The Astrophysical Journal, 999:181 (13pp), 2026 March 10 https://doi.org/10.3847/1538-4357/ae3153

TTC: Transformer-based TDE Classifier for the Wide Field Survey Telescope (WFST)

Ranfang Zheng[1,2,10], Zheyu Lin[1,2,10], Xu Kong[1,2,3,10], Dezheng Meng[1,2], Zelin Xu[1,2], Lulu Fan[1,2,3], Ji-an Jiang[1,2,4], Ning Jiang[1,2], Jie Lin[1,2], Tinggui Wang[1,2,3], Qingfeng Zhu[1,2,3], Feng Li[5], Ming Liang[6], Hao Liu[5], Zheng Lou[7], Wentao Luo[3], Jinlong Tang[8], Hairen Wang[7], Jian Wang[3,5], Yongquan Xue[1,2], Dazhi Yao[7], Hong-fei Zhang[5], Wen Zhao[1,2], Xianzhong Zheng[9], and Yingxi Zuo[7]

[1] Department of Astronomy, University of Science and Technology, Hefei 230026, People's Republic of China; rfzheng@mail.ustc.edu.cn, linzheyu@mail.ustc.edu.cn, xkong@ustc.edu.cn
[2] School of Astronomy and Space Sciences, University of Science and Technology of China, Hefei 230026, People's Republic of China
[3] Institute of Deep Space Sciences, Deep Space Exploration Laboratory, Hefei 230026, People's Republic of China
[4] National Astronomical Observatory of Japan, National Institutes of Natural Sciences, Tokyo 181-8588, Japan
[5] State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei 230026, People's Republic of China
[6] National Optical Astronomy Observatory (NSF's National Optical-Infrared Astronomy Research Laboratory) 950 N Cherry Ave. Tucson, AZ 85726, USA
[7] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, People's Republic of China
[8] Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, People's Republic of China
[9] Tsung-Dao Lee Institute and Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education, Shanghai Jiao Tong University, Shanghai 201210, People's Republic of China
Received 2025 August 22; revised 2025 December 6; accepted 2025 December 21; published 2026 March 2

Abstract

We propose the Transformer-based Tidal disruption events (TDE) Classifier (`TTC`), specifically designed to operate effectively with both real-time alert streams and archival data of the Wide Field Survey Telescope (WFST). It aims to minimize the reliance on external catalogs and find TDE candidates from pure light curves, which is more suitable for finding TDEs in faint and distant galaxies. `TTC` consists of two key modules that can work independently: (1) A light-curve parametric fitting module and (2) a Transformer (`Mgformer`) based classification network. The training of the latter module and the evaluations for each of the modules utilize a light-curve dataset of 7413 spectroscopically classified transients from the Zwicky Transient Facility (ZTF). The `Mgformer`-based module is superior in performance and flexibility. Its representative recall and precision values are 0.79 and 0.76, respectively, and can be modified by adjusting the threshold. It can also efficiently find TDE candidates within 30 days from the first detection. For comparison, the parametric fitting module yields values of 0.72 and 0.40, respectively, while it is >10 times faster in average speed. Hence, the setup of modules allows a trade-off between performance and time, as well as precision and recall. `TTC` has successfully picked out all spectroscopically identified TDEs among ZTF transients in a real-time classification test, and selected $\sim$20 TDE candidates in the deep field survey data of WFST. The discovery rate will greatly increase once the differential database for the wide-field survey is ready.

Unified Astronomy Thesaurus concepts: Tidal disruption (1696); Transient sources (1851); Light curve classification (1954); Wide-field telescopes (1800)

1. Introduction

A tidal disruption event (TDE) occurs when a star passes sufficiently close to a black hole and gets torn apart by tidal forces. Approximately half of the stellar debris is bound to the black hole, forming an accretion disk, while the remaining material is ejected into space (M. J. Rees 1988; S. Gezari 2021). TDEs exhibit distinctive features across various electromagnetic wavelengths. While TDEs were initially discovered in the soft X-ray sky survey in the late 1990s (N. Bade et al. 1996; S. Komossa & N. Bade 1999) through archived data, the optical band has taken the lead in the discovery number and rate in the recent decade, thanks to the advanced wide-field optical sky surveys. Real-time detection in the optical band, combined with spectroscopic confirmation, has become the primary approach for identifying TDEs at present. Currently, more than 100 optical-bright TDEs have been found, mostly contributed by ZTF (e.g., S. van Velzen et al. 2019, 2021; E. Hammerstein et al. 2023; Y. Yao et al. 2023). Their identification heavily relies on their unique spectroscopic features, including the rise and fall of a blue continuum, and the possibly adherent broad H, He and Bowen fluorescent emission lines (A. Zabludoff et al. 2021).

