

Flare Transformer: 磁場画像と物理特徴量を用いた太陽フレア予測

Flare Transformer: Solar Flare Prediction using Magnetograms and Sunspot Physical Features

兼田 寛大*¹ 飯田 紘*¹ 西塚 直人*² 久保 勇樹*² 杉浦 孔明*¹
Kanta Kaneda Tsumugi Iida Naoto Nishizuka Yûki Kubo Komei Sugiura

*¹慶應義塾大学
Keio University

*²国立研究開発法人情報通信研究機構
National Institute of Information and Communications Technology

The prediction of solar flares is essential for reducing the potential damage to social infrastructures that are vital to society. However, predicting solar flares accurately is a very challenging task. In this paper, we propose a solar flare prediction model, Flare Transformer, which handles both images and physical features through the Magnetogram Module and the Sunspot Feature Module. We introduce the transformer attention mechanism to model the temporal relationships. We also introduce a new differentiable loss function to balance the two major metrics of the Gandin–Murphy–Gerrity score and Brier skill score. Comparative experiments using Gandin–Murphy–Gerrity score and true skill statistics as metrics showed that the proposed method achieves better performance than baseline methods and human experts.

1. はじめに

太陽フレアによって放出される X 線や高エネルギー粒子などは、電波障害や停電、宇宙飛行士への健康的被害などをもたらす [Bhattacharjee 20]. また、Carrington 級フレアによる経済的損失は北米で約 1,630 億ドルと推定されている [SwissRe 16]. そのため、宇宙天気予報による太陽フレアの予測技術が重要である。しかし、太陽フレアの発生を正確に予測することは非常に難しい。実際、専門家でも太陽フレアを予測するのは非常に困難である。例えば、2000–2015 年の期間における専門家予測は $GMGS = 0.48$ (Gandin–Murphy–Gerrity score) および $TSS_{\geq M} = 0.50$ (true skill statistics) であった [Kubo 17]. ただし、これらの指標は完璧な予測に対してスコア 1 を返す。

このような背景から、本論文では 24 時間以内に発生する最大の太陽フレアクラスを予測するタスクを扱う。図 1 に提案手法の概要図を示す。入力、Helioseismic and Magnetic Imager (HMI) [Scherrer 12] によって撮影された磁場画像、および領域レベルの物理特徴量であり、出力は予測クラスである。

既存手法の最大のボトルネックは、実際に発生するフレアよりも小さなクラスのフレアを誤って予測してしまうことである。例えば、DeFN [Nishizuka 18] は X クラスフレアの 89% を M クラスと誤って予測した。

以上のことから、本論文では時系列画像と物理特徴量を扱い、正確で信頼性の高い太陽フレア予測を行う Flare Transformer (FT) を提案する。

本研究の独自性は以下である。

- Magnetogram Module (MM) と Sunspot Feature Module (SFM) により、磁場画像と物理特徴量の両方を扱う太陽フレア予測モデル、FT を提案する。(3.1 章を参照)
- 入力特徴量間の時間的関係をモデル化するため、transformer 型注意機構を導入する。
- 太陽フレア予測において重要な指標である GMGS と Brier Skill Score (BSS) のバランスをとるため、GMGS 損失と BSS 損失を導入する。(3.2 章を参照)

連絡先: 兼田寛大, 慶應義塾大学, 神奈川県横浜市港北区日吉 3-14-1, k.kaneda@keio.jp

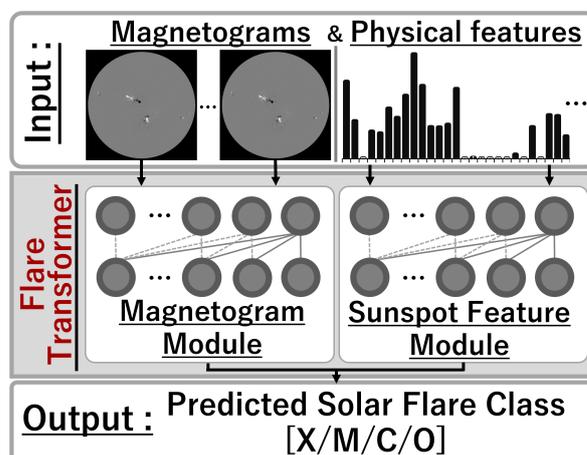


図 1: Flare Transformer の概要図。

2. 問題設定

本論文では、時刻 t から 24 時間以内に発生する最大の太陽フレアクラス、 $\mathbf{y}_t = \text{flareclass}(\max\{p_{t+1}, p_{t+2}, \dots, p_{t+24}\})$ を予測するタスクを扱う。ここで、 p_t と $\text{flareclass}(\cdot)$ はそれぞれ時刻 t から 1 時間以内の X 線強度と、4 種類の太陽フレアクラス、X, M, C, O についての 1-of-K 表現を表す。

