Michine-learning in Heliophysics @ Amsterdam, Netherland, 17 Sep 2019, oral, 11:10-11:50 am (30+10 min)

- Nishizuka et al. 2018 ApJ 858, 113
- Nishizuka et al. 2017 ApJ 835, 156

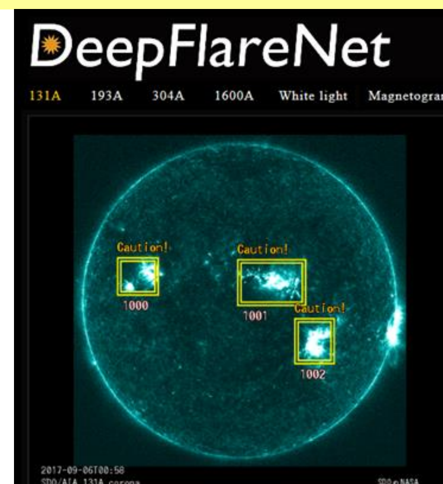
Solar Flares and Eruptions Predicted by Deep Neural Networks: *Deep Flare Net (DeFN)*

○ N. Nishizuka[1],

K. Sugiura[2], Y. Kubo[1], M. Den[1], M. Ishii[1]

[1]Space Environment Laboratory, AER, NICT

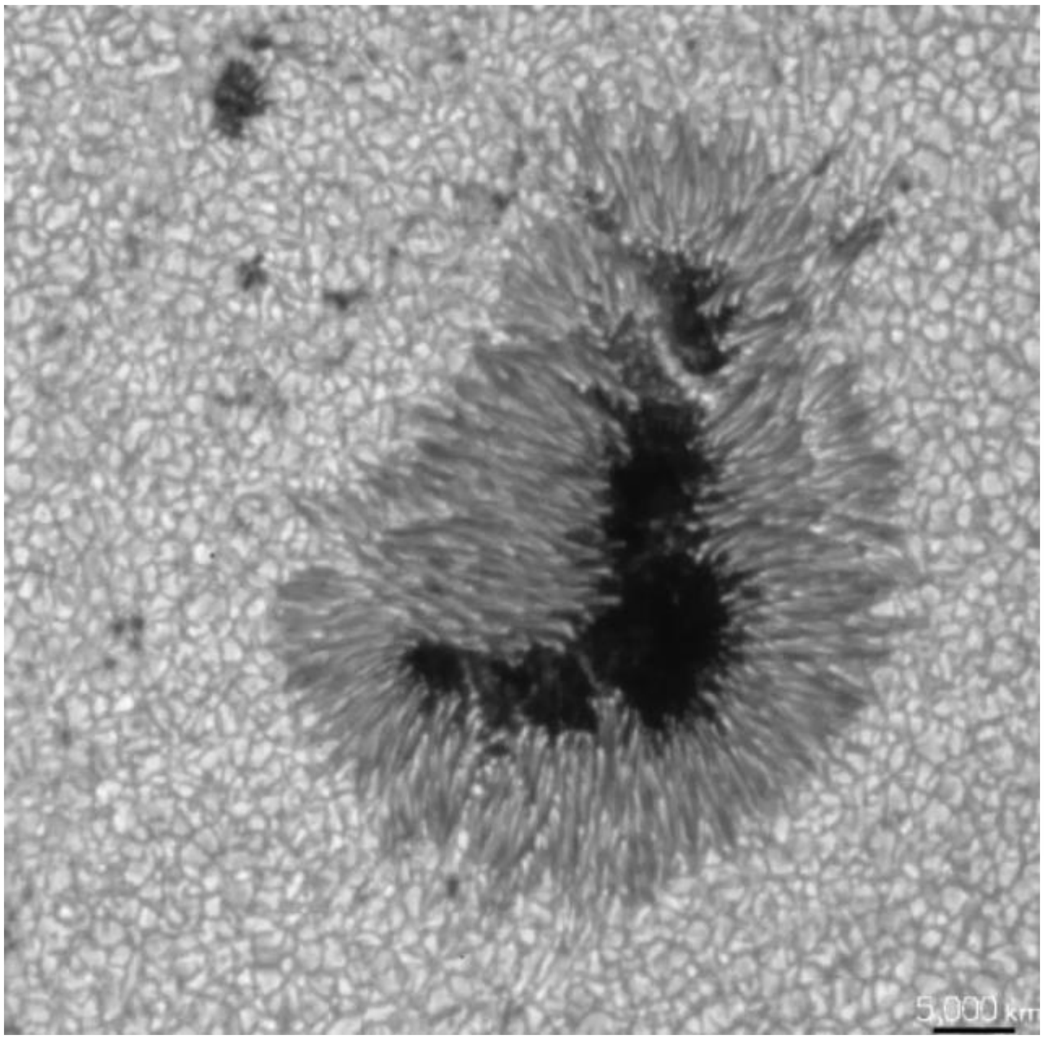
[2]Advanced Translation Technology Lab, ASTREC, NICT



https://defn.nict.go.jp

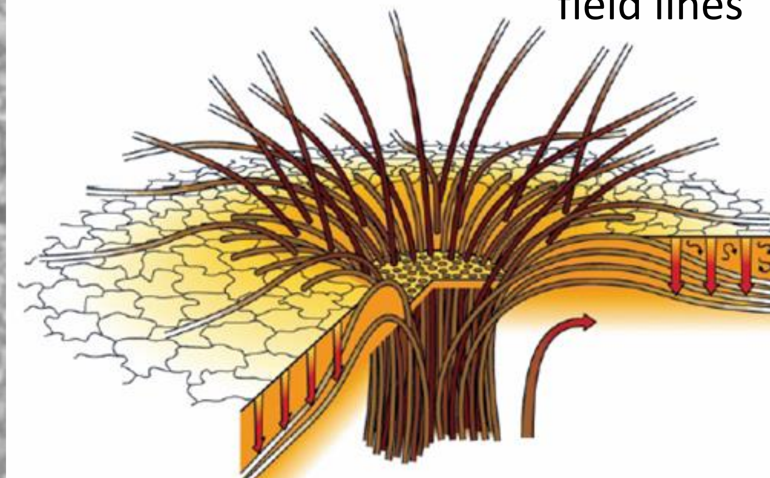
Space observation of sunspots

G-band filter, Hinode/SOT, made by Okamoto



Convection
(granule)

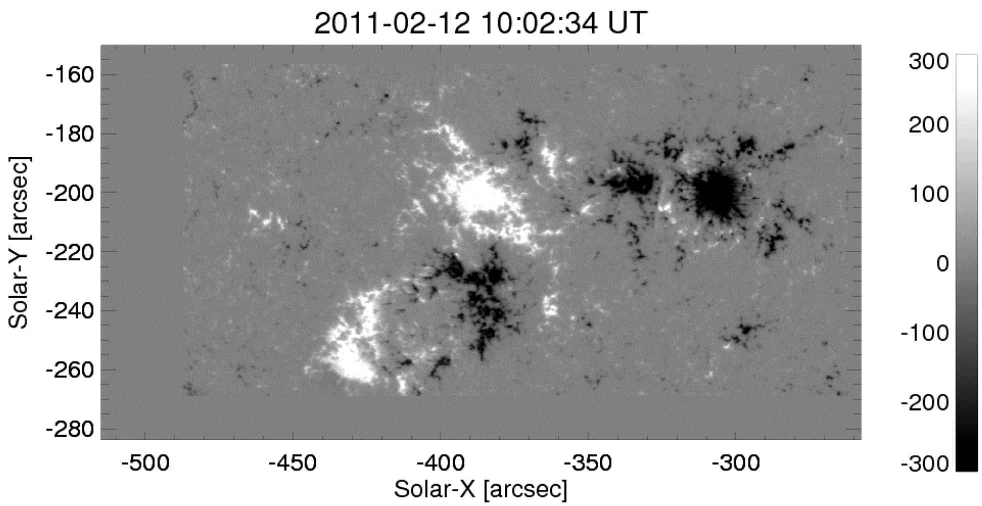
Stable quality, more precise
& higher spatial resolution



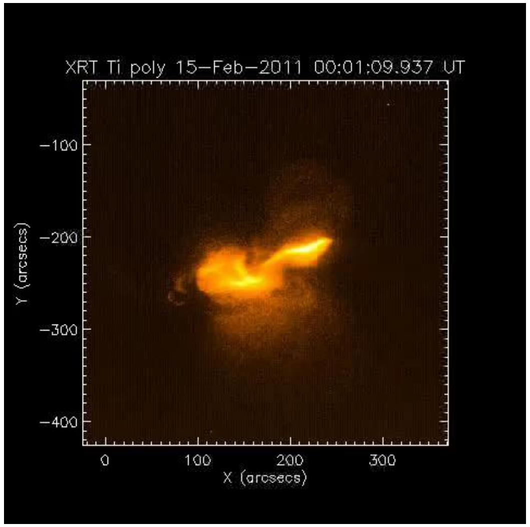
[Thomas & Weiss 2002 ARAA]

