Processing Solar Images to Forecast Coronal Mass Ejections using Artificial Intelligence

Savvas Raptis
KU Leuven, Centre for Mathematical Plasma Astrophysics

Supervisors:

Giovanni Lapenta, Jorge Amaya

Preventing a disaster

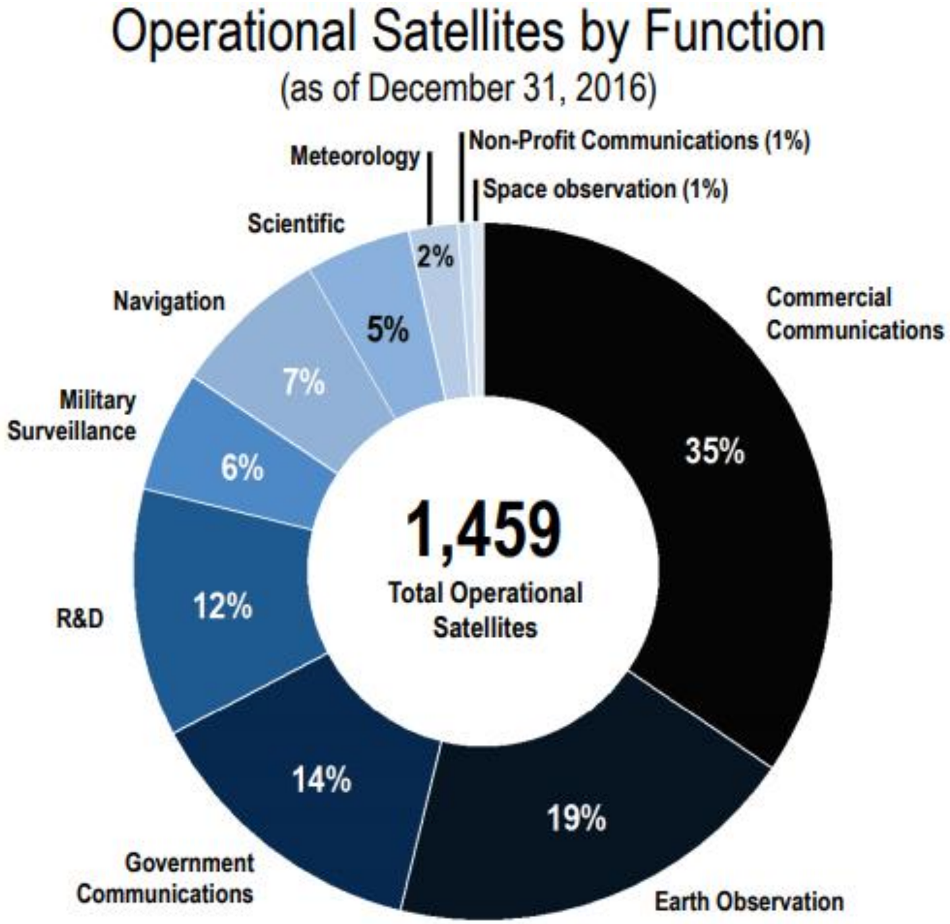


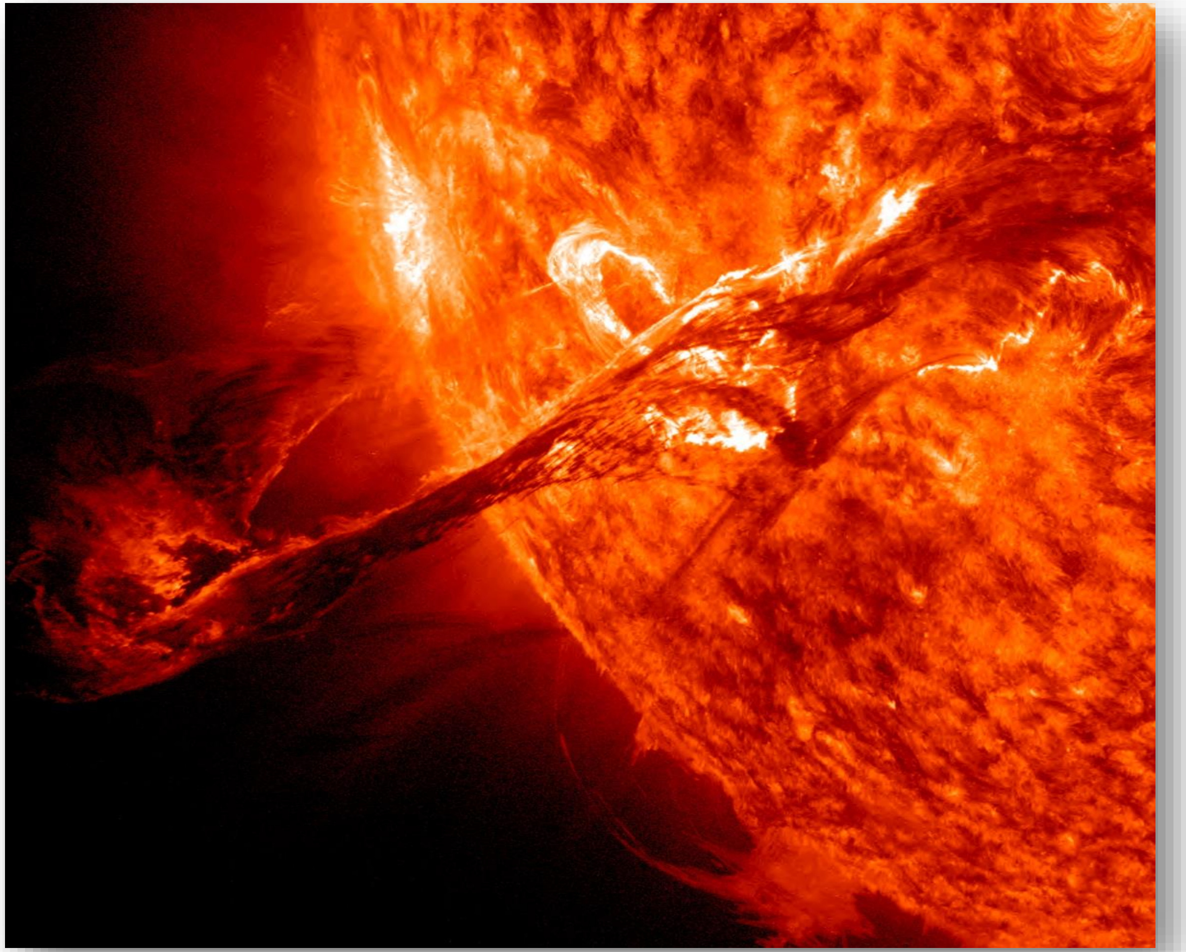
*Figure Courtesy: SIA (Satelite Industry Association)

If Satellites Stop:

- No Telecommunications
- No Military surveillance
- No Weather forecast
- No GPS

What can cause this disaster?

Coronal Mass Ejections (CMEs)