本タスクにおいて、以下の入出力を想定する。

- 入力: 磁場画像および物理特徴量。
- 出力: 各太陽フレアクラスに関する予測確率を示す 4 次元ベクトル。

本タスクでは、モデル出力 $p(\hat{\mathbf{y}}_t)$ が \mathbf{y}_t にできるだけ近いことが望ましい。ただし、太陽フレアは各クラスの発生確率が不均衡である現象のため、すべてのフレアを O クラスと予測することは不適切である。このような予測を避けるためには、GMGS [Gandin 92] や BSS [Nishizuka 20] などの太陽フレア予測の分野において標準的な指標を最大化する $p(\hat{\mathbf{y}}_t)$ を出力することが望ましい。

また、本論文では以下の理由から太陽フレア予測タスクを回帰問題として扱わない。

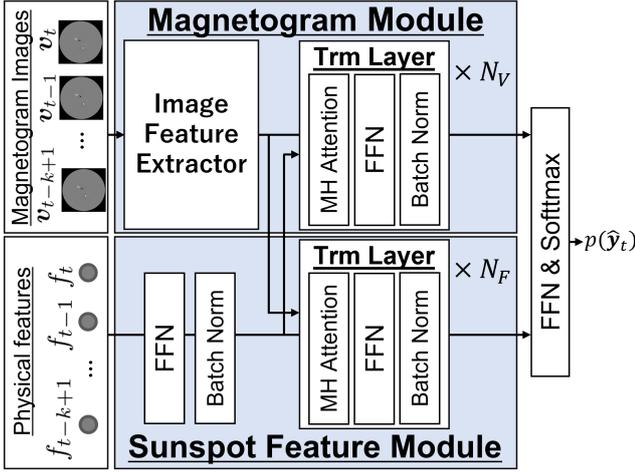


図 2: 提案手法のネットワーク図.

- 上記の分類タスクが太陽フレア予測の分野において標準的な設定であるため. (例: [Nishizuka 18] [Park 18])
- 専門家予測が上記の設定となっているため.

以上の理由から, 本タスクを分類問題として扱うことは合理的である.

本論文では, 太陽全体の画像から太陽フレアクラスを予測することを前提とする. そのため, 黒点の領域レベルにおける予測 [Nishizuka 18] は扱わない. また, SDO/HMI [Scherrer 12] で撮影された太陽画像を磁場画像と定義する.

タスクの評価尺度には, GMGS, TSS_{≥M} [Kubo 17] および BSS_{≥M} を使用する. ここで, “≥M” は太陽フレアクラスについて, M クラス以上または M クラス未満の二値として評価したことを示す.

3. 提案手法

3.1 Flare Transformer

ネットワーク入力 \mathbf{x} を以下のように定義する.

$$\mathbf{x} = (V_{t-k+1:t}, F_{t-k+1:t})$$

$$V_{t-k+1:t} = (\mathbf{v}_{t-k+1}, \mathbf{v}_{t-k+2}, \dots, \mathbf{v}_t)$$

$$F_{t-k+1:t} = (\mathbf{f}_{t-k+1}, \mathbf{f}_{t-k+2}, \dots, \mathbf{f}_t)$$

ここで, $\mathbf{v}_t \in \mathbb{R}^{512 \times 512}$ と $\mathbf{f}_t \in \mathbb{R}^{90}$ はそれぞれ時刻 t における磁場画像および物理特徴量を示す.

ネットワークの構造を図 2 に示す. 図において, “MH Attention” は multi-head attention 層, “Trm Layer” は transformer 層, そして “FFN” は順伝播層を示す. 提案手法は主に 2 つのモジュールからなり, Magnetogram Module (MM) および Sunspot Feature Module (SFM) と呼ぶ.

