Stellar flare detection methods in TESS data: application and performance study (postprint)

Authors: Min Li, Liang Wang, Zhiqiang Zou, Ali Luo, Bo Qiu, Peng Jia, Ying Shan, Liang Wang, Zhiqiang Zou

Date: 2025-12-03T00:00:00+00:00

Abstract

The detection of stellar flares is crucial to understanding dynamic processes at the stellar surface and their potential impact on surrounding exoplanetary systems. Extensive time series data acquired by the Transiting Exoplanet Survey Satellite (TESS) offer valuable opportunities for large-scale flare studies. A variety of methods is currently employed for flare detection, with machine learning (ML) approaches demonstrating strong potential for automated classification tasks, particularly for the analysis of astronomical time series. This review provides an overview of the methods used to detect stellar flares in TESS data and evaluates their performance and effectiveness. It includes our assessment of both traditional detection techniques and more recent methods, such as ML algorithms, highlighting their strengths and limitations. By addressing current challenges and identifying promising approaches, this manuscript aims to support further studies and promote the development of stellar flare research.

Full Text

Preamble

Astronomical Techniques and Instruments, Vol. 2, September 2025, 310–318
Review
Open Access

Stellar flare detection methods in TESS data: application and performance study

Min Li[1,2], Liang Wang[4,5,6], *Zhiqiang Zou*[1,2,3], Ali Luo[3,6,7], Bo Qiu[8], Peng Jia[9], Ying Shan[1,2]

[1]College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[2]Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing 210023, China
[3]University of Chinese Academy of Sciences, Nanjing 211135, China
[4]Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences, Nanjing 210042, China
[5]CAS Key Laboratory of Astronomical Optics & Technology, Nanjing Institute of Astronomical Optics & Technology, Nanjing 210042, China
[6]University of Chinese Academy of Sciences, Beijing 100049, China
[7]CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Beijing 100101, China
[8]School of Intelligent Science and Technology, University of Science and Technology Beijing, Beijing 100083, China
[9]College of Electronic Information and Optical Engineering, Taiyuan University of Technology, Taiyuan 030024, China

*Correspondences: liangwang@niaot.ac.cn; zouzq@njupt.edu.cn

Received: March 12, 2025; Accepted: April 10, 2025; Published Online: May 16, 2025
https://doi.org/10.61977/ati2025028; https://cstr.cn/32083.14.ati2025028

Citation: Li, M., Wang, L., Zou, Z. Q., et al. 2025. Stellar flare detection methods in TESS data: application and performance study. *Astronomical Techniques and Instruments*, 2(5): 310–318. https://doi.org/10.61977/ati2025028.

Abstract

The detection of stellar flares is crucial to understanding dynamic processes at the stellar surface and their potential impact on surrounding exoplanetary systems. Extensive time series data acquired by the Transiting Exoplanet Survey Satellite (TESS) offer valuable opportunities for large-scale flare studies. A variety of methods is currently employed for flare detection, with machine learning (ML) approaches demonstrating strong potential for automated classification tasks, particularly for the analysis of astronomical time series. This review provides an overview of the methods used to detect stellar flares in TESS data and evaluates their performance and effectiveness. It includes our assessment of both traditional detection techniques and more recent methods, such as ML algorithms, highlighting their strengths and limitations. By addressing current challenges and identifying promising approaches, this manuscript aims to support further studies and promote the development of stellar flare research.

Keywords: Stellar flare detection; TESS light curve; ML; Automatic classification

1. INTRODUCTION

Stellar flares are explosive phenomena at the surface and in the atmosphere of stars [?]. They are caused by a release of energy induced by the star' s magnetic field and are a very common manifestation of stellar activity. Flares are most common in late-type M dwarfs [?, ?], but they also occur frequently in main-sequence stars and, more rarely, evolved giant stars [?]. The study of stellar flares is essential to better understand stellar evolution and planetary habitability [?]. For example, intense stellar flare activity can strongly affect the surrounding planets [?, ?], notably by heating their atmosphere [?, ?] or causing ionospheric disturbances [?, ?].

Research on stellar flares has traditionally relied on observational data collected by ground-based and space-borne telescopes [?]. Recently, the development of larger-scale space missions, such as the TESS [?], has considerably transformed flare studies. To date, TESS has provided an extensive, high-quality dataset that has already markedly accelerated progress in stellar flare research and improved our understanding of flare properties, energy release mechanisms, and role in stellar activity cycles.

Flare detection methods can be categorized into traditional and ML approaches. Traditional methods typically involve detrending and statistical analysis of stellar observations. For example, the WARPFINDER algorithm developed by Pietras et al. [?] employs a three-step approach (detrending, differencing, and curve fitting) to detect flare events automatically. Although effective, such methods frequently require manual parameter adjustment and cannot process large datasets efficiently.

ML methods have been increasingly applied to the analysis of astronomical datasets, including TESS data. Recent advances include the application of convolutional neural networks (CNNs) such as Stella, developed by Feinstein et al. [?], to study the relationship between stellar properties and flare activity; recurrent neural networks (RNNs) by Vida et al. [?] for flare detection; and CNNs with ensemble learning by Tu et al. [?] to identify superflares (stellar burst phenomena with longer duration and larger energy than typical stellar flares) in TESS data.