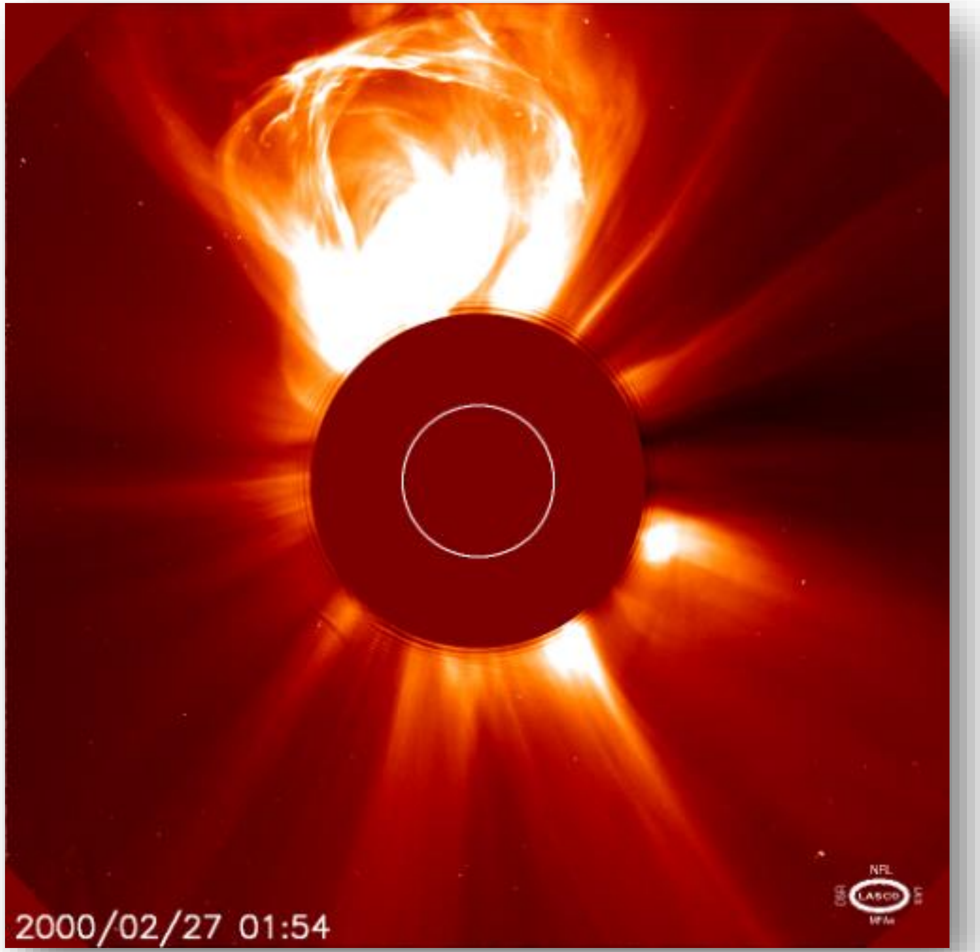
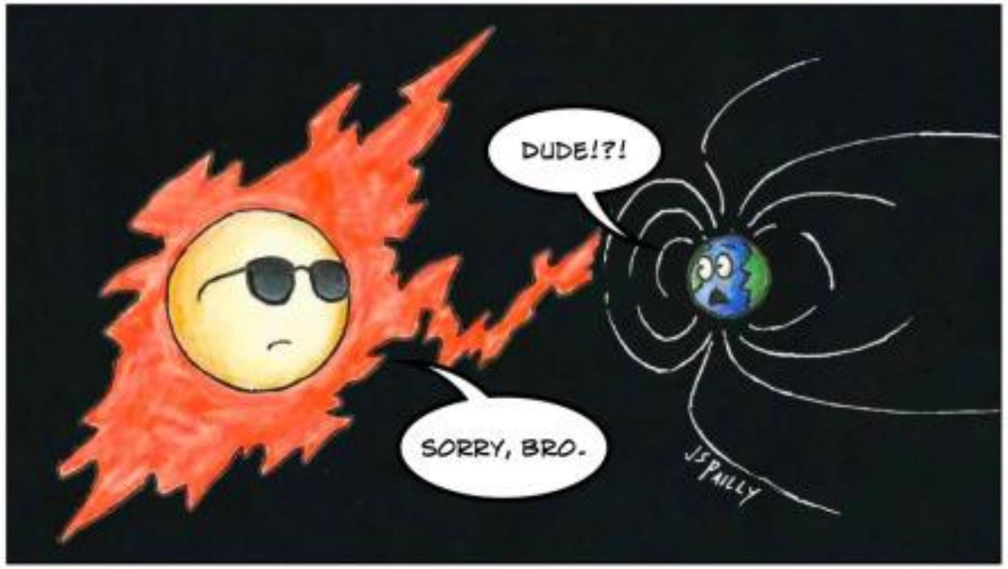


*Figure Courtesy: NASA/ESA, SDO and SOHO satellites



Theory

*Figure Courtesy: https://planetpailly.com/

Coronal Mass Ejections – CMEs

<u>CMEs</u>
Energetic events from the Sun

⇩

Particles in space
+
Disturbances to Earth's magnetic field

⇩

Problems in
Satellites/Communications/Grids/infrastructures etc.

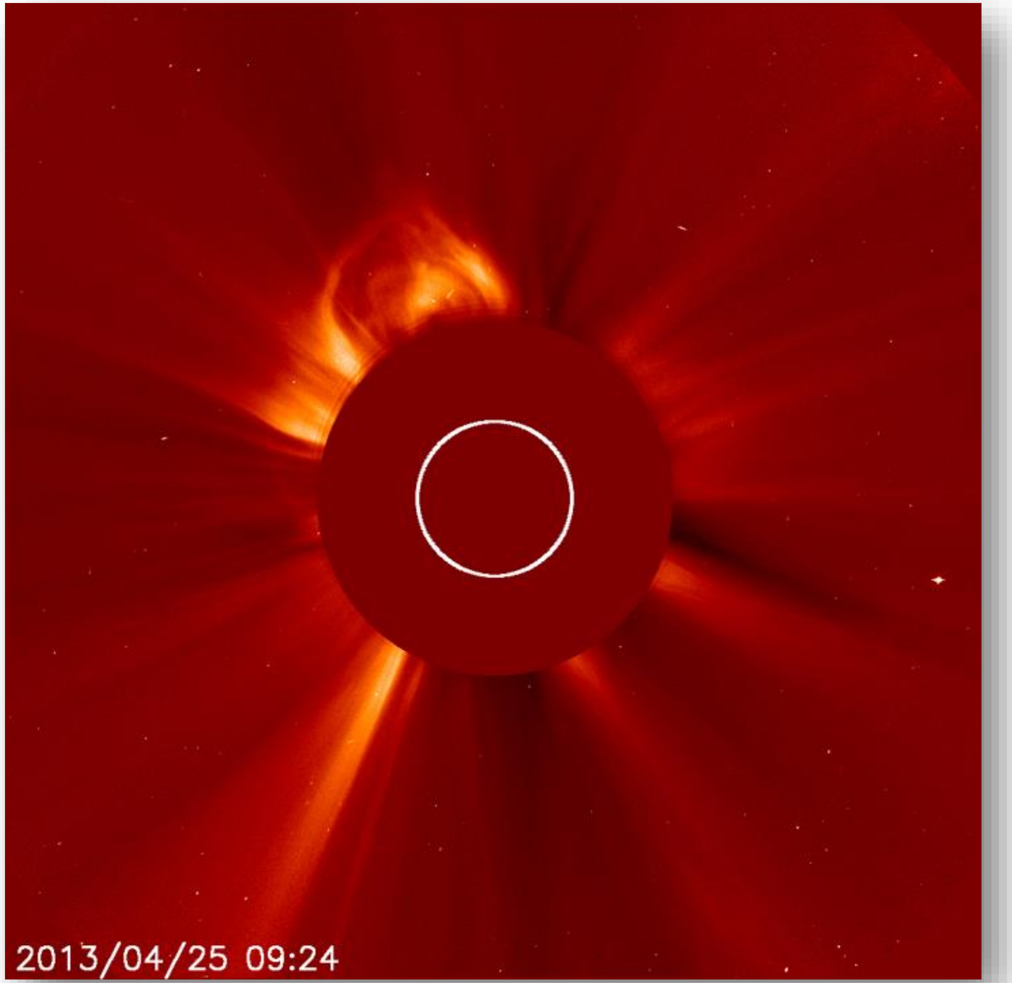


*Figure Courtesy: NASA/ESA, SOHO satellite

Halo CMEs

<u>Halo CMEs</u>
Earth-directed CMEs. Can be seen from coronagraph.