MM では, はじめに磁場画像 $V_{t-k+1:t}$ が以下のように計算される.

$$\mathbf{h}_V = f_{FE}(V_{t-k+1:t})$$

ここで, f_{FE} は図 3 に示す Image Feature Extractor を示す. 図において, “Conv”, “Batch Norm” および “Avg pooling” はそれぞれ畳み込み層, 正規化層, および平均プーリング層を示す. 同様に, SFM において物理特徴量 $F_{t-k+1:t}$ が以下のように計算される.

$$\mathbf{h}_F = f_{BN}(f_{FFN}(F_{t-k+1:t}))$$

ここで, f_{BN} は正規化層, f_{FFN} は FFN を示す. その後, \mathbf{h}_V と \mathbf{h}_F を結合することで \mathbf{h}_{VF} を得る.

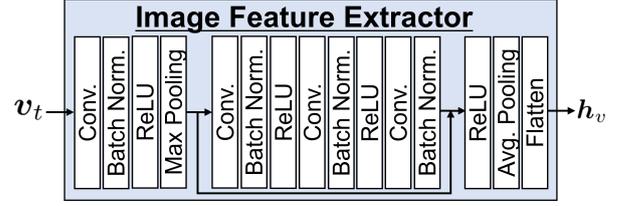


図 3: Image Feature Extractor の構造.

次に, N_V 層の transformer 層において入力画像および物理特徴量の時系列関係が計算される. Multi-head attention 層では, \mathbf{h}_V および \mathbf{h}_{VF} がそれぞれ $\mathbf{h}_V^{(i)} \in \mathbb{R}^{k \times d}$ および $\mathbf{h}_{VF}^{(i)} \in \mathbb{R}^{k \times 2d}$ ($i = 1, \dots, N_{\text{head}}$) に分割される. ここで $d = H/N_{\text{head}}$ であり, H は隠れ層の数, N_{head} はヘッド数を示す.

クエリ $Q^{(i)} \in \mathbb{R}^{k \times d}$, キー $K^{(i)} \in \mathbb{R}^{k \times 2d}$ そしてバリュー $V^{(i)} \in \mathbb{R}^{k \times 2d}$ が i 番目のヘッドについて以下のように計算される.

$$Q^{(i)} = W_q^{(i)} \mathbf{h}_V^{(i)}, K^{(i)} = W_k^{(i)} \mathbf{h}_{VF}^{(i)}, V^{(i)} = W_v^{(i)} \mathbf{h}_{VF}^{(i)}$$

ここで, $W_q^{(i)}, W_k^{(i)}, W_v^{(i)}$ はそれぞれ $Q^{(i)}, K^{(i)}, V^{(i)}$ の重みを表す.

Transformer 層の出力 \mathbf{h}_{trm} は以下のように計算される.

$$\mathbf{h}_{\text{trm}} = f_{BN}(f_{FFN}(\mathbf{h}_{\text{mha}}))$$

$$\mathbf{h}_{\text{mha}} = (\mathbf{f}_{\text{attn}}^{(1)}, \dots, \mathbf{f}_{\text{attn}}^{(N_{\text{head}})})$$

$$\mathbf{f}_{\text{attn}}^{(i)} = \text{softmax}\left(\frac{Q^{(i)} K^{(i)\top}}{\sqrt{d}}\right) V^{(i)}$$

MM の出力 \mathbf{h}_{MM} は N_V 層の transformer 層によって計算され, SFM の出力 \mathbf{h}_{SFM} も同様に N_F 層の transformer 層によって計算される.

最後に, 予測クラス y_t^* は以下のように計算される.

$$y_t^* = \text{argmax}_i(p(\hat{y}_{ti}))$$

$$p(\hat{y}_t) = \text{softmax}(f_{FFN}(\mathbf{h}_{\text{MM}}; \mathbf{h}_{\text{SFM}}))$$

ここで, $p(\hat{y}_{ti})$ は i 番目のクラスに関する予測確率を示す.

3.2 損失関数

FT では, GMGS [Gandin 92] と BSS [Nishizuka 20] のバランスをとるため, GMGS 損失および BSS 損失を導入する.

既存手法で用いられている損失関数は, GMGS とは無関係な重みを用いてクラス間のバランスを調整しているため, GMGS の改善には有効ではない. 一方で, GMGS のスコア行列を重みとして用いることで, GMGS を効果的に向上させることができると考えられる. 以上のことから, GMGS 損失を以下のように定義する.

$$\mathcal{L}_{\text{GMGS}} = -\frac{1}{NI} \sum_{n=1}^N s_{i^*j^*} \sum_{i=1}^I y'_{ni} \log(p(\hat{y}_{ni}))$$

$$i^* = \text{argmax}_i(y_{ni}), j^* = \text{argmax}_j(p(\hat{y}_{nj}))$$

ここで, N, I および $s_{i^*j^*}$ はそれぞれサンプル数, クラス数, および GMGS のスコア行列 [Gandin 92] を示す. また, $p(\hat{y}_{ni}), y_{ni}$, そして y'_{ni} はそれぞれ n 番目のサンプルの i 番目のクラスに関する予測確率, 正解ラベル, そしてラベル平滑化された \mathbf{y}_n を示す.

