Grand Challenge
UNIVERSITY OF COLORADO BOULDER
SPACE WEATHER CENTER

University of Colorado Boulder
COLLEGE OF ENGINEERING AND APPLIED SCIENCE

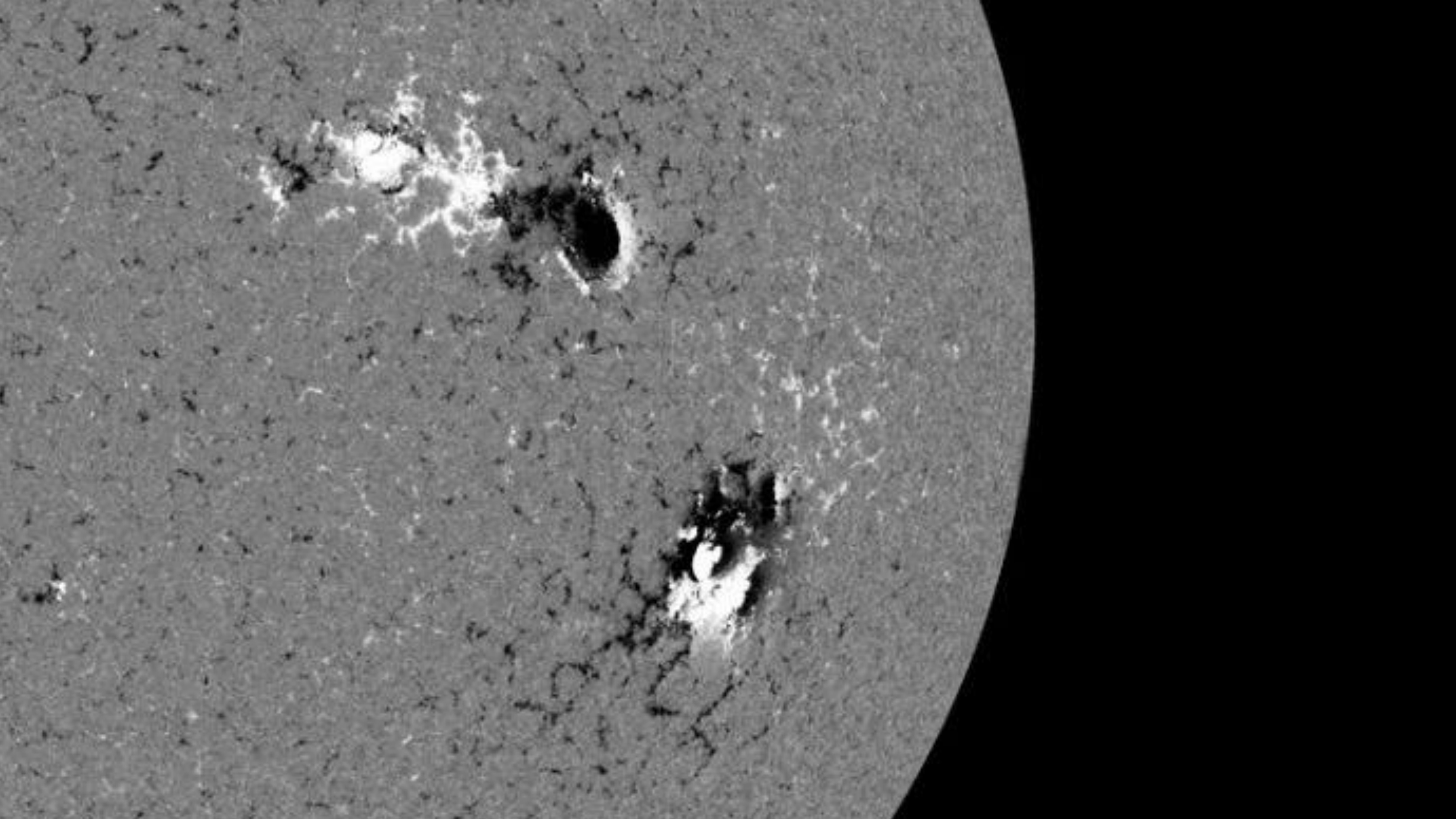
Leveraging Topogical Data Analysis and Deep Learning for Solar Flare Prediction

T. E. Berger[1], V. Deshmukh[1], E. Bradley[1], J. Meiss[1], N. Nishizuka[2]

1. University of Colorado at Boulder
2. NICT, Japan

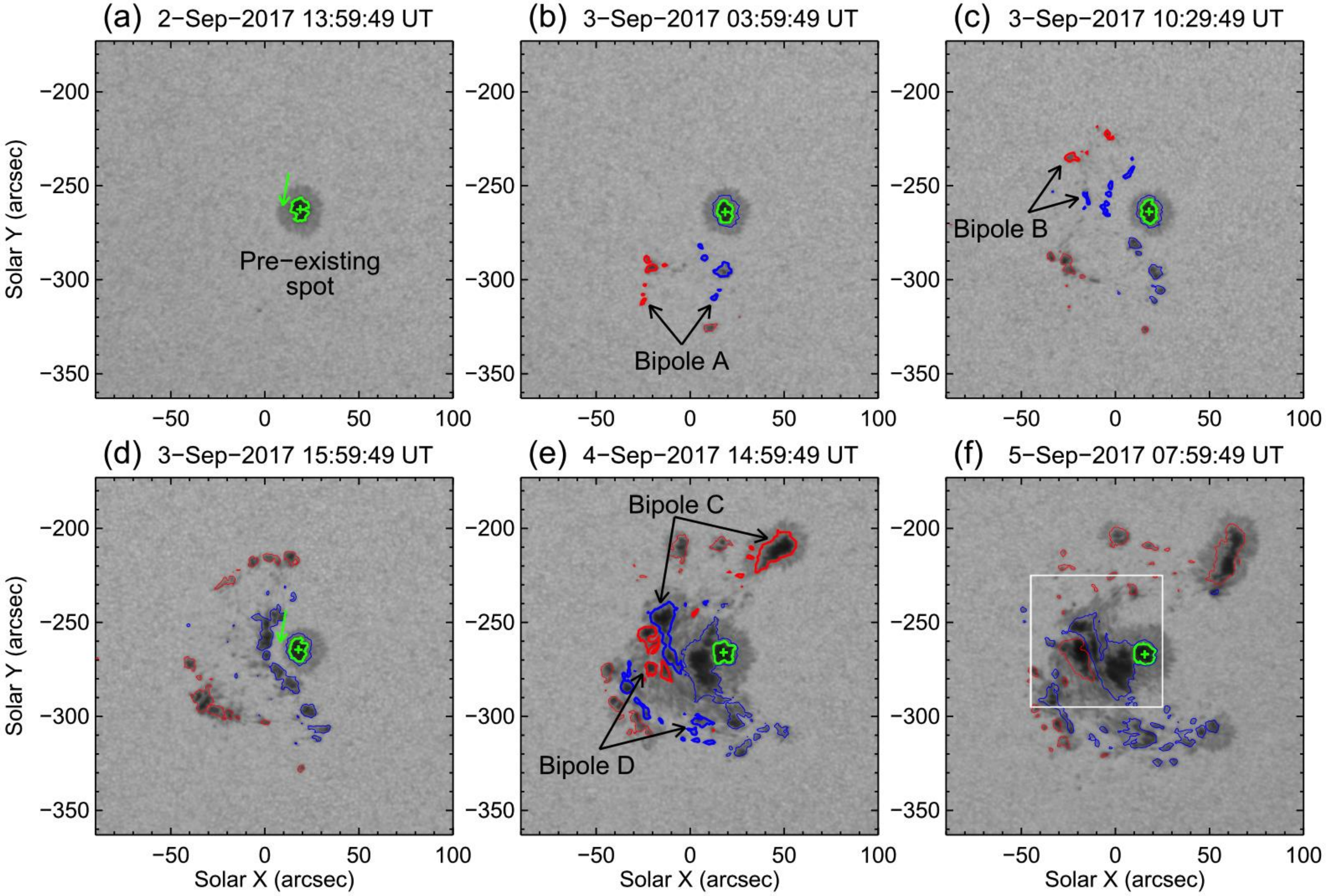
SWx TREC

Courtesy NASA Goddard SVS: svs.gsfc.nasa.gov



Magnetic Evolution to Flaring State

From simple…



…to complex shapes!

Yang et al., ApJ, **834**, 150, 2017

Operational Flare Forecasting

Input data

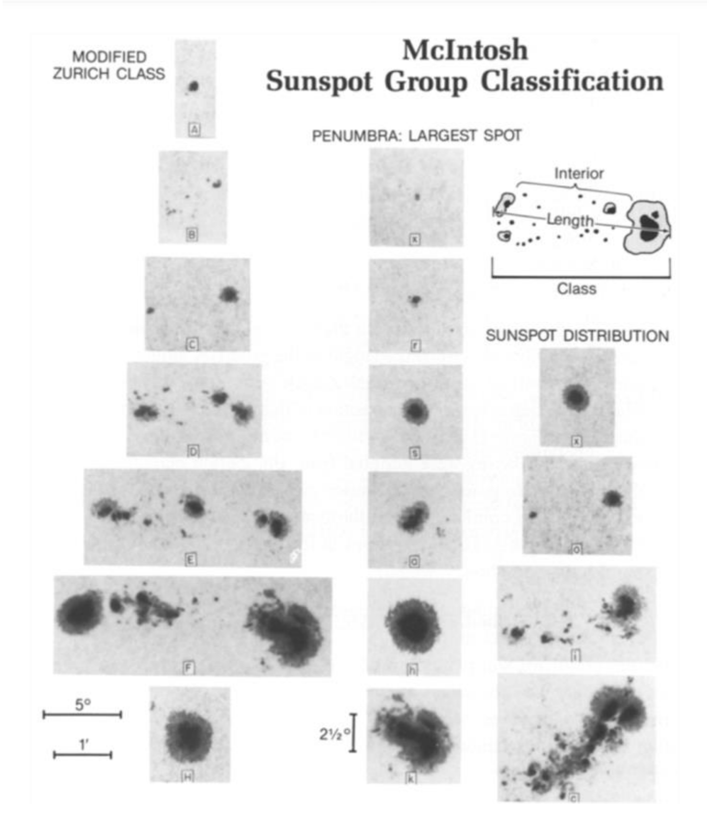
ISOON

GONG

SDO/HMI

Continuum & Magnetogram Images

Classification System



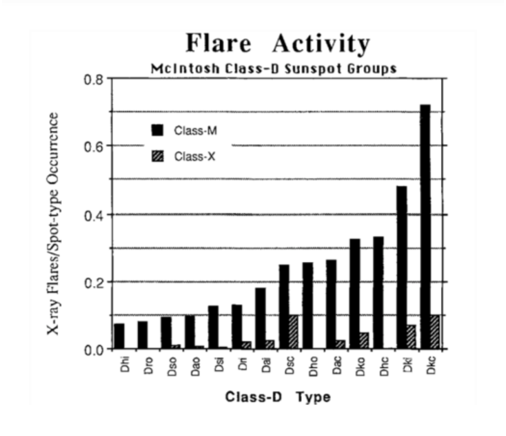
McIntosh, SolPhys, **125**, 251, 1990

McIntosh and/or Mt. Wilson Classification

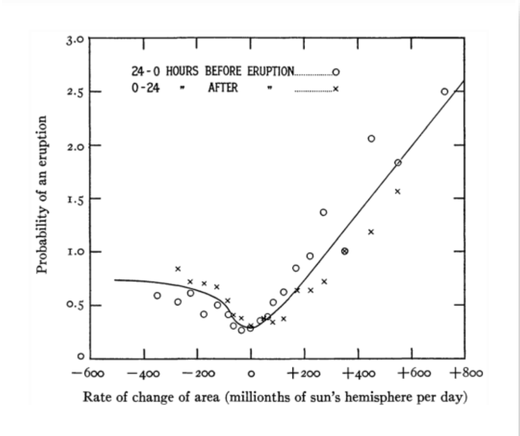
Analysis

Human Forecaster Processes

- Climatology look-up table: P_f for given class
- Growth/decay in spot & total AR area
- Flaring History
- Forecaster expertise



McIntosh, SolPhys, **125**, 251, 1990



e.g., Giovanelli, ApJ, **89**, 555, 1939

Output

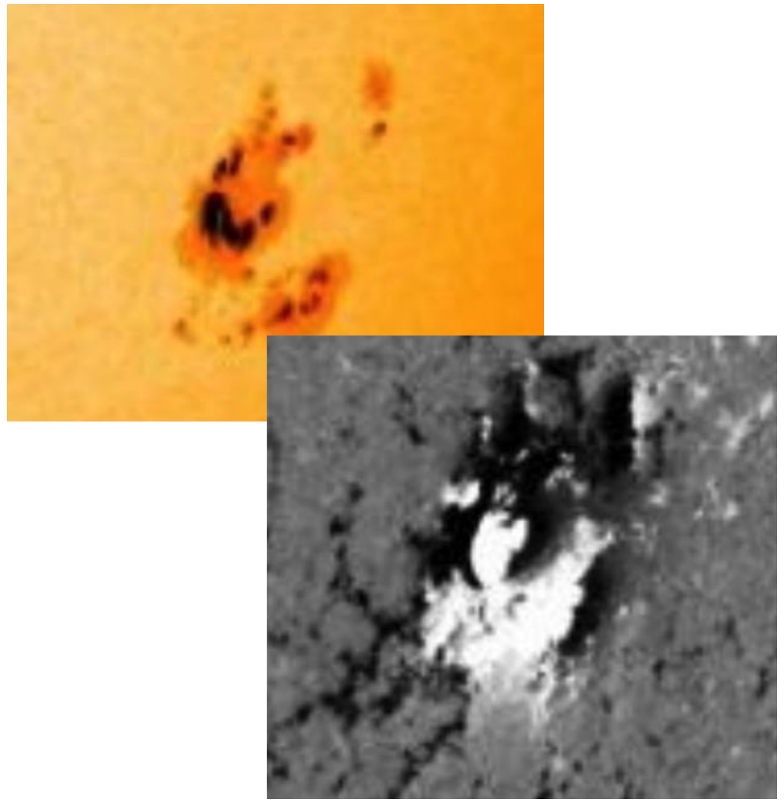
Probability of X-ray flare of magnitude C,M,X