<u>Why important?</u>
Going to Earth

⇩

More effects on mankind

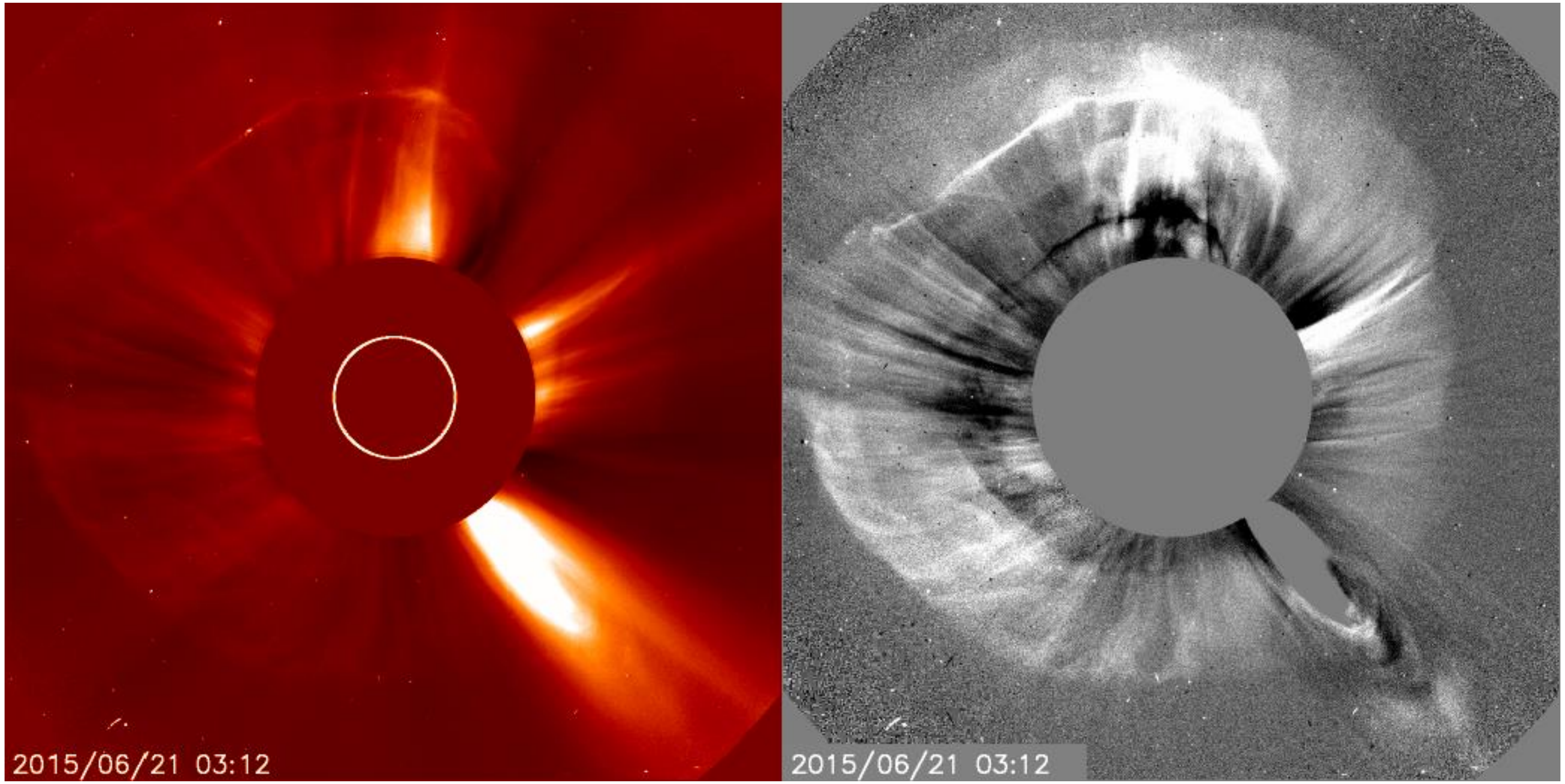
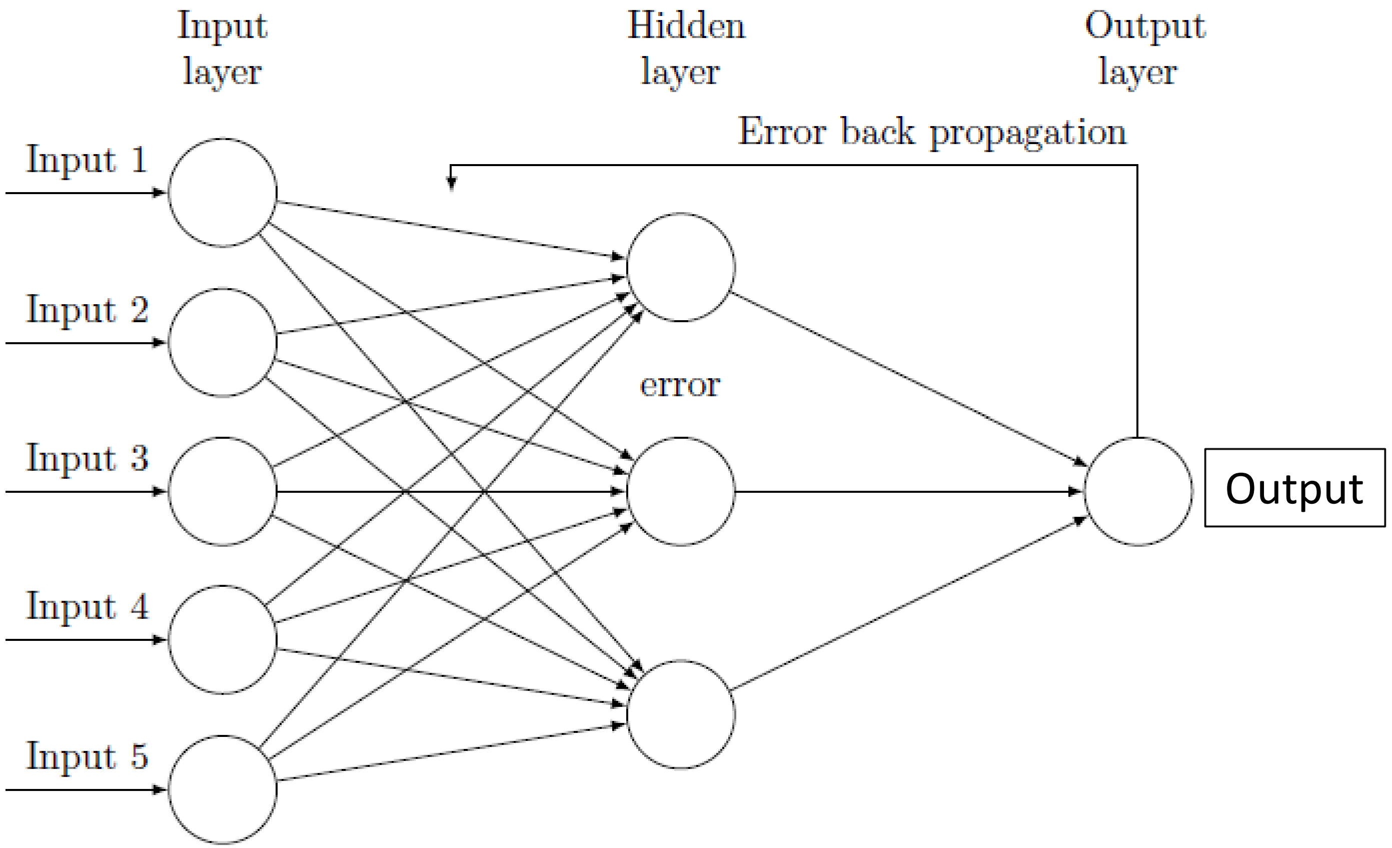


*Figure Courtesy: NASA/ESA, SOHO satellite

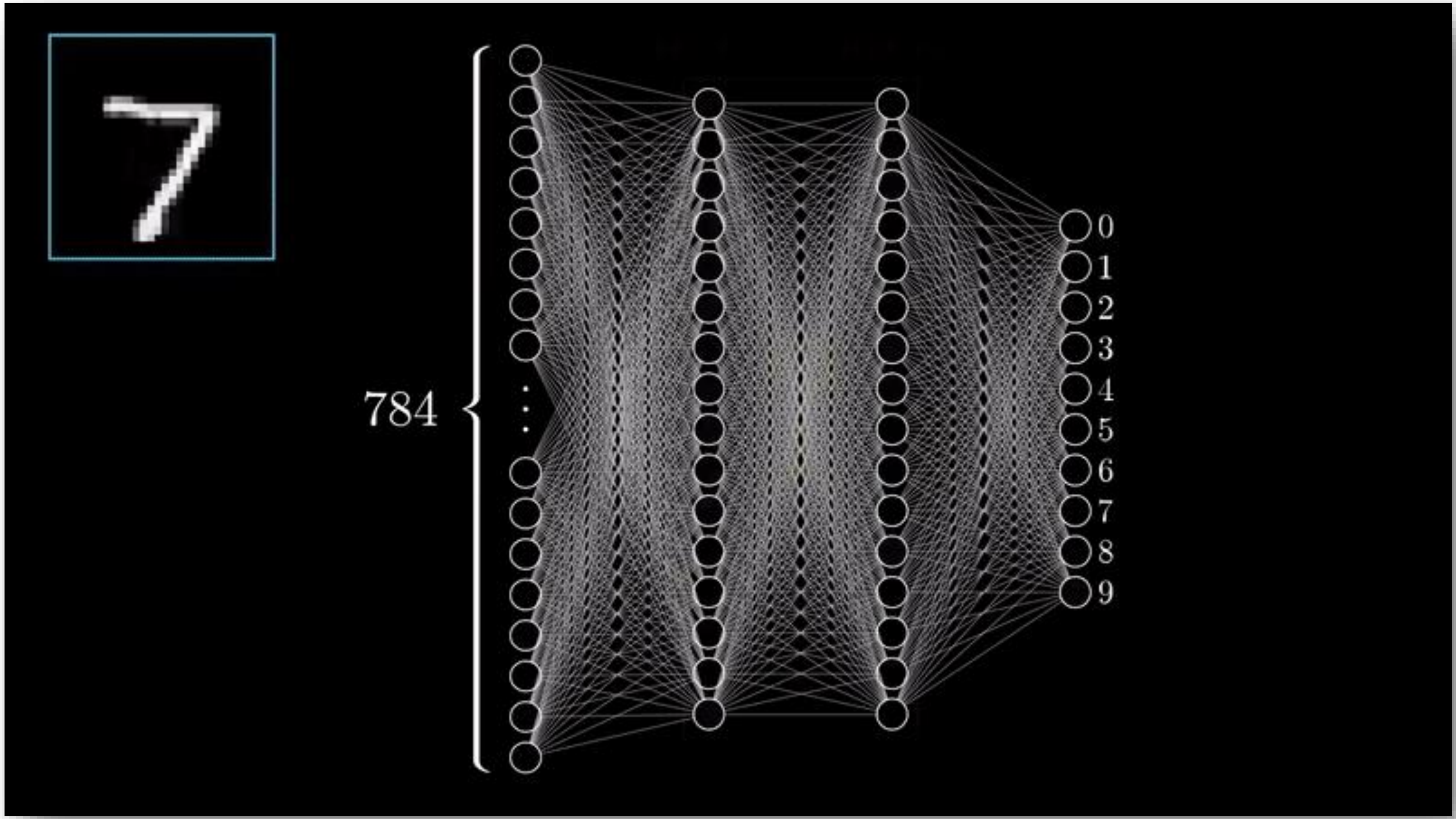
What is machine learning & A.I ?

