Operational Verification of Deep Learning - Based Solar Flare Forecasts

Son Hoang, Robin Chataut, Chetraj Pandey
Department of Computer Science, Texas Christian University

Introduction

- Solar flares are critical space weather events, requiring reliable and continuously updated forecasting systems based on high-resolution solar observations.
- Recent advances in deep learning have improved flare prediction, but most models remain confined to research settings with limited real-time deployment and operational validation.
- This project develops an operational, near-real-time forecasting system, integrating automated data ingestion, deep learning inference, and continuous performance evaluation within an interactive application.

Data and Model

- We utilize full-disk line-of-sight magnetogram images from the Solar Dynamics Observatory (SDO), retrieved in near real-time through the Helioviewer API [1]. The system automatically downloads JP2 images, converts them to JPG format, and resizes them to a standardized resolution (512 × 512) for model input (Fig. 1.1).

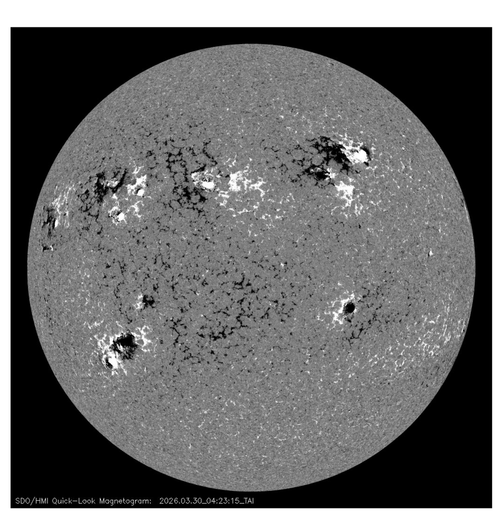


Fig. 1.1: Latest full disk resized 512x512 solar magnetogram [1]

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

$$HSS = 2 \times \frac{TP \times TN - FN \times FP}{((P \times (FN+TN) + (TP+FP) \times N))},$$

where $N = TN + FP$ and $P = TP + FN$.

Fig. 1.2: Formula for TSS and HSS [2]

- A custom AlexNet-based CNN [3,4] processes grayscale magnetogram images to produce softmax probabilities, enabling binary flare prediction (Flare vs. No Flare) within a 12-hour forecast window. These predictions then flow through an automated pipeline—ingestion, preprocessing, inference, and logging—before being evaluated against observed events using TSS and HSS (Fig. 1.2), supporting continuous monitoring and drift detection via an interactive Streamlit interface.

Methodology: Operational Pipeline and Forecast Evaluation

- To ensure robust and consistent predictions, we implement an automated pipeline that retrieves solar magnetogram images, applies preprocessing (format conversion and resizing), and standardizes inputs for model inference (Fig. 2.1).
- For each processed image, the model generates probabilistic flare predictions (Fig. 2.2), which are stored with timestamps and forecast windows to enable structured tracking of prediction outcomes over time.
- We evaluate model performance by aligning predictions with observed flare events [5] and computing skill metrics (TSS, HSS) [2].