Although optical surveys currently play a dominant role in the discovery of TDEs, significant limitations remain, due to observational depth, that hinder a comprehensive understanding of the TDE population. The faint-end cutoff of the TDE luminosity function, for instance, remains poorly constrained, owing to the small number of detections, despite theoretical predictions suggesting a larger population of faint TDEs (S. van Velzen 2018; Z. Lin et al. 2022b; Y. Yao et al. 2023), and we have indeed found the existence of this group at present (P. Charalampopoulos et al. 2023, 2024; W. B. Hoogendam et al. 2024). In addition, current optically selected TDE samples are strongly biased toward quiescent host galaxies, with very few detections in star-forming environments. TDEs in such galaxies may be obscured by intense star formation activity or dust extinction (Y. Wang et al. 2022; N. Jiang et al. 2023; M. Masterson et al. 2024). Overcoming these limitations will require deeper observations and time-domain surveys with higher sensitivity and cadence, to construct a more complete and unbiased sample.

[10] Corresponding author.

With the commissioning of next-generation sky survey facilities such as WFST (T. Wang et al. 2023) and the Vera Rubin Observatory (LSST Science Collaboration et al. 2009; Ž. Ivezić et al. 2019), hundreds to thousands of optical-bright TDEs are predicted to be discovered each year (K. Bricman & A. Gomboc 2020; Z. Lin et al. 2022a). This signals the era of large-sample analysis, but also poses challenges regarding the spectroscopic resources. Automatic classification based on photometric data can serve as a plausible solution, as the current number of optical TDE samples is enough for training and testing.

Several machine learning (ML) algorithms have been specifically designed for multidimensional time-series classification. These models are either trained on simulated datasets like PLAsTiCC or ELAsTiCC (e.g., Avocado (K. Boone 2019), ATAT (G. Cabrera-Vives et al. 2024), and ORACLE (V. G. Shah et al. 2025; Y. Tang et al. 2025)), or trained on real survey datasets like ZTF (e.g., FLEET (S. Gomez et al. 2023), tdescore (R. Stein et al. 2024), NEEDLE (X. Sheng et al. 2024), and ALeRCE (R. Carrasco-Davis et al. 2021; F. Förster et al. 2021; P. Sánchez-Sáez et al. 2021; M. Perez-Carrasco et al. 2023; M. Pavez-Herrera et al. 2025)), while M. L. Lanza et al. (2026) have focused on attempting to quickly identify TDEs based on ascending stage of light curves. Most of these algorithms incorporate additional external data, such as the properties of host galaxies. In particular, redshift enables the determination of a source's luminosity, and angular separation between the transient and host galaxy aids in classification, as different types of transients have different preferences from the central to outskirt regions of the galaxy.

Although external data are indeed useful, they sometimes can exclude peculiar but important transients. In particular, an off-center TDE, or a TDE occurring around a much smaller black hole than that in the center of galaxy, will be unfortunately excluded by a strict angular separation cut (e.g., $<0\farcs6$ in S. van Velzen et al. 2019). Although they are estimated to occur at a rate of merely $\sim$1% of that of a central TDE, some reliable candidates have been identified (e.g., D. Lin et al. 2018; C. C. Jin et al. 2025; Y. Yao et al. 2025). They can signal the existence of intermediate-mass black holes (IMBHs) or runaway BHs that ejected from binary supermassive black hole (SMBHs) mergers or gravitational-wave recoil events (e.g., G. Fragione et al. 2018; A. Ricarte et al. 2021; J. N. Y. Chang et al. 2025), and provide a cleaner environment than dense nuclear regions for probing the BH accretion processes (J. E. Greene et al. 2020; K. C. Patra et al. 2025; Y. Yao et al. 2025). Therefore, our goal is to develop a method for identifying TDE candidates solely based on their light curves, minimizing reliance on external cross-matching with other datasets.