Making the computer "learn" from data without being explicitly programmed

Neural Networks & Backpropagation



Visualization of Neural Network



*Video Courtesy: **3Blue1Brown** (Check him on YouTube!)

Convolution Neural Network (CNN) Layers

Convolution
Extract features & Keep spatial relationship

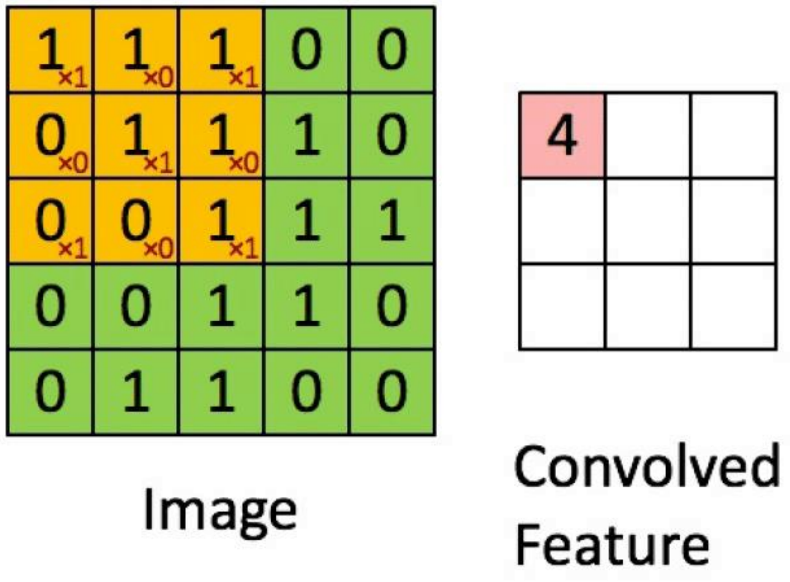


*Figure Courtesy: Erik Reppel

Pooling/Subsampling
Reduce dimensionality & retain information

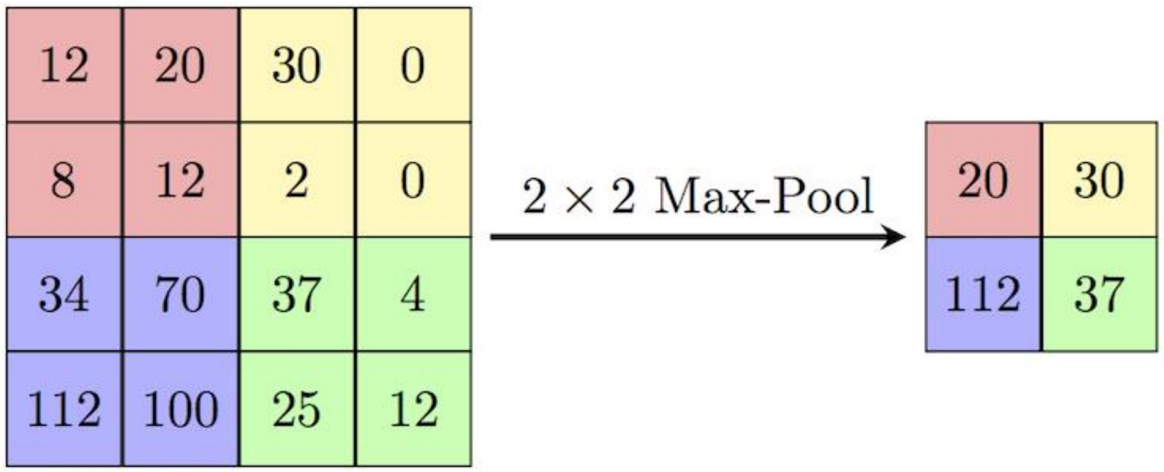


*Figure Courtesy: Cambridge Spark Ltd

Example of CNN

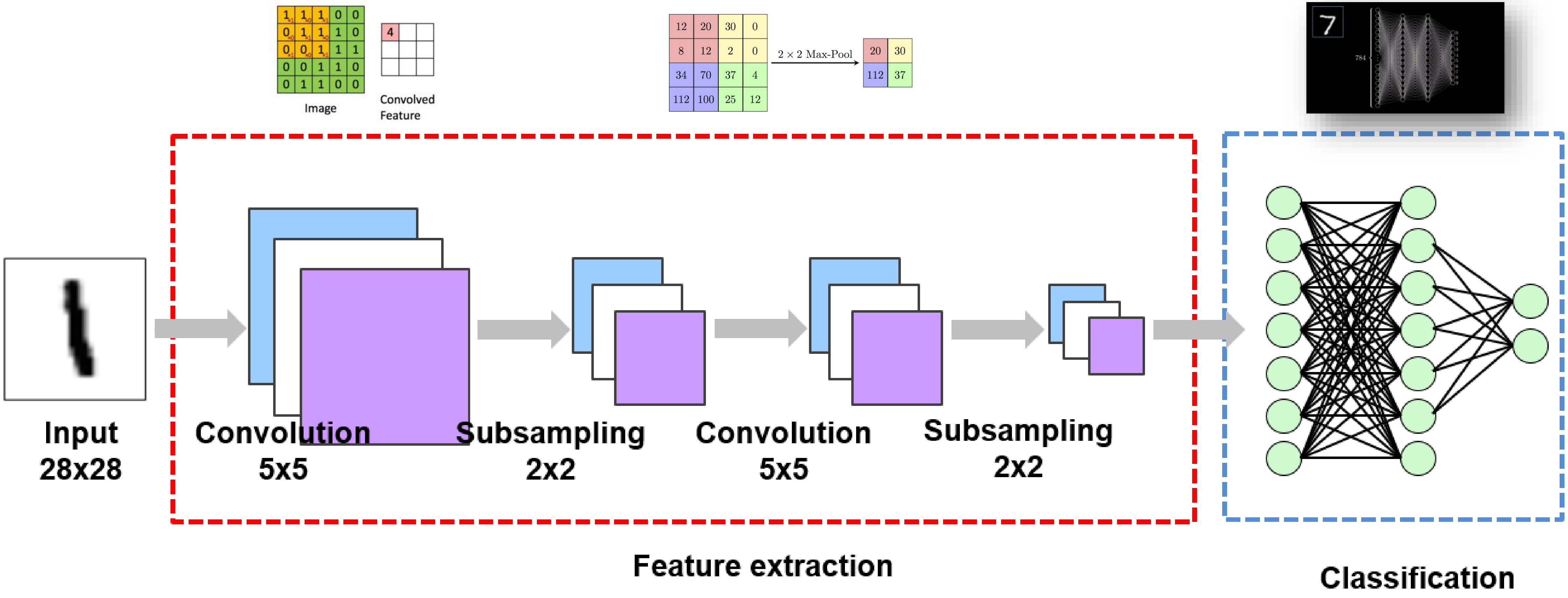


*Figure Courtesy: Suhyun Kim iSystems Design Labs

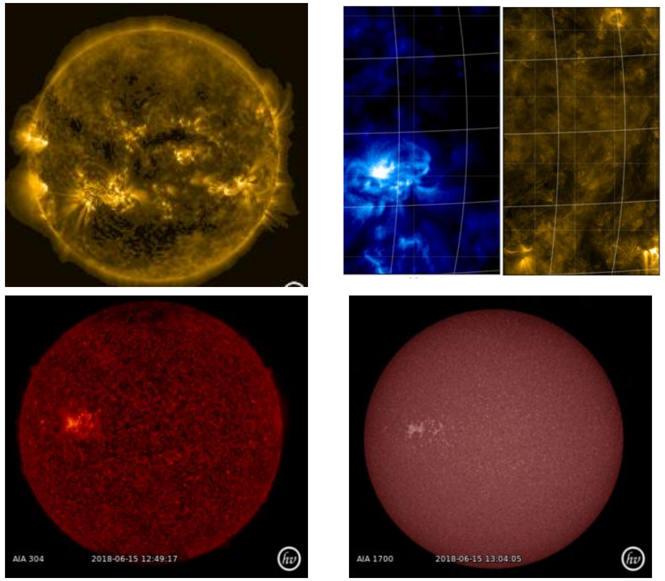
Analysis

Main Goal

Forecast the emerged CMEs using solar images taken from SDO and CNN

Input

SDO Images during 2014



Output

LASCO/CACTUS Catalogs

Date		Characteristics					
2014/01/02	13:48:06	184	57	894	959	825	711
2014/01/03	00:24:05	264	18	225	272	169	0
2014/01/03	02:24:06	51	24	657	637	674	720
2014/01/03	03:47:08	61	44	1132	1303	961	965
2014/01/03	07:36:05	62	17	250	193	306	615
2014/01/03	10:36:05	65	21	316	273	358	627
2014/01/03	12:36:05	265	25	277	287	267	34
2014/01/03	18:00:06	154	60	208	114	295	430
2014/01/03	18:48:05	90	31	89	179	0	0
2014/01/03	19:36:05	222	112	286	331	237	0

Machine Learning Project

1st Part Data Enhancement

2nd Part CNN implementation

Improving Input Project

3rd Part Pre-processing Tool & History Maps

1st Part Data Enhancement

2nd Part CNN implementation

3rd Part Pre-processing Tool & History Maps

The Data Enhancement Project

Data cleaning (Multiple CME) + Data enhancement (Halo CME) + Data integration (No CME dates) = Enhanced CACTUS/LASCO catalogs

~~0025,1997/09/21 02:07,01,244,014,0484,0020~~
0026,1997/09/21 02:07,01,288,018,0384,0027

$\theta > 90^{\circ}$

0006,1997/09/09 20:06,03,279,108,0609,0124,0316,0813, ✓
0007,1997/09/13 06:25,01,258,014,0349,0771,0237,1922, ✗

Add non-CME date when:

$$t_{i+1} - t_i \geq 3\ [h]$$

Date	**CME** [0/1]	**Halo CME** [0/1]
2014/01/01 00:12:05	1	0
2014/01/04 23:12:05	1	1
2014/05/03 20:42:00	0	0

1st Part Data Enhancement

2nd Part CNN implementation

3rd Part Pre-processing Tool & History Maps

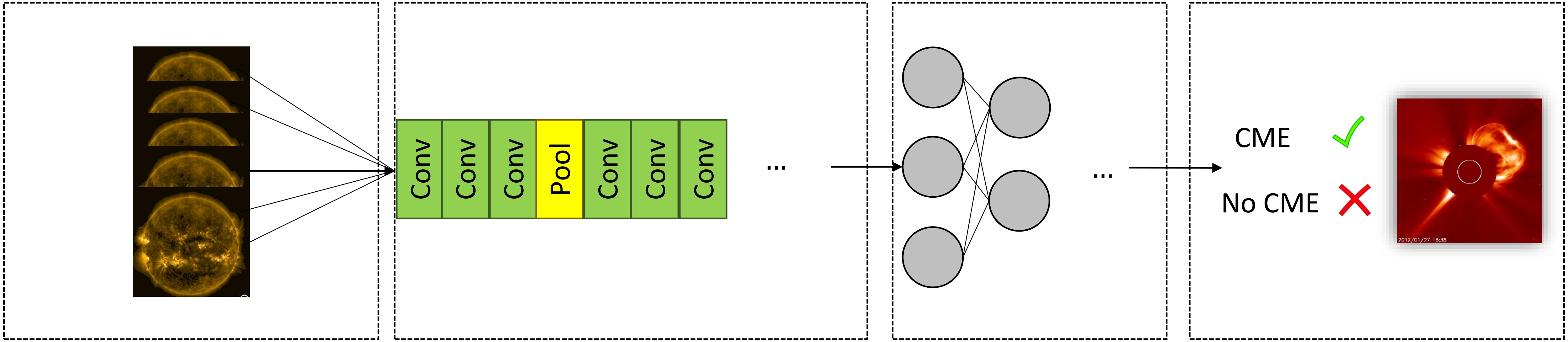
The Machine Learning Project

<u>Input</u>
[13,512,512]

<u>Convolution Neural Network</u>
Feature Extraction

<u>Dense Neural Network</u>
Classification

<u>Output</u>
CME / No CME

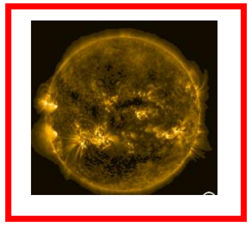


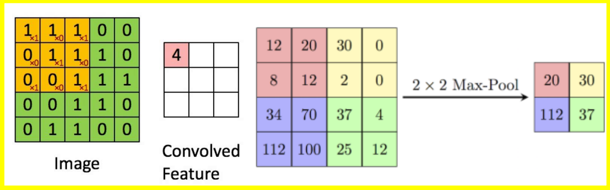
Input = 13 SDO images, 2 [h] history before the event.

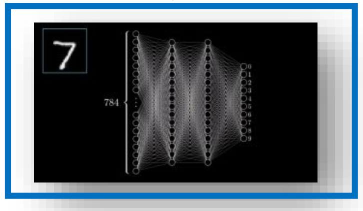
Output = 1/0

Final CNN architecture

Layer	Details & Operations	Output shape
Input	-	[512,512,13]
Convolution	Convolution [14] & 3x3 Kernel	[510,510,14]
Convolution	Convolution [16] & 3x3 Kernel	[508,508,16]
Convolution	Convolution [18] & 3x3 Kernel	[506,506,18]
Max Pooling	Max Pooling with 2x2 Kernel	[253,253,18]
Dropout	20 % Dropout	[253,253,18]
Convolution	Convolution [20] & 3x3 Kernel	[251,251,20]
Convolution	Convolution [28] & 3x3 Kernel	[249,249,28]
Convolution	Convolution [36] & 3x3 Kernel	[247,247,36]
Max Pooling	Max Pooling with 2x2 Kernel	[247,247,36]
Dropout	20 % Dropout	[123,123,36]
Convolution	Convolution [40] & 3x3 Kernel	[121,121,40]
Convolution	Convolution [56] & 3x3 Kernel	[119,119,56]
Convolution	Convolution [72] & 3x3 Kernel	[117,117,72]
Max Pooling	Max Pooling with 2x2 Kernel	[58,58,72]
Dropout	40 % Dropout	[253,253,18]
Convolution	Convolution [80] & 3x3 Kernel	[56,56,80]
Convolution	Convolution [112] & 3x3 Kernel	[54,54,112]
Convolution	Convolution [144] & 3x3 Kernel	[52,52,144]
Max Pooling	Max Pooling with 2x2 Kernel	[26,26,144]
Flatten	Flattening of the input	97344
Fully Connected	400 Neuron - Dense layer	400
Fully Connected	200 Neuron - Dense layer	200
Fully Connected	2 Neuron - Dense layer	2
Output	Classifier, 0.5 Threshold Sigmoid	2