These models, trained on large volumes of TESS data, have enhanced detection efficiency and minimized the need for manual intervention. ML methods provide robust and systematic solutions to process large observational datasets, enabling automatic identification and classification of flare events.

In this review, we provide a comprehensive overview of stellar flare detection methods from TESS data that covers traditional techniques and ML approaches.

We evaluate the advantages and limitations of each method, focusing on accuracy and computational efficiency. Additionally, we discuss the challenges and opportunities of applying ML to stellar flare detection. The article structure is as follows: in Section 2, we introduce TESS data; in Section 3, we describe traditional and ML flare detection methods; in Section 4, we present experimental results and evaluate method performance; in Section 5, we discuss the challenges of applying ML to astronomical research; and in Section 6, we summarize flare detection methods and suggest future research directions.

2.1. TESS Mission and Scientific Objectives

The primary objective of the TESS mission [?] is the search for exoplanets, especially small telluric planets located in the habitable zone of their star. These planets are identified with the transit method, which consists of observing a slight decrease in stellar brightness caused by a planet passing in front of its star. Additionally, the high precision of TESS data is well-suited to study stellar physics. By analyzing light curves (LCs), scientists can study stellar properties such as rotation, pulsation, and magnetic activity that provide a basis to characterize stellar structure and evolution.

During its primary mission, TESS conducted a two-year survey of the solar neighborhood. On July 4, 2020, its primary mission ended; TESS is now in its extended mission. The planets discovered by TESS range from small telluric planets to gas giants, clearly demonstrating the diversity of planets in our galaxy. These exoplanets and, additionally, active stars identified by TESS during its primary mission have been defined as priority observation targets for subsequent missions, such as the James Webb Space Telescope [?], to improve our understanding of planetary atmospheres and stellar activity. During its primary mission (Section 2.2), TESS imaged approximately 75% of the sky, discovered 623 confirmed exoplanets, and listed nearly 7,643 additional candidates, which are currently awaiting confirmation by independent instruments or methods [?].

2.2. TESS Data Acquisition and Processing

TESS is equipped with four wide-field cameras, which are fixed to the satellite platform. To achieve pointing toward different regions of the sky, the entire spacecraft is rotated. During its primary mission phase, TESS executed 26 distinct pointings, with each pointing maintained for approximately two orbital cycles. By the end of the primary mission, the full southern and northern celestial hemispheres had been surveyed.

To maximize sky coverage, these pointings were evenly distributed along the ecliptic longitude. Each individual camera provides a field of view of $24° \times$

24°, and the four cameras together form a combined synthetic field of view of 24° × 96°, covering approximately 2,300 square degrees. One of the cameras is directed toward the ecliptic pole, while the remaining three are spaced at equal intervals between the ecliptic pole and an ecliptic latitude of 6°, thereby optimizing overall sky coverage.

During operation, each camera nominally acquires an image every 2 s. To optimize storage and processing time, single images are accumulated onboard by sets of 60 to generate composite images with equivalent exposure times of 2 min that are stored in the onboard solid-state buffer (SSB). The 2-min equivalent images, or "postage stamps", focus on areas surrounding the target source; their usual size is 10 × 10 pixels. Their main use is to characterize light variability near a target for subsequent analysis. Then, each image is analyzed by aperture photometry to generate a flux array (the LC), ultimately stored in data files labeled "LC files" [?]. An example of generated photometric data is shown in Fig. 1 [Figure 1: see original paper]. In addition, aggregated FFIs are collected every 30 min and also stored in the SSB for larger-scale light variation analysis and for stellar research. Although it covers a considerably larger field of view than the postage stamps, each FFI remains highly sensitive to small signal variations, providing high-resolution, high-sensitivity, large-scale data for exoplanet and stellar research.

Every 13.7 days, from 8 h before to 8 h after perigee, data stored in the SSB, including all postage stamps and FFIs, are transmitted to the ground segment through the National Aeronautics and Space Administration Deep Space Network [?].

2.3. Characteristics of TESS LCs

Two types of LC data are acquired by TESS for its target stars (440,000 targets as of February 2023) at intervals of 2 min (standard LCs) and 20 s (fast-LCs) [?]. Because the number of target stars with available fast-LCs is small, we restrict our review to the TESS 2-min LC data. Their characteristics are described in this section.

Standard LCs of some target stars exhibit gradual brightness increases or decreases that can be attributed to slow alterations of the internal structure or physical state of these stars. In other cases, small amplitude fluctuations can be discerned in the LCs. These fluctuations may be associated with small-scale activity at the star' s surface or with star vibrations. Conversely, sporadic events, such as stellar flares or planetary transits, induce sudden brightness changes materialized by peaks or troughs in the LC. Such events provide essential data on stellar atmospheric activity and magnetic field strength.