The transformer model excels at capturing long-range dependencies within time-series data (T. Uchiyama 2023; Q. Wen et al. 2023). By leveraging positional encoding, transformer models can capture temporal information in a manner akin to long short-term memory and gated recurrent units, while also offering improved scalability, flexibility, and advantages in parallel computation. While transformer models have shown promising results for transient source classification using simulated datasets, such as the ELAsTiCC dataset (G. Cabrera-Vives et al. 2024), their performance on real-world data remains largely untested. In this study, we introduce a variant of the transformer model, Mgformer, (J. Wen et al. 2024), as the base algorithm of our TDE filter, we developed the Transformer-based TDE Classifier (TTC) to search for TDE candidates merely relying on the light curves, and apply this filter onto the ZTF and WFST light-curve datasets.

In addition to the machine learning approach, we also propose a complementary parametric method. This parametric framework enables a rapid reduction in the number of candidates requiring further inspection, thereby decreasing the computational cost and mitigating the screening workload of the machine learning pipeline. Moreover, the parametric method itself possesses intrinsic discriminatory power, allowing it to serve as an independent preliminary filter. The balance between the parametric and machine learning components can be flexibly adjusted according to the specific objectives and requirements of the task.

The structure of the paper is as follows. In Section 2, we briefly introduce the construction of our ZTF light curve dataset for training and testing. In Section 3, we introduce our integrated search methodology in details. In Section 4, we demonstrate the performance of this method on WFST survey data and present the representative TDE candidates that pass the filter. In Section 5, we discuss the discoverability of TDEs at early stages, the performance and computational efficiency of modules, and the special TDE types that will be taken into consideration in the future updates. In Section 6, we conclude the study with a summary of our findings.

2. Construction of the ZTF Light-curve Dataset

The light curves utilized in this paper are based on the PSF photometry on the differential images (F. J. Masci et al. 2019). We introduce the construction of the ZTF light curve dataset in this section, and leave the introduction of the WFST light curve dataset to Section 4.

To evaluate the accuracy and completeness of our TDE filter, our dataset includes TDE and other 14 transient types, following the settings from PLAsTiCC (A. I. Malz et al. 2019). We further group these types into six categories, and use them for our later labeling:

1. TDE.
2. SN Ia = SN Ia, SN Iax, and SN Ia (91bg-like).
3. SN Ib/c = SN Ib, SN Ic, SN Ib/c, and SN Ibn.
4. SN II = SN II, SN IIP, SN IIn, and SN IIb.
5. SLSN (superluminous supernova) = SLSN-I and SLSN-II.
6. AGN (active galactic nucleus).

We limit the discovery date of the targets to the period from 2019 to 2024 (first discovery epoch in ZTF), and only select the targets that have at least five data points in each of the ZTF g and r bands. For TDEs, we further refine the samples by requiring magnitude variations in both the g and r bands to be greater than 0.5 mag, and excluding sources that are too close to the detection limit, i.e., those fainter than 19.5 mag at the peak. As a result, 50 TDEs and 7363 other types of transients are selected. The TDE sample includes 21 TDEs from the ZTF-I sample (E. Hammerstein et al. 2023), 16 additional TDEs from the subsequent publication (Y. Yao et al. 2023), and 13 more TDEs in TNS. More information for this sample can be found in Table A1. The constitution of the dataset is