P_f (n)

n = 24, 48, 72 hours

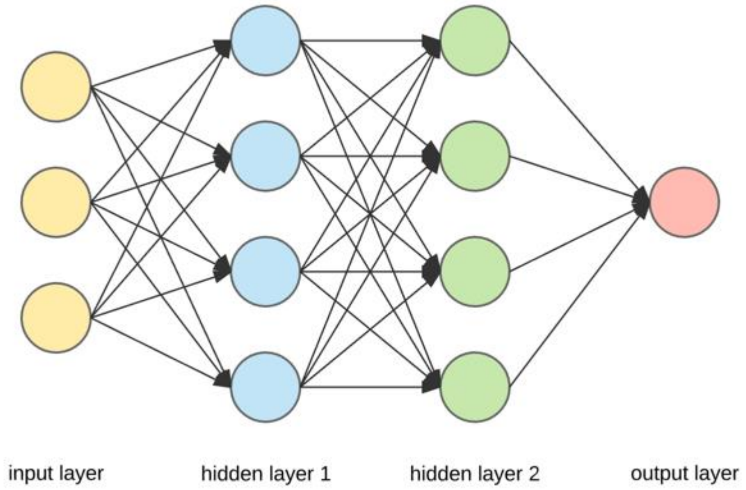
Data-Driven Approach to Flare Forecasting

Input data | **"Feature Vector"** | **Machine Learning Model** | **Output**



Continuum, Magnetogram, Hα, EUV, X-ray, … Images

Typically physics-based characteristics of the active region



SVM, decision tree, MLP network, CNN, etc…

Probability of X-ray flare of magnitude C,M,X

$P_f(n)$

$n = n_1, n_2, n_3, \ldots$ hours

We propose a feature engineering approach based on formalizing the shape of the photospheric magnetic structure for flare prediction using <u>Computational Geometry</u> and <u>Topological Data Analysis</u>

Approach: Interaction and Shape Persistence

Geometry-based Approach

- Can we quantify the interaction of polarity structures in Active Regions for ML input?
- Background/Motivation:
  - Hale and McIntosh Classifications
  - Fractal dimension for flare prediction (McAteer et al., 2009)

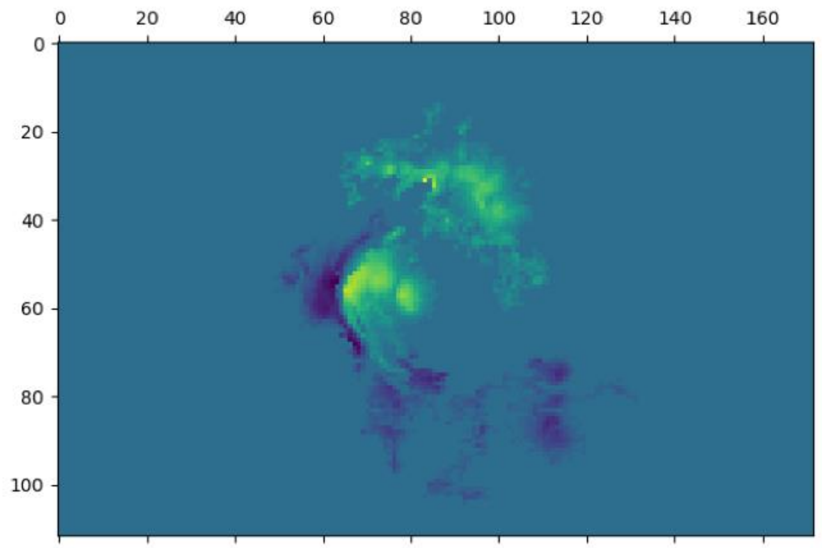
Topology-based Approach

- Can we quantify the "landscape" of **radial magnetic flux density maps**?
- Background/Motivation:
  - "Magnetic Charge Topology" (e.g., Barnes et al., 2005)
  - Euler Characteristic for flare indication (Knyazeva et al., 2014)

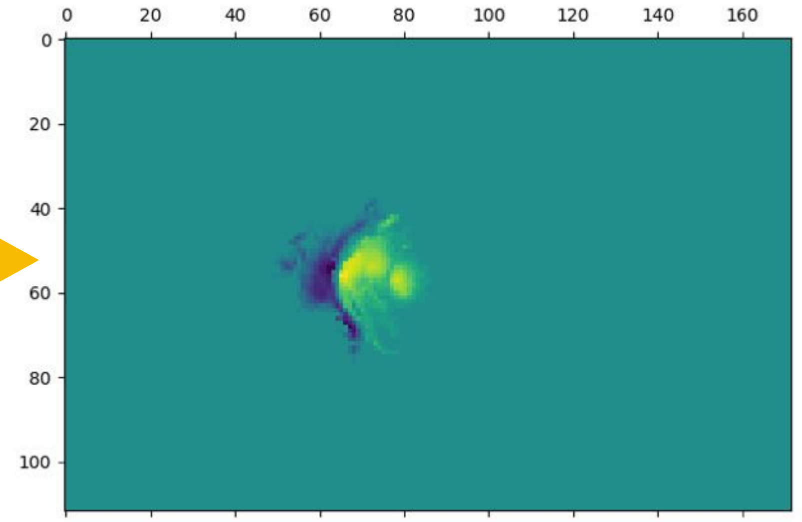
Computational Geometry: Polarity Interaction

1. Compute interaction factor (aka "Ising Energy", Ahmed et al., 2010) between all pairwise positive/negative components:

$$I = \frac{B_r^+ B_r^-}{r_{min}^2}$$



Radial Magnetic Field map
$|B_r|$ > 200 gauss



Most Interacting Pair (MIP)
highest I value

Name	Units	# of features
Total number of positive (negative) clusters	*integer*	2
Size of largest positive (negative) clusters	*arcseconds*	2
Interaction factor of MIP	$G^2/arcseconds^2$	1
COM distance between positive and negative elements of MIP	*arcseconds*	1
Smallest distance between elements of MIP	*arcseconds*	1
COM distance to smallest distance ratio	*arcseconds*	1
Total magnetic flux of each element of MIP	G	2
Size of each element of MIP	$arcseconds^2$	2
Total magnetic flux per unit area of each element of MIP	$G/arcseconds^2$	2
Total magnetic flux of largest elements in the magnetogram	G	2

Geometric Feature Vector
(16 elements)

2. Derive secondary geometric properties of MIP (center of mass distance, area, unsigned flux, minimum distance, …) and overall active region.

3. Assemble feature vector.

Topology Approach: Shape of Magnetic Field

- Topology = fundamental mathematics of shape: invariant under continuous transformations (homeomorphisms)
- Abstract, metric-free, description of sets: number of pieces β_0, number of holes β_1, number of voids β_2, …. “***Betti numbers***”

Same Topology:

$B_0 = 1, \beta_1 = 1$

$B_0 = 1, \beta_1 = 1$

Different Topology:

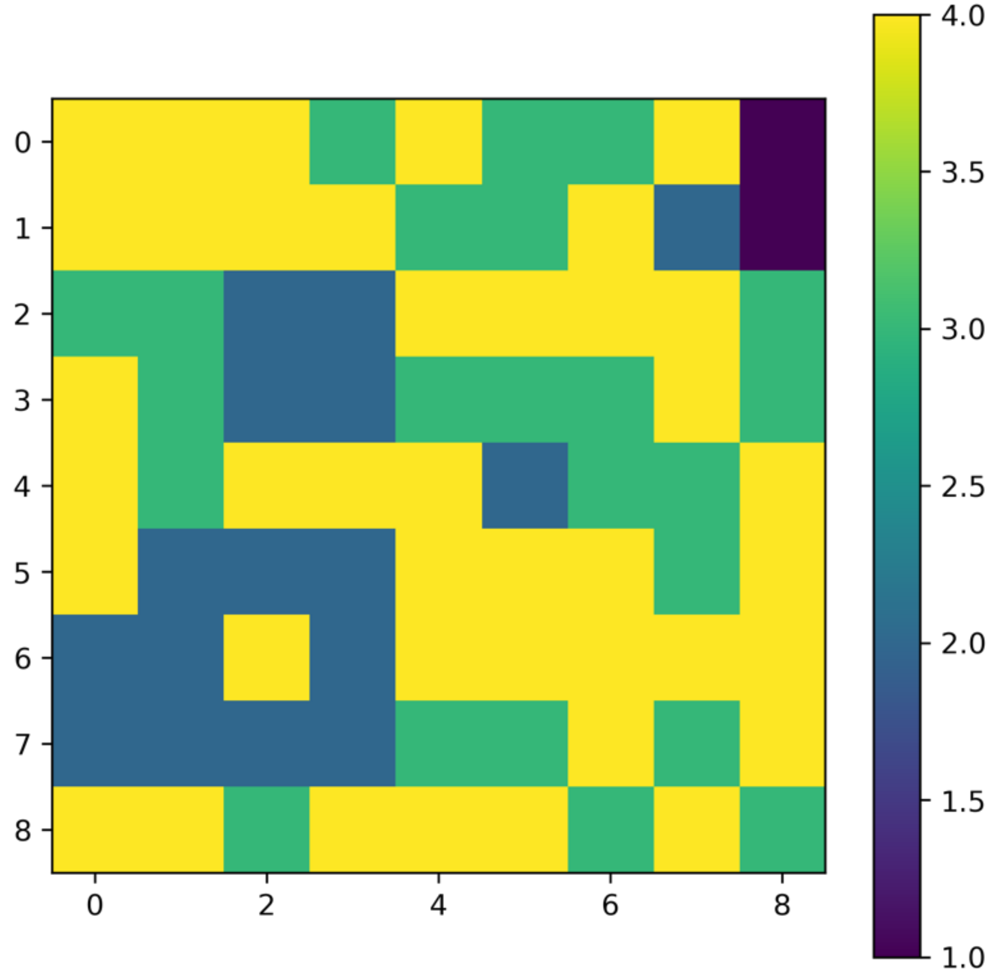
$B_0 = 1, \beta_1 = 100$

$B_0 = 1, \beta_1 = 0$

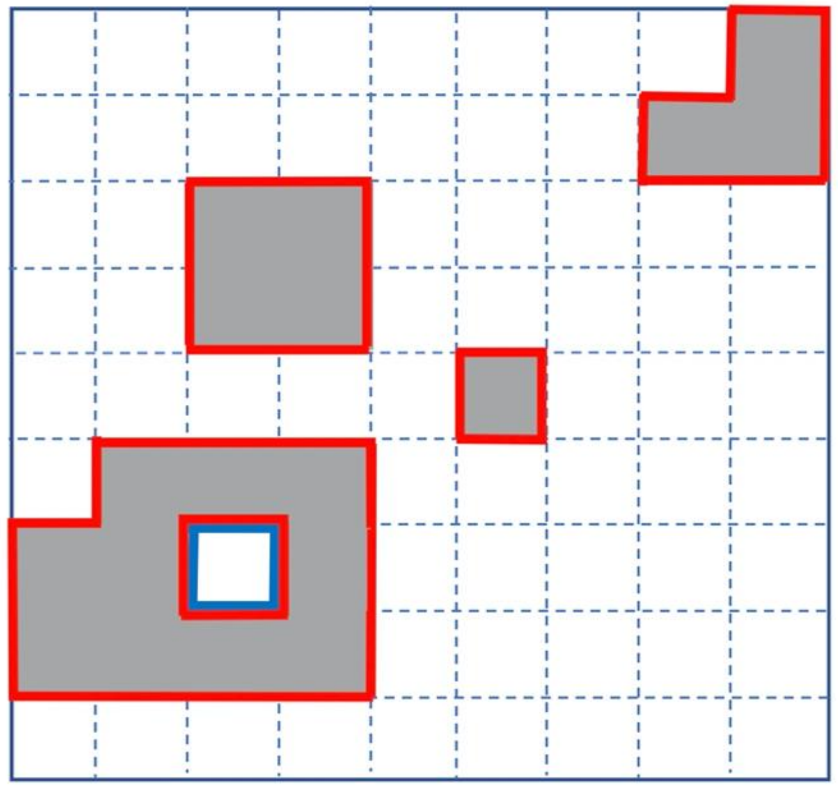
- **Topological Data Analysis (TDA)**: application of topology to sampled (digitized) sets