次に, 予測の信頼性を向上させるため, BSS 損失を以下のように定義する. ここで, BSS は微分可能であるため, BSS

表 1: 各手法における定量的結果.

Method	GMGS \uparrow	TSS $\geq M\uparrow$	BSS $\geq M\uparrow$
DeFN [Nishizuka 18]	0.375 \pm 0.141	0.413 \pm 0.150	-0.022 \pm 0.782
DeFN-R [Nishizuka 20]	0.302 \pm 0.055	0.279 \pm 0.162	0.036 \pm 0.982
Ours (FT)	0.503\pm0.059	0.530\pm0.112	0.082\pm0.974
Human [Kubo 17] [Murray 17]	0.48	0.50	0.16

損失に使用することができる.

$$\mathcal{L}_{\text{BSS}} = -\frac{1}{NI} \sum_{n=1}^N \sum_{i=1}^I (p(\hat{y}_{ni}) - y_{ni})^2$$

損失関数として、以下を使用する.

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda_{\text{GMGS}} \mathcal{L}_{\text{GMGS}} + \lambda_{\text{BSS}} \mathcal{L}_{\text{BSS}}$$

ここで、 \mathcal{L}_{CE} は \mathbf{y}_n と $\hat{\mathbf{y}}_n$ に関する交差エントロピー損失を示し、 λ_{GMGS} および λ_{BSS} は重みを示す.

4. 実験

4.1 データセットと実験設定

実験では、SDO のウェブアーカイブ *1 および物理特徴量データベース *2 から収集したデータセットを用いて本手法を評価した.

本データセットには、2010 年 6 月から 2017 年 12 月における 61315 のサンプルが含まれている. また本データセットにおいて、正解ラベルが X, M, C, O クラスとなるサンプルはそれぞれ 492, 4745, 19736, 36342 となっている. ここで、X や M クラスの太陽フレアは発生する可能性が極めて低い事象である. 実際に、2017 年において X クラスの太陽フレアは全体の約 2.9%であった.

本研究では時系列予測において標準的な手法である時系列交差検証 [Tashman 00] に基づき、訓練集合とテスト集合を分割した. ここで時系列交差検証では、訓練サンプルはテスト集合を構成するサンプルよりも時系列的に前のサンプルで構成される.

ハイパーパラメータは、 N_V を 1, N_F を 2, λ_{GMGS} を 0.01, λ_{BSS} を 10 とした. 最適化には Adam を使用し、学習率は 7×10^{-7} , バッチサイズは 32 とした. 学習にはメモリ 11GB 搭載の GeForce RTX 2080 および Intel Core i9-9900K を使用した.

4.2 実験結果

表 1 にベースライン手法と提案手法の定量的な結果を示す. ベースライン手法は、物理特徴量のみを入力とする DeFN [Nishizuka 18] および DeFN-R [Nishizuka 20] とした. また、表には各スコアの平均値と標準偏差を示した.

タスクの評価尺度には、GMGS [Gandin 92], TSS $\geq M$ [Kubo 17] および BSS $\geq M$ [Nishizuka 20] を使用した. ここで、“ $\geq M$ ” は太陽フレアクラスが M クラス以上または M クラス未満の二値として評価したことを示す. 本論文では、GMGS と BSS $\geq M$ を主要尺度とした. GMGS は各クラスの発生確率が不均衡なクラス分類タスクに適した尺度である. [Kubo 17]. また、BSS は太陽フレア予測における標準的な指標であり、予測の信頼性を評価する尺度である [Nishizuka 20]. なお、2016 年は X クラスに太陽フレアが発生しなかったため、GMGS については近似値を用いた.