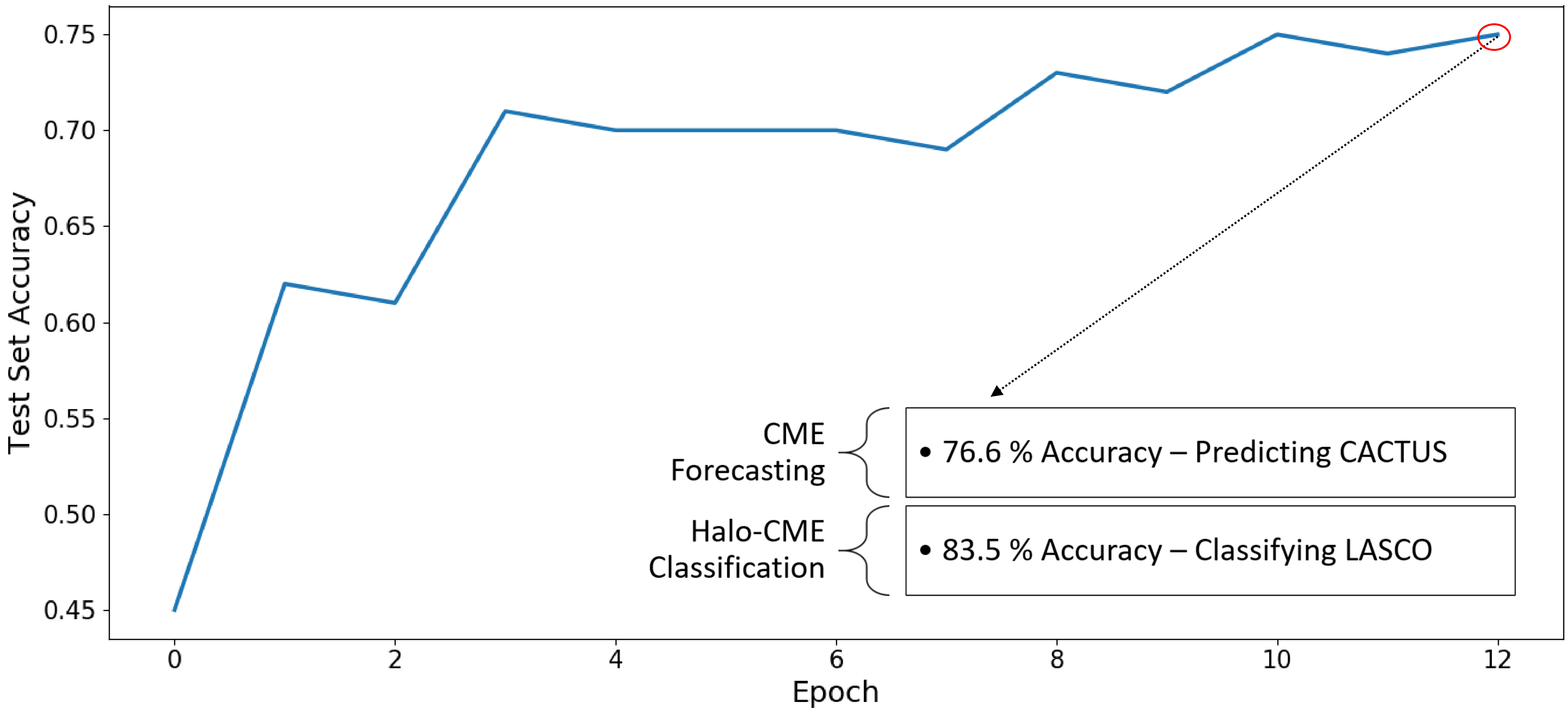




CME ✓

No CME ✗

Result of CNN



1st Part Data Enhancement

2nd Part CNN implementation

3rd Part Pre-processing Tool & History Maps

Pre-processing Tool – Motivation

Previous input

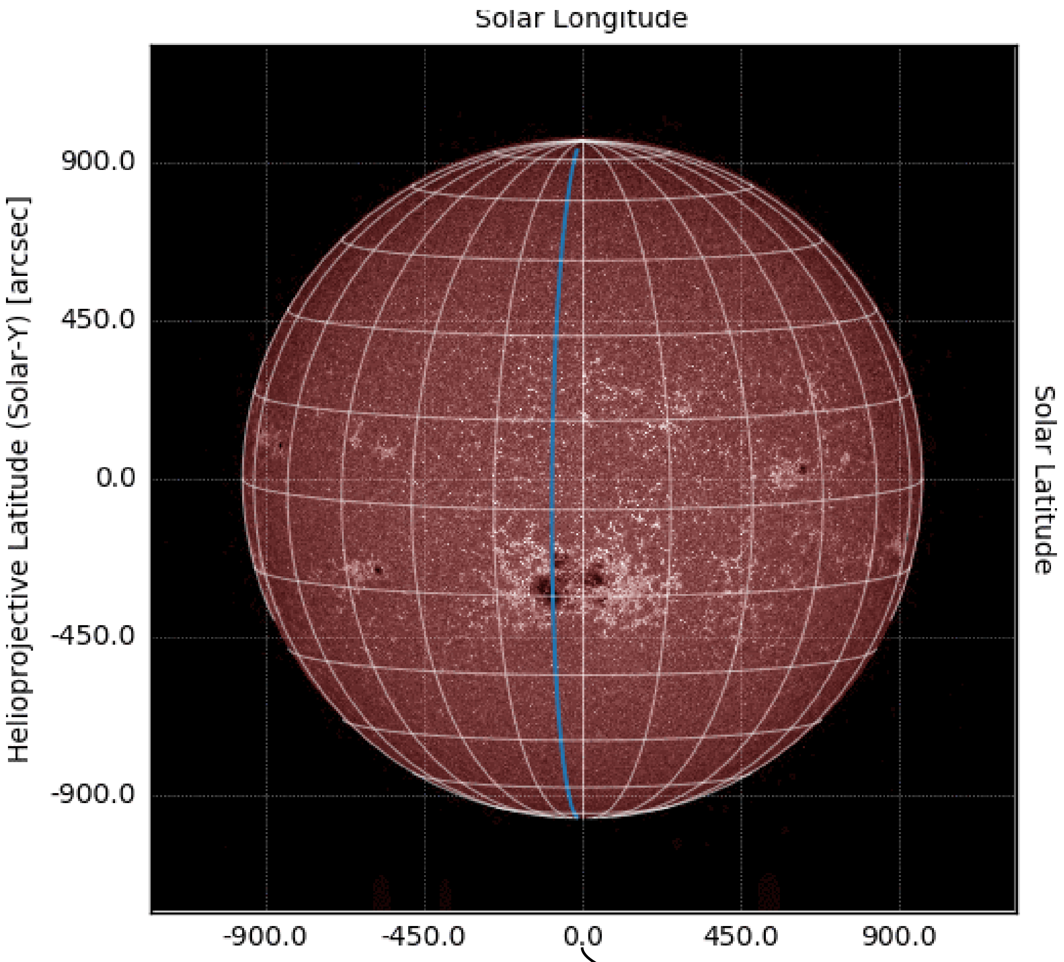
(+) Promising results.

(-) Expensive computationally and memory wise.

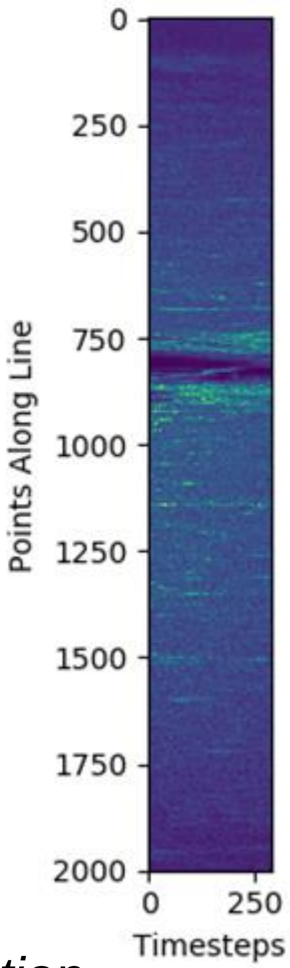
Using the pre-processing tool

(+) **New input** → less computational time & memory consumption.
(?) Better results.