Earth size

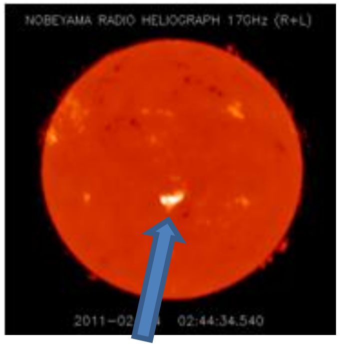
Evolution of Sunspot & Prediction



Evolution of sunspots (Hinode/SOT), by Y. Iida

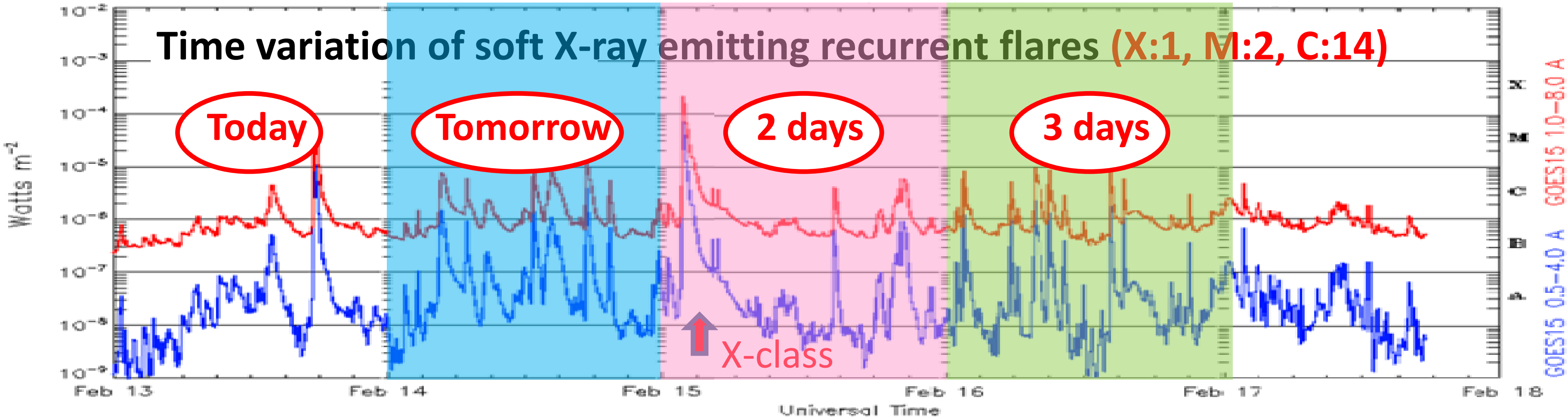


Intermittent Flares (Hinode/XRT)



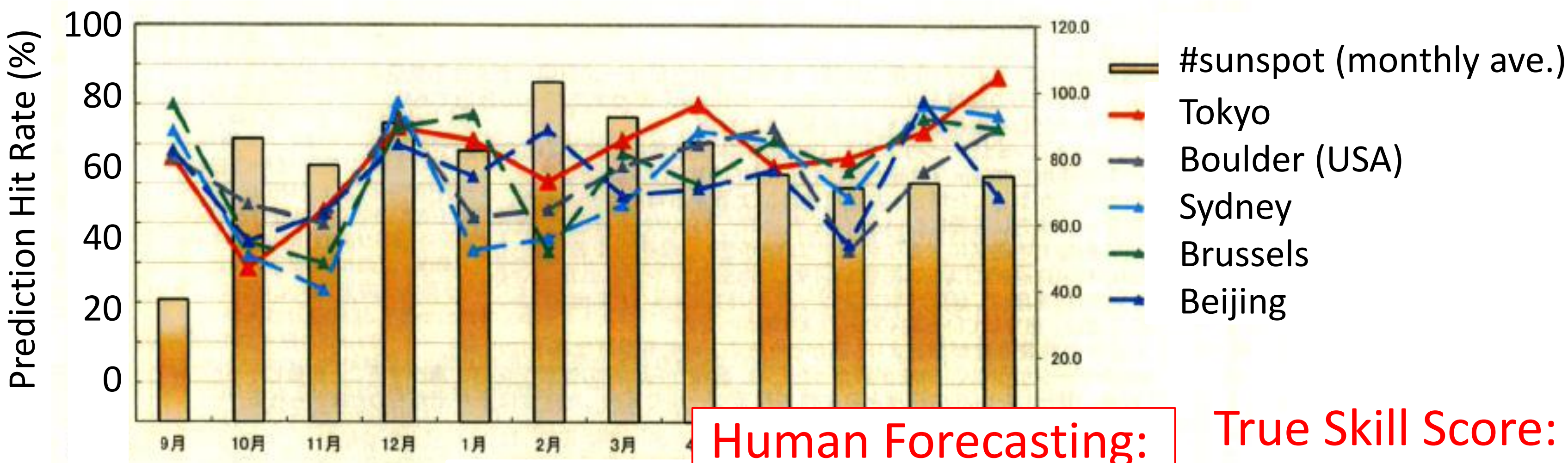
Solar flare
(NoRH data)
radio

What is the maximum class of flares within the following 24 hours ?



Daily Space Weather Forecasting

Sunspot Num. & Hit rate of flare predictions in 2013~2014



Month (2013 Sep – 2014 Aug)

Human Forecasting:
Hit rate~30-80%

True Skill Score:
TSS~0.5
(-1.0 < TSS < 1.0)

① Empirical forecast
② Statistical method
③ Machine-learning
④ Numerical simulation

Check points by Human Forecasting

①**White light**: Sunspot Area, Shape (αβγδ)

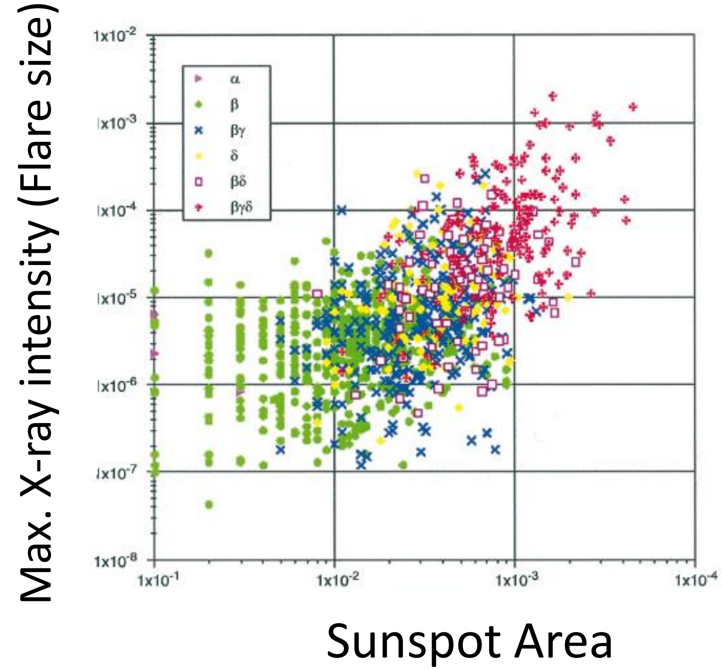
②**Soft X-ray**: Flare history in a sunspot

③**Magnetic field in the photosphere**:
Magnetic neutral line
Complexity of the structure
Shear angle, Flux Emergence

④**Bottom Chromosphere**: UV Brightening

⑤Limb obs.: ARs hidden by the east limb.

[Sammis et al. 2000]



- Observation data is too huge to deal with by human forecasting.
- Real-time operation (<24 hrs) in an automated method.
- Better feedback of daily operation results to the next.
- New approach to reveal the mechanism of solar flares.

Application of ML to Solar Flare Prediction

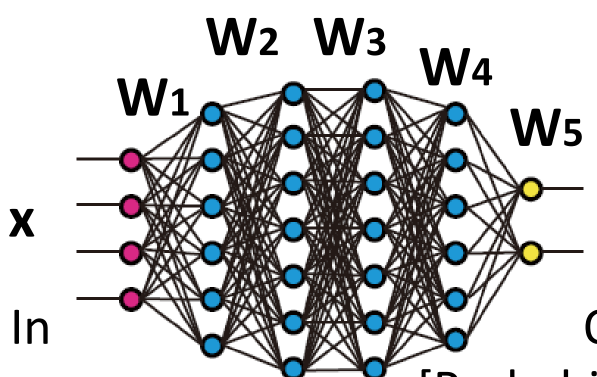
Machine-Learning

(1) To construct algorithms that can learn from and automatically make classification or prediction on known/unknown data.

(2) To classify and predict the complex data, beyond the human processing capability.

[Nishizuka+2018 ApJ]

★ Neural Networks (NN)



- Repeating linear & non-linear conversions of the input data at each layer.

$$y = f(Wx + b) = a_0x^n + a_1x^{n-1} + \dots$$

Linear (matrix)

non-linear: to separate the dataset by curves. to distort the space for an easy split.

- Parameters W_i are optimized to minimize the cost function ($\approx \Sigma(y - y_{real})^2$, cross entropy).

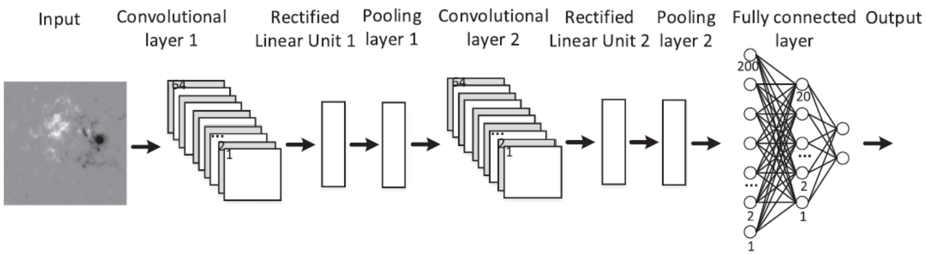
⇒ Similar to a **Polynomial fitting**.
If dim(x) is large, **over-fitting** is a problem.