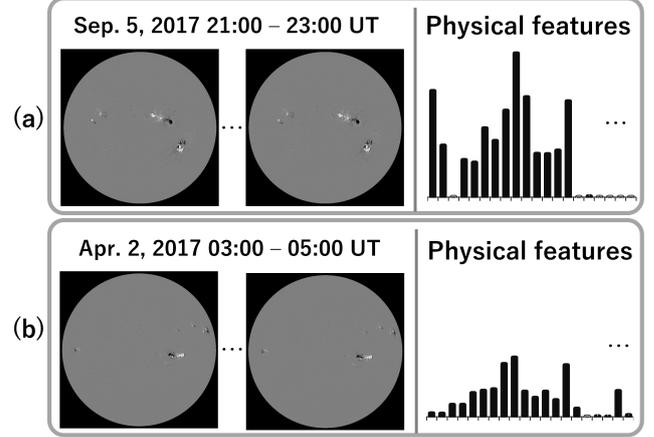


図 4: 定性的結果. (a) X クラスの太陽フレア. (b) M クラスの太陽フレア.

表 1 より、DeFN, DeFN-R, および FT の GMGS はそれぞれ 0.375, 0.302, 0.503 ポイントであった. したがって、FT は DeFN より GMGS において 0.128 ポイント優れていた.

表 1 は専門家予測の値も示している. 表より、専門家予測に対する GMGS と TSS $\geq M$ は、2000–2015 年においてそれぞれ 0.48 と 0.50 であった [Kubo 17]. 本手法は GMGS と TSS $\geq M$ において人間の専門家を上回っており、このことから FT が太陽フレア予測において有効であると考えられる.

図 4 に定性的結果を示す. 図 4(a) は 2017 年 9 月 5 日 21:00–23:00 の磁場画像および物理特徴量である. 注意としてスペースの都合上、 \mathbf{f}_t については 90 のうち最初の 20 の特徴量について正の値のみを表示している. 予測は $y_t^* = \text{“X”}$ である. ここで、2017 年 9 月 5 日 23 時から 24 時間以内である 2017 年 9 月 6 日 12 時 2 分に X クラスの太陽フレアが発生した. したがって、FT は正しい最大太陽フレアのクラスを予測することができた. 同様に、M クラスの太陽フレアを正しく予測したサンプルを図 4(b) に示す.

4.3 Ablation studies

Ablation studies として、以下の条件を定めた.

- $V_{t-k+1:t}$ を入力とした場合の性能を調べるため、 $F_{t-k+1:t}$ を削除した.
- N_F を小さくした場合の性能を調べるため、SFM の transformer 層を削除した.
- N_V を大きくした場合の性能を調べるため、MM の transformer 層を追加した.

表 2 に ablation studies の定量的結果を示す. GMGS は条件 (a) で 0.283 ポイント減少した. この結果から、SFM を導入し画像と物理特徴量の両方を扱うことで、性能が向上したと考えられる. また、条件 (b) と (c) では、各評価尺度のスコアに若干の変動が見られた. 本論文では、各スコアのバランスが取れていたため条件 (d) を提案手法として選択した.

*1 <https://sdo.gsfc.nasa.gov/data/>

*2 <https://wdc.nict.go.jp/IONO/wdc/solarflare/index.html>

表 2: Ablation studies における定量的結果.

Conditions	GMGS↑	TSS _{≥M} ↑	BSS _{≥M} ↑
(a) w/o $F_{t-k+1:t}$	0.220±0.116	0.198±0.371	-1.77±0.225
(b) $(N_V, N_F) = (1, 1)$	0.516±0.089	0.485±0.082	0.052±1.05
(c) $(N_V, N_F) = (2, 2)$	0.563±0.070	0.551±0.123	0.011±0.965
(d) $(N_V, N_F) = (1, 2)$	0.503±0.059	0.530±0.112	0.082±0.974

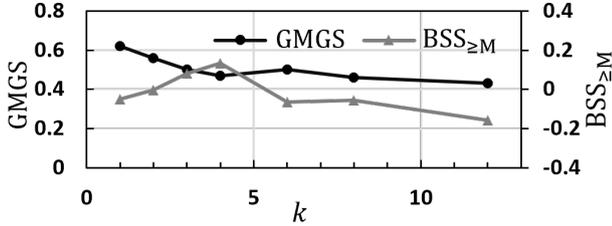


図 5: GMGS および BSS_{≥M} に対する k の影響.

表 3: 2017 年テスト集合における, GMGS_{Influence} によるエラー分析の結果.

