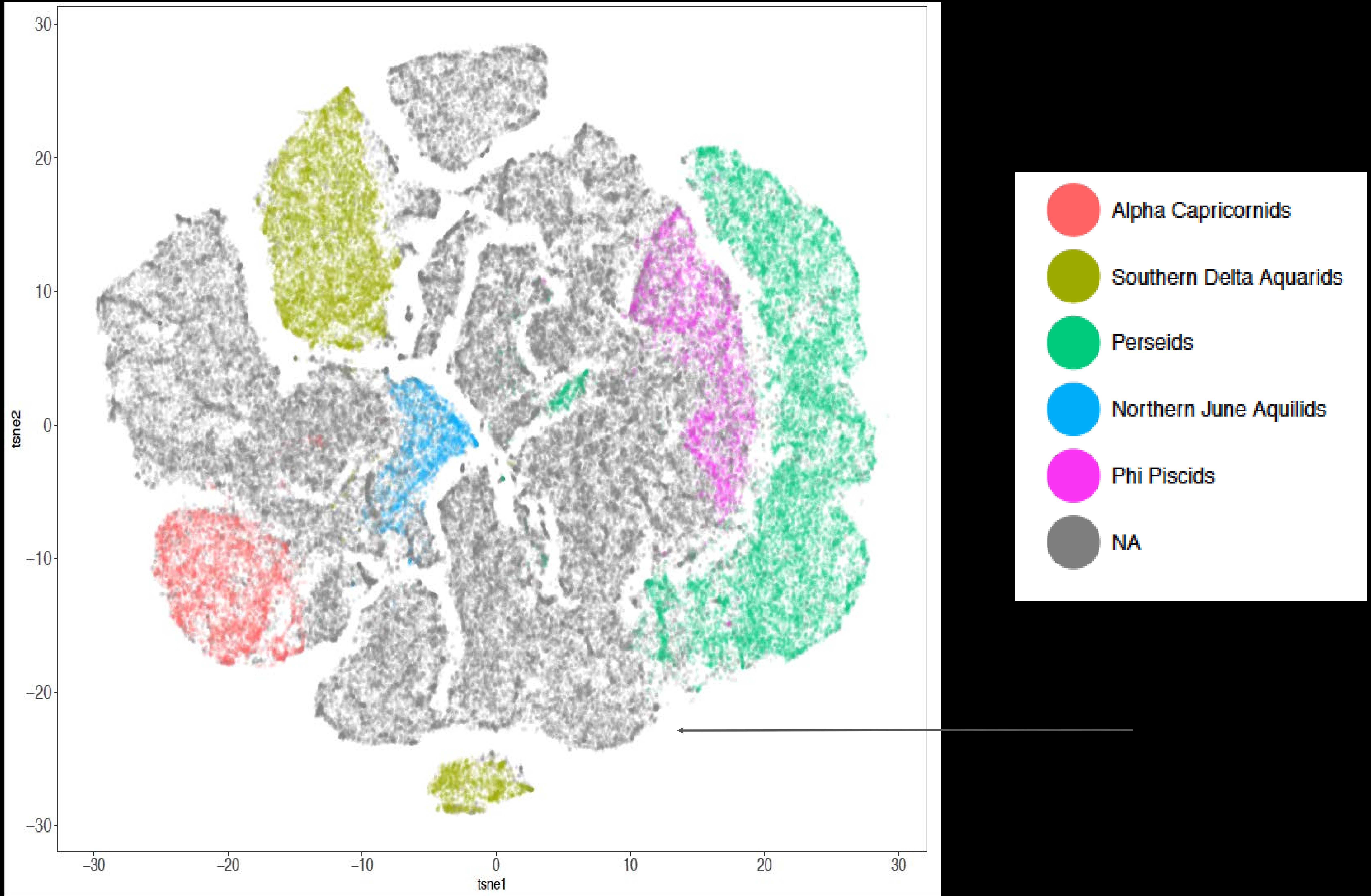


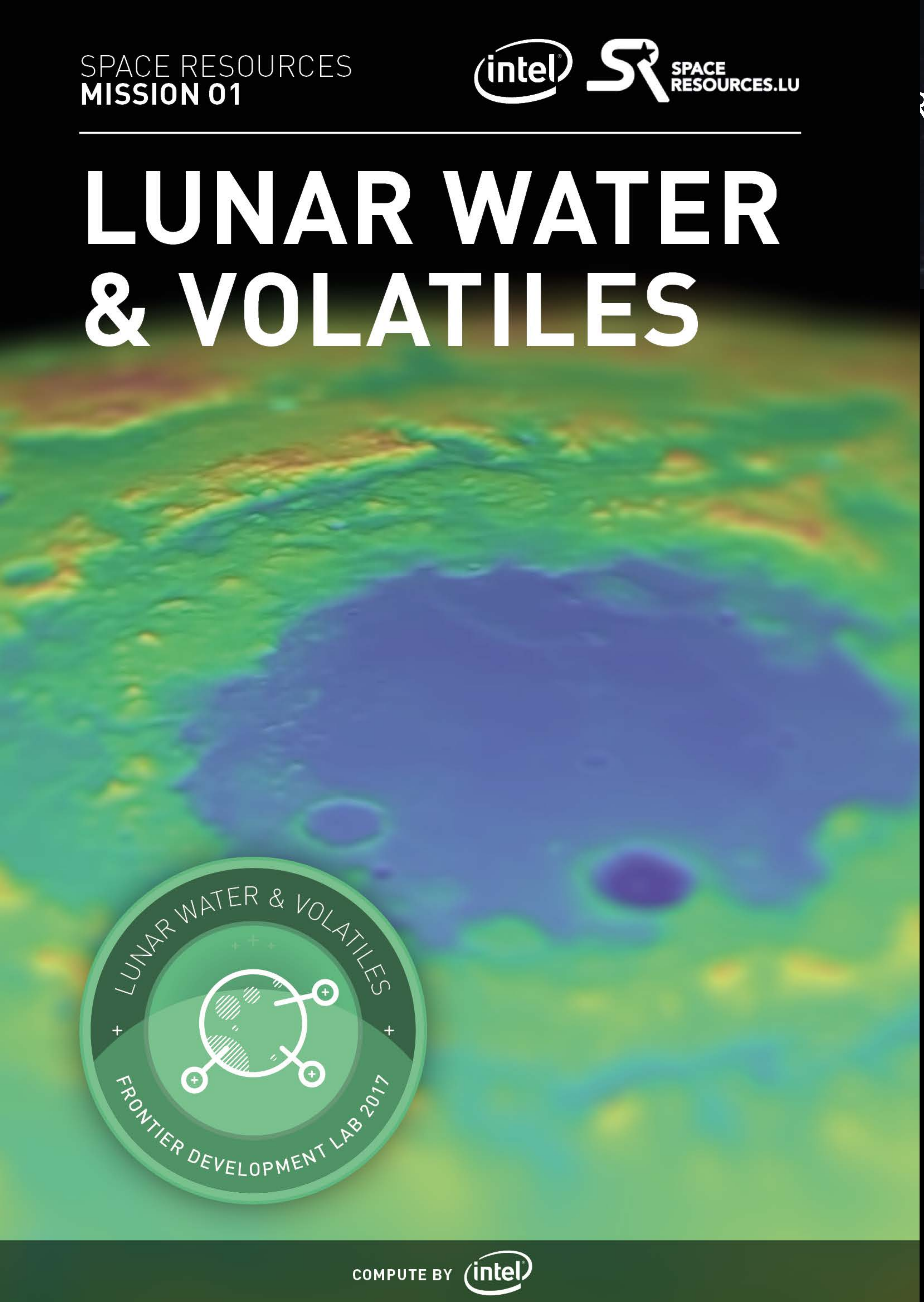
Start a new network?
Contact us.

Today's date: 2017-08-09

Most meteors today:
Last submission by:

Established meteor showers





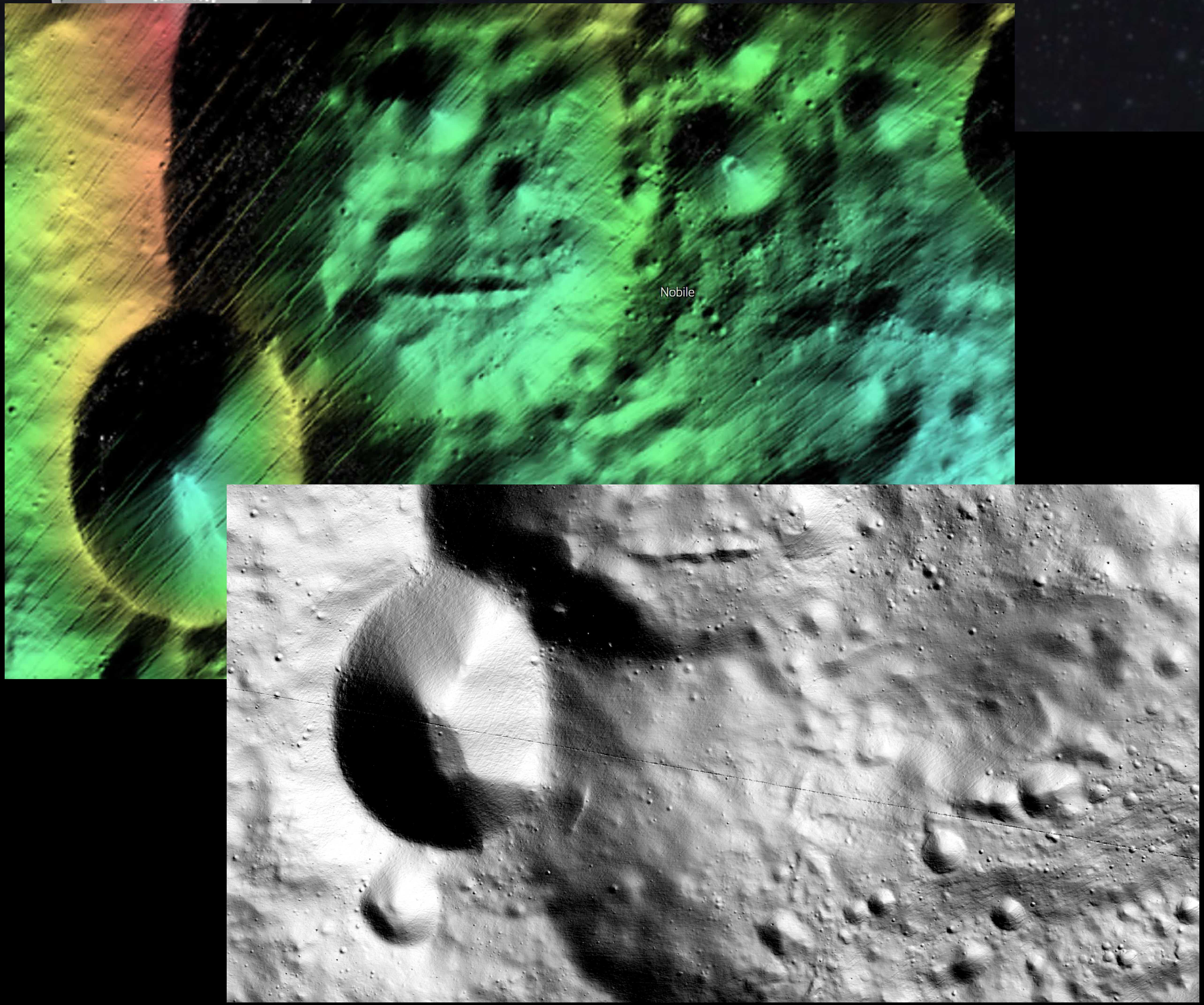
RADAR 3D SHAPE MODELING

- **Maps that detail the regions of interest in the dark polar regions are plagued by artefacts and shadow variability that severely hamper the planning of future prospecting missions.**