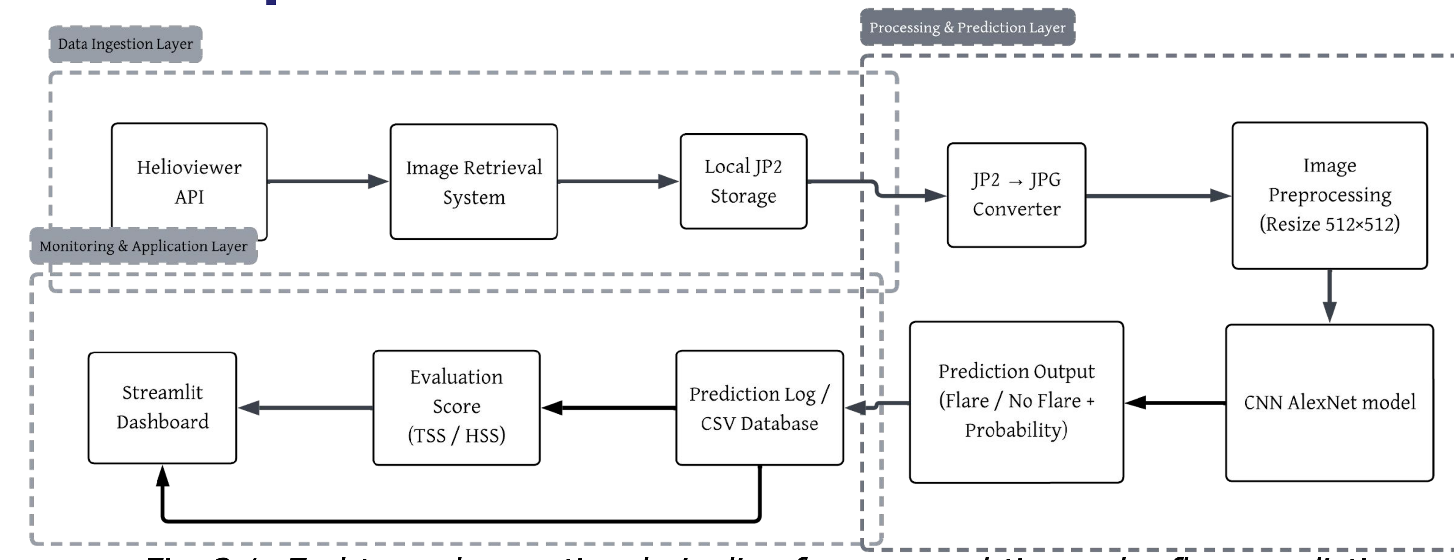


Fig. 2.1: End-to-end operational pipeline for near-real-time solar flare prediction