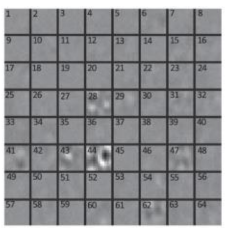
Deep Neural Network (DNN)

- RNN (LSTM)
- CNN, GoogleNet, Residual Net
- GAN, SimGAN

Solar Flare Prediction using DNN

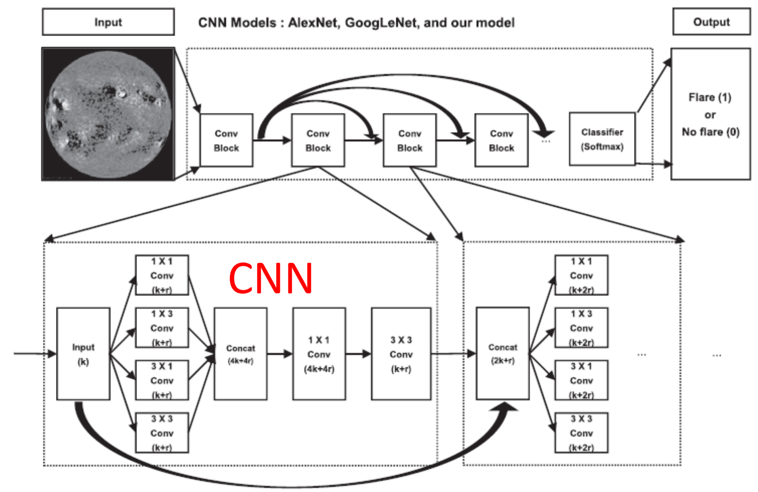
ML became popular after Bobra & Couvidat 2015 (using SDO data: 2010-)





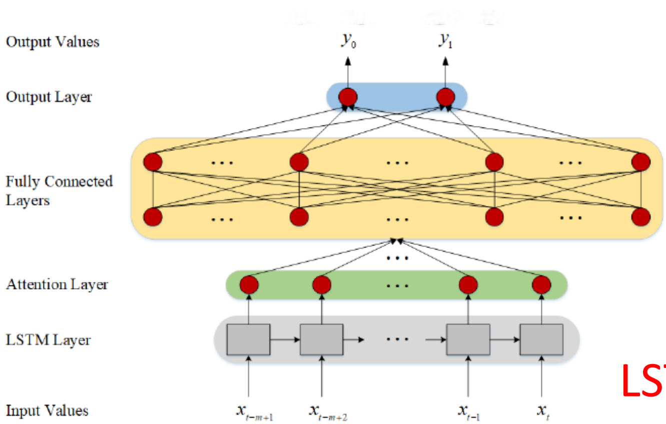
Convolutional Neural Networks (CNN)

[X. Huang et al. 2018 ApJ]



CNN

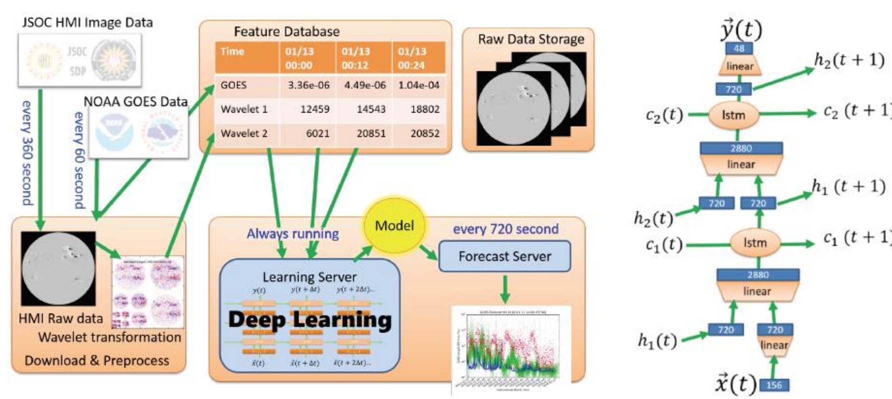
[E. Park et al. 2018 ApJ]



LSTM

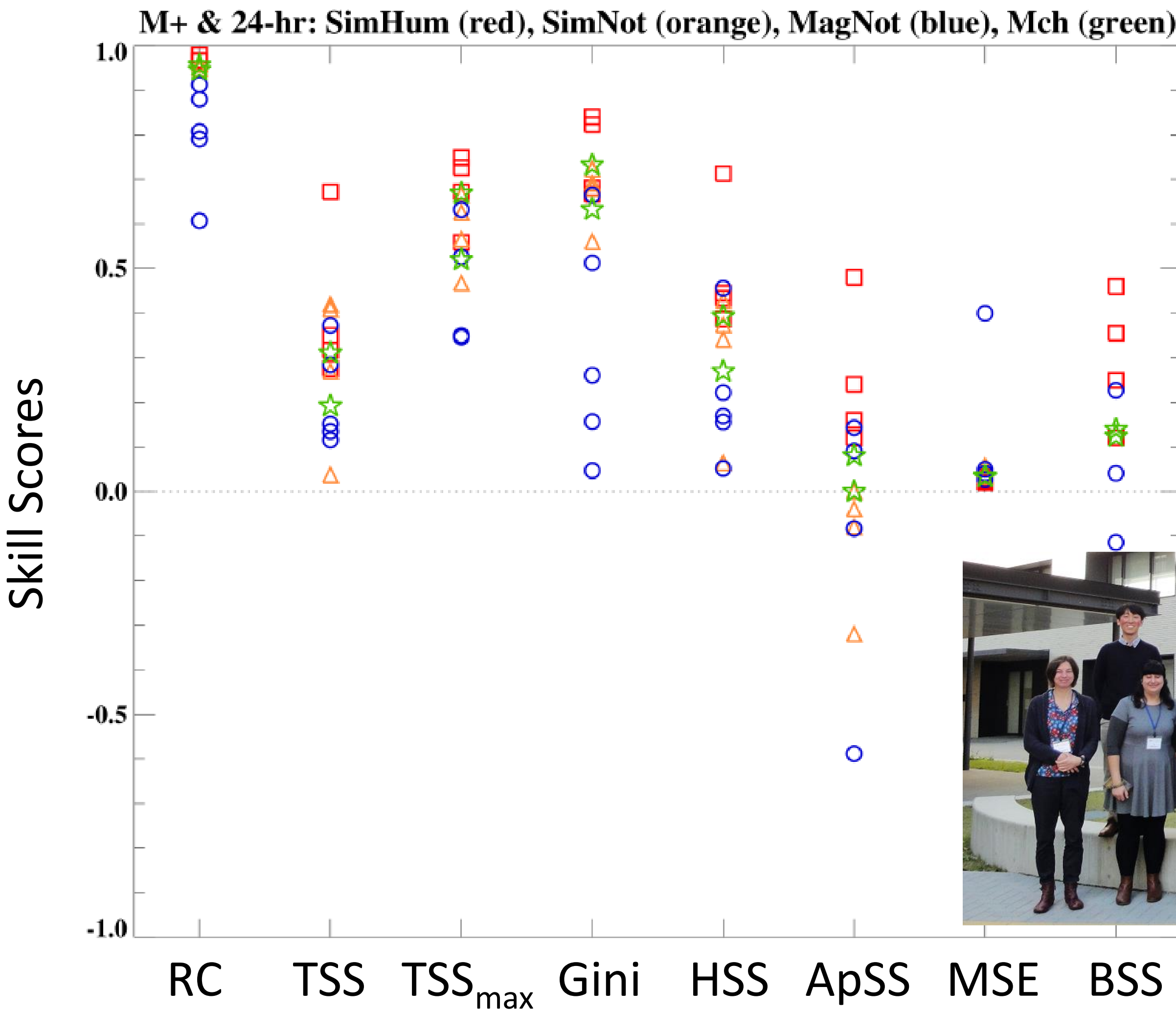
[H. Liu et al. 2019 ApJ]

Long-Short Term Memory (LSTM) / RNN



[T. Muranushi et al. 2016 arXiv]

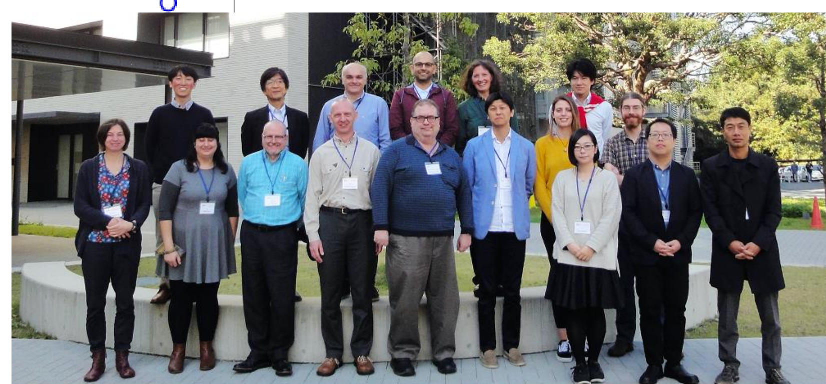
Comparison of Operational Forecasting of solar flares



[Leka, S.–H. Park, et al. 2019 ApJ]

The ranking of models depends on metric selection.

[Benchmark of flare predictions in Nagoya, Japan, in Nov 2017]

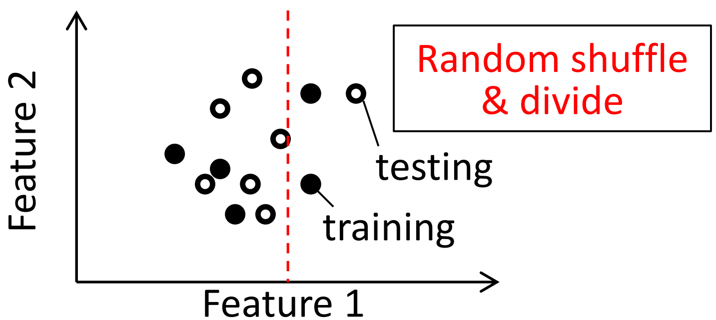


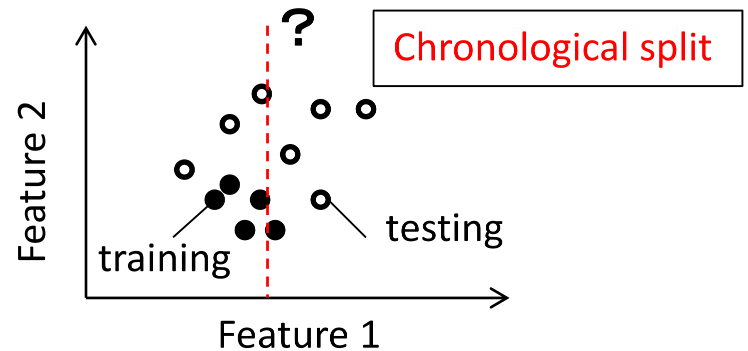
cf) Barnes+2016 ApJ

Tasks of Machine-learning Models

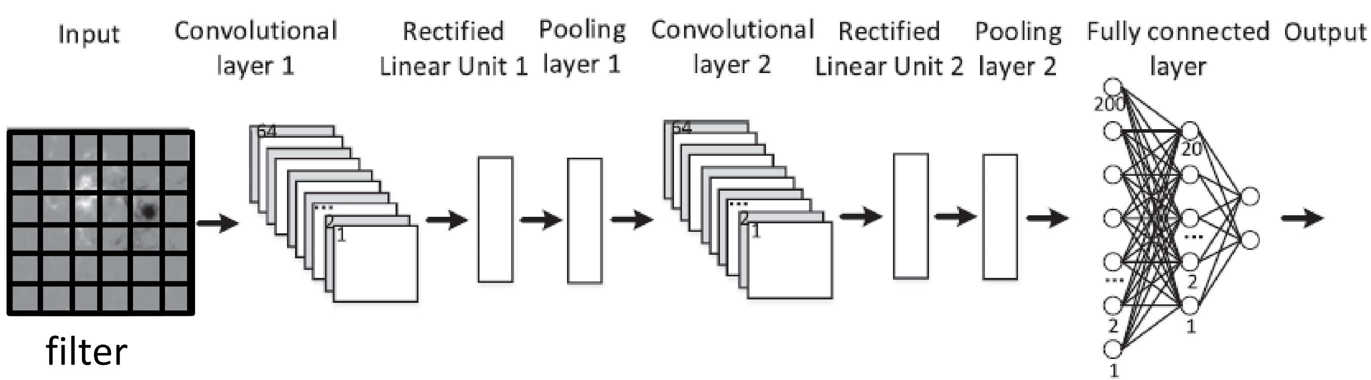
(1) Prediction using the Real-time Data :

- Previously we evaluated our model using the past data. [Nishizuka+2017 ApJ]
- But when evaluated with the real-time data, it was found that the performance was not enough (TSS=0.3).