History Map (HM) – Single Line Example

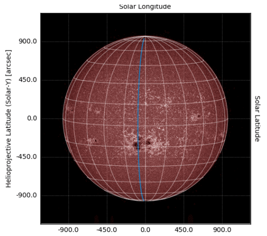


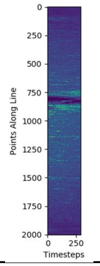
HM



Same blue line, different representation

Pre-processing Tool – Procedure & Output





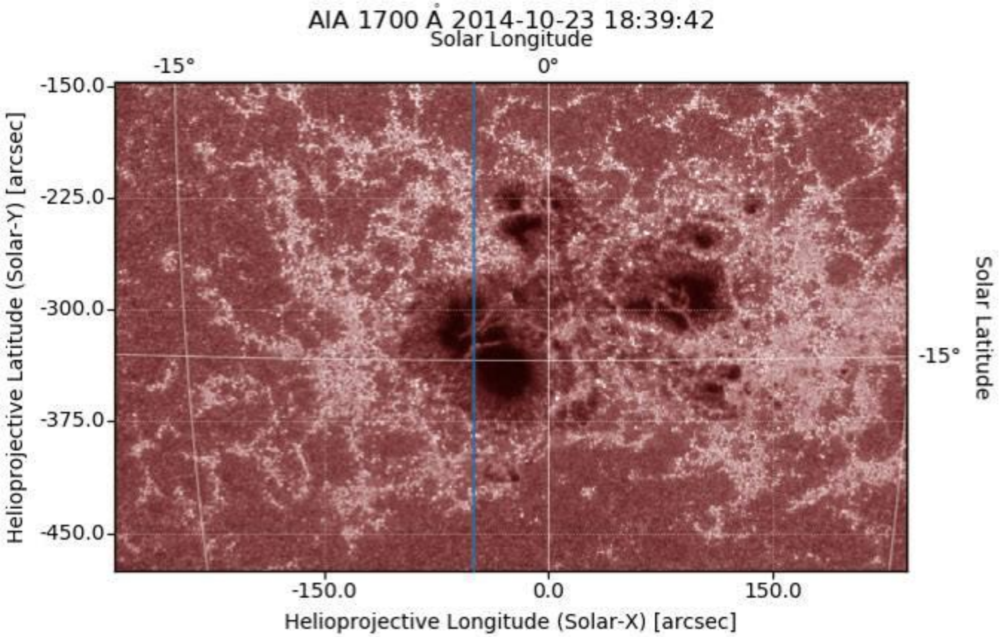
Procedure

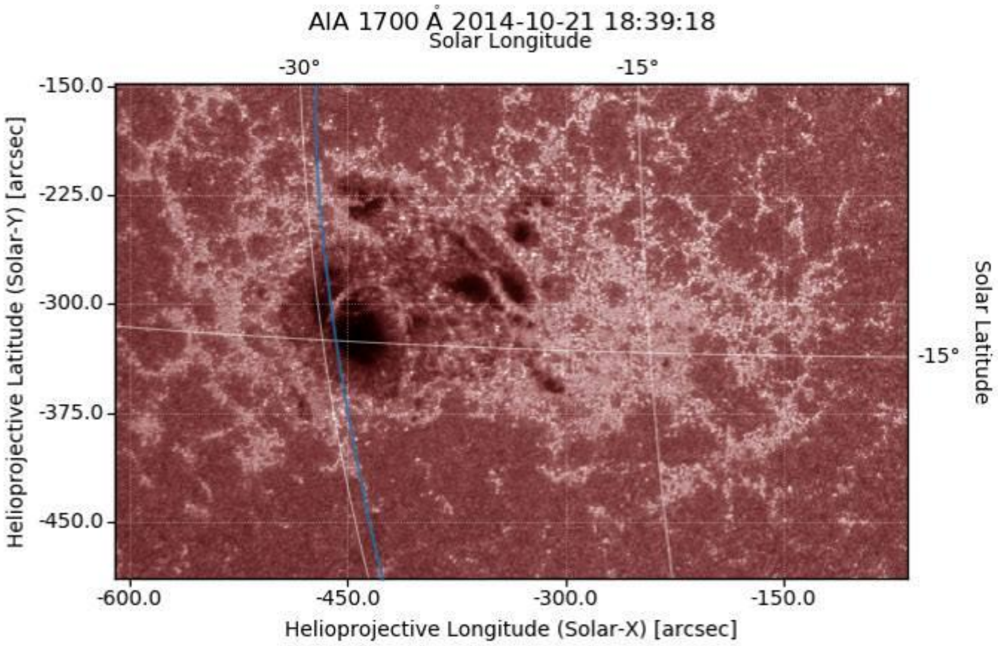
- Download data
- Track Sun’s differential rotation for every longitude line
- Go to next date on Catalog
- Repeat

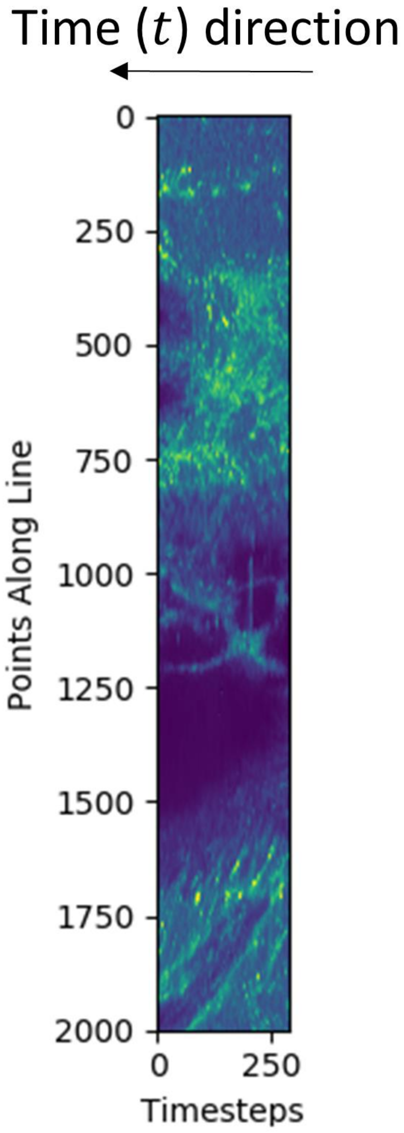
Output

1) Database file (.sql)
2) Scientific images (.FITS)
3) History maps (.png)
4) NN input (.npy)
5) Animation (.gif)

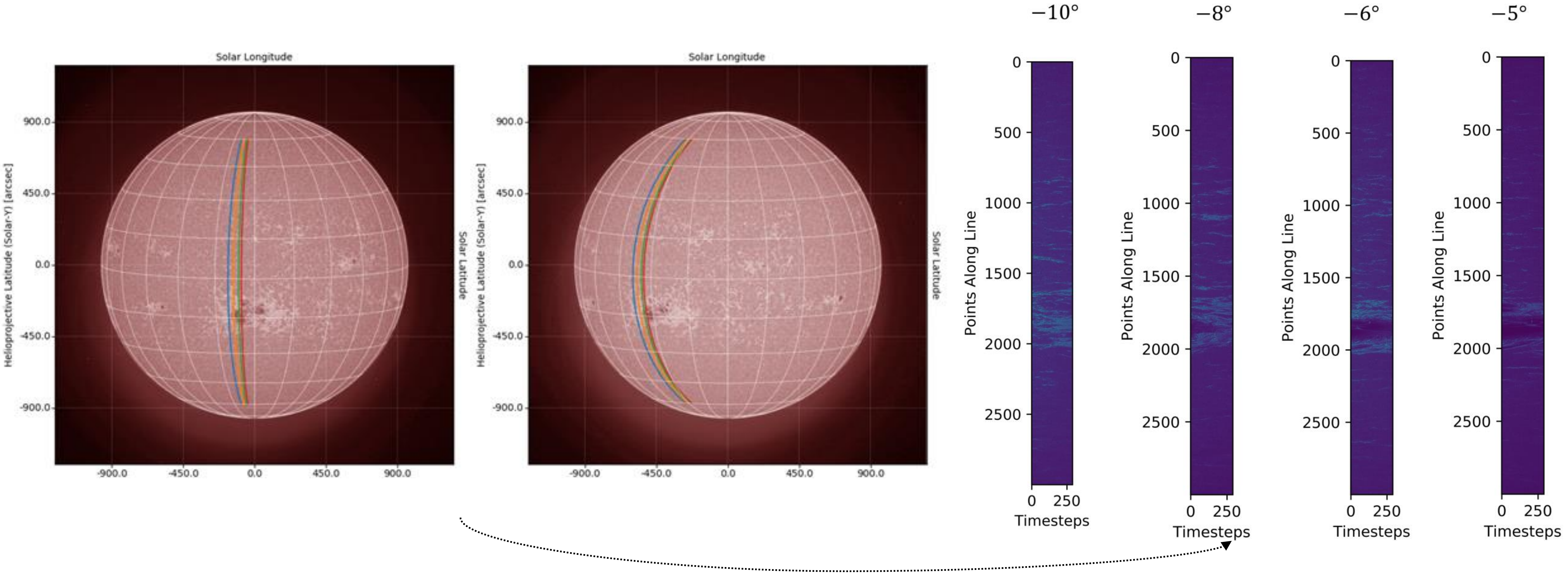
Pre-processing Tool – Sunspot







History Map – Multi Line Example



Why History Maps ?

- Substantial decrease of data and computational time (order of magnitude).

  E.g. Originally: [512,512,13], now [x,y,13]. In practice, $y \approx 10 \ll 512$

- Structures are shown in a frame that is co-moving → time evolution is shown.

- Possibly useful for forecasting other phenomena such as Solar flares or Sunspots.