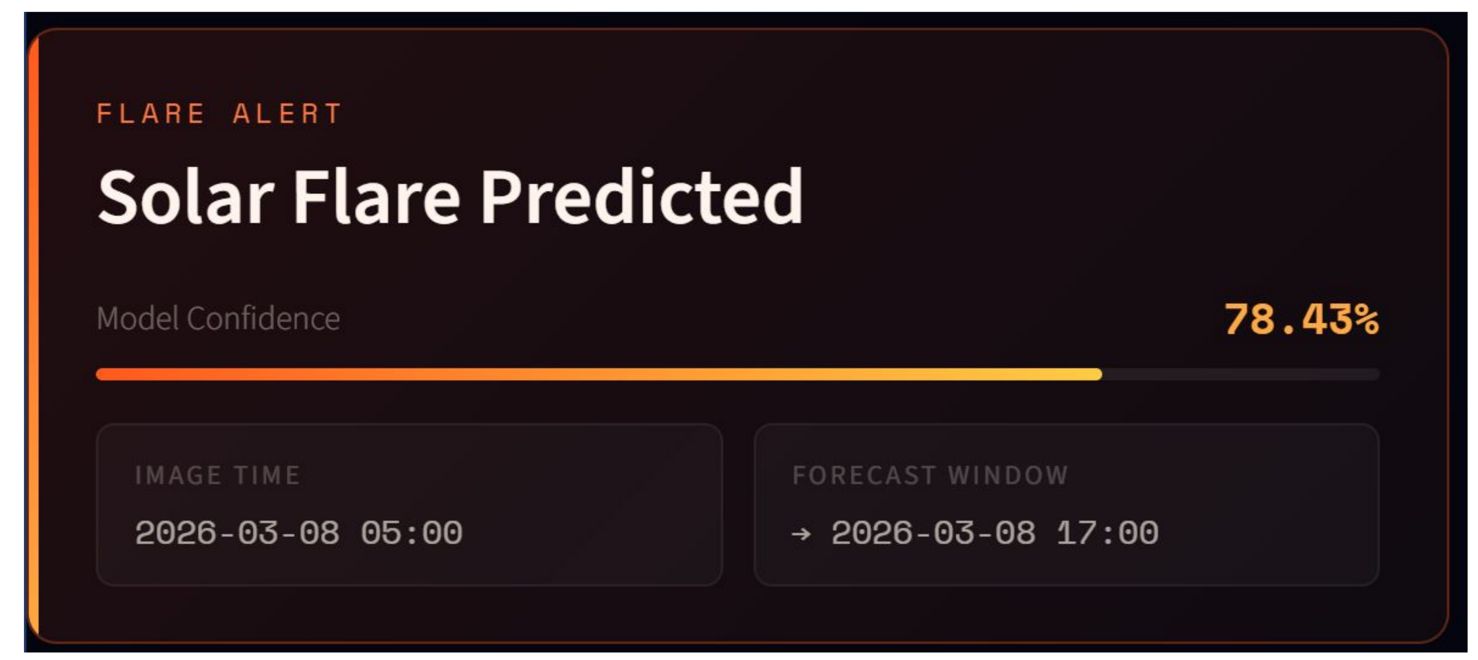


Fig. 2.2: Prediction probability associated with the most recent input image

Application Interface and Deployment Pipeline

- The dashboard presents real-time flare predictions, recent forecast history (Fig. 3.1), observed GOES flare events (Fig. 3.2), confusion matrix outcomes (Fig. 3.3), and evaluation metrics (TSS, HSS) (Fig. 3.4).
- We observe that among 846 predictions, the model identifies 189 (22.3%) true positives and 291 (34.4%) true negatives, while producing 336 (39.7%) false positives and 30 (3.5%) false negatives; although it captures most flare events (POD = 0.8630), the high false alarm rate (FAR = 0.5359) results in moderate skill (Fig. 3.3). Due to API maintenance, last week's evaluation is unavailable; however, performance improves over longer windows, with TSS increasing from +0.1686 (1 month) to +0.3271 (since Feb 1) and HSS from +0.0747 to +0.2249 (Fig. 3.4).

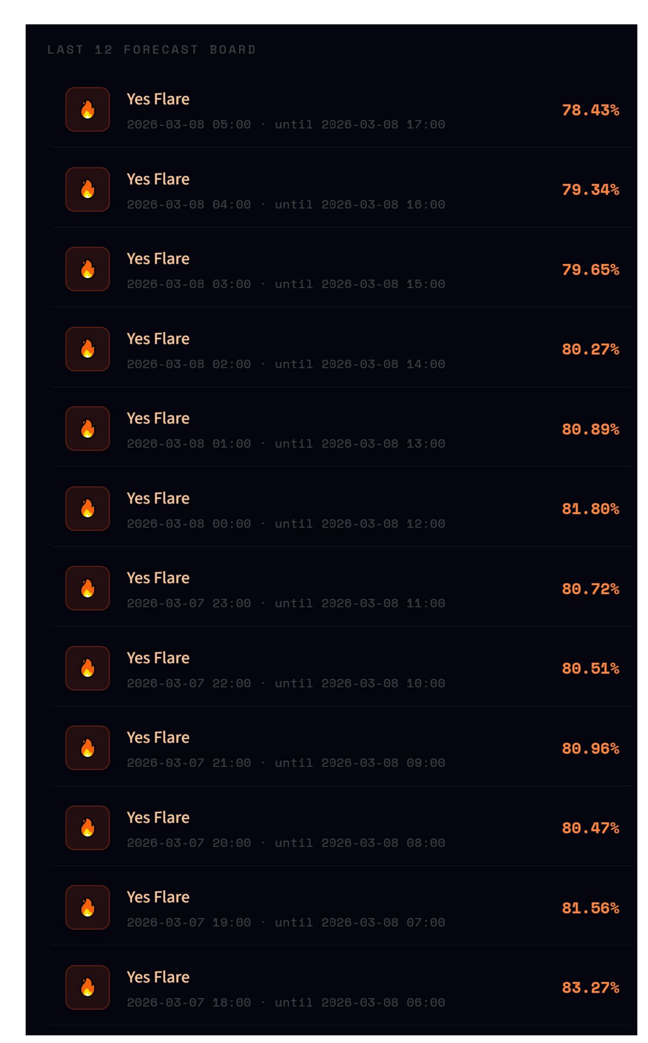


Fig. 3.1: Last 12 model predictions with associated probabilities.