TDA of digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

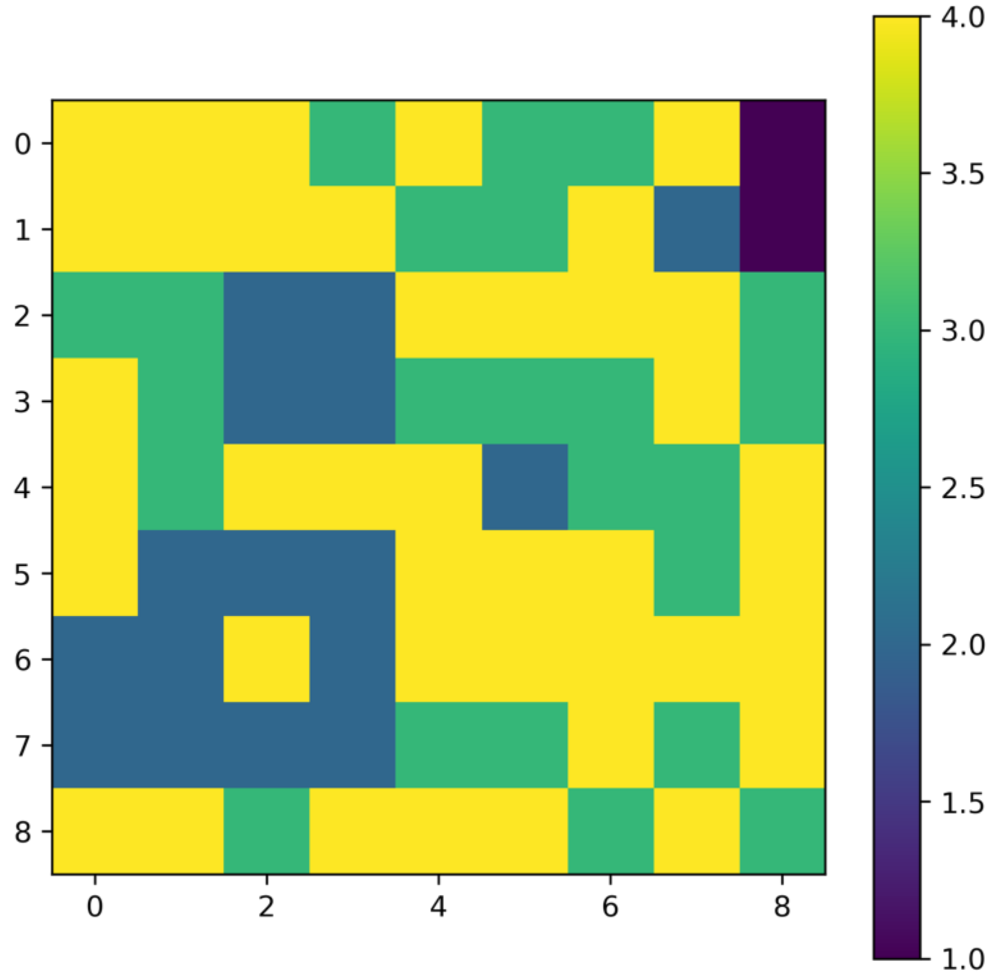


Cubical complex for threshold = 2

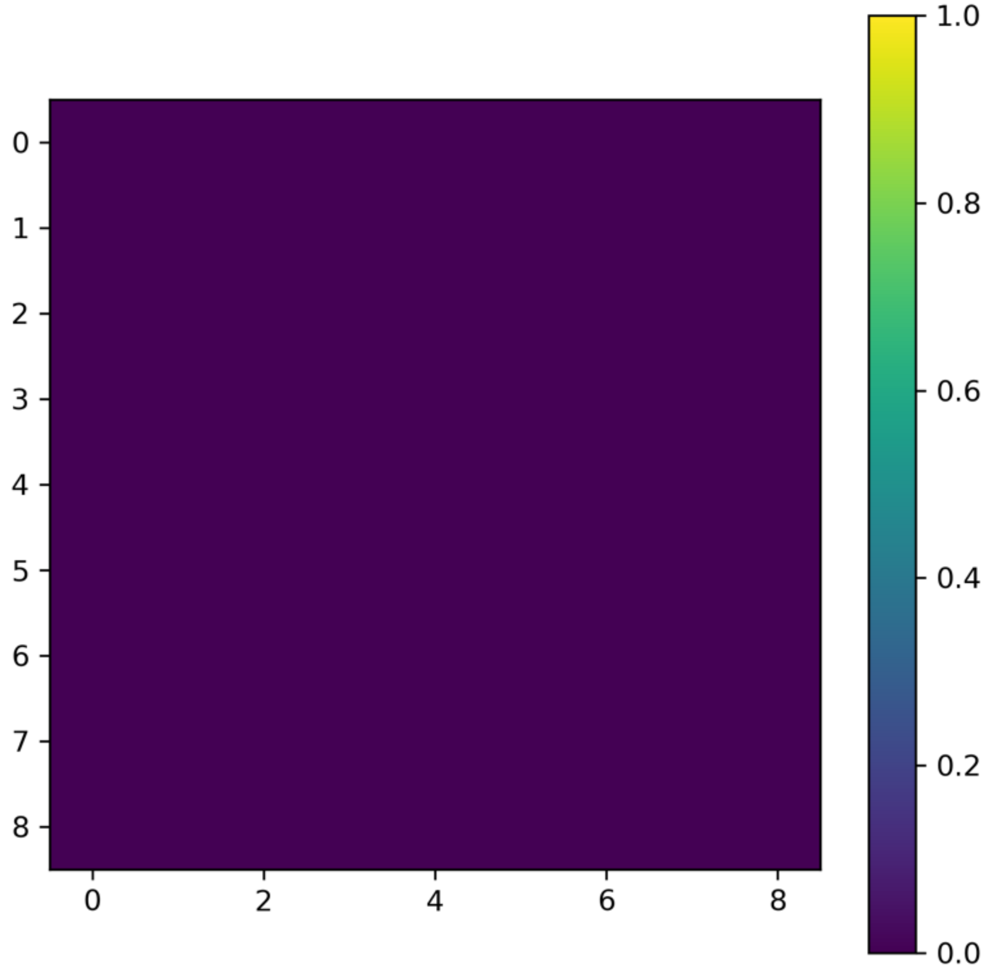
$\beta_0 = 4$, $\beta_1 = 1$

TDA for digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

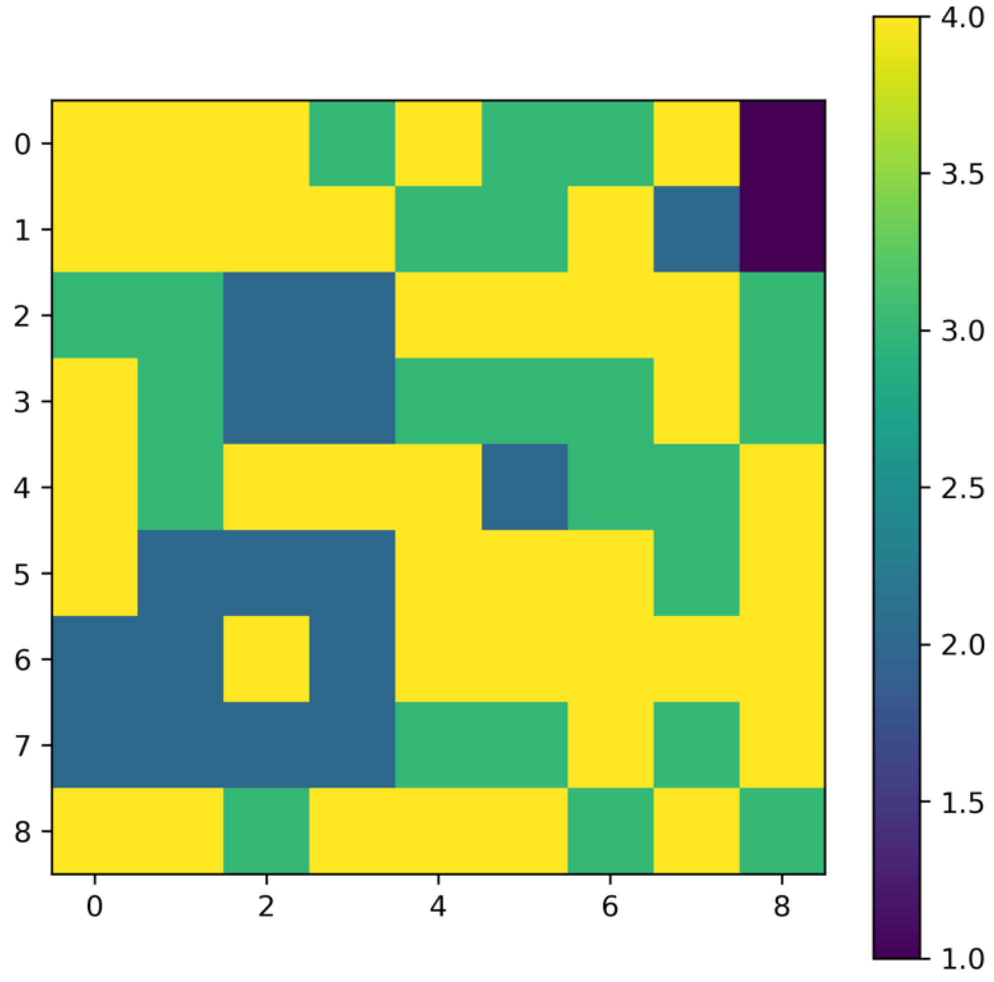


Cubical complex for threshold = 0

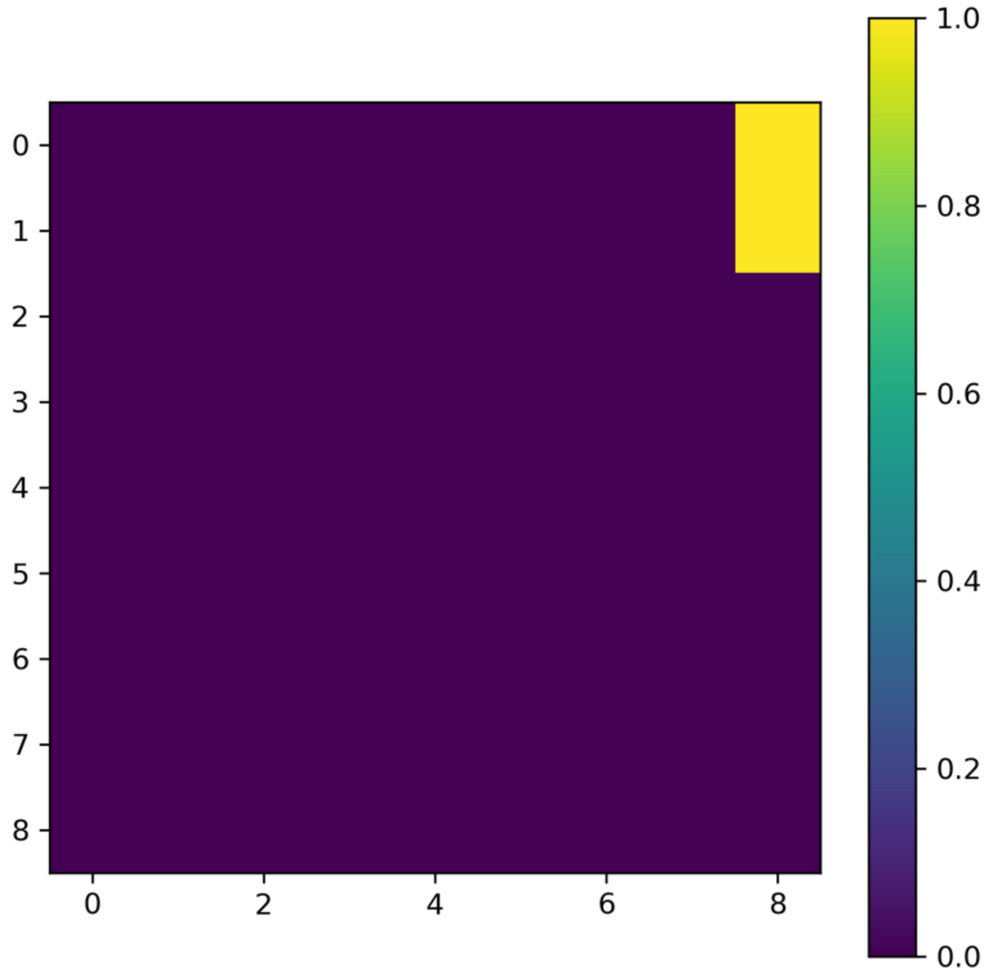
$\beta_0 = 0$, $\beta_1 = 0$

TDA for digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

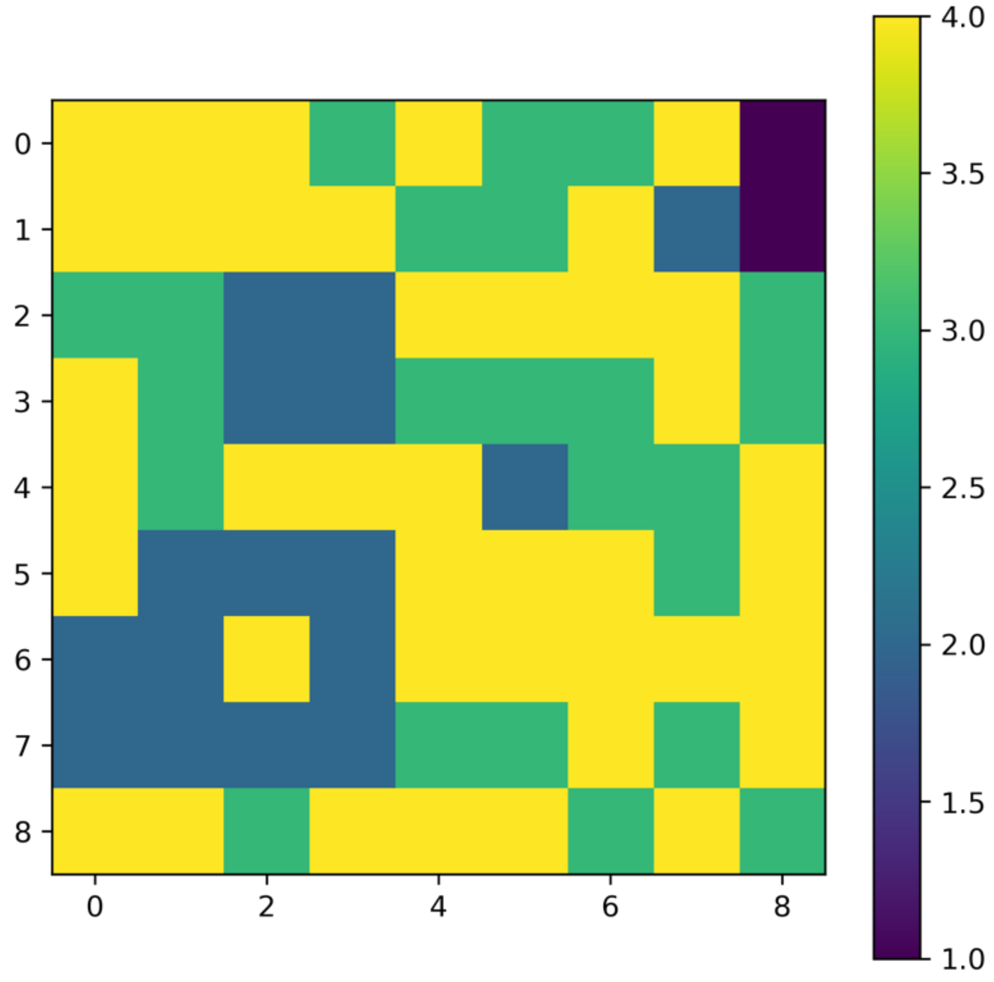


Cubical complex for threshold = 1

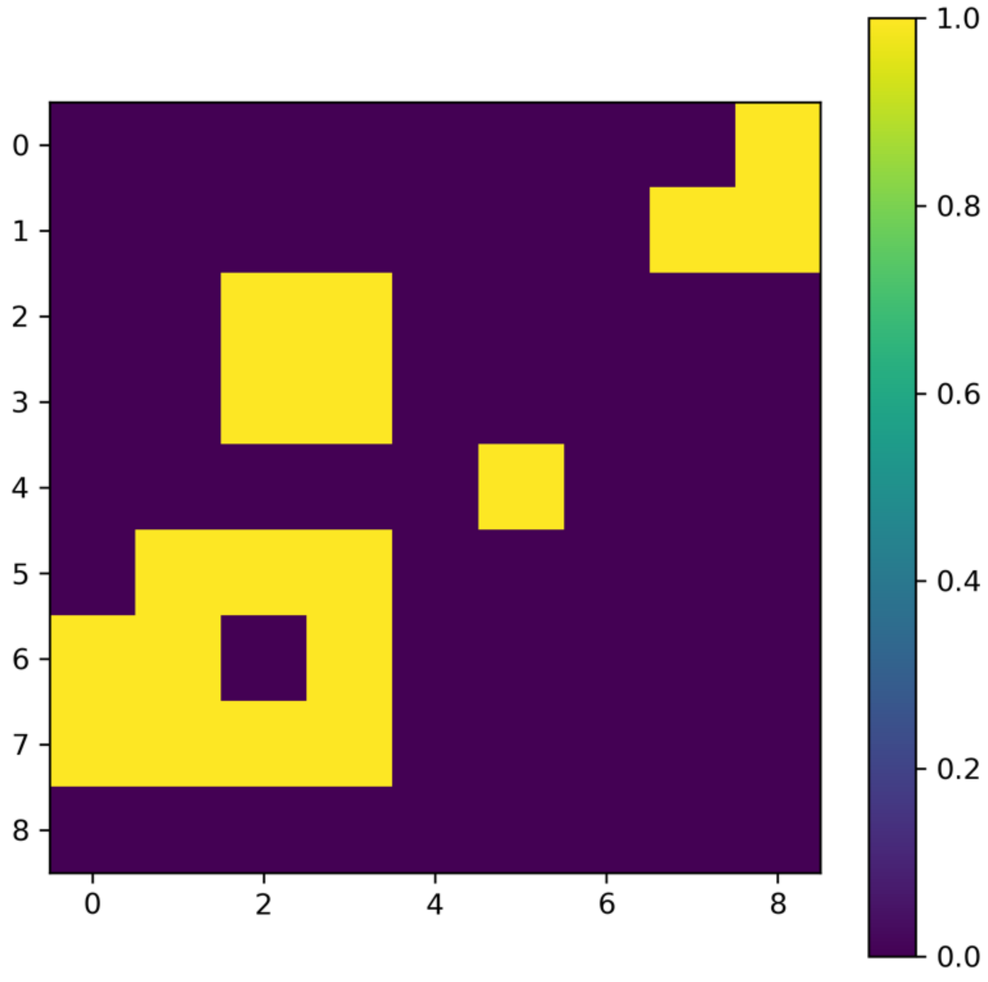
$\beta_0 = 1$, $\beta_1 = 0$