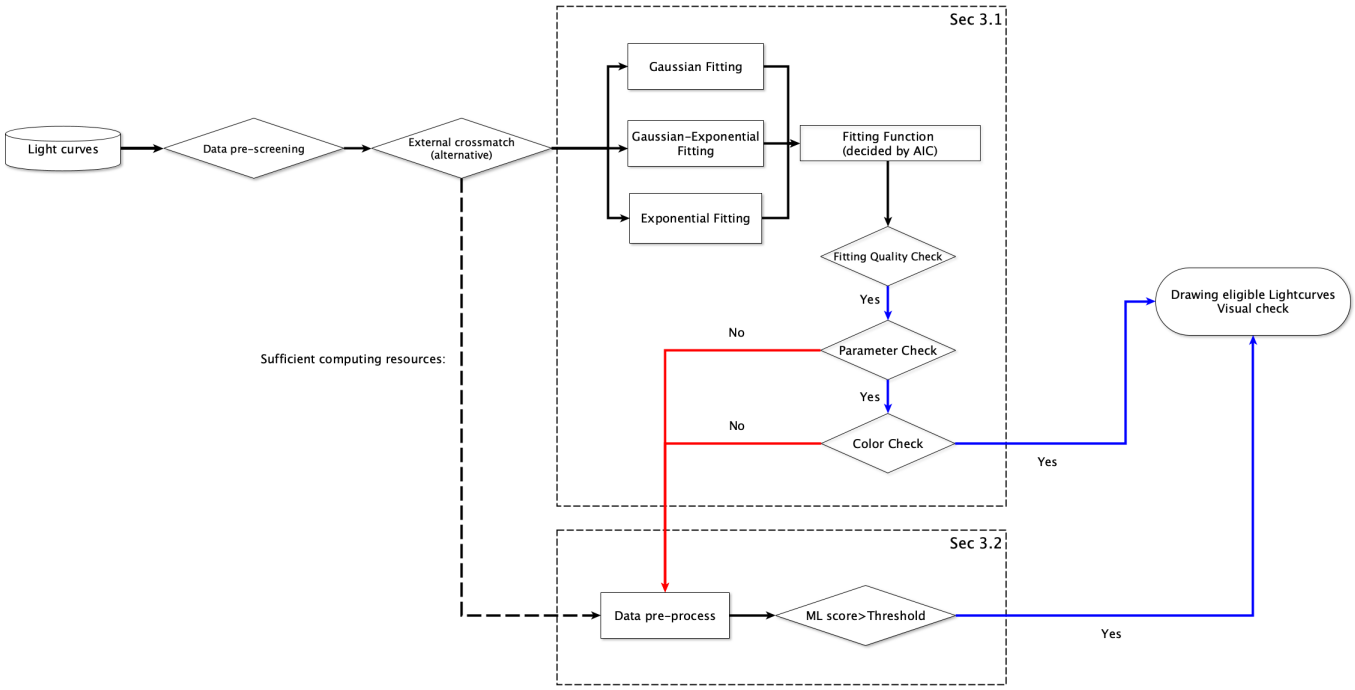

Figure 1. The flowchart of the whole process. The blue arrow indicates cases where the classification result is Y, while the red arrow represents cases where the classification result is N. The dashed line indicates the scenario in which ample computational resources allow for the direct execution of machine learning classification tasks. The prescreening of the light curve and the external cross-match will be introduced at the beginning of Section 3, the parameter fitting method will be introduced in Section 3.1, and the machine learning classification will be introduced in Section 3.2.

Table 1
The Constitutions of the ZTF Light-curve Dataset

TDE	SN Ia	SN Iax	SN Ia(91bg-like)	SN Ib	SN Ic	SN Ib/c	SN Ibn
50(0.7%)	4910(66.2%)	24(0.32%)	69(0.9%)	160(2.1%)	181(2.4%)	36(0.49%)	40(0.54%)
SN II	SN IIP	SN IIn	SN IIb	SLSN-I	SLSN-II	AGN	Total
1128(15.2%)	151(2.0%)	273(3.7%)	116(1.6%)	129(1.7%)	77(1.0%)	69(0.9%)	7413

Note. With the exception of TDEs, all other data were obtained through APIs. TDEs are in bold to emphasize their central role in our study.

displayed in Table 1. We retrieve the light-curve data of the above 7413 transients through the bulk download service of the `Lasair`[11] API.

3. Integrated Filter

Our filter makes use of two complementary methods, based on the parametric fitting and `Mgformer` model, respectively. We will show more details in Sections 3.1 and 3.2. Both methods possess the capability to independently identify TDE candidates and can be applied either sequentially or independently, depending on the specific needs of the analysis. Each approach offers distinct advantages, as detailed in Section 5.2. The choice of implementation can be flexibly adapted based on the requirements of the task.