(2) Prediction using Deep Neural Network (DNN) & Convolutional NN :



- XAI
- Trusted AI

[Huang+2018 ApJ]

cf) E. Jonas+2018 Sol. Physics.

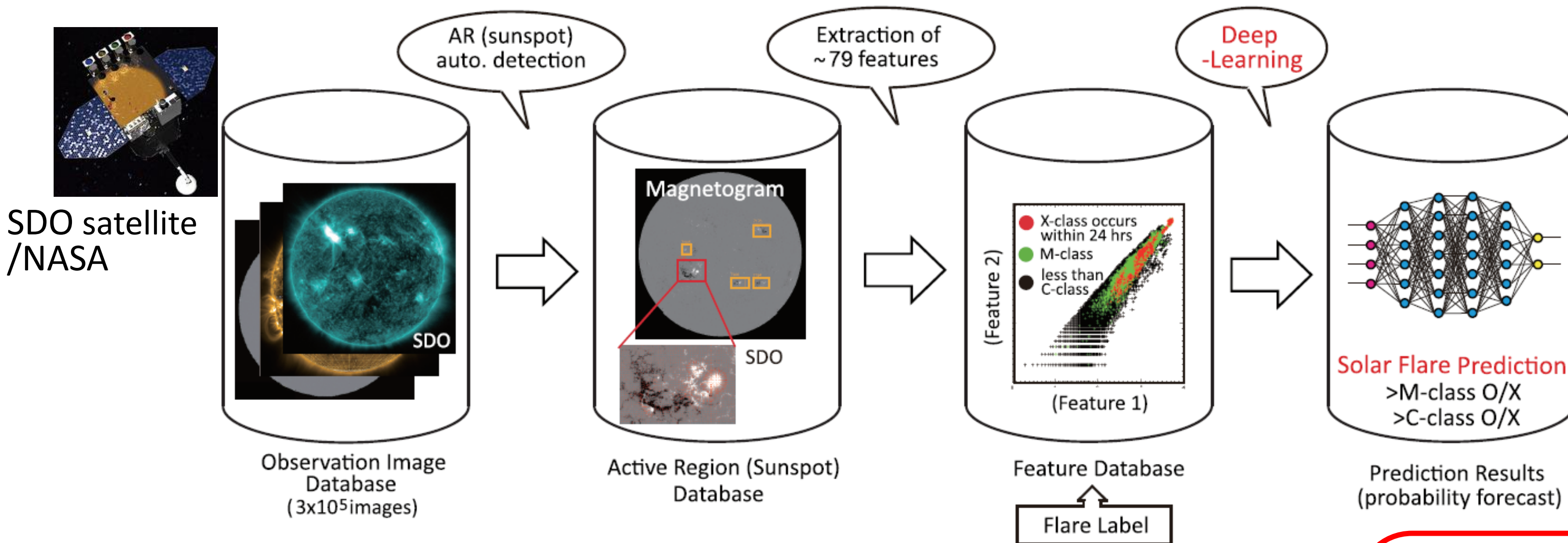
Features are automatically detected, but they are black boxes.

Deep Flare Net (DeFN) model

- We developed a flare prediction model using DNN, to increase TSS.
- The evaluation was done by the real-time data in an operational setting.
- Deep Flare Net (**D**<u>**e**</u>**FN**) = **Excellent** + <u>**N**</u>ippon/<u>**N**</u>ICT/….

↑ advice by Robert Steenburgh-san

[Nishizuka+2018 ApJ]



Period: 2010~ 2015
X class ~40 events
M class ~460 events
C class ~4000 events
Reduction to 1 hr cadence

- Magnetogram (B_{LOS}) 3TB, 10^5 images
- Vector magnetogram (B_x, B_y, B_z) 12TB
- Photospheric BP (UV 1600Å) 3TB
- Coronal hot BP (EUV 131Å) 3TB

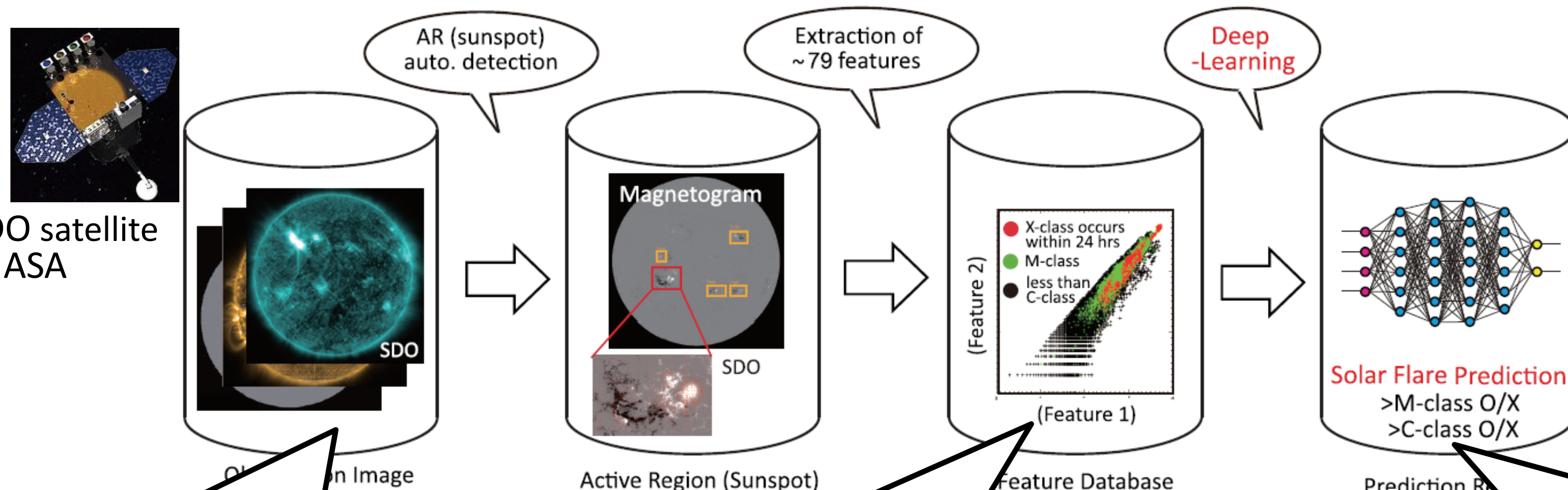
Flare probability
≧M ○%
< M ○%
≧C ○%
< C ○%

Deep Flare Net (DeFN) model

- We developed a flare prediction model using DNN, to increase TSS.
- The evaluation was done by the real-time data in an operational setting.
- Deep Flare Net (DeFN) = **Excellent** + Nippon/NICT/….

↑ advice by Robert Steenburgh-san

[Nishizuka+2018 ApJ]



SDO satellite /NASA

+ New wavelength
EUV 131Åimages
Coronal hot BP
($T>10^7$ K)

60 → 79 features
※19 new features

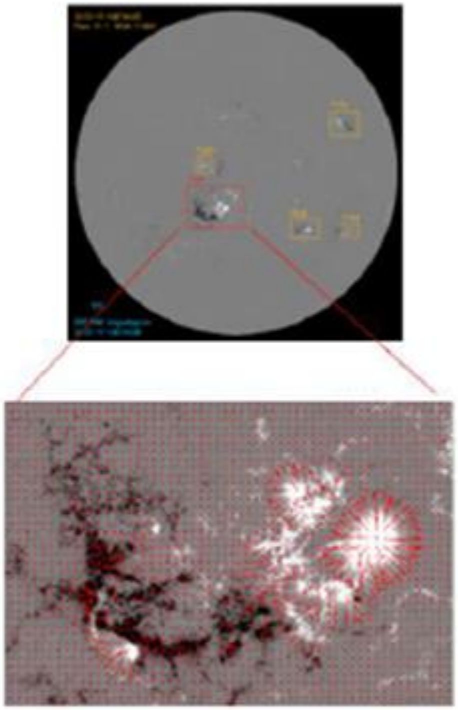
SVM, kNN, ERT → DNN
[baseline, NN+2017]

M class ~460 events
C class ~4000 events
Reduction to 1 hr cadence

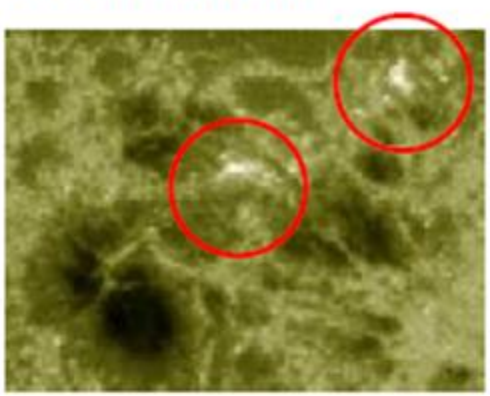
- Magnetogram (B_{LOS}) 3TB, 10^5 images
- Vector magnetogram (B_x, B_y, B_z) 12TB
- Photospheric BP (UV 1600Å) 3TB
- Coronal hot BP (EUV 131Å) 3TB