TDA for digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

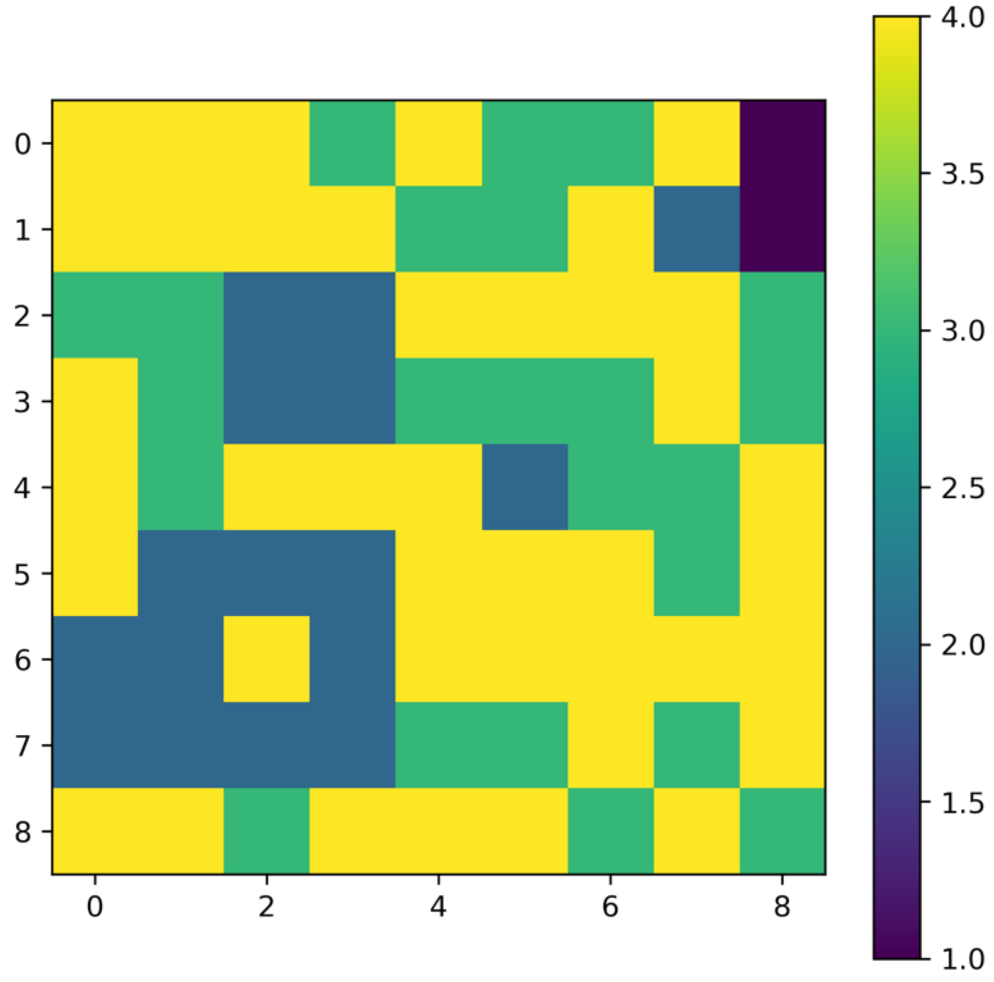


Cubical complex for threshold = 2

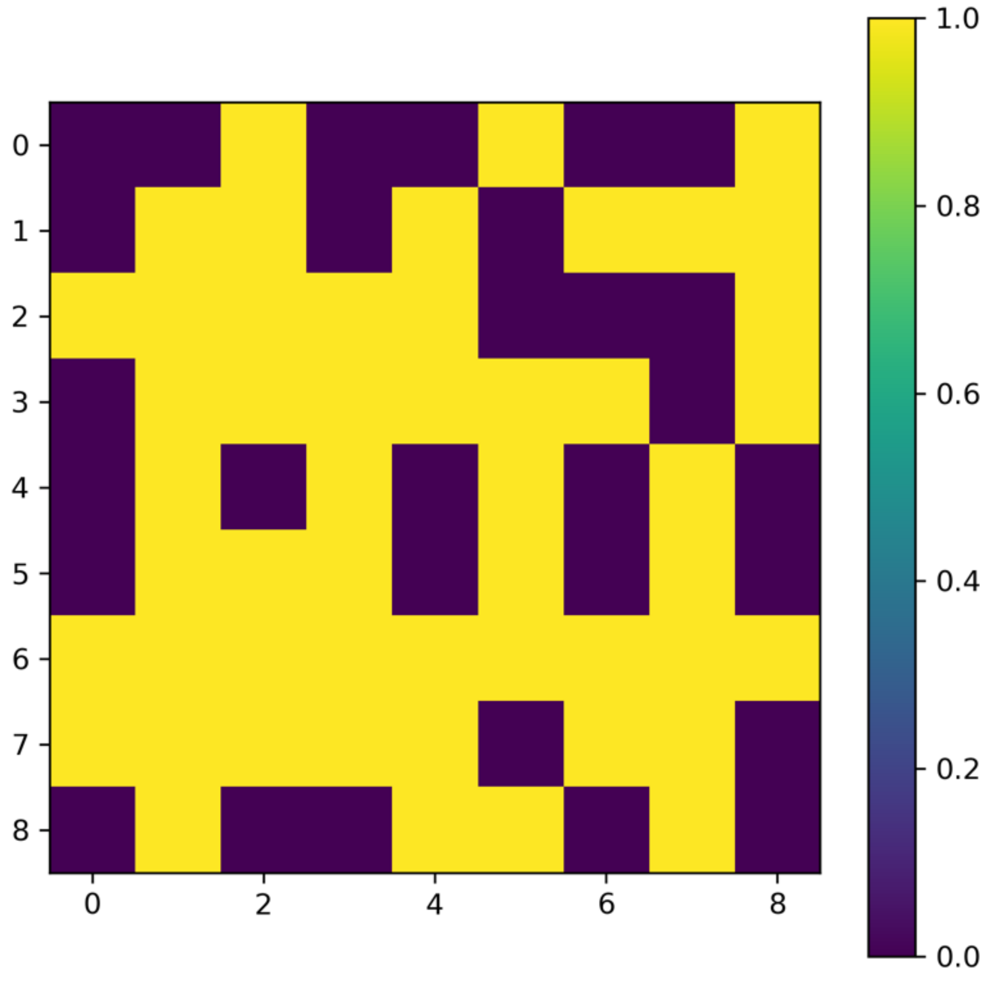
$\beta_0 = 4$, $\beta_1 = 1$

TDA for digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

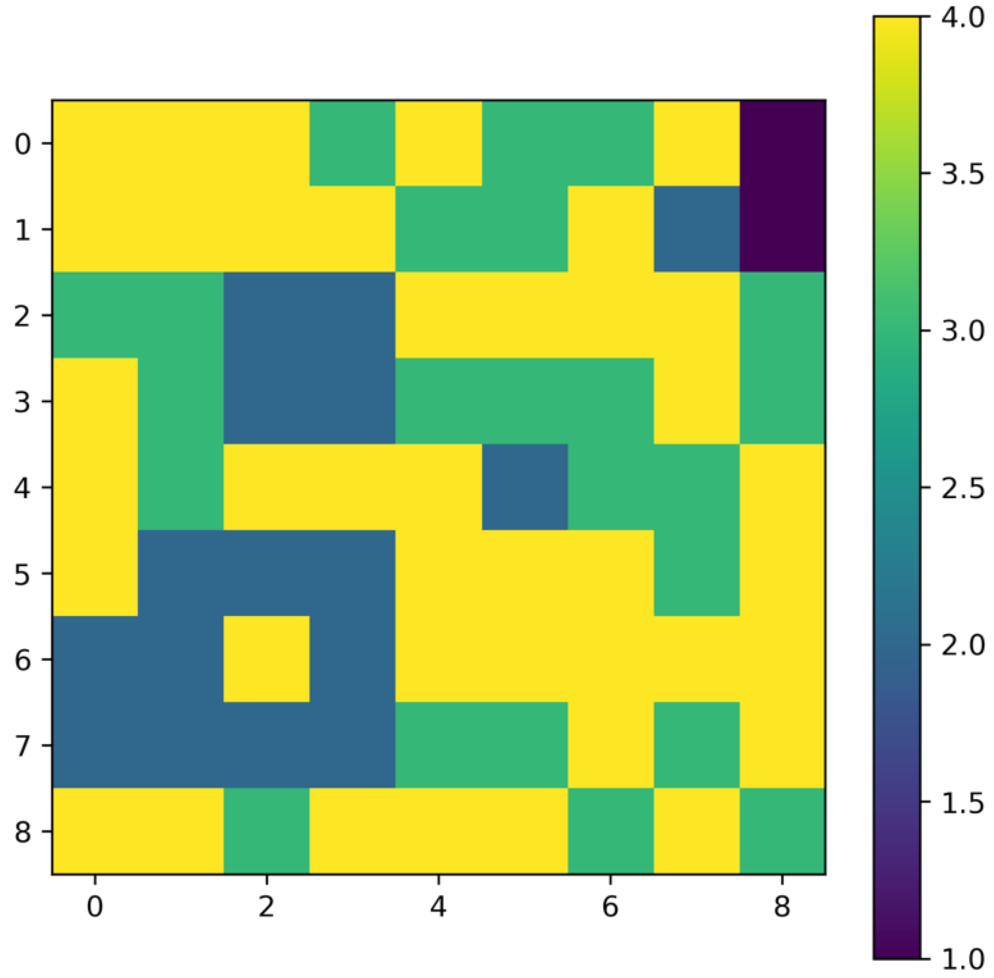


Cubical complex for threshold = 3

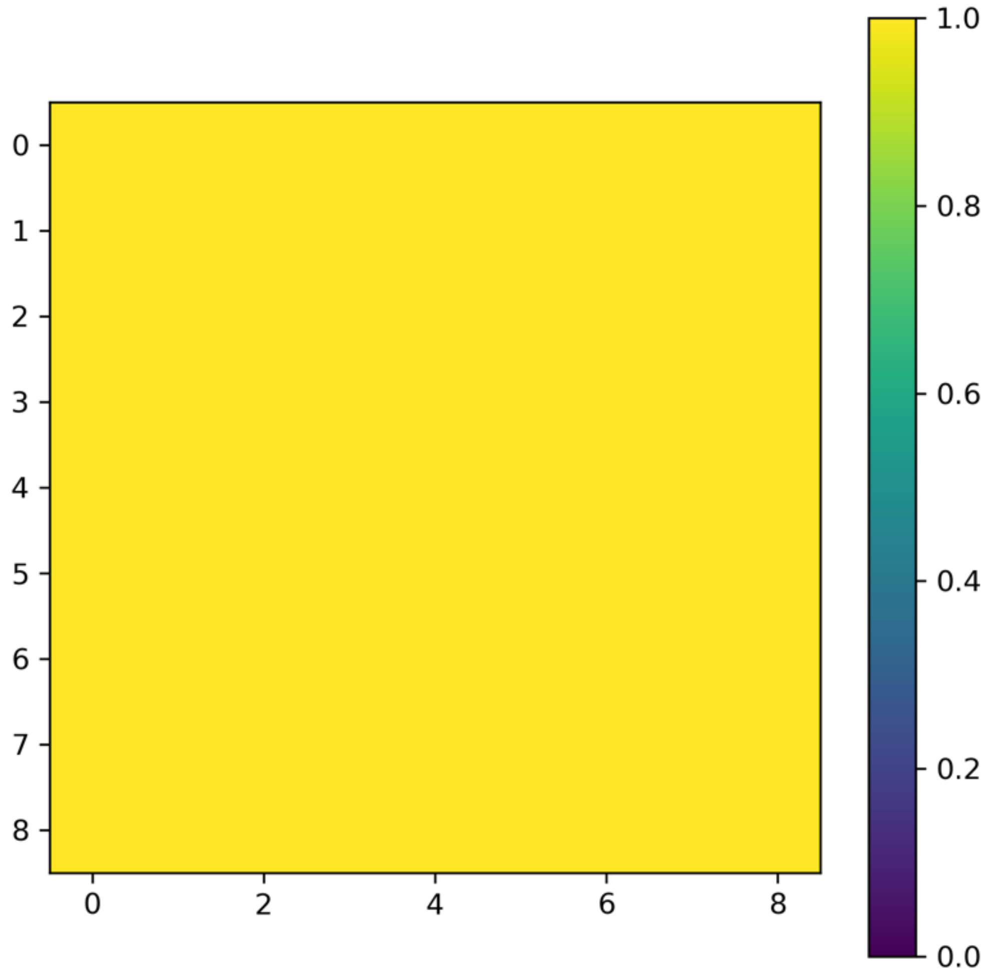
$\beta_0 = 1$, $\beta_1 = 5$

TDA for digital images

- Cubical complex: conceptually replace each pixel with a “cube” with height set by magnetic flux density value
- Grow a cubical complex by sub-level thresholding the magnetogram at higher and higher values.