観測クラス	予測クラス	GMGS _{Influence}
X	M	0.1335
M	C	0.0885
C	O	0.0578
M	O	0.0442
X	O	0.0114
O	C	0.0093
O	M	0.0030
C	X	0.0019
M	X	0.0018
O	X	0.0015
C	M	0.0012

次に, 過去の画像が予測性能に与える影響を調べるため, 様々な k に対して GMGS および BSS_{≥M} を評価した. ここで, $k = 4$ は $\mathbf{x}_{t-3:t}$ をモデル入力として使用したことを意味する. その結果を図 5 に示す. 図より, GMGS と BSS_{≥M} の値はそれぞれ $k = 1$ と $k = 4$ で最も高くなるという結果が得られた. このことは, GMGS を最大化するためには $k = 1$ で十分であることを示していると考えられる. 一方で, BSS_{≥M} とのバランスをとるためには, 適切な k を選択する必要があることも示している.

4.4 エラー分析

最後に, エラー分析を行う. まず, 誤って予測したサンプルが GMGS に与える影響を以下のように定義する.

$$\text{GMGS}_{\text{Influence}} = \frac{c_{ij}(s_{ii} - s_{ij})}{N}$$

ここで, c_{ij} および s_{ij} はそれぞれ混同行列および GMGS のスコア行列 [Gandin 92] の (i, j) 成分を示す. 表 3 に GMGS_{Influence} に関する結果を示す. 表より, ボトルネックは観測クラスが X クラスのフレアについて, 誤って M クラスのフレアと予測することであった. そのため, このボトルネックを改善することで性能が効率的に向上すると考えられる.

5. おわりに

本論文では, 時刻 t から 24 時間以内に発生する最大の太陽フレアクラス予測モデル, Flare Transformer を提案した. 提案手法による貢献は以下である.

- MM と SFM により, 磁場画像と物理特徴量の両方を扱う太陽フレア予測モデル FT を提案した.
- 入力特徴量間の時間的關係をモデル化するため, transformer 型注意機構を導入した.
- 太陽フレア予測において重要な指標である GMGS と BSS のバランスをとるため, GMGS 損失および BSS 損失を導入した.
- GMGS と TSS について, FT が専門家予測よりも優れた予測を行うことを示した.

謝辞

本研究の一部は, JSPS 科研費 20H04269, JST ムーンショット, NEDO の助成を受けて実施されたものである.

参考文献

- [Bhattacharjee 20] Bhattacharjee, S., Alshehhi, R., Dhuri, D., Alshehhi, R., Dhuri, D., et al.: Supervised Convolutional Neural Networks for Classification of Flaring and Nonflaring Active Regions using Line-of-sight Magnetograms, *The Astrophysical Journal*, Vol. 898, No. 2, p. 98 (12pp) (2020)
- [Gandin 92] Gandin, L. and Murphy, A.: Equitable Skill Scores for Categorical Forecasts, *Monthly Weather Review*, Vol. 120, No. 2, pp. 361–370 (1992)
- [Kubo 17] Kubo, Y., Den, M., and Ishii, M.: Verification of Operational Solar Flare Forecast: Case of Regional Warning Center Japan, *Journal of Space Weather and Space Climate*, Vol. 7, p. A20 (16pp) (2017)
- [Murray 17] Murray, S., Bingham, S., Sharpe, M., et al.: Flare forecasting at the met office space weather operations centre, *Space Weather*, Vol. 15, No. 4, pp. 577–588 (2017)
- [Nishizuka 18] Nishizuka, N., Sugiura, K., et al.: Deep flare net (DeFN) model for solar flare prediction, *The Astrophysical Journal*, Vol. 858, No. 2, p. 113 (8pp) (2018)
- [Nishizuka 20] Nishizuka, N., Kubo, Y., Sugiura, K., Den, M., et al.: Reliable Probability Forecast of Solar Flares: Deep Flare Net-Reliable (DeFN-R), *The Astrophysical Journal*, Vol. 899, No. 2, p. 150 (8pp) (2020)
- [Park 18] Park, E., Moon, Y.-J., Shin, S., et al.: Application of the Deep Convolutional Neural Network to the Forecast of Solar Flare Occurrence Using Full-disk Solar Magnetograms, *The Astrophysical Journal*, Vol. 869, No. 2, p. 91 (6pp) (2018)
- [Scherrer 12] Scherrer, P., Schou, J., Bush, R., Kosovichev, A., Bogart, R., et al.: The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO), *Solar Physics*, Vol. 275, No. 1, pp. 207–227 (2012)
- [SwissRe 16] SwissRe, : Solar storm; How to calculate insured / reinsured losses?, Space Weather Workshop (2016)
- [Tashman 00] Tashman, L.: Out-of-sample tests of forecasting accuracy: an analysis and review, *International Journal of Forecasting*, Vol. 16, No. 4, pp. 437–450 (2000)