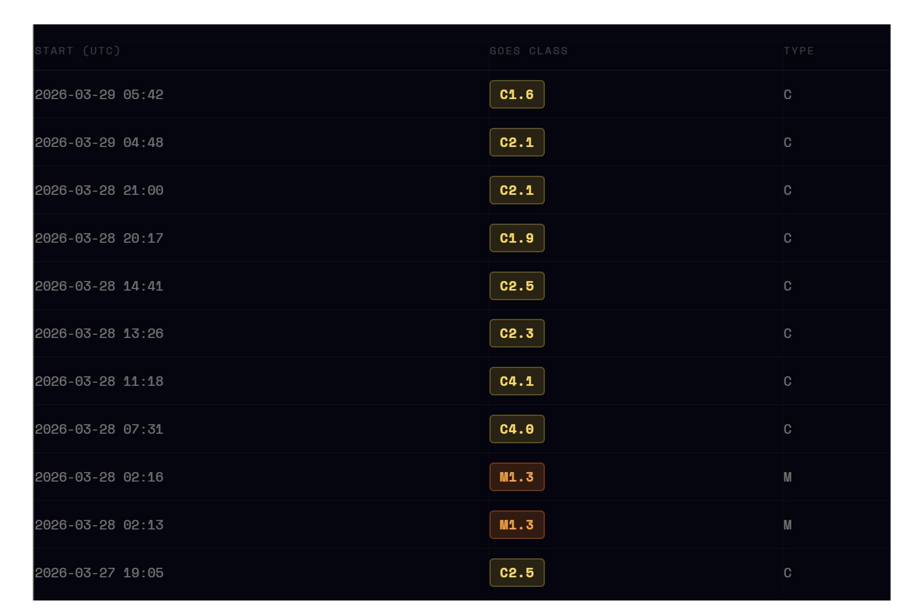


Fig. 3.2: Recent GOES-classified flare events are utilized as ground-truth observations.

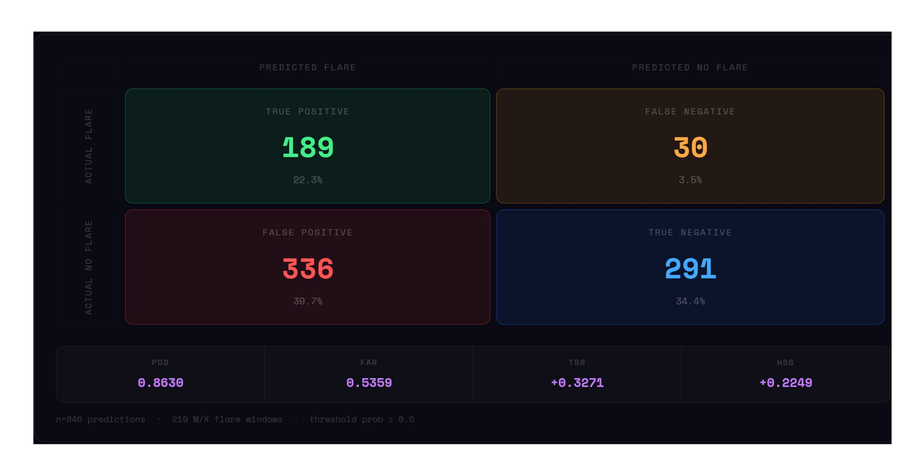
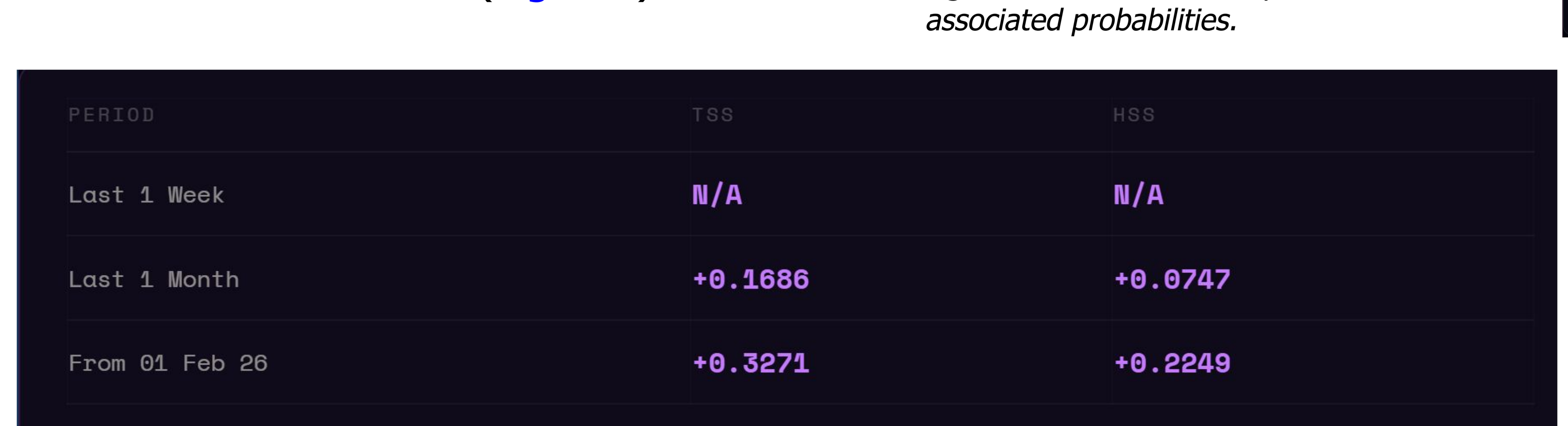