“Magnetogram” with 4 flux density levels

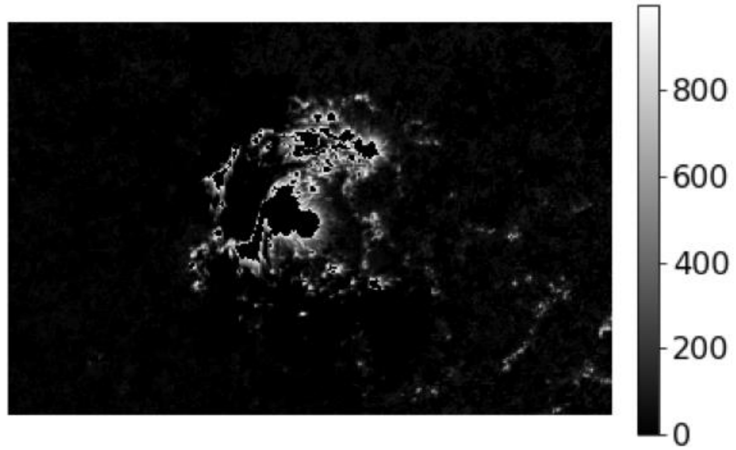


Cubical complex for threshold = 4

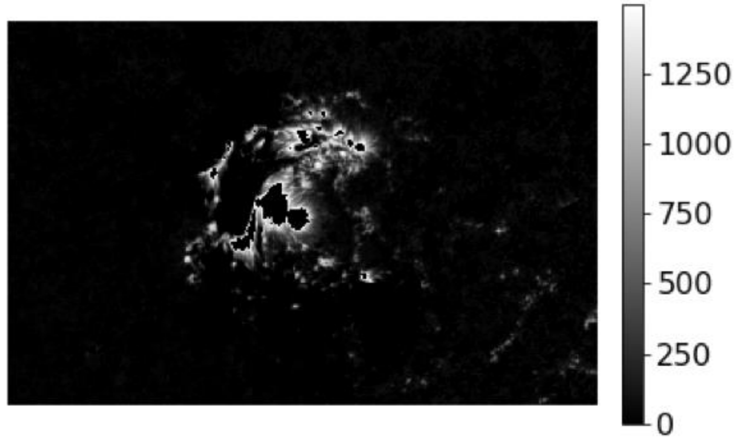
$\beta_0 = 1$, $\beta_1 = 0$

Persistence Diagrams

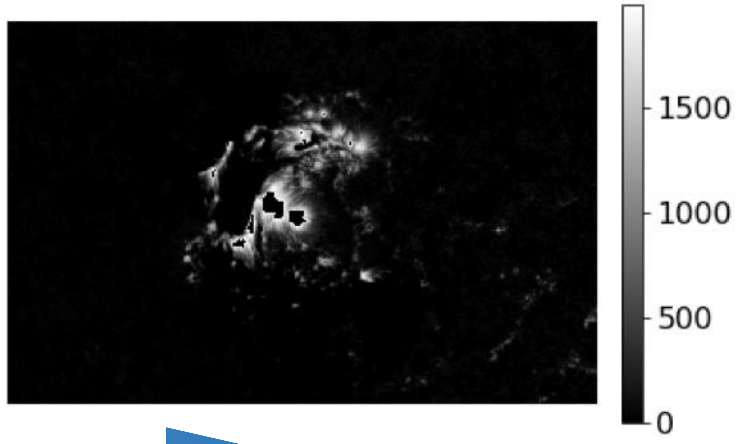
Threshold = 1000 gauss



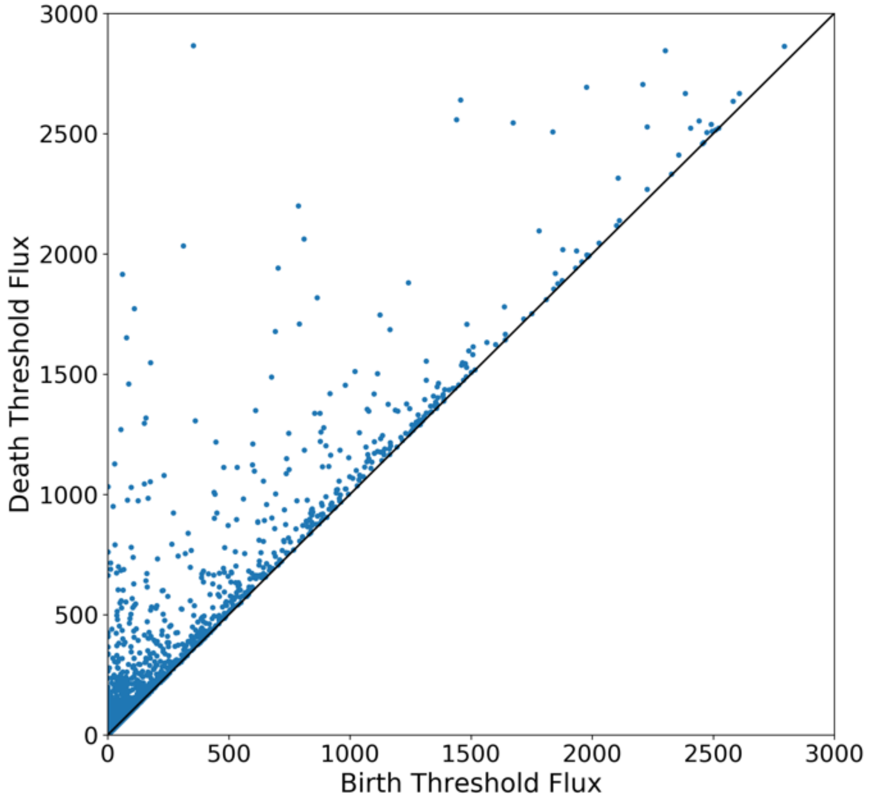
Threshold = 1500 gauss



Threshold = 2000 gauss



Adding more cubes as sub-level threshold increases

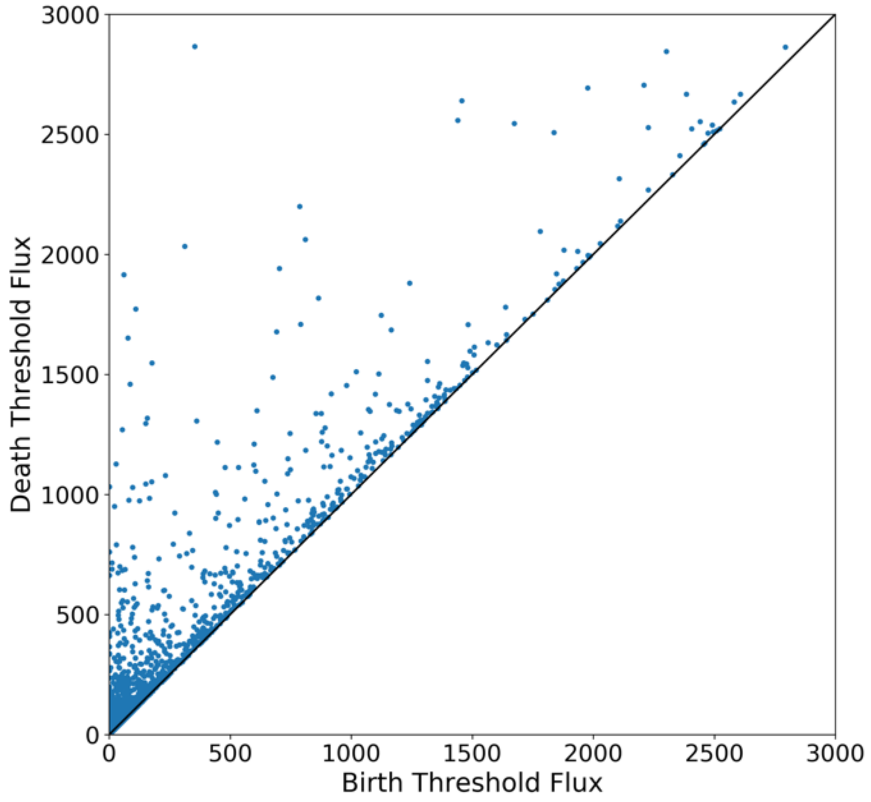


β_1 persistence diagram

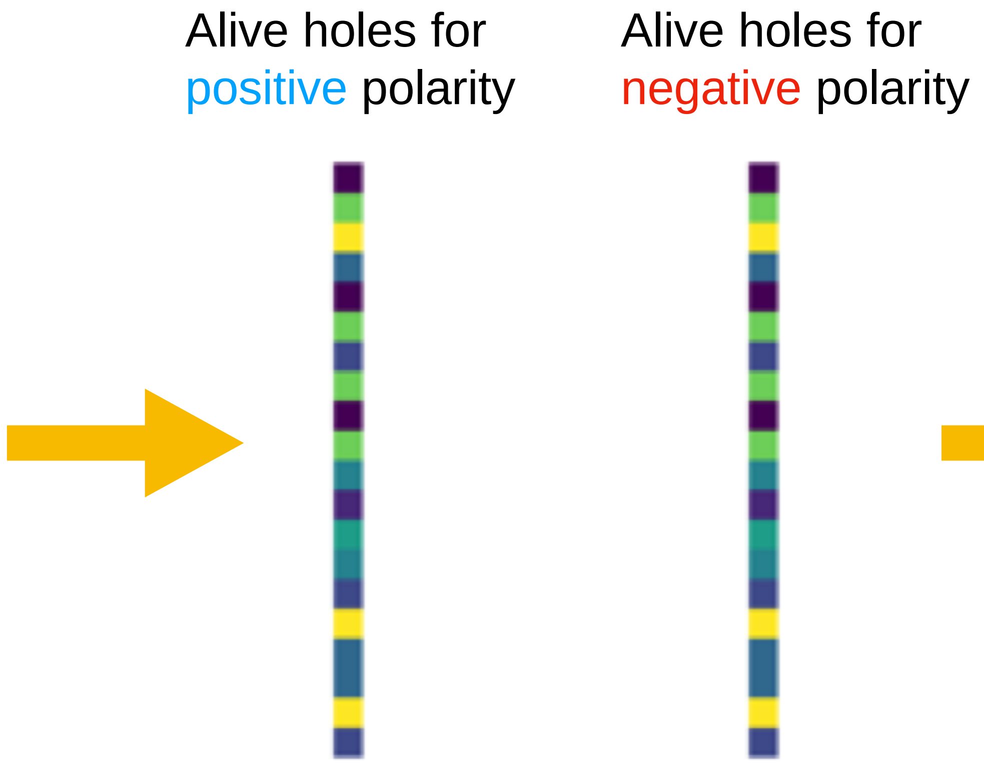
Birth and death of “holes” in the magnetogram.
Perform analysis for positive and negative polarities separately.