Conclusion

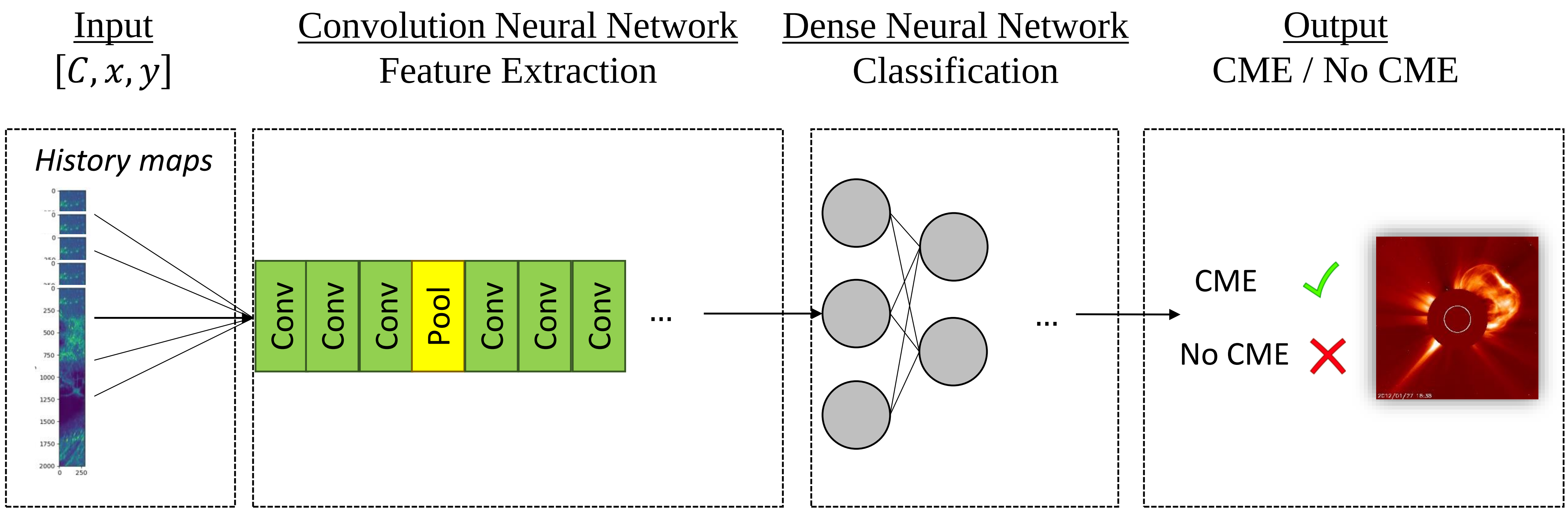
Summary

1. Enhanced, clean and processed SDO data and CACTUS/LASCO catalogs

2. Created multiple CNN models, with the best obtaining 76.6% prediction on CMEs and 83.5% classification between CME and halo-CME.

3. Created a pre-processing tool that derives “History Maps” (HM). Possibly useful in future Machine learning research and Solar data analysis.

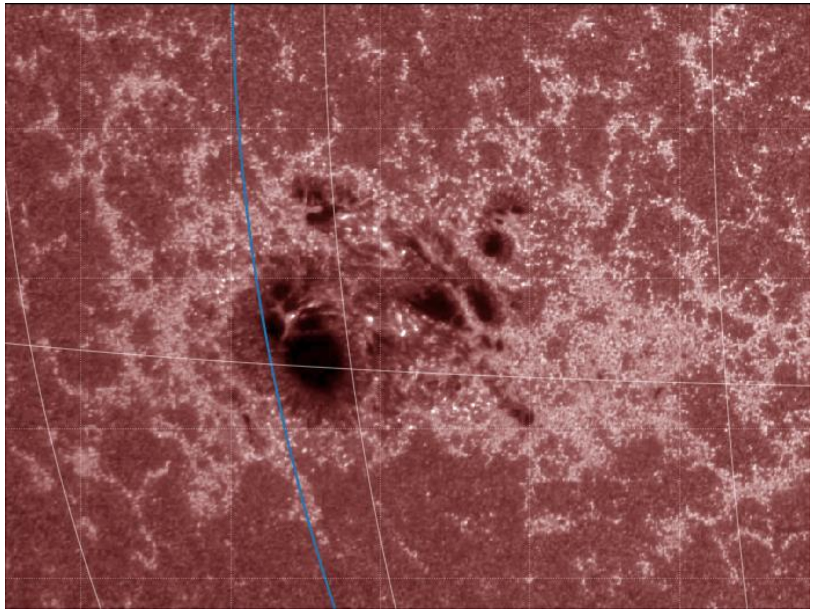
Future Proposal



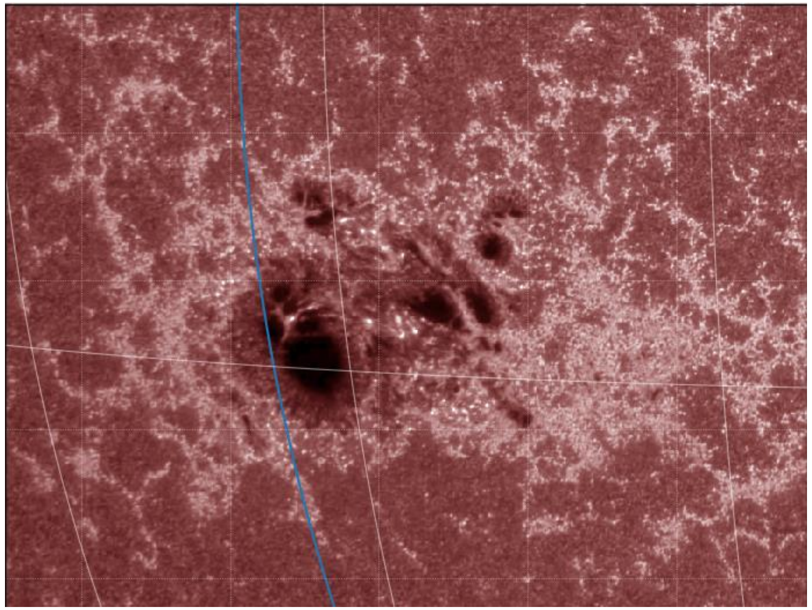
Extras

Differential Rotation Models

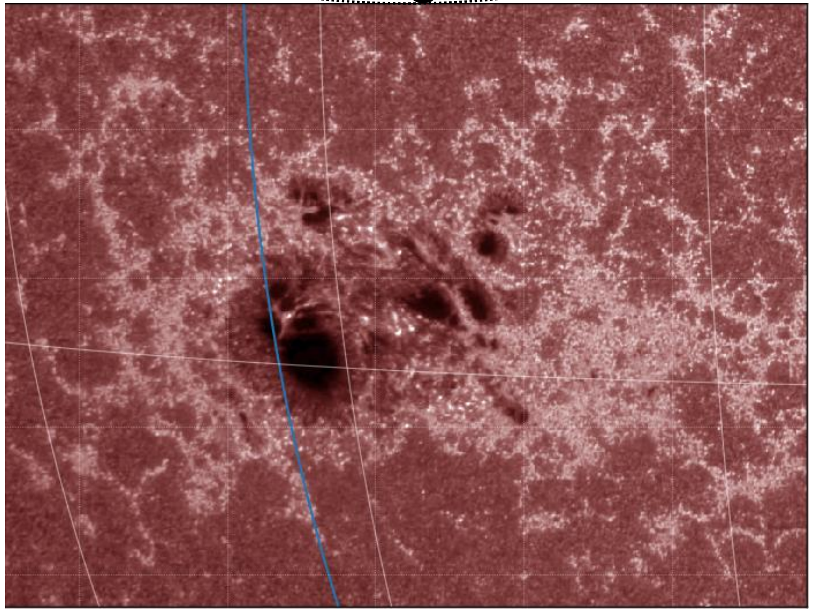
Allen



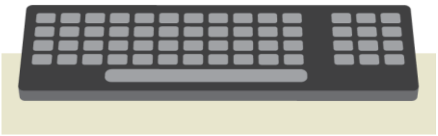
Howard

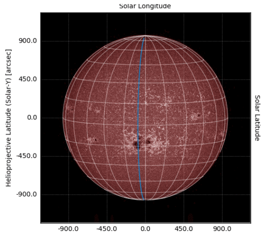


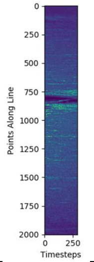
Snodgrass ✓



Pre-processing Tool – Design







Input	Procedure	Output
1) Date – Date of event 2) x – Points on line 3) y – Longitude lines 4) dt – Time-step 5) T – Total time 6) λ – Wavelength 7) C – Catalog	• Download data for T [h] before event • Track Sun's differential rotation for every line (y) using *dt* step • Go to next *date* on Catalog (c) • Repeat	1) Database file (.sql) 2) Scientific images (.FITS) 3) History maps (.png) 4) NN input (.npy) 5) Animation (.gif)