In addition, periodic brightness variations are frequently observed in the LCs of TESS targets; they usually reflect the star's rotation, surface activity, or possible planetary transits. The star's rotation period can be inferred from these periodic LC patterns. The amplitude of the brightness variations reflects the intensity of stellar activity, especially if they are caused by magnetic activity or stellar rotation. Stellar spots on the star' s surface can also affect the amplitude of the fluctuations.

In TESS data, flare events can be identified by the shape of the LC features: they are usually short and characterized by a rapid and steep brightness increase followed by a slow decrease, forming a typical "spike-shaped" peak, as illustrated in Fig. 2 [Figure 2: see original paper], with a maximum brightness largely exceeding background brightness fluctuations. In some active stars, multiple flare "bursts" can be observed, appearing as independent peaks in the LC. These multiple bursts are important to study the periodicity of stellar magnetic activity.

---

3. FLARE DETECTION METHODS

In this section, we present an overview of stellar flare detection methods, which are broadly categorized into traditional threshold-based approaches and ML techniques.

3.1. Traditional Threshold Detection Methods

In traditional threshold detection methods, candidate flare events are identified by LC brightness variations exceeding a preset statistical threshold (e.g., 2.5 standard deviations or higher) [**?**, **?**]. Here, the standard deviation (σ) represents the statistical dispersion of stellar brightness measurements in an LC when the observed star is in its quiescent, non-flaring state. It quantifies the variability in the star brightness induced by instrumental noise, stellar activity, or other non-flare-related fluctuations. Mathematically, if F_i represents the brightness measurements and $\bar{F}$ is the mean quiescent flux, σ is given by

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - \bar{F})^2} \tag{1}$$

The flare start and end times are determined from the set of consecutive data points exceeding the preset threshold \cite{25-32}. Standard steps of threshold detection methods include LC detrending, noise removal, threshold filtering, and manual validation.

3.1.1. Detrending methods

The objective of detrending is to eliminate long-term variations in the LC, such as stellar spot modulation and instrument drift, while preserving short-term flare signals [?]. To achieve this, common techniques such as polynomial fitting, spline smoothing, and iterative denoising are employed [?]. Polynomial fitting uses low-order polynomials (e.g., cubic polynomials) to model the long-term LC trend. This trend is then subtracted from the original data. This effectively removes low-frequency variations from the LC [?]. Spline smoothing and Gaussian filtering are two common techniques used to smooth LCs, aiming to reduce low-frequency noise while preserving short-term flare signals [?]. Finally, iterative denoising removes outliers or large flares through multiple iterations and gradually converges toward a stable, detrended LC [?].

In summary, detrending methods are effective in removing undesired, long-term brightness variations from observed LCs, thereby enhancing the observation sensitivity for flare signal detection. However, such methods also present limitations: for example, overfitting might occur, potentially resulting in the misidentification of large flares as long-term trends; moreover, their overall effectiveness strongly depends on the selection of appropriate parameters.

3.1.2. Difference methods

Different methods use a threshold detection approach relying on brightness differences to identify flare events [?, ?]. Standard difference calculation methods include "running differences" and "resting model" differences. The running difference method calculates flux differences between temporally adjacent LC points and compares them with a predefined threshold. Differences that exceed the threshold are then marked as flare candidates [?]. The resting model difference method globally subtracts a "resting model" (e.g., the detrended LC) from the original LC. This process results in a different LC that can effectively capture rapid brightness changes unrelated to long-term trends [?]. Different methods are particularly effective in identifying sudden brightness increases during flares, hence to detect flares on short timescales. Additionally, they are computationally simple and can be easily parallelized. However, such methods are sensitive to noise, which might cause misidentification of cosmic rays or short-term instrumental errors as flares. Therefore, combining different methods with detrending techniques is essential to reduce the likelihood of erroneously interpreting cosmic rays or transient instrumental effects as real flaring events.

3.2. ML Methods

The core idea in applying ML methods to stellar flare detection is to apply data-driven models automatically to "learn" stellar flare features, thereby enhancing the model' s detection accuracy and generalization ability. A crucial first step in this process is data preprocessing, which often includes LC normalization [?, ?]—a technique that rescales the data to a common baseline to reduce observational noise. Additionally, the Synthetic Minority Over-sampling Technique (SMOTE) [?] is commonly applied to address class imbalance, a situation where flare (rare events) images are considerably outnumbered by non-flare images, by synthetically generating more examples of the minority image class (flares). In terms of feature extraction, some methods use manual feature definition, e.g., flare duration, peak amplitude, or equivalent duration [?], while others rely on automatic extraction of time or frequency-domain features, for example, using CNNs [?] for pixel-level data [?]. For model selection, commonly used supervised learning methods include random forests (RFs) [?, ?] and extreme gradient boosting (XGBoost) [?], while deep learning approaches include CNNs [?], RNNs \cite{48-50}, and deep neural networks (DNNs) \cite{51-53}.

Finally, model ensemble techniques [?] improve detection stability and accuracy through multi-algorithm voting (i.e., several models are applied to the same data and detection is confirmed if at least two models concurrently identify a flare event) or enhance the robustness of the detection system through stacking [?] or weighted fusion [?]. A combination of ML techniques can thus provide a more reliable and comprehensive solution than traditional methods for flare detection.