In addition to the light-curve analysis model, we also explore the use of external data. These data include the angular separation between the transient source and the center of its host galaxy, the redshift and spectral classification of the host galaxy obtained from surveys such as SDSS or DESI, historical photometric baselines from time-domain surveys including ZTF and ATLAS, and the mid-infrared color index $W1 - W2$. They are only introduced after initial filtering, as we aim to minimize the dependence on external data. Although they can increase the reliability of our judgments, such steps are not necessary in our work.

Our entire flowchart is shown in Figure 1. First, we perform a prescreening of the data by requiring at least five data points with $\mathrm{SNR} > 5$ in both of the g and r bands. Subsequently, an optional external data cross-matching can be conducted. When computational resources are sufficient, we recommend applying machine learning algorithms directly for classification (Section 3.2), to take advantage of the early-time light curves. If this is insufficient, the parametric method could be used, as it is effective in excluding false light curves, with much less computational resources (Section 3.1). After the preliminary fitting quality check, further inspections of the fitting parameters and color evolution are performed. To avoid missing genuine candidates due to the strict partitioning of the parametric method, sources filtered out in these steps are still allowed to proceed to the machine learning classification stage. Finally, all resulting candidates are presented on a web interface for visual inspection.

[11] https://lasair-ztf.lsst.ac.uk/

Table 2
Constraints for Parametric Fitting in Each Band

Parameter	Initial Parameter Setting	Range for Fitting
σ/day	10	[1, 50]
τ/day	20	[1, 1000]
$t_{\rm peak}$/day	$t_{\rm max}$	$[t_{\rm max}-20, t_{\rm max}+20]$
$F_{\rm peak}$/mJy	$F_{\rm max}$	$[F_{\rm max}, 10\times F_{\rm max}]$
C/mJy	0	[−0.1, 0.1]

Notes. σ: the rise timescale. τ: the fading timescale. $t_{\rm max}$: the observation time at which $F_{\rm max}$ occurs. $F_{\rm max}$: the maximum observed flux in the band.

3.1. Method Based on Parametric Fitting

We employ a rise-and-fall function to model the light curve:

$$F_\nu(t) = C_\nu + F_{\nu,\rm peak}\times\begin{cases} e^{-(t-t_{\rm peak})^2/2\sigma^2}, & t\leqslant t_{\rm peak} \\ e^{-(t-t_{\rm peak})/\tau}, & t > t_{\rm peak}\end{cases}. \quad (1)$$

For the rising stage, a Gaussian function is employed, while for the descending stage, an exponential function is used. This function is similar as those used on ZTF TDE samples (e.g., S. van Velzen et al. 2021; E. Hammerstein et al. 2023), while a baseline correction term C_ν has been added to compensate for inaccuracies in the data calibration.

In our analysis, these baseline signals are generally attributed to residual system artifacts in the difference images, but they may also arise from imperfect calibration. The former typically appears as a small number of scattered data points preceding the main rise, whereas the latter often leads to a global offset in the differential light curve, shifting it upward or downward. This effect is present in in the light curves derived from both ZTF and WFST.

We employ the `curve_fit` from the `scipy` package for curve fitting. This method is based on maximum likelihood estimation and allows simultaneous fitting of multiparameter functions. The parameter constraints and initial settings used in the fitting process are listed in Table 2, primarily following the approach of E. Hammerstein et al. (2023).

The quality of the fit is assessed by using two parameters: the R^2 score and the mean absolute percentage error (MAPE), along with the timescale corresponding to the fitting parameters. The definitions of the MAPE and R^2 scores are

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{F_t - A_t}{A_t}\right|, \quad (2)$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(A_t - F_t)^2}{\sum_{t=1}^{n}(A_t - \bar{A})^2}, \quad (3)$$

where A_t and F_t represent the observed value and model-predicted value, while $\bar{A}$ represents the mean of A_t.

For each individual source, we perform curve fitting using the Gaussian component, the exponential component, and the full piecewise function separately. The fitting is conducted independently for the g-band and r-band light curves. The optimal model is selected based on the Akaike information criterion (AIC), which is computed for each fit as follows:

$$\mathrm{AIC} = 2\,k - 2\ln\hat{L}, \quad (4)$$

where k is the number of free parameters in the model and $\hat{L}$ is the maximum value of the likelihood function for the model. A lower AIC value indicates a better model fit, as it reflects a more favorable trade-off between goodness of fit and model complexity.