≧M ○%
< M ○%
≧C ○%
< C ○%

Extraction of 79 Solar Features



Vector magnetic field in a detected AR



Area
Total unsigned mag. Flux
Flux imbalance fraction
max /average Bz
max/average grad Bz
Max. length of **Neutral lines**
Total length of NLs
The number of NLs

Flare history (X, M)@AR
Flare history@1d before
soft X-ray 2hrs/4hrs average
max soft X-ray@1d before

X-ray/EUV131 data **1 & 2 hr before an image**

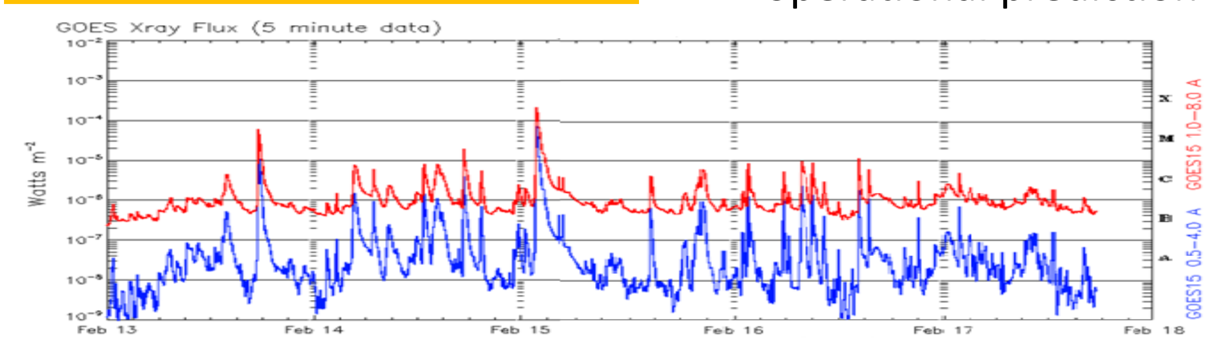
< Vector magnetogram >
Current helicity ($\Sigma Bz \cdot Jz$)
Lorentz force (ΣB^2)
Vertical current (Jz)

Chromospheric BP Area
Chm. BP max intensity
Chm. BP total intensity

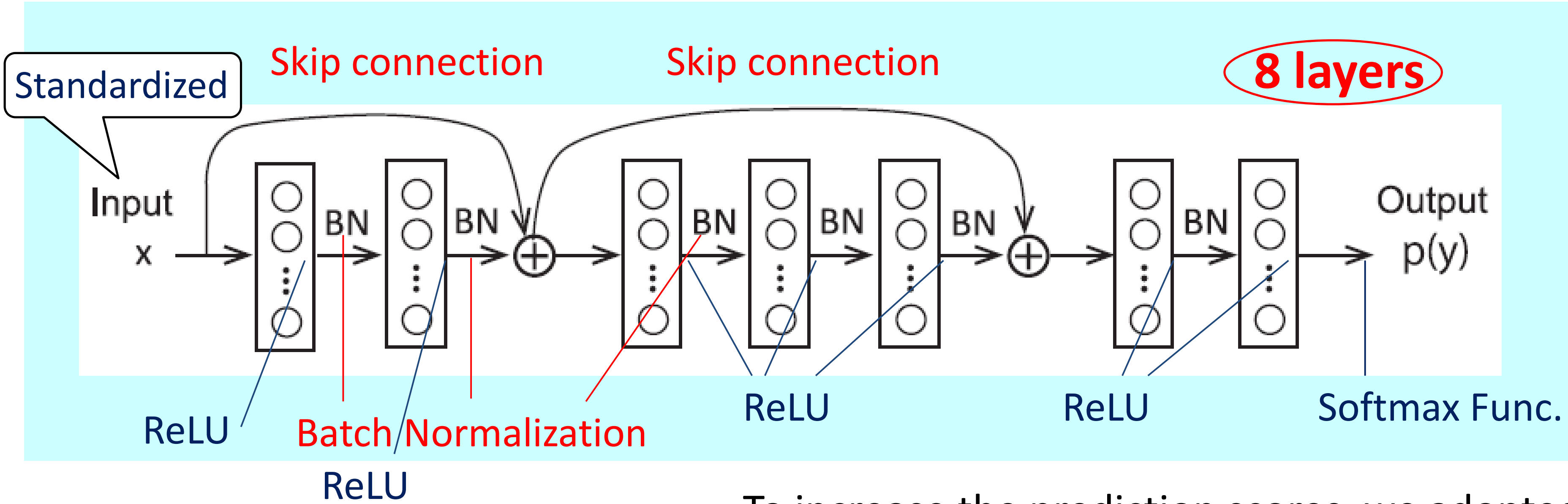
Time derivative (2,12,**24 hrs**)

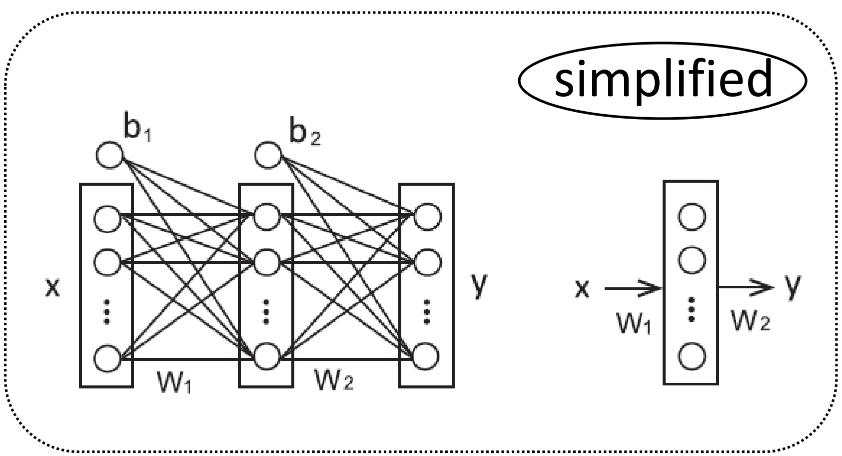
Coronal hot BP Area (131Å)
Coronal BP max intensity
Coronal BP total intensity

New features added for an operational prediction !!



Structure of Deep Flare Net





To increase the prediction scores, we adopted

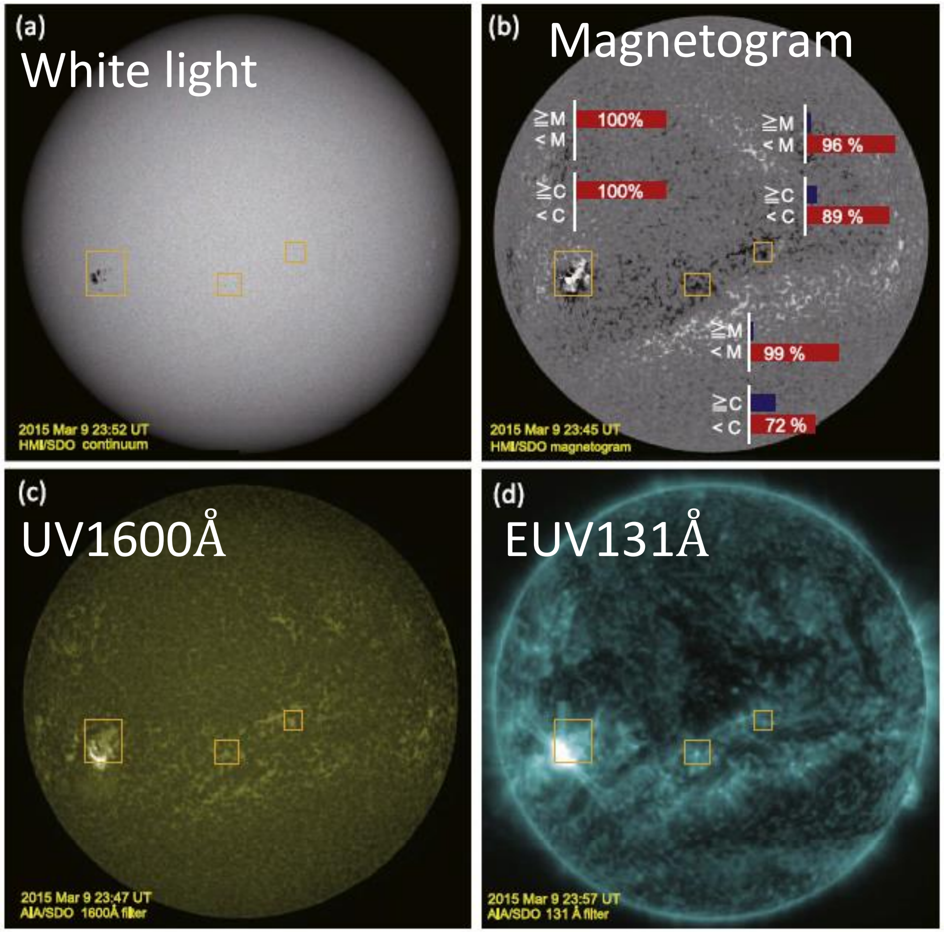
- ReLU (activation function)
- Skip connection (Residual Net)
- (mini-)Batch Normalization (BN)
- Weighted cross entropy (loss function)

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} w_k y^*_{nk} \log p(y_{nk}).$$

W_k = (1, 50) for >M
= (1, 4) for >C

* Liu+2019 used the same loss function.

Probability Forecast at each AR



At the last layer, softmax function,

$$p(yi) = \frac{\exp(yi)}{\sum_{j=1}^{N} \exp(yi)}$$

gives the probability of flare occurrence.

P(y_1) : ≧M-class flare probability
P(y_2) : < M-class flare probability

are used for two-class prediction.