---

4. PERFORMANCE EVALUATION

4.1. Performance Evaluation Metrics

In ML methods, the stellar flare detection task can be formulated as a binary classification problem, because its primary objective is to differentiate between two distinct states: "flare" and "non-flare." The confusion matrix is an essential tool to evaluate model performance for such binary classification tasks; it provides detailed characterization of the model' s predictive abilities using complementary metrics [?]. As illustrated in Table 1 , the confusion matrix is a 2 $\times$ 2 table that includes the four possible outcomes of binary classification: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

Table 1. Confusion matrix

	Actual true	Actual false
Predicted true	TP	FP
Predicted false	FN	TN

In this study, we use the confusion matrix to quantify the models' predictive accuracy for flare detection by calculating key performance metrics: accuracy, recall, and precision.

(1) Accuracy

Accuracy is the proportion of correctly predicted flare events among the total sample set [?],

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

For stellar flare detection, high accuracy indicates that the model correctly identifies flare events, with fewer FPs.

(2) Recall

Recall describes the model' s ability to identify positive samples, i.e., the proportion of detected flare events relatively to the total number of real flare events,

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

High recall indicates that the model can effectively identify a large proportion of real flare events and avoids FNs [?].

(3) Precision

Precision is the proportion of correctly predicted positive samples among all positive predictions,

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

It measures the "positive prediction accuracy" of the model [?].

4.2. Comparison Between Traditional and ML Methods

We designed an experiment to evaluate and compare the stellar flare detection performance of traditional threshold detection methods and ML methods. The dataset used for this experiment consisted of the TESS 2-min data for M dwarf stars. To ensure data comparability under identical evaluation conditions, we preprocessed LC data by applying normalization, binary segmentation into flare and non-flare samples, and annotation of the corresponding labels.

Specifically, for standard threshold methods, we used the approach proposed by Yang et al. [?] as a detrending method and the brightness difference detection technique introduced by Shibayama et al. [?] as a different method. For ML, we used a CNN developed by Feinstein et al. [?].

Although the implementation of the traditional threshold detection and ML methods differs substantially, their outputs are consistent: both types of methods predict a binary class label ("flare" or "non-flare") for each sample. This consistency allows for direct comparison between the results of both approaches and the true labels and for calculation of the accuracy, precision, and recall. We also compared the relative computation time required by all methods to assess their practical efficiency.

The experimental results (Fig. 3 [Figure 3: see original paper]) demonstrated that, in terms of accuracy, precision, and recall, ML methods clearly outperformed traditional methods, illustrating the strong ability of CNNs to handle complex LC data. Although a single CNN architecture was tested in our analysis, its performance suggests that all deep learning models, particularly those with convolutional structures, are well suited to capture local and hierarchical features in LC data. This result is consistent with prior studies that reported similar advantages of CNNs for time-series and astrophysical signal classification tasks, indicating the potential generalizability of our results. In our experiment, traditional methods—particularly the difference method—performed better in terms of detection speed, mostly because of the computational efficiency of simple brightness difference calculations. In contrast, the detrending and ML approaches required additional processing steps such as polynomial fitting and model training, resulting in higher computational costs. The traditional methods evaluated here represent commonly used techniques reported in the literature; although they do not include all existing traditional approaches, they provide a practical comparison framework. Despite their processing speed advantage, these methods exhibited lower accuracy. The considerable improvement in detection performance achieved by our CNN-based model highlights the higher suitability of ML techniques—especially deep learning models—for reliable stellar flare identification in complex LC data.

With this experiment, we validated the effectiveness of ML methods for stellar flare detection while highlighting the computational advantages of traditional methods. Future directions of research can include investigating combined approaches to develop more efficient and accurate flare detection algorithms.

4.3. ML Performance

For stellar flare detection tasks, the performance variability of ML algorithms evaluated using accuracy, recall, and precision reflects their strengths and limitations for data feature extraction, temporal information capture, and overall detection accuracy. Table 2 summarizes the performance of well-established ML algorithms (RF, XGBoost) and deep learning algorithms (CNN, RNN, DNN) in terms of accuracy, recall and precision.

Table 2. Reported performance of established ML methods

Method	Application datasets	Accuracy/(%)	Recall/(%)	Precision/(%)
Feinstein et al. (CNN) [?]	TESS LCs	97	97	92
Tu et al. (CNN) [?]	TESS pixel-level images	96	96	91
Vida et al. (RNN) [?]	Kepler and LCs combined with simulated data	80	80	70
Tu et al. (ensemble deep learning, Stacking) [?]	TESS pixel-level images	97	96	96
Tu et al. (ensemble deep learning, Voting) [?]	TESS pixel-level images	99	99	99
Lin et al. (DNN) [?]	TESS LCs	95	96	94
Lin et al. (RF) [?]	TESS LCs	97	97	96
Lin et al. (XG-Boost) [?]	TESS LCs	98	97	98

Feinstein et al. [?] and Vida et al. [?] used the LC flux as primary input feature. Feinstein et al. [?] applied CNNs to LC data from TESS, achieving accuracy, recall, and precision values of 97%, 97%, and 92%, respectively; their results indicated that, although the input data consisted only of raw LC flux data, their CNNs effectively captured flare patterns with high predictive ability. Similarly, Vida et al. [?] applied RNNs to TESS LC data and achieved accuracy, recall, and precision values of 80%, 80%, and 70%, respectively. Compared with the results of Feinstein et al. [?], the RNN performance was notably lower, possibly because of the model complexity and limitations in handling time-series data, particularly for flare detection over long timescales.