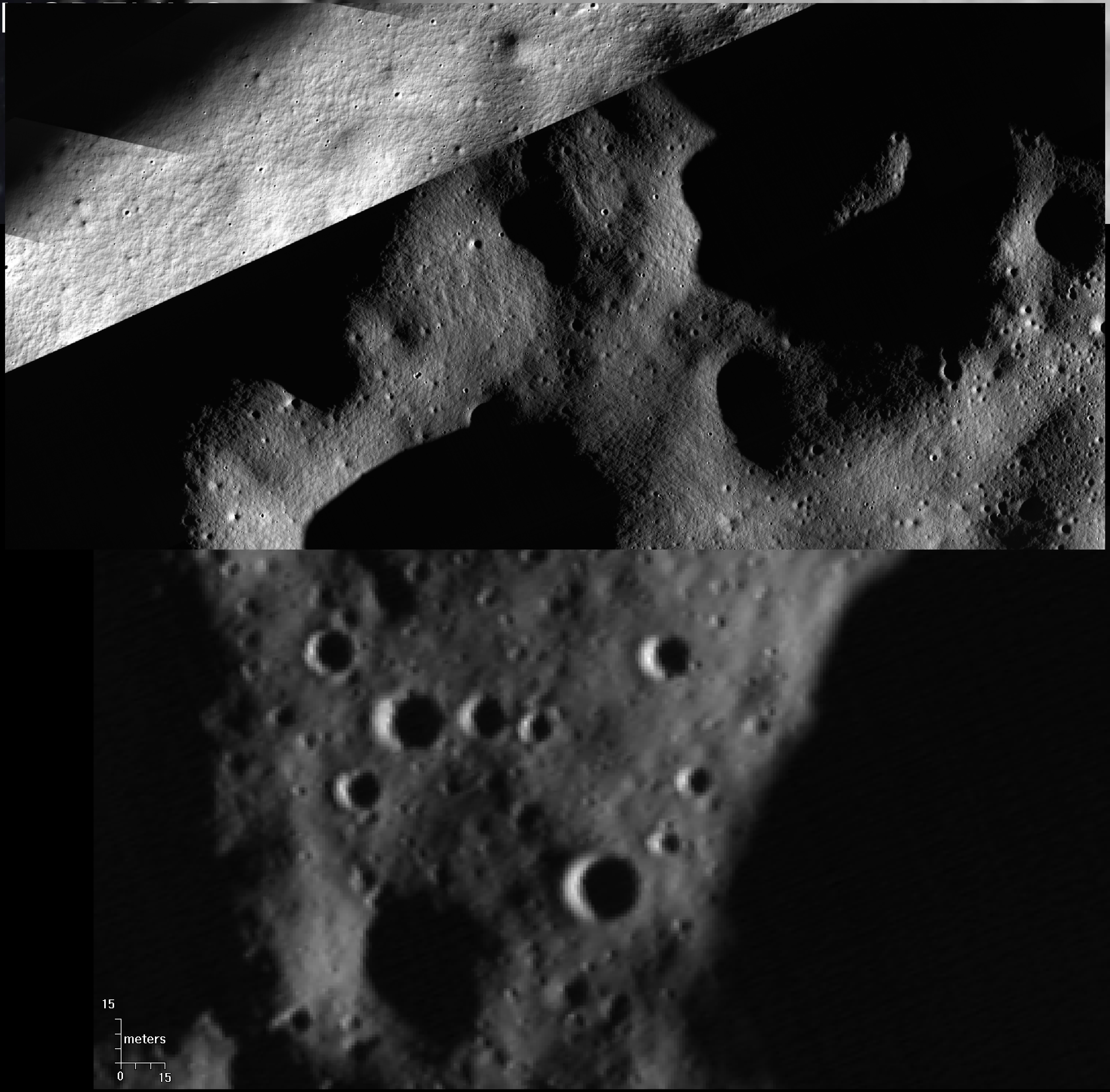
- **A large dataset was compiled for the south polar region and high-level feature extraction was performed. Results showed an impressive speed-up of 100x compared to human experts, with more than 98.4% agreement when approaching a crater labelling.**

- **This work represents a potential keystone to facilitate accessing water on the Lunar surface and future traverse planning.**