Persistence Diagrams to Feature Vectors

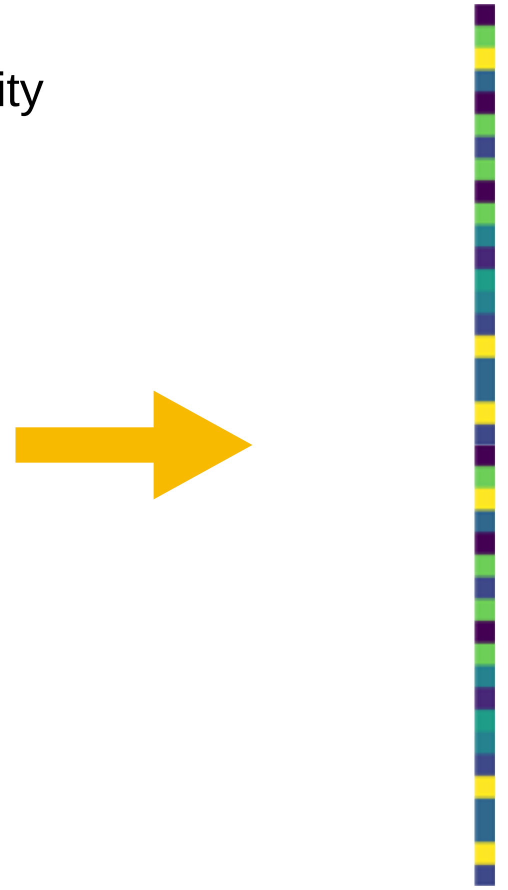
- Count the number of “alive” holes as a function of 10 threshold intervals on SDO/HMI B_r magnetograms.
- Perform for positive and negative polarities separately.
- Combine to create single feature vector.



β_1 persistence diagram



Polarity analysis



Topological Feature Vector
(20 elements)

Deep Learning Architecture

“physics-based”

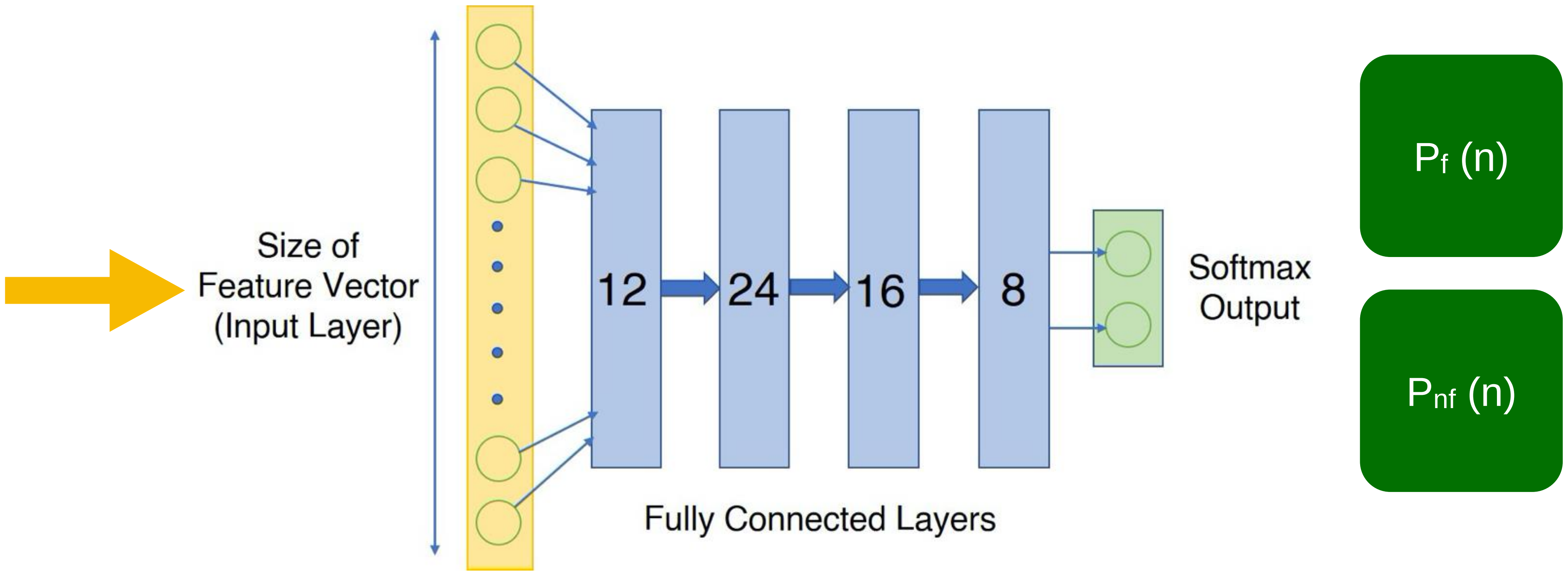
SHARP (19)

Geometry (16)

Topology (20)

Combinations (35, 36, 39)

Feature Sets



Model Parameters	Settings
Total hidden layers	4
Hidden layer activation function	Rectified Linear Unit
Output layer activation function	Softmax
Weight regularization	L2 with a penalty of 10^{-2}
Training batch size	64
Loss function	Weighted Binary Cross-Entropy (weights: $w_{no-flare} = 1$, $w_{flare} = 100$)
Optimization Method	AdaGrad (Parameters: Learning rate: 5×10^{-4})
Number of Epochs	50

Experimental Setup

Data

438,539 SDO/HMI vector magnetograms (1 hour cadence) from 3,691 SHARPs Active Regions Between 2010—2016

Binary Classification

M1.0+ flares in the next 24 hours are labeled as 1 (5,538 magnetograms*), else 0

Data Shuffling

Random 70/30 training/testing split based on active regions.

Performance Metric

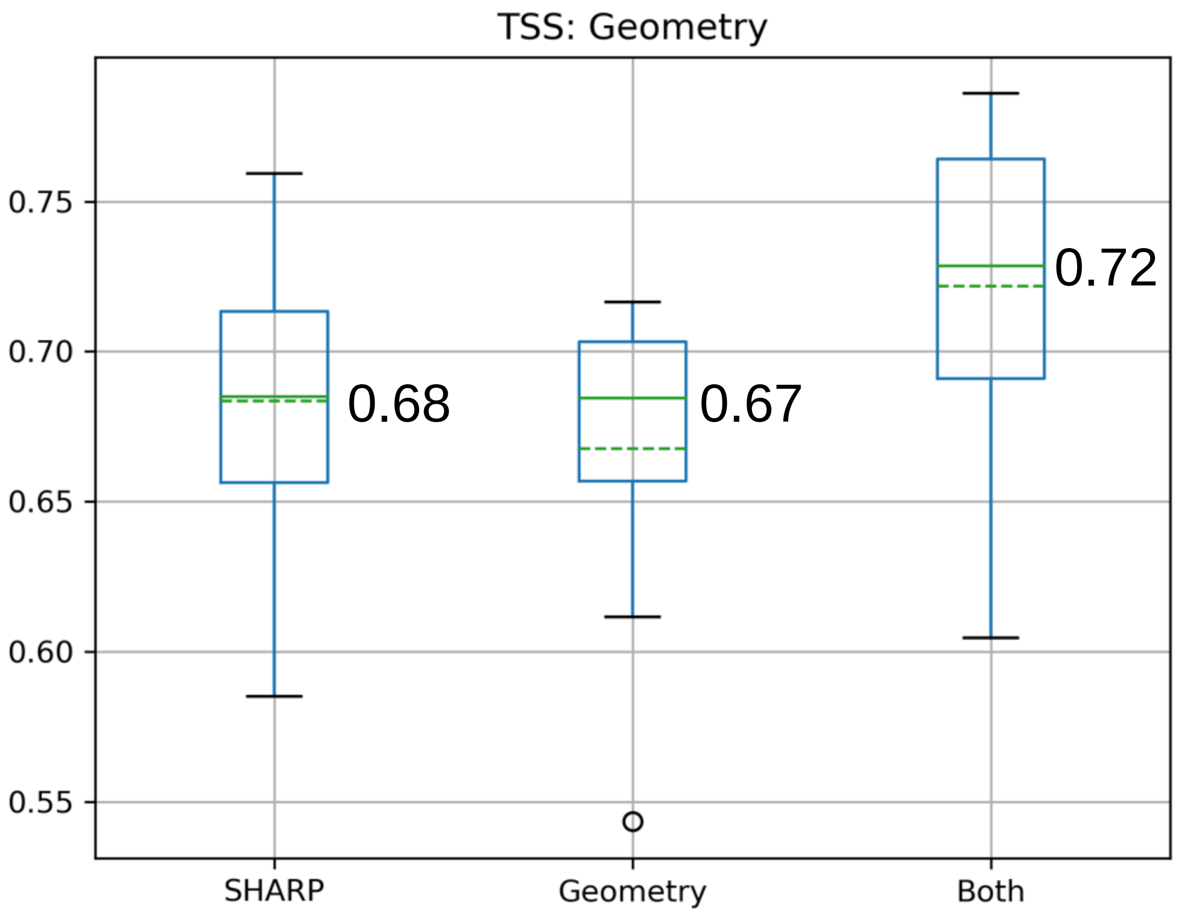
$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}$$

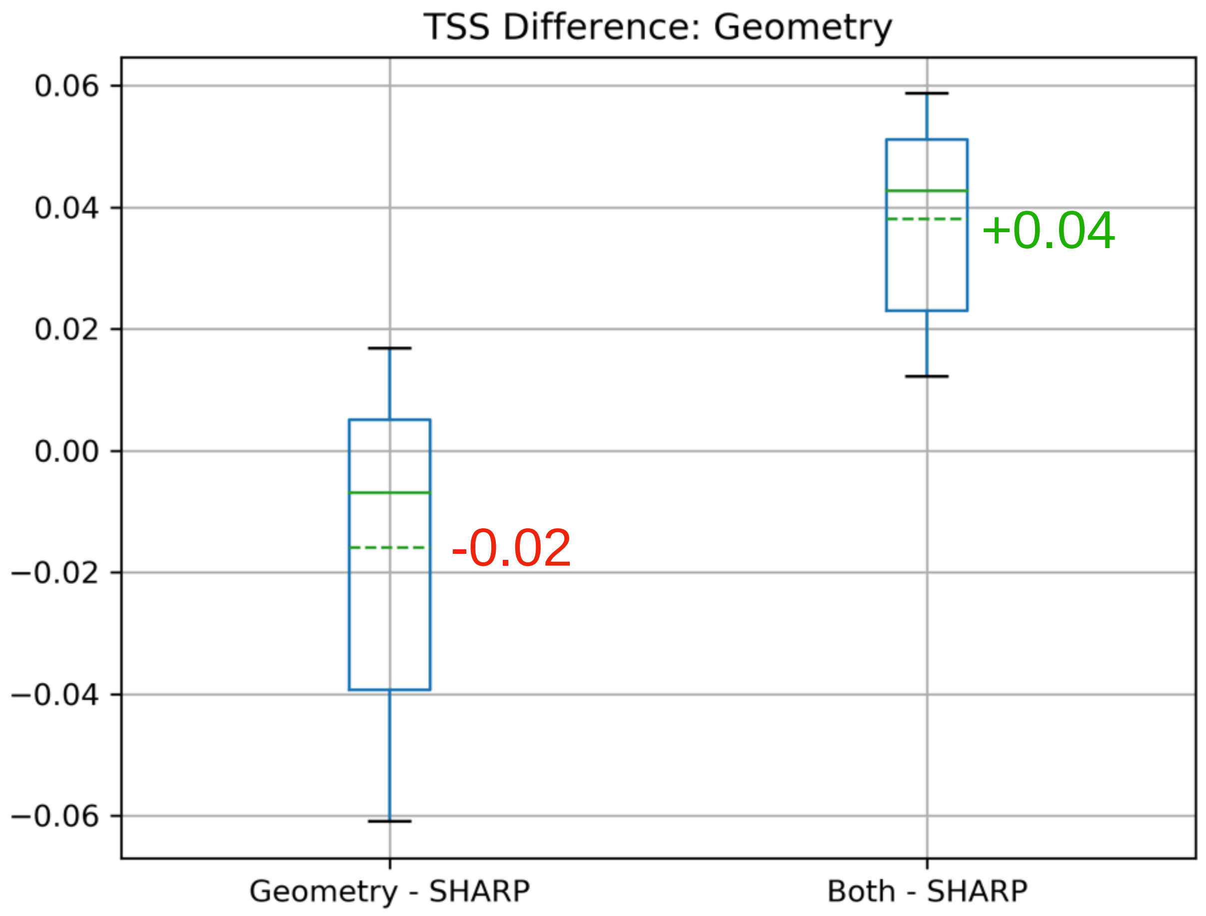
Iterations

10 splits with different random seeds used to shuffle active regions.

*At least one M1.0 or larger flare in the next 24 hours following the magnetogram time stamp. 1.2% event rate implies typical **class imbalance issue** for binary classification machine learning problem.