In contrast, Tu et al. [?] used pixel-level TESS image data as input features for automatic feature extraction. Their CNN achieved accuracy, recall, and precision values of 96%, 96%, and 91%, respectively, demonstrating the advantage of CNNs for image data processing, with high efficiency for direct feature extraction and flare identification. Moreover, Tu et al. [?] further improved model performance by applying ensemble learning methods. Using the stacking method, they achieved accuracy, recall, and precision values of 97%, 96%, and 96%, respectively. This ensemble learning method enhances overall performance by combining the strengths of multiple models. With the ensemble learning voting method, they achieved their best results, with equally high accuracy, recall, and precision values of 99%. This indicates that combining predictions from multiple models strongly improves flare identification accuracy.

Lin et al. [?] used LC flare features manually extracted from TESS data, including total flare duration, shock phase duration, decay phase duration, equivalent duration, and peak amplitude, as input features for a DNN and achieved accuracy, recall, and precision values of 95%, 96%, and 94%, respectively; this indicates that combining manually computed features with deep learning models can also yield high identification performance. Additionally, they evaluated RF and XGBoost algorithms and achieved accuracy, recall, and precision values of 97%, 97%, and 96%, respectively, for RF and 98%, 97%, and 98%, respectively, for XGBoost. The XGBoost algorithm exhibited outstanding (and the best overall) performance in handling data with structured features, particularly for such classification tasks.

In summary, the ensemble learning voting method of Tu et al. [?], relying on image data and combining predictions from multiple models, achieved highest performance, with strongly enhanced flare detection accuracy. In contrast, deep learning models using only LC flux features, such as those of Feinstein et al. [?] and Vida et al. [?], achieved comparatively poor performance, likely explained by known limitations of RNNs in capturing complex patterns. Finally, Lin et al. [?] demonstrated that a combination of manually extracted features and established ML methods could also yield excellent performance.

---

5. DISCUSSION

Traditional methods for stellar flare detection primarily rely on threshold techniques and LC detrending. Albeit simple and efficient, such methods are susceptible to noise interference. Astrophysical phenomena (such as star spots) or instrumental noise may produce FPs [?]. Low-amplitude flares may not reach the detection threshold and produce FNs [?].

With recent advances in ML and deep learning, methods relying on feature engineering and data balancing have markedly improved flare detection accuracy. For example, Lin et al. [?], combining RF and XGBoost algorithms with SMOTE oversampling, achieved a flare recovery rate of 92% and additionally detected 2,000 small flare events [?]. Deep learning methods, such as CNNs or "long short-term memory" neural networks, have further expanded model capabilities and overcome the limitations of traditional methods [?, ?, ?]. Finally, ensemble learning methods have provided, to date, the largest accuracy improvement for superflare detection, enhancing pixel-level classification accuracy to 99% [?].

Despite the considerable potential of ML methods for stellar flare detection, there remain several challenges and limitations, described hereafter.

Data scarcity and imbalance: Stellar flare data represent a very small proportion of the large observational datasets currently available, resulting in sample scarcity in the ML model training data, particularly for low-energy flares. This sample imbalance can affect the model' s generalization performance, with good detection results for high-energy flares but difficult detection of low-energy events [**?**, **?**].

High rate of FP: LCs are affected by noise and interfering astronomical phenomena (such as brightness peaks from pulsating stars) that can be misclassified as flares. Even after data processing, such interferences can still affect deep learning models during flare detection, increasing the rate of FPs. This problem particularly affects large observational datasets with a substantial number of LCs [**?**].

Long processing time and high computational cost: The computational cost of deep learning models, especially long short-term memory neural networks and CNNs, is high when processing large-scale LC data because of the time-consuming, but necessary, training and debugging, especially for parameter adjustment and hyperparameter optimization [**?**].

Noise interference and instrumental errors: The CCD sensors and data processing chains of spaceborne instruments introduce noise in the observation data, occasionally perceived erroneously as flare signals. Although ML models can be trained to filter this noise, the filtering is frequently incomplete and can cause misclassification of noise features as flares [**?**].

Challenges in cross-task data transfer: Observational data properties differ between space telescopes (such as Kepler and TESS), e.g., in terms of sampling rates and wavelength ranges. Therefore, a model trained on a specific dataset, if applied to another dataset, frequently exhibits limited performance and requires adjustment or retraining to adapt to the new data [**?**].

Subjectivity in data labeling: Stellar flare event labeling typically relies on visual inspection. However, this manual approach can be inconsistent, particularly when LCs exhibit low signal-to-noise ratios, which hampers human judgment. In such cases, flare labels may be assigned to ambiguous features, effectively introducing incompletely characterized labels—referred to as "noisy labels" —into the training dataset. These noisy labels can negatively affect the learning process and limit the model' s generalization ability [**?**].