Building lunar maps at the poles is problematic

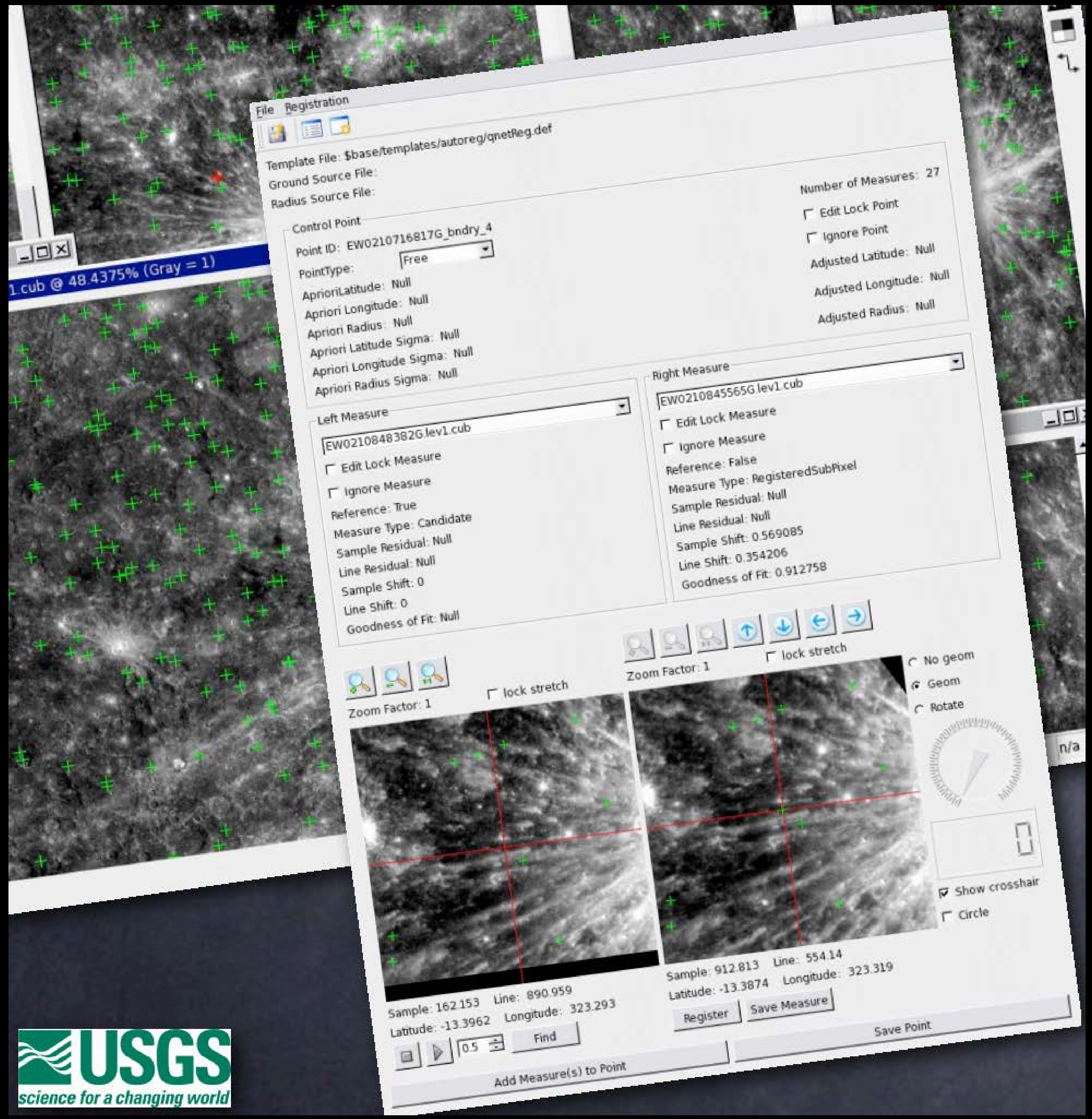


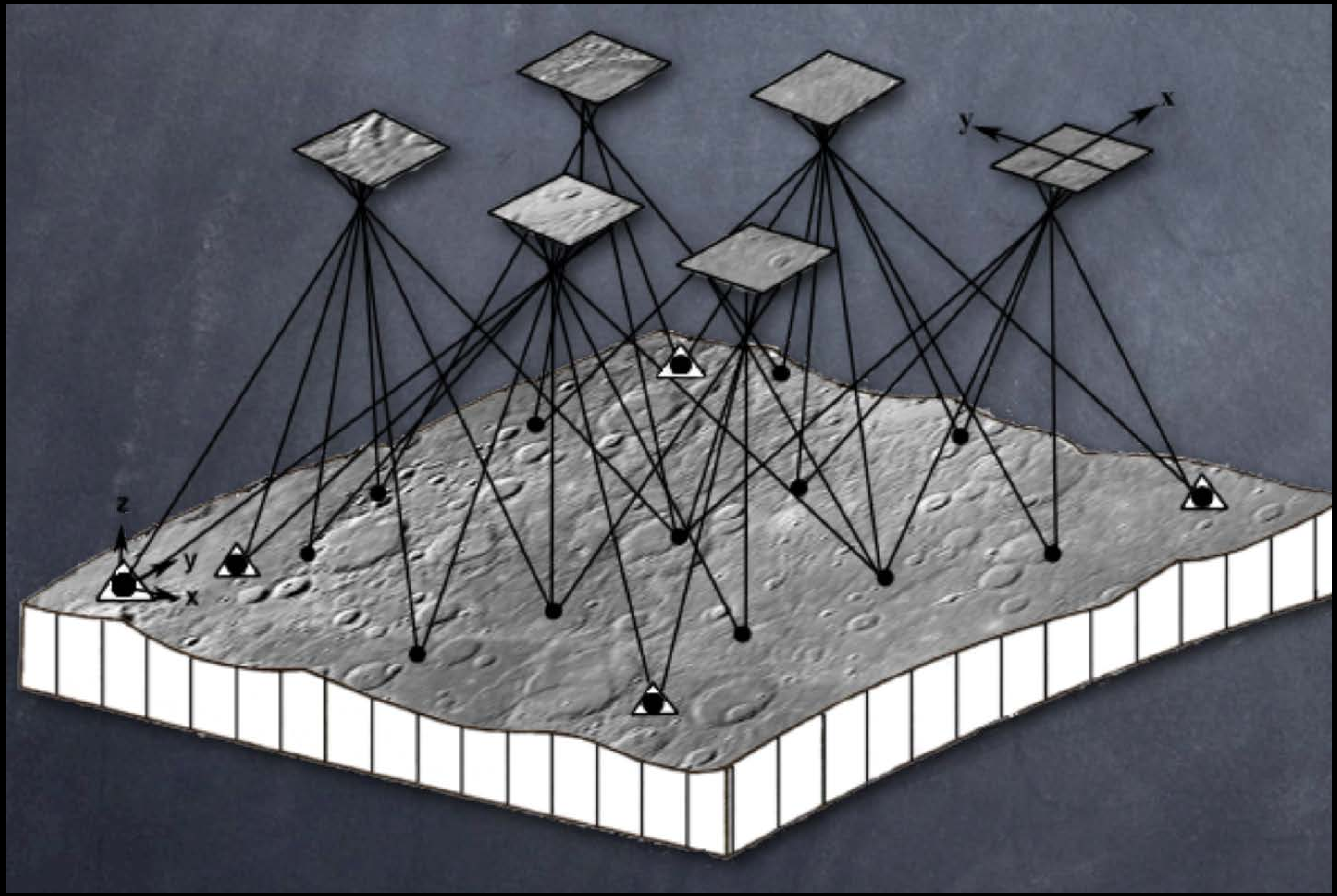
Lunar Orbiter Laser Altimeter
Digital Elevation Model (LOLA DEM), 20 m resolution



Narrow Angle Camera (NAC)
Optical images, 0.5 m resolution

Improving Maps Conventionally.
Most algorithms remain unused.