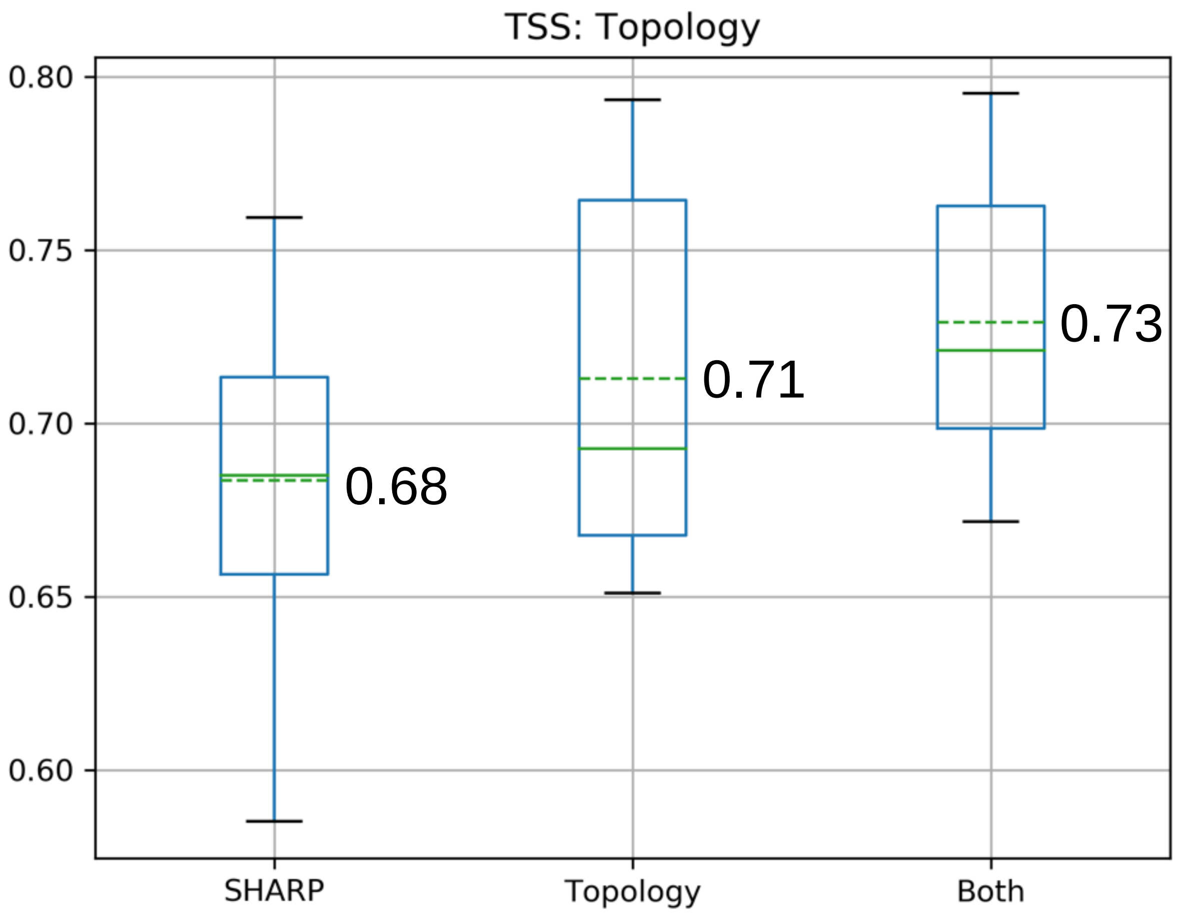
Results: Geometric Features

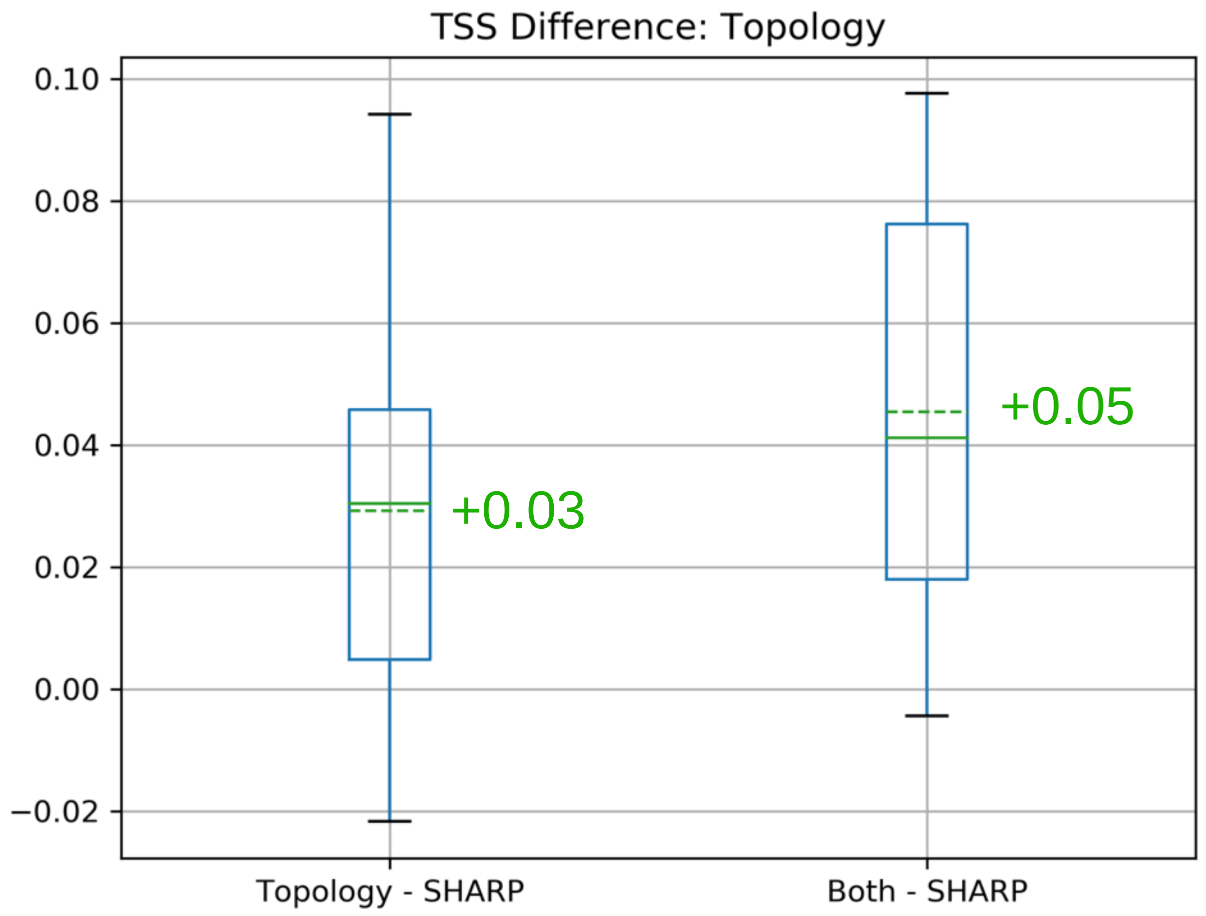




- Geometric features alone do no better than SHARP feature set.
- SHARP + Geometry feature set shows a mean TSS improvement of 0.04 over SHARP feature set alone.

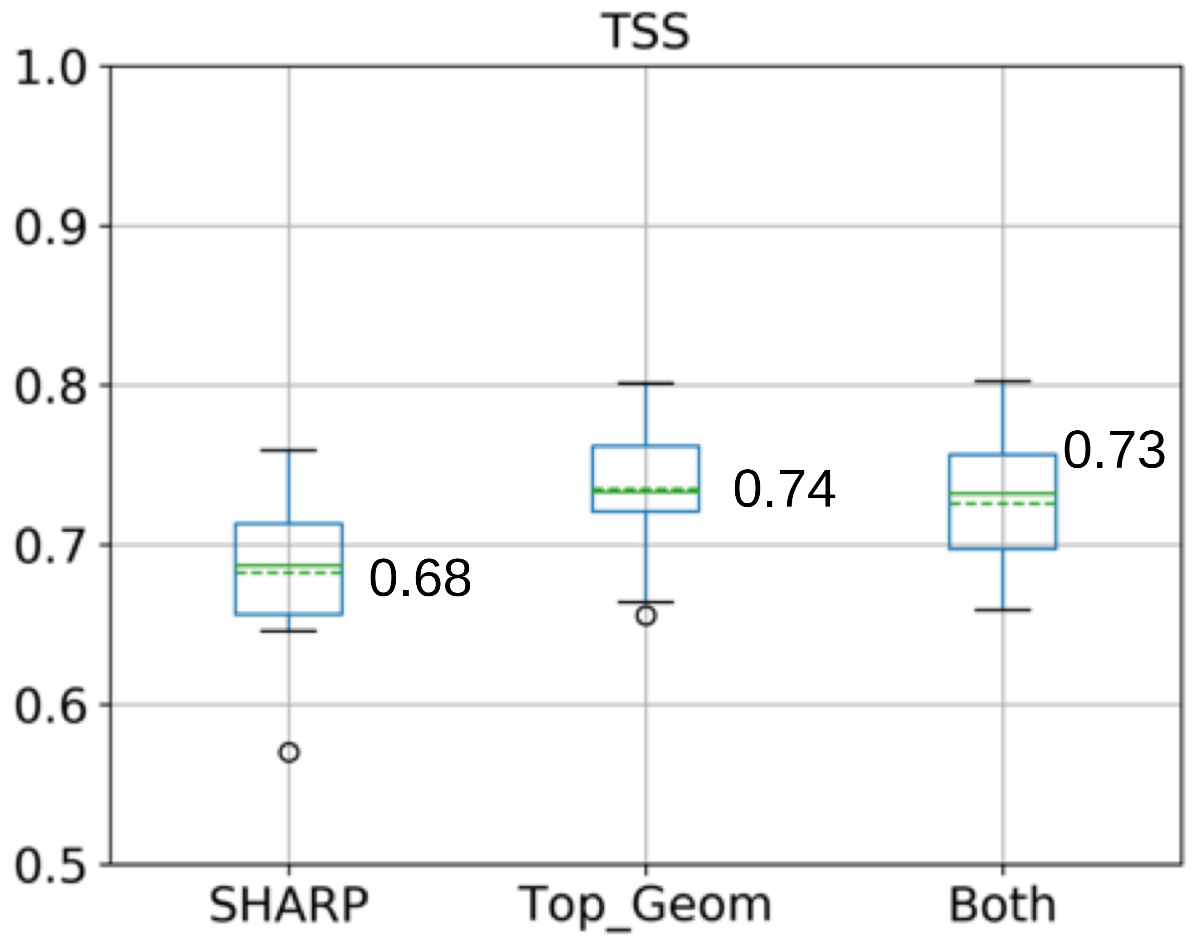
Results: Topological Features

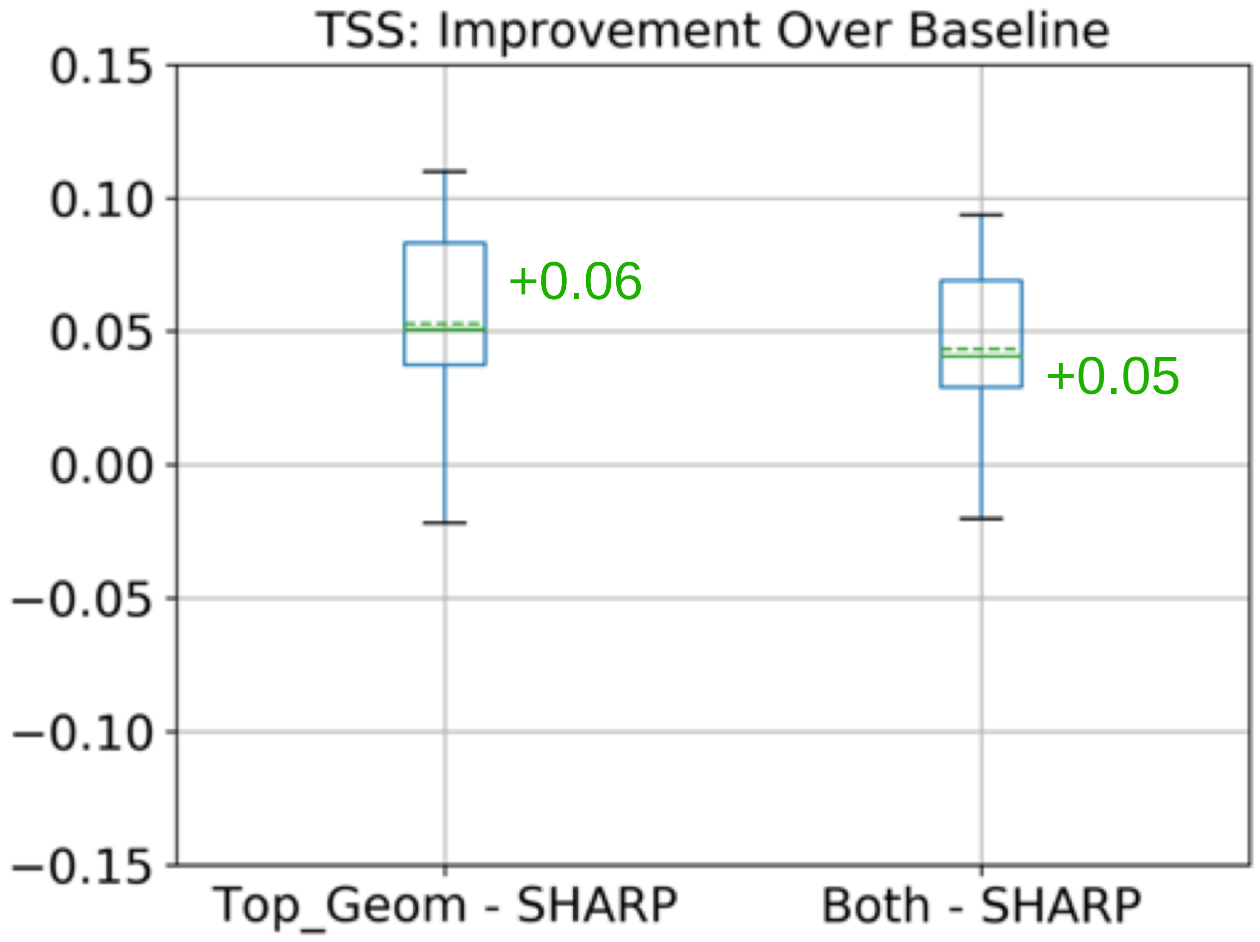




- Topological feature set shows a mean TSS improvement of 0.03 over SHARP feature set alone.
- SHARP + Topology feature set shows a mean TSS improvement of 0.05 over SHARP feature set alone.