These challenges reflect both the potential and the limitations of ML methods for stellar flare detection; they also indicate new research directions to improve models and data processing methods.

6. CONCLUSION

Stellar flare detection is an important astronomical research topic, investigated with either traditional or ML methods whose suitability depends on the application.

Before stellar flare catalogs became available or sufficiently extensive, early flare detection methods relied primarily on visual identification. Such traditional methods use mathematical, unlabeled approaches; they are time-efficient and not labor-intensive. However, despite their simplicity and efficiency, their performance is considerably influenced by the accuracy of their initial settings (detection threshold and detrending). ML can be used to optimize parameter selection (e.g., dynamic threshold adjustment) and improve detection sensitivity.

Because of their pattern recognition properties, ML methods can efficiently identify complex flare shapes and partly alleviate misclassification caused by human judgment errors. For example, they have demonstrably outperformed traditional methods for the processing of large datasets, such as those from TESS or Kepler. However, their practical application still presents substantial challenges. Model training relies on a large number of high-quality labeled samples; consequently, labeling of stellar flare events requires large volumes of observational data and expert cross-validation, which increases data acquisition costs.

Therefore, to further generalize the application of ML techniques and enhance the accuracy of stellar flare detection, we propose the following research directions:

(1) Intelligent transformation of detection methods

Empirical threshold-based detection criteria are progressively replaced by ML algorithms, such as CNNs or XGBoost models, that improve identification accuracy and enhance generalization capabilities to weaker flares through autonomous learning of complex LC features, concurrently allowing for the automated processing and systematic analysis of extensive datasets.

(2) Multisource data fusion and collaborative detection

By combining high-quality data from multiple sources, e.g., from TESS for its high-precision photometry data, Kepler for its long-term monitoring, and XMM-Newton for its multiwavelength X-ray data [**?**], researchers can improve the signal-to-noise ratio of flare signals through cross-band feature correlation. Furthermore, multitask collaborative analysis frameworks allow for simultaneous investigation of flare activity, stellar magnetic fields, and stellar spot evolution, thereby maximizing the scientific output.

(3) High temporal resolution for dynamic process analysis

Current telescopes achieve very short exposure times, e.g., image acquisition by TESS at a 20-s cadence [**?**]; consequently, substructures of the stellar flare energy release can be resolved by analyzing these short-timescale data. Combined with rapid-response spectroscopic and multiwavelength observations, this technological advance has been essential to characterize the temporal evolution of microphysical processes such as magnetic reconnection and particle acceleration, with essential theoretical breakthroughs on stellar eruption mechanisms.

Pursuing research on these combined topics will potentially induce considerable progress in stellar flare research and strongly expand our understanding of energy interactions between stars and their planetary systems.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (12473104 and U2031144).

AI DISCLOSURE STATEMENT

AI-assisted tools such as StarWhisper and ChatGPT (OpenAI) were employed to translate the original Chinese text into English and refine the linguistic quality of this section. These tools were used solely for language improvement and did not contribute to data analysis, interpretation, or the formulation of scientific conclusions. The authors carefully reviewed, edited, and revised the AI-generated texts to their own preferences, assuming ultimate responsibility for the content of the publication.

AUTHOR CONTRIBUTIONS

Min Li conceived and designed the study, performed the experiments, analyzed the data, and wrote the paper. Liang Wang provided key ideas, background knowledge, and contributed to data interpretation. Ying Shan guided and supported the experimental work. Zhiqiang Zou, Ali Luo, Bo Qiu, and Peng Jia reviewed the content and provided valuable feedback on manuscript writing and academic standards. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Benz, A. O., Güdel, M. 2010. Physical processes in magnetically driven flares on the sun, stars, and young stellar objects. *Annual Review of Astronomy and Astrophysics*, 48(1): 241–287.
2. Walkowicz, L. M., Basri, G., Batalha, N., et al. 2011. White-light flares on cool stars in the Kepler quarter 1 data. *The Astronomical Journal*, 141(2): 50.
3. Vida, K., Kriskovics, L., Oláh, K., et al. 2016. Investigating magnetic activity in very stable stellar magnetic fields-Long-term photometric and spectroscopic study of the fully convective M4 dwarf V374 Pegasi. *Astronomy & Astrophysics*, 590: A11.
4. Oláh, K., Seli, B., Kővári, Z., et al. 2022. Characteristics of flares on giant stars. *Astronomy & Astrophysics*, 668: A101.
5. West, A. A., Hawley, S. L., Bochanski, J. J., et al. 2008. Constraining the age-activity relation for cool stars: the sloan digital sky survey data release 5 low-mass star spectroscopic sample. *The Astronomical Journal*, 135(3): 785.
6. Khodachenko, M. L., Ribas, I., Lammer, H., et al. 2007. Coronal mass ejection (CME) activity of low mass M stars as an important factor for the habitability of terrestrial exoplanets. I. CME impact on expected magnetospheres of Earth-like exoplanets in close-in habitable zones. *Astrobiology*, 7(1): 167–184.
7. Yelle, R., Lammer, H., Ip, W. H. 2008. Aeronomy of extra-solar giant planets. *Space Science Reviews*, 139: 437–451.
8. Vida, K., Kővári, Z., Pál, A., et al. 2017. Frequent flaring in the TRAPPIST-1 system–unsuited for life? *The Astrophysical Journal*, 841(2): 124.
9. Roettenbacher, R. M., Kane, S. R. 2017. The stellar activity of TRAPPIST-1 and consequences for the planetary atmospheres. *The Astrophysical Journal*, 851(2): 77.
10. Scalo, J., Kaltenegger, L., Segura, A., et al. 2007. M stars as targets for terrestrial exoplanet searches and biosignature detection. *Astrobiology*, 7(1): 85–166.
11. Yelle, R. V. 2004. Aeronomy of extra-solar giant planets at small orbital distances. *Icarus*, 170(1): 167–179.
12. Ricker, G. R., Vanderspek, R., Winn, J., et al. 2016. The transiting exoplanet survey satellite. *In Proceedings of SPIE*, 9904: 767–784.
13. Pietras, M., Falewicz, R., Siarkowski, M., et al. 2022. Statistical analysis of stellar flares from the first three years of TESS observations. *The*