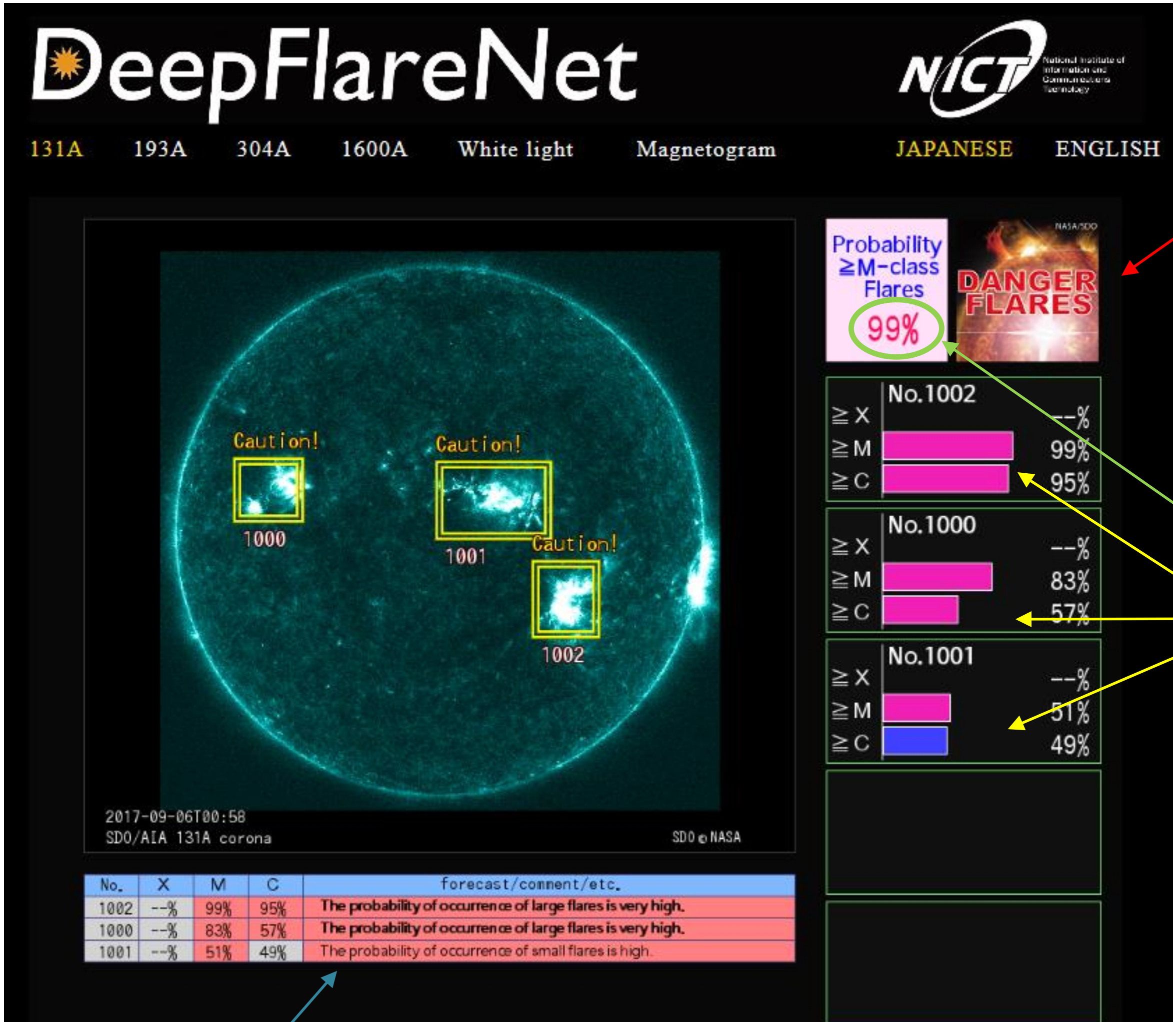
Finally, DeFN predicts flares by taking the max. probability

$$\hat{y} = \underset{k}{\operatorname{argmax}}\, p(y_k).$$

※ For **2-class** classification, thresh.=**50%**.

※ This can be easily extended to **4-class** classification. (X/M/C/O)

Operational Forecasting by DeFN



Text comments

URL https://defn.nict.go.jp

Alert mark

Danger / Warning / Quiet
like weather forecast (☀☁☂)

Flare Occurrence Probability

- Total prob. on disk
- Prob. at each region

- Near-real time data (nrt) from JSOC system (Stanford U., LMSAL, NASA)
- We started operating a web of Deep Flare Net in April. (Internal operation from Dec 2018.)

Prediction Results & Evaluation

- We evaluated the prediction results in an operational setting. We divided the database with a chronological split: 2010-2014 for training and 2015 for testing. We used TSS for evaluation.

		Observation: flare	Observation: no
Prediction	flare	TP	FP
	no	FN	TN

● **True Skill Statistic**
[by Hanseen & Kuipers '65]

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

TSS = 1 means 100%
Hit of flare prediction

[Nishizuka+2018 ApJ]

≧M-class

		Observation: flare	Observation: no
Prediction	flare	963	4382
	no	54	25937

TSS= **0.80**

≧C-class

		Observation: flare	Observation: no
Prediction	flare	4967	4420
	no	1171	20778

TSS= 0.63

- We achieved TSS=0.80, though flares are overestimated. (large FP)

cf) Huang+2018 ApJ: TSS=0.66 (≧M), 0.49 (≧C) DNN

Prediction Results & Evaluation

- We evaluated the prediction results in an <u>operational setting</u>. We divided the database with a chronological split: 2010-2014 for training and 2015 for testing. We used TSS for evaluation.

		Observation	
		flare	no
Prediction	flare	TP	FP

● **True Skill Statistic**
[by Hanseen & Kuipers ’65]

TSS = TP — FP

TSS = 1 means 100% Hit of flare prediction

★Evaluation in the Real-time Operation (Jan-May 2019)

≧M-class

		Observation	
		flare	no
Prediction	flare	0	25
	no	0	491

TSS= **none**

≧C-class

		Observation	
		flare	no
Prediction	flare	26	24
	no	4	463

TSS= <u>0.82</u>

By changing the probability threshold from 50%, we can control TSS & FP. Users can select it, depending on their purposes/demands.

Importance Ranking of Features (from ERT)

Ranking	Features	Importance
1.	Xhis	0.0519
2.	Xmax1d	0.0495
3.	Mhis	0.0365
4.	TotNL	0.0351
5.	Mhis1d	0.0342
6.	NumNL	0.0341
7.	Usflux	0.0332
8.	CHArea	0.0235
9.	Bave	0.0230
10.	Xhis1d	0.0224
11.	TotBSQ	0.0199
12.	VUSflux	0.0196
13.	Bmax	0.0193
14.	MeanGAM	0.0179
15.	dt24SavNCPP	0.0171

Total 50 features

- Flare history (total, 1day),
- Max X-ray intensity 1 day before

- Total length of Neutral Lines
- Number of NLs
- Unsigned magnetic flux,
- averaged/max Bz

Chromospheric Bright Area

- Total magnitude of Lorentz force
- Mean angle of field from radial
- Sum of the modules of the net current per polarity

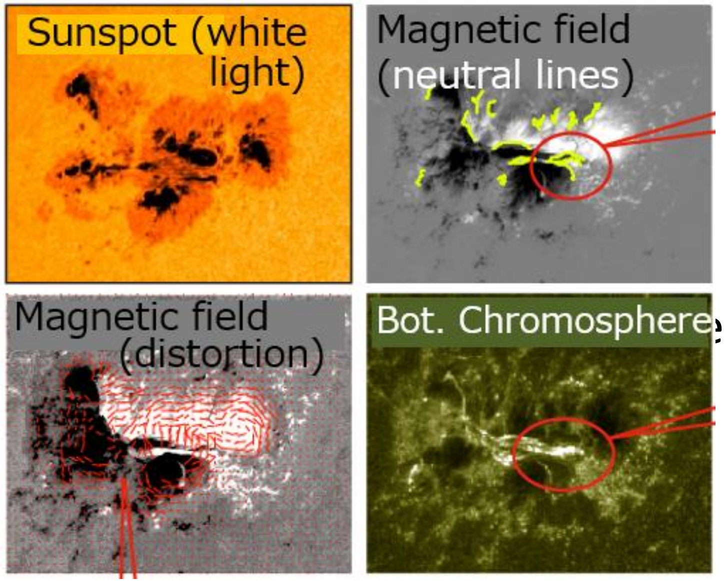
Importance Ranking of Features (from ERT)

Ranking	Features	Importance
1.	Xhis	0.0519
2.	Xmax1d	0.0495
3.	Mhis	0.0365
4.	TotNL	0.0351
5.	Mhis1d	0.0342
6.	NumNL	0.0341
7.	Usflux	0.0332
8.	CHArea	0.0235
9.	Bave	0.0230
10.	Xhis1d	0.0224
11.	TotBSQ	0.0199
12.	VUSflux	0.0196
13.	Bmax	0.0193
14.	MeanGAM	0.0179
15.	dt24SavNCPP	0.0171

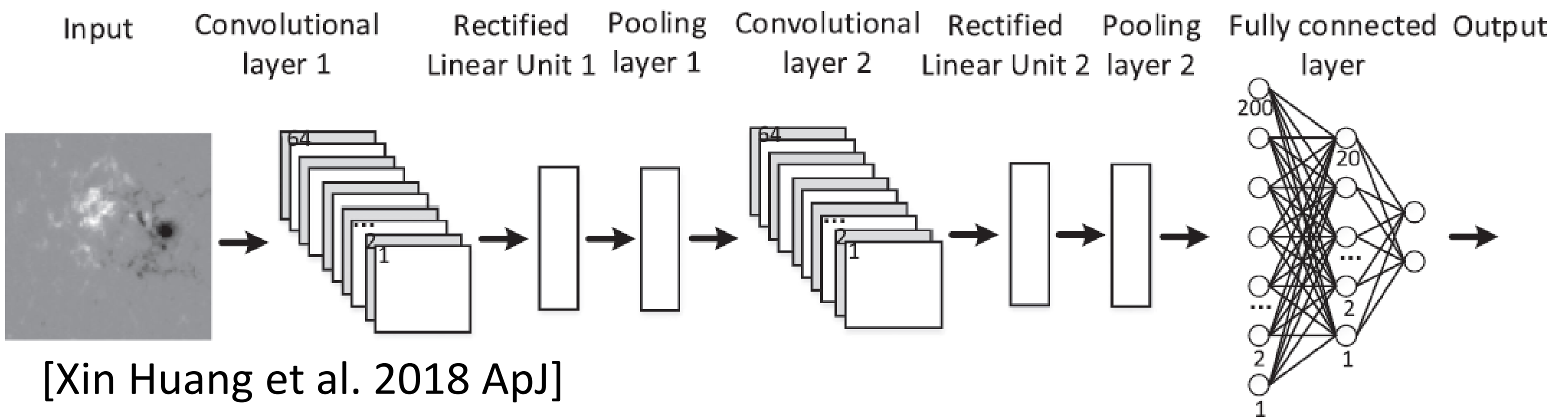
Total 50 features

Ranking:
① Random Forest, **ERT** (Gini Impurity)
② **LASSO** (SVM) → Feature selection
③ Fisher-score