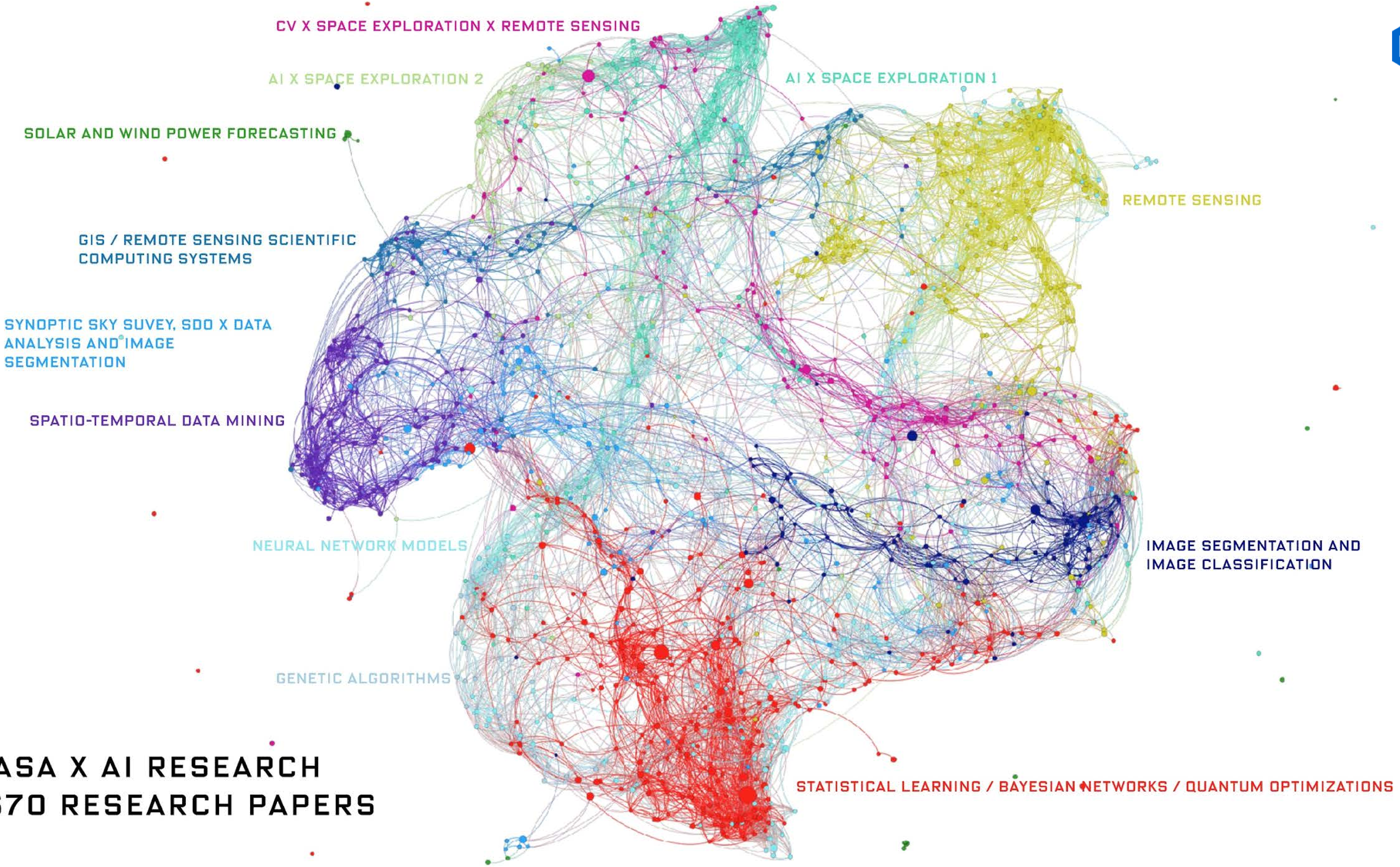
Timing Comparison of FDL Technique

Group	Human	Single-Layer	CNN
Accuracy	-	Poor	98.4%
Time (1000 Images)	1-3 hours	10 hours	1 minute
Person-hours	1-3 hours	-	-

AI & SPACE SCIENCES
FRONTIER DEVELOPMENT LAB 2017

AI & SPACE SCIENCES

NASA
FRONTIER
DEVELOPMENT LAB



MISO

NASA X AI RESEARCH
1670 RESEARCH PAPERS

NASA X AI RESEARCH

Closing Thoughts

- Focus on applied AI solutions using mainstream deep learning tools, thereby complementing and informing the research into novel AI technology being undertaken by other NASA teams.

- Strong incentive for the private sector to participate due to commercial opportunities that are implicit in the outcome;

- Clear risk/cost reduction benefit to manned activities beyond LEO, and for cis-lunar operations in particular;

- Problem definitions for which relevant data has already been collected and is available for use under an open license.

By way of example, consider the application of AI to Space Weather

- Solar flares and associated proton storms pose a significant risk to astronauts beyond LEO, and offer little or no warning. The Apollo “near miss” of the August 1972 solar flare provides a dramatic example of this concern.

- Multiple industry sectors have a vested commercial interest in seeing improvements to solar flare predictions and better heliophysics modeling in general. Examples include the power utilities, insurance companies, communications and satellite operators, and the military.

- There are hundreds terabytes of well structured heliophysics data highly suited to deep learning applications, including the archives from SDO/AIA, ACE, and SOHO.

- The image-centric nature of solar data (e.g. SDO – HMI and AIA) makes it easy to leverage the rapid advances in image analysis that the AI community has contributed into open source.

- There are tantalizing indications that machine learning techniques can offer better predicative capabilities over physics-based models, which leads many to believe that the use of neural net deep learning will prove to be even more effective.