Astrophysical Journal, 935(2): 143.

14. Feinstein, A. D., Montet, B. T., Ansdell, M., et al. 2020. Flare statistics for young stars from a convolutional neural network analysis of TESS data. *The Astronomical Journal*, 160(5): 219.
15. Vida, K., Bódi, A., Szklenár, T., et al. 2021. Finding flares in Kepler and TESS data with recurrent deep neural networks. *Astronomy & Astrophysics*, 652: A107.
16. Tu, Z. L., Wu, Q., Wang, W., et al. 2022. Convolutional neural networks for searching superflares from pixel-level data of the transiting exoplanet survey satellite. *The Astrophysical Journal*, 935(2): 90.
17. Kalirai, J. 2018. Scientific discovery with the James Webb space telescope. *Contemporary Physics*, 59(3): 251−290.
18. Guerrero, N., Seager, S., Huang, C. X., et al. 2021. The TESS objects of interest catalog from the TESS prime mission. *The Astrophysical Journal Supplement Series*, 254(2): 39.
19. Huang, C. X., Vanderburg, A., Pal, A., et al. 2020. Photometry of 10 million stars from the first two years of TESS full frame images: part I. *Research Notes of the AAS*, 4(11): 204.
20. Mullally, S. 2018. TESS Archive Manual, in TESS Archive Documentation Center, Space Telescope Science Institute. Available from https://outerspace.stsci.edu/display/TESS [Accessed 2024−10−25].
21. Oelkers, R., Stassun, K. 2018. Precision light curves from TESS full-frame images: A different imaging approach. *The Astronomical Journal*, 156(3): 132.
22. Hatt, E., Nielsen, M. B., Chaplin, W. J., et al. 2023. Catalogue of solar-like oscillators observed by TESS in 120-s and 20-s cadence. *Astronomy & Astrophysics*, 669: A67.
23. Jackman, J. A., Shkolnik, E., Loyd, R. P. 2021. Stellar flares from blended and neighbouring stars in Kepler short cadence observations. *Monthly Notices of the Royal Astronomical Society*, 502(2): 2033−2042.
24. Meng, G., Zhang, L. Y., Su, T., et al. 2023. Properties of flare events on M stars from LAMOST spectral survey based on Kepler and TESS light curves. *Research in Astronomy and Astrophysics*, 23(5): 055001.
25. Davenport, J. R., Hawley, S. L., Hebb, L., et al. 2014. Kepler flares. II. The temporal morphology of white-light flares on GJ 1243. *The Astrophysical Journal*, 797(2): 122.
26. Gao, Q., Xin, Y., Liu, J. F., et al. 2016. White-light flares on close binaries observed with Kepler. *The Astrophysical Journal Supplement Series*, 224(2): 37.

27. Van Doorsselaere, T., Shariati, H., Debosscher, J. 2017. Stellar flares observed in long-cadence data from the Kepler mission. *The Astrophysical Journal Supplement Series*, 232(2): 26.

28. Lu, H. P., Zhang, L. Y., Shi, J., et al. 2019. VizieR Online Data Catalog: magnetic activities from M-type star LAMOST & Kepler. *VizieR Online Data Catalog*.

29. Huang, L. C., Ip, W. H., Lin, C. L., et al. 2020. M-dwarf Eclipsing Binaries with Flare Activity. *The Astrophysical Journal*, 892(1): 58.

30. Günther, M. N., Zhan, Z. C., Seager, S., et al. 2020. Stellar flares from the first TESS data release: exploring a new sample of M dwarfs. *The Astronomical Journal*, 159(2): 60.

31. Ilin, E. 2021. AltaiPony-Flare science in Kepler, K2 and TESS light curves. *Journal of Open Source Software*, 6(62): 3277.

32. Yang, Z. L., Zhang, L. Y., Meng, G., et al. 2023. Properties of flare events based on light curves from the TESS survey. *Astronomy & Astrophysics*, 669: A15.