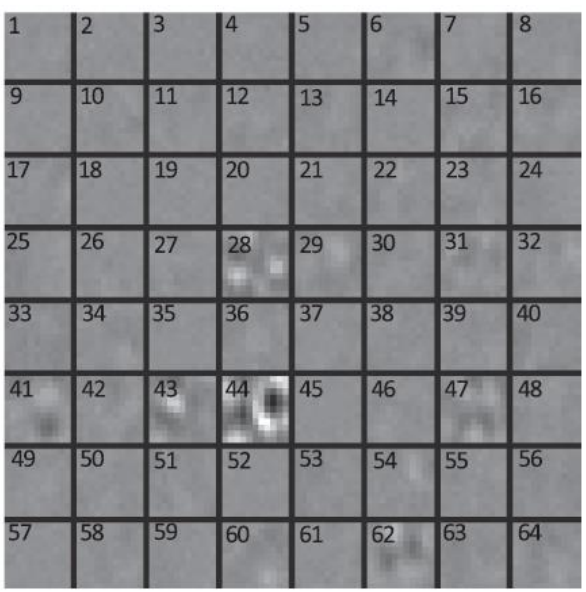
- Total length of Neutral Lines
- Number of NLs
- Unsigned magnetic flux,

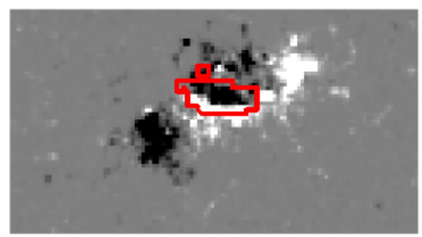


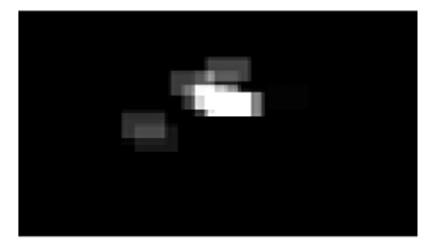
Comparison with Convolutional Neural Network (CNN) used for prediction



[Xin Huang et al. 2018 ApJ]







- (Left) convolutional kernel
- The pattern indicates important information
- CNN also focus on the area of strong magnetic field and around the magnetic neutral lines.

→ Comparison between geometrical and physical parameters. Is there a parameter which has never been considered?

Standard Evaluation method ?

Agreement / Consensus

- Event selection
  - positive/negative event ratio
  - including near limb events? (not scientific, but operational)

- The way to split the training/testing datasets
  - Random shuffle & divide (10-fold CV)
  - week shuffle & divide, AR shuffle & divide
  - chronological split (Time series CV)
  - one leave-out (most operational)

Radar chart

Star ranking

- Selection of metrics
  - User dependent? Users select models by themselves? How?
  - Easy indicators to show the robust performance of models.
  - AI model needs a clear rule/purpose for design/optimization.
  - We should have a competition by determining rules, to accelerate the AI model development. (like Kaggle, or in CCMC)

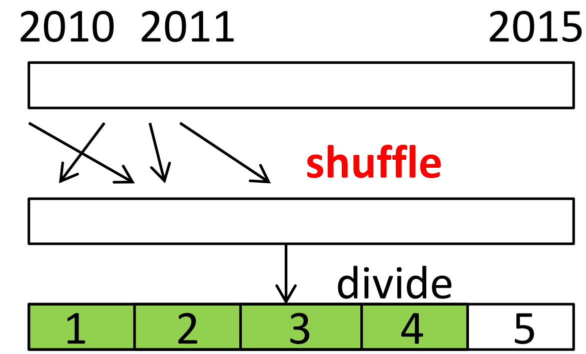
kaggle

The way to split the training/testing datasets

① K-fold Cross Validation= shuffle & divide

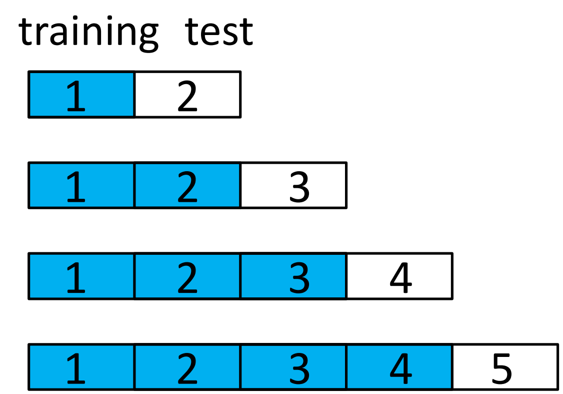
(ー) separating data before a flare to both training/testing datasets

(ー) training by future data.



② Time-series Cross Validation

good for evaluation in an operational setting



③ Forecasting

Time series CV ≠ forecasting

(＋) final destination

(ー) hard to compare methods A & B

(ー) It takes time

or 2011-2016

2012-2017

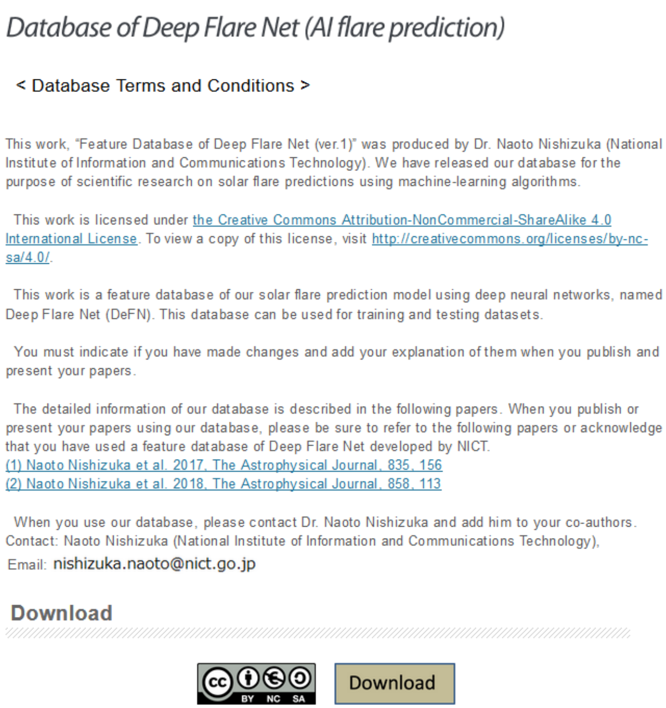
2009-2018

Leave-one-out CV

Database & Code of DeFN

◆ Database Release

http://wdc.nict.go.jp/IONO/wdc/solarflare/index.html



Database of Deep Flare Net (AI flare prediction)

< Database Terms and Conditions >

This work, "Feature Database of Deep Flare Net (ver.1)" was produced by Dr. Naoto Nishizuka (National Institute of Information and Communications Technology). We have released our database for the purpose of scientific research on solar flare predictions using machine-learning algorithms.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/.

This work is a feature database of our solar flare prediction model using deep neural networks, named Deep Flare Net (DeFN). This database can be used for training and testing datasets.

You must indicate if you have made changes and add your explanation of them when you publish and present your papers.

The detailed information of our database is described in the following papers. When you publish or present your papers using our database, please be sure to refer to the following papers or acknowledge that you have used a feature database of Deep Flare Net developed by NICT.
(1) Naoto Nishizuka et al. 2017, The Astrophysical Journal, 835, 156
(2) Naoto Nishizuka et al. 2018, The Astrophysical Journal, 858, 113

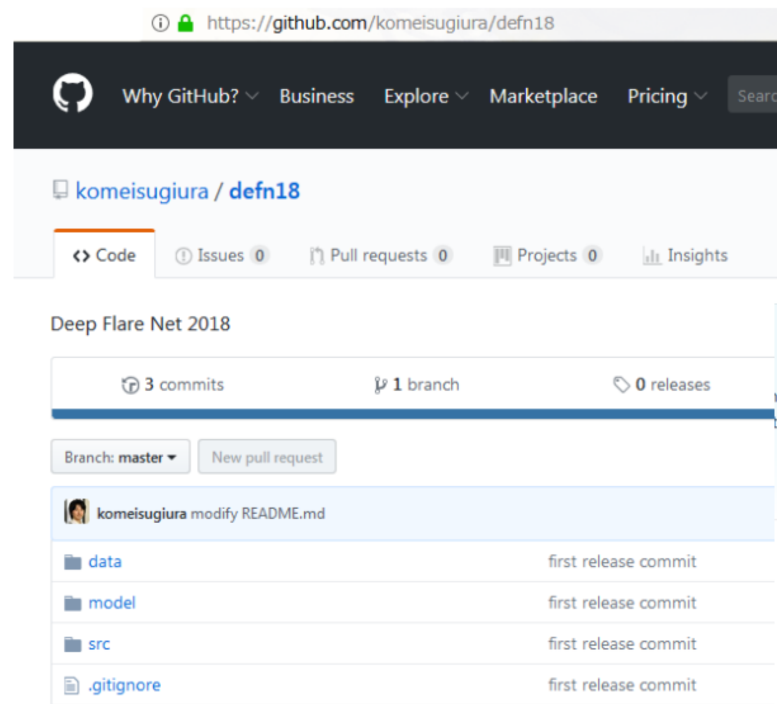
When you use our database, please contact Dr. Naoto Nishizuka and add him to your co-authors.
Contact: Naoto Nishizuka (National Institute of Information and Communications Technology),
Email: nishizuka.naoto@nict.go.jp

Download

BY NC SA Download

◆ Code Release

https://github.com/komeisugiura/defn18



For the purpose of 1) the evaluation of our model by others, 2) the comparison of different models, 3) acceleration of the new model development, and 4) standardization of feature database.