Results: Geometry+Topology Features



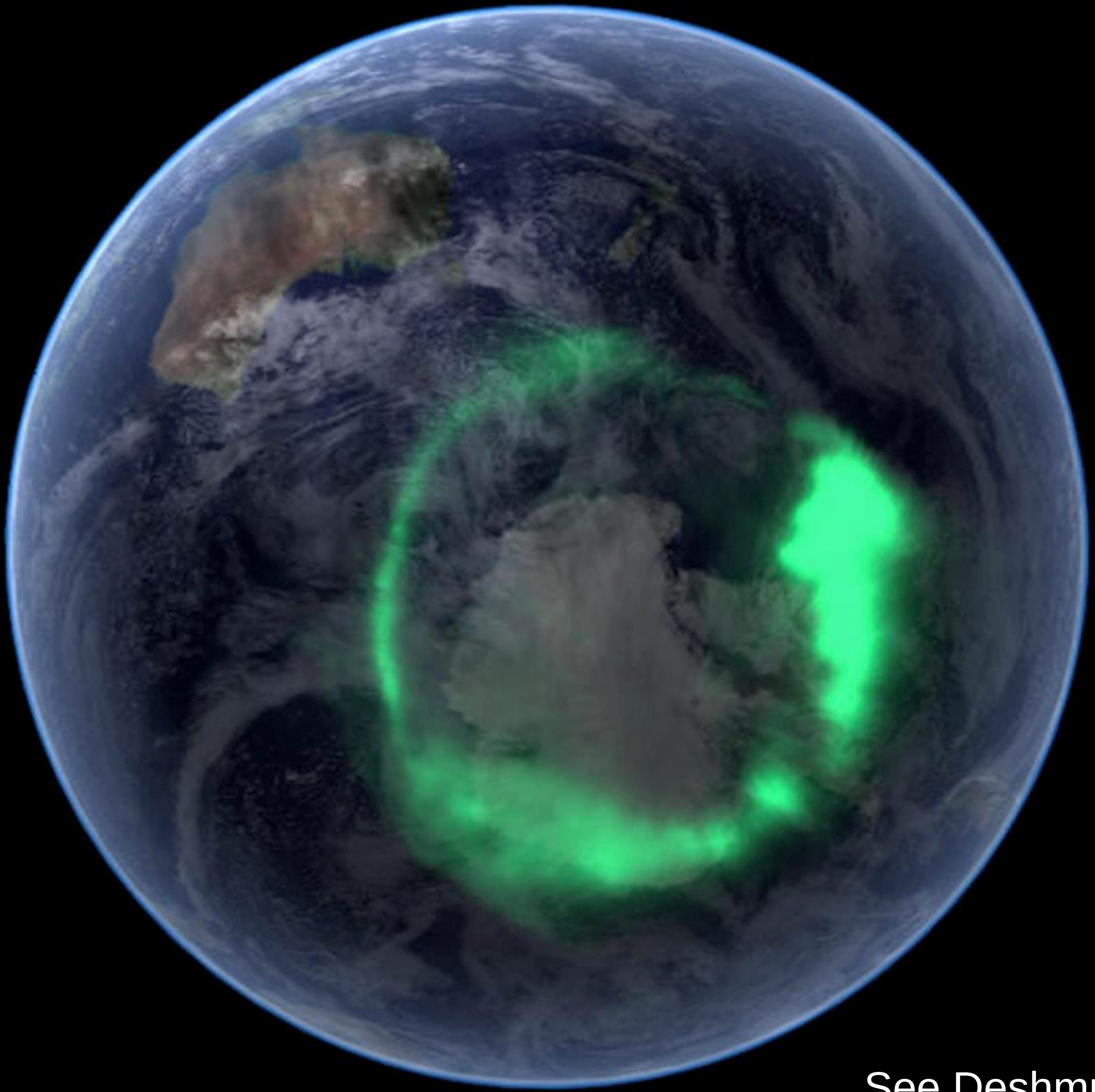


- Topological + Geometric features show a mean TSS improvement of 0.06 over SHARP feature set alone.
- SHARP + Topology + Geometry shows a mean TSS improvement of 0.05 over SHARP feature set alone.

Conclusions

- **Geometric and topological features are important for ML flare forecasts**
  - Geometrical features combined with “physics-based” features show improvement over physics-based features alone.
  - Topological features alone out-perform SHARPs physics-based features.

- **Future work:**
  - Add topological features from continuum images.
  - Experiment with geometry and topology feature inputs to other flare models (e.g. DeepFlareNet)
  - Investigate the evolution of features using recurrent neural network technology.



Thank You!

See Deshmukh et al., Journal of Space Weather and Space Climate, submitted, for more details.

Other Metrics

Geometry-based features:

Feature Set	Accuracy	Precision	Recall	F-1 Score	HSS
SHARP	0.89 ± 0.01	0.09 ± 0.01	0.79 ± 0.06	0.15 ± 0.01	0.13 ± 0.01
Geometry	0.88 ± 0.01	0.08 ± 0.01	0.79 ± 0.06	0.14 ± 0.01	0.12 ± 0.01
SHARP + Geometry	0.90 ± 0.01	0.09 ± 0.01	0.82 ± 0.06	0.17 ± 0.01	0.15 ± 0.01
Geometry Improvement	-0.01 ± 0.00	-0.01 ± 0.00	0.00 ± 0.03	-0.01 ± 0.01	-0.02 ± 0.01
SHARP + Geometry Improvement	0.01 ± 0.00	0.01 ± 0.00	0.03 ± 0.02	0.01 ± 0.00	0.01 ± 0.00

Topology-based features:

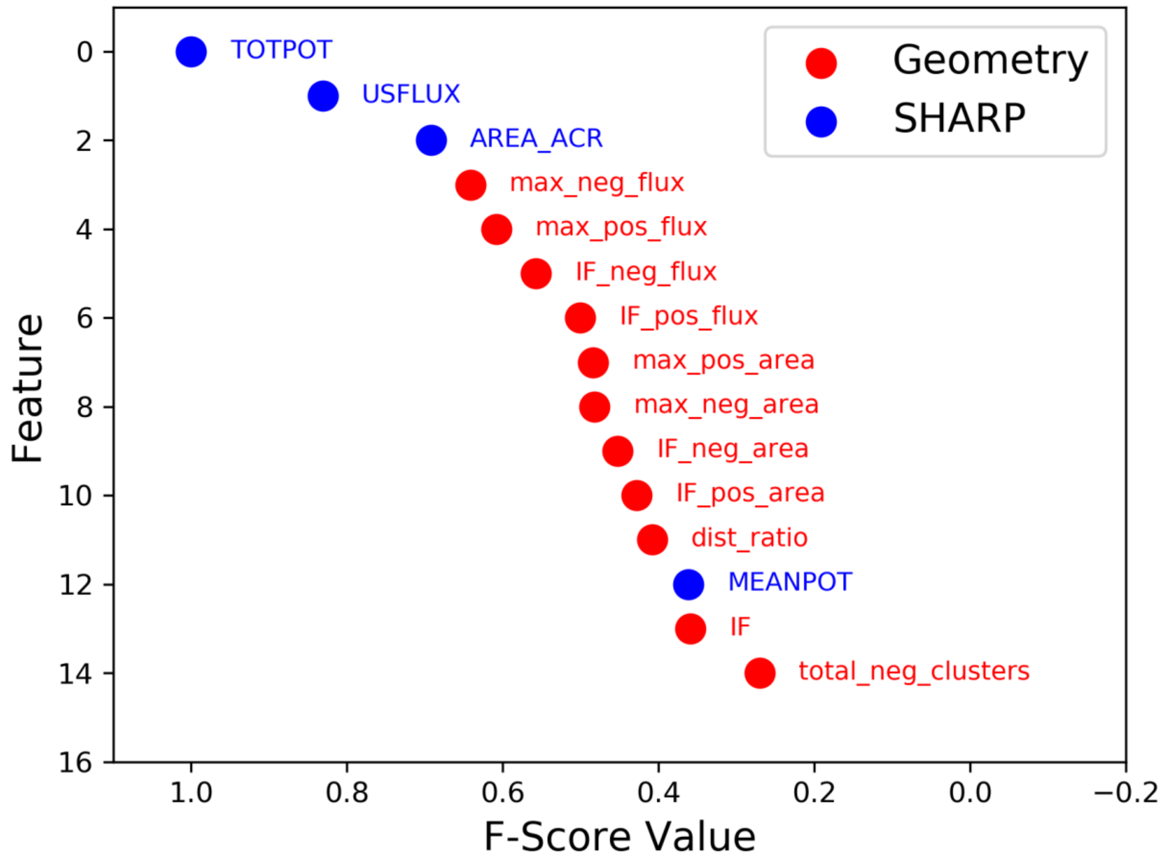
Feature Set	Accuracy	Precision	Recall	F-1 Score	HSS
SHARP	0.89 ± 0.01	0.09 ± 0.01	0.79 ± 0.06	0.15 ± 0.01	0.13 ± 0.01
Topology	0.90 ± 0.01	0.10 ± 0.01	0.81 ± 0.06	0.18 ± 0.02	0.16 ± 0.02
SHARP + Topology	0.90 ± 0.01	0.09 ± 0.01	0.83 ± 0.05	0.17 ± 0.02	0.15 ± 0.02
Topology Improvement	0.01 ± 0.01	0.01 ± 0.01	0.02 ± 0.05	0.02 ± 0.01	0.02 ± 0.01
SHARP + Topology Improvement	0.01 ± 0.01	0.01 ± 0.00	0.04 ± 0.04	0.01 ± 0.01	0.02 ± 0.01

Probability of Detection

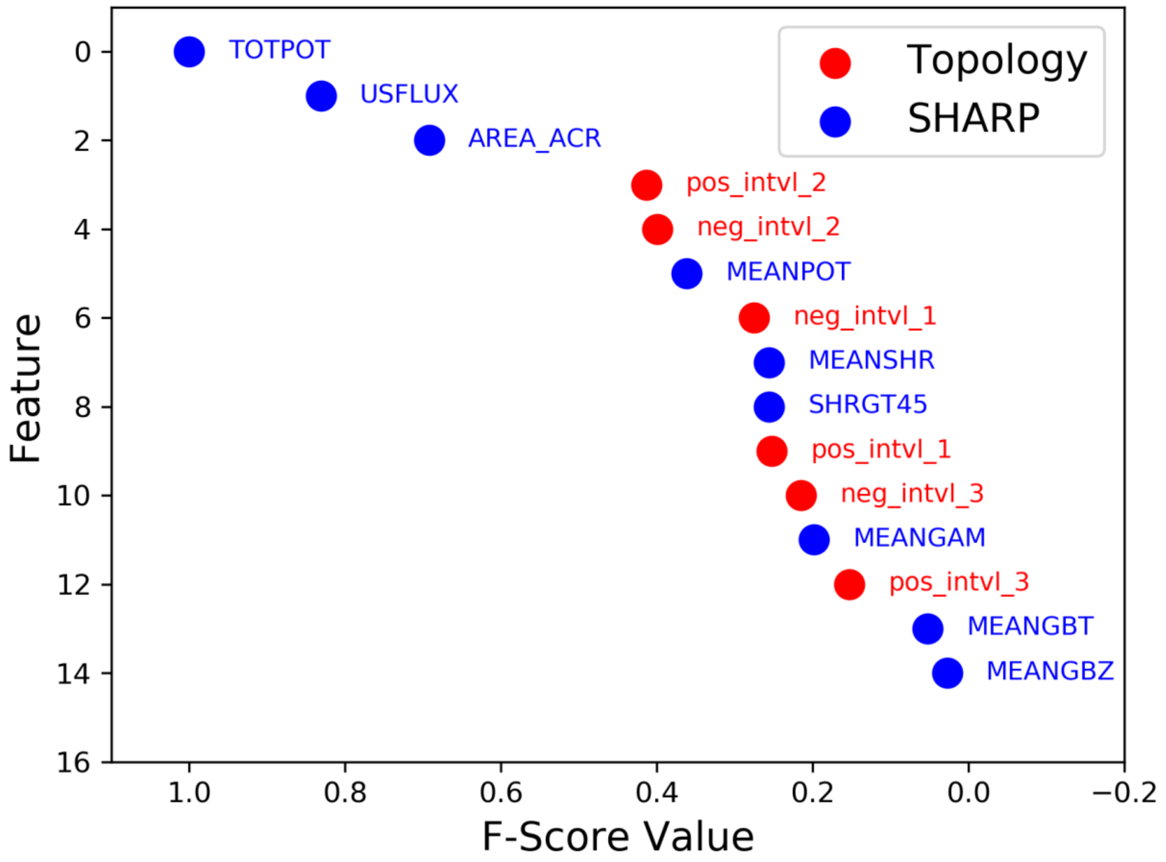
HSS_2 (Bobra 2015)

Feature Ranking: Fisher score analysis

Top 15 most “effective” features



SHARP + geometry features



SHARP + topology features