33. Davenport, J. R. 2016. The Kepler catalog of stellar flares. *The Astrophysical Journal*, 829(1): 23.

34. Yang, H. Q., Liu, J. F. 2019. The flare catalog and the flare activity in the Kepler mission. *The Astrophysical Journal Supplement Series*, 241(2): 29.

35. Brasseur, C., Osten, R. A., Tristan, I. I., et al. 2023. Constraints on stellar flare energy ratios in the NUV and optical from a multiwavelength study of GALEX and Kepler flare stars. *The Astrophysical Journal*, 944(1): 5.

36. Lurie, J. C., Davenport, J. R., Hawley, S. L., et al. 2015. Kepler flares III: stellar activity on GJ 1245A and B. *The Astrophysical Journal*, 800(2): 95.

37. Shibayama, T., Maehara, H., Notsu, S., et al. 2013. Superflares on solar-type stars observed with Kepler. I. Statistical properties of superflares. *The Astrophysical Journal Supplement Series*, 209(1): 5.

38. Okamoto, S., Notsu, Y., Maehara, H., et al. 2021. Statistical properties of superflares on solar-type stars: results using all of the Kepler primary mission data. *The Astrophysical Journal*, 906(2): 72.

39. Wu, C. J., Ip, W. H., Huang, L. C. 2014. A study of variability in the frequency distributions of the superflares of G-type stars observed by the Kepler mission. *The Astrophysical Journal*, 798(2): 92.

40. Howard, W. S., MacGregor, M. A. 2022. No such thing as a simple flare: substructure and quasi-periodic pulsations observed in a statistical sample of 20 s cadence TESS flares. *The Astrophysical Journal*, 926(2): 204.

41. Chawla, N. V., Bowyer, K. W., Hall, L. O., et al. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16: 321−357.

42. Lin, C. L., Apai, D., Giampapa, M. S., et al. 2024. Scalable, Advanced Machine Learning Based Approaches for Stellar Flare Identification: Application to TESS Short-cadence Data and Analysis of a New Flare Catalog. *The Astronomical Journal*, 168(6): 234.

43. Yamashita, R., Nishio, M., Do, R. K. G., et al. 2018. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9: 611−629.

44. Breiman, L. 2001. Random forests. *Machine Learning*, 45: 5−32.

45. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011. Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12: 2825−2830.

46. Chen, T. Q., Guestrin, C. 2016. XGBoost: A scalable tree boosting system. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

47. Sun, Z. Y., Bobra, M. G., Wang, X. T., et al. 2022. Predicting solar flares using CNN and LSTM on two solar cycles of active region data. *The Astrophysical Journal*, 931(2): 163.

48. Liu, H., Liu, C., Wang, J. T., et al. 2019. Predicting solar flares using a long short-term memory network. *The Astrophysical Journal*, 877(2): 121.

49. Kag, A., Saligrama, V. 2021. Training recurrent neural networks via forward propagation through time. *In Proceedings of the 38th International Conference on Machine Learning*.

50. Platts, J., Reale, M., Marsh, J., et al. 2022. Solar flare prediction with recurrent neural networks. *The Journal of the Astronautical Sciences*, 69(5): 1421−1440.

51. LeCun, Y., Bengio, Y., Hinton, G. 2015. Deep learning. *Nature*, 521(7553): 436−444.

52. Baron, D. 2019. Machine learning in astronomy: A practical overview. *arXiv: 1904.07248*.

53. Shrestha, A., Mahmood, A. 2019. Review of deep learning algorithms and architectures. *IEEE Access*, 7: 53040−53065.

54. Mahajan, P., Uddin, S., Hajati, F., et al. 2023. Ensemble learning for disease prediction: A review. *Healthcare (Basel)*, 11(12): 1808.

55. Mohammadifar, S. 2023. Stacking-and voting-based ensemble deep learning models (SEDL and VEDL) and active learning (AL) for mapping

land subsidence. *Environmental Science and Pollution Research*, 30: 26580−26595.

56. Luque, A., Carrasco, A., Martín, A., et al. 2019. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91: 216−231.
57. Zhou, Z. H. 2021. *Machine Learning*. Singapore: Springer Nature.
58. Yang, H. Q., Liu, J. F., Gao, Q., et al. 2017. The flaring activity of M dwarfs in the Kepler field. *The Astrophysical Journal*, 849(1): 36.
59. Althukair, A., Tsiklauri, D. 2023. Main sequence star superflare frequency based on entire Kepler data. *Research in Astronomy and Astrophysics*, 23(8): 085017.
60. Feinstein, A. D., Seligman, D. Z., France, K., et al. 2024. Evolution of flare activity in GKM stars younger than 300 Myr over five years of TESS observations. *arXiv: 2405.12038*.
61. Stelzer, B., Caramazza, M., Raetz, S., et al. 2022. The Great Flare of 2021 November 19 on AD Leonis-Simultaneous XMM-Newton and TESS observations. *Astronomy & Astrophysics*, 667: L9.

Note: Figure translations are in progress. See original paper for figures.