Database & Code of DeFN

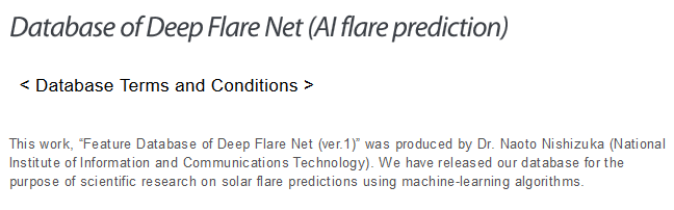
◆ Database Release

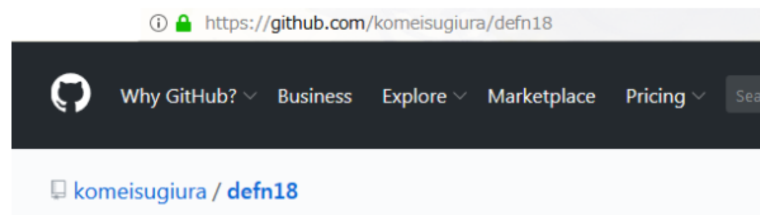
http://wdc.nict.go.jp/IONO/wdc/solarflare/index.html

WORLD DATA SYSTEM ICSU

◆ Code Release

https://github.com/komeisugiura/defn18

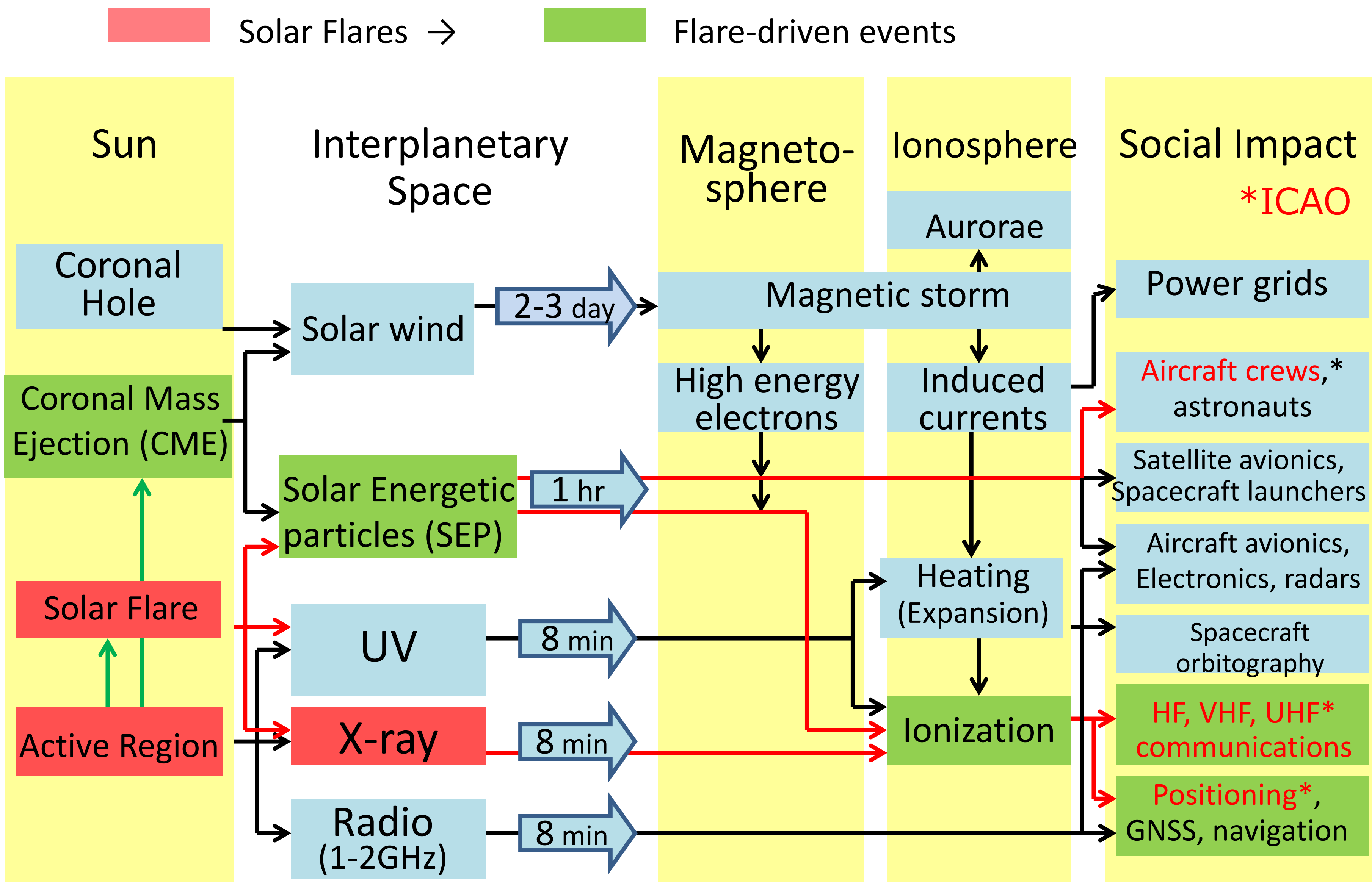




- Without a natural database (flare event ratio = chronological base rate, not arbitrary selected), it's better to use TSS for comparison.
- With a **natural database**, we can discuss several skill scores like BSS and HSS as well as TSS, to compare each model in a fair way.
- Let's share natural database of DeFN and others if possible.

For the purpose of 1) the evaluation of our model by others, 2) the comparison of different models, 3) acceleration of the new model development, and 4) standardization of feature database.

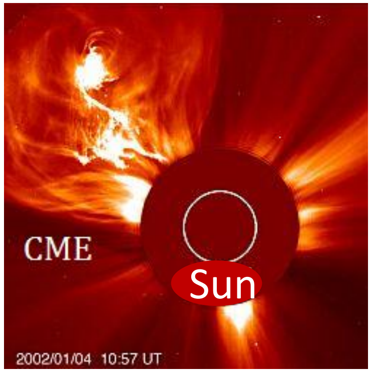
Application to other SWx Forecasting



[Revised from the original: Observatorie de Paris]

Application to CME Prediction

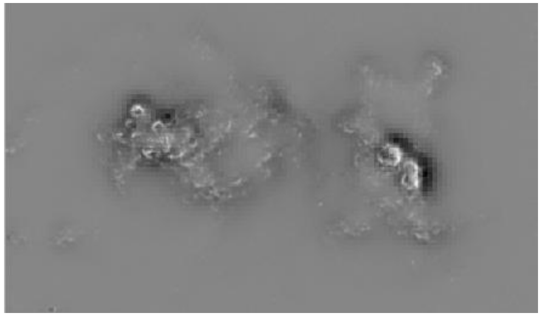
- Our flare prediction (DeFN) model can be applied to a CME occurrence prediction model, by extending training database and labels with a CME list.



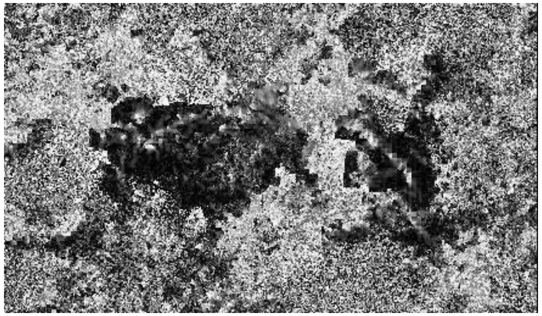
Flare & CME
(LASCO/SOHO)

- **New global features**:
  Magnetic helicity & free energy (global twist),
  Shear & Dip angles (non-potentiality),
  Poynting flux (Energy injection), sigmoid/loop length.

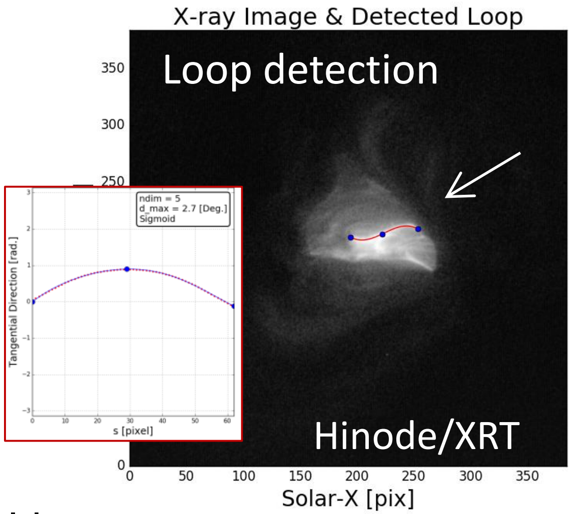
- 1,240 CMEs with velocity >500km/s, angle width>30°



Magnetic free energy
$E=(Bx^2+By^2)-(Bxp^2+Byp^2)$



Shear angle
(stress from the potential field
In the horizontal direction)



[Iida & Kawabata]

Summary

- We developed a solar flare prediction model using machine-learning/DNN, which we named Deep Flare Net (DeFN). The model can predict the flare probability at each region. For 24hr prediction, TSS=0.80 for ≧M, and TSS=0.63 for ≧C.
- We started operating a Web site of DeFN model in April. In the real-time operation, TSS=0.82 for ≧C flares (thresh=50%).
- DeFN model can be used to predict other SWx phenomena, such as CMEs from ARs (30%). (Transfer-learning)
- Users can determine the threshold of probability, depending on their purposes/demands. We would like to have a useful feedback from users to further improve our model.
- We opened our database and code, so we would like lots of people to use them and have lots of discussion.