Fig. 3.3: Confusion matrix summarizing model performance: true positives, false positives, false negatives, and true negatives.



PERIOD	TSS	HSS
Last 1 Week	N/A	N/A
Last 1 Month	+0.1686	+0.0747
From 01 Feb 26	+0.3271	+0.2249

Fig. 3.4: Evaluation metrics (TSS and HSS) across different time windows.

Remarks and Future Directions

- Our results indicate that the operational model is capable of generating continuous, near-real-time flare forecasts, with performance improving over longer evaluation windows despite the inherent rarity of M/X-class events.
- These findings suggest that the deep learning model captures meaningful flare-related signals but remains sensitive to noisy solar features, leading to false positives and highlighting the challenge of threshold selection in rare-event prediction.
- Overall, the system provides an end-to-end framework for real-time solar flare prediction, integrating data acquisition, model inference, and continuous performance evaluation within a unified pipeline.
- **Future directions** include improving model robustness through threshold tuning and class imbalance handling, enhancing API robustness & fallback handling, and extending evaluation across longer time horizons and alternative modeling approaches.

References

1. Stanford University, Joint Science Operations Center (JSOC), "HMI Latest Data," [Online]. Available: https://jsoc1.stanford.edu/hmi_latest
2. C. Pandey, A. Ji, T. Nandakumar, and R. A. Angryk, "Exploring Deep Learning for Full-disk Solar Flare Prediction with Empirical Insights from Guided Grad-CAM Explanations," IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA), 2023, doi: 10.1109/DSAA60987.2023.10302639.
3. A. Alex Krizhevsky, I. Ilya Sutskever, and G. E. Geoffrey Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NeurIPS), 2012.
4. A. Krizhevsky, "One Weird Trick for Parallelizing Convolutional Neural Networks," 2014.
5. Lockheed Martin Solar and Astrophysics Laboratory (LMSAL), "Latest Events Archive," [Online]. Available: https://www.lmsal.com/solarsoft/latest_events_archive.html

Acknowledgements

This work is supported by the Department of Computer Science at Texas Christian University, Fort Worth, Texas. The data used in this study is courtesy of NASA/SDO, the AIA, EVE, and HMI science teams and NOAA (NGDC).

Contact Information

Son Hoang
Undergraduate, Data Science major
Department of Computer Science
Davis College of Science and Engineering
Texas Christian University
Email: son.t.hoang@tcu.edu
Website: https://github.com/hoangthaison2306/