To trigger this selection, both the g and r bands are required to have more than five valid detections in which the signal-to-noise ratio >5. If so, the functions will be applied to the g- and r-band light curves to obtain the rise timescales σ_g and σ_r, peak epochs $t_{g,\rm peak}$ and $t_{r,\rm peak}$, peak fluxes $F_{g,\rm peak}$ and $F_{r,\rm peak}$ and baseline offsets C_g and C_r. We present in Figure B1 the distributions of the fitted parameters for all samples. TDEs typically exhibit rising timescales σ exceeding approximately 10 days and declining timescales τ generally greater than 20 days.

We require that $\sigma_g > 10$ days and $\sigma_r > 10$ days, In addition, to better limit the fitting results in both bands, we simultaneously require during the fitting process that $|\sigma_g - \sigma_r| < 5$ days.

Meanwhile, both the fitting parameters and the fitting quality are used for screening. Specifically, we require the R^2 scores for both the g and r bands to exceed 0.7, while the MAPE values for both bands must be less than 0.3. However, it should be noted that the fitting quality criteria provided here are intended primarily for reference and are most applicable in situations where substantial fractions of bogus detections are present. For the majority of light curves that correspond to real detections, we recommend adopting a more relaxed criterion or even omitting the fitting quality check altogether.

Additional color-related screening is conducted. First, the $g-r<0$ color near the peak value is examined. Second, based on observational experience, some blue SNe near the peak exhibit rapid cooling, leading to a "crossover" point. Specifically, when $r<g$ after the crossover, this intersection is set to occur no earlier than 20 days after the peak. Meanwhile, $d(g-r)/dt < 0.01\ \mathrm{day}^{-1}$ within 1 month after the first observation (K. Bučar Bricman et al. 2023) is also checked. Figure B2 presents the peak colors of TDEs and non-TDEs, together with the color evolution within 30 days after the peak. It is evident that the vast majority of TDEs exhibit bluer peak colors and relatively slower color changes following maximum light.

Following the screening procedure based on goodness-of-fit and color variation criteria, 36 out of 50 TDEs were successfully identified, yielding a recall of 0.72. Among the remaining 7363 non-TDE sources, 53 were incorrectly classified as TDEs, resulting in a precision of 0.40. Without the fitting quality check, recall will rise to 0.79, while precision will only drop to 0.3. For definitions of recall and precision, see Section 3.2.2. We observed that, among the TDEs that were not successfully selected (for example, ZTF22aadesap and ZTF22aacgcwv), the majority failed due to the g-band magnitude at peak brightness being slightly higher than that of the r band. The remaining cases, such as ZTF20acqoiyt and ZTF20abgwfek, exhibited relatively pronounced color evolution during the declining phase.

3.2. Method Based on the Mgformer Model

3.2.1. Data Preprocessing

To facilitate model convergence and enhance generalization performance, we first apply min-max normalization to the g- and r-band flux values. For each individual source, we identify its maximum and minimum flux values. Both photometric bands of the same source are then normalized using a common

scaling factor, such that $\mathrm{flux}_{\mathrm{max}} = \max(\mathrm{flux}_{\mathrm{g,max}}, \mathrm{flux}_{\mathrm{r,max}})$ and $\mathrm{flux}_{\mathrm{min}} = \min(\mathrm{flux}_{\mathrm{g,min}}, \mathrm{flux}_{\mathrm{r,min}})$, in order to preserve the intrinsic color and color-evolution characteristics of the source.

To further enhance the model's sensitivity to early-phase variability and improve its robustness, we employ a truncated data augmentation strategy.

Before performing data augmentation, we first divide the dataset into training and test subsets to ensure that the same source does not appear in both, thereby avoiding overly optimistic evaluation results. A stratified sampling strategy is adopted, allocating 65% of each class to the training set and the remaining to the test set. After the division, data augmentation is applied independently to the training and test sets.

Specifically, each light curve is segmented at multiple quantiles of its temporal length, namely at 0.2, 0.4, 0.6, 0.8, and 1.0, starting from the initial observation. Only segments containing more than four valid observations in both the *g* and *r* bands are retained, to ensure data quality. Incorporating light curves of varying completeness as a form of data augmentation enables the model to learn features across different evolutionary stages of transient events. This approach improves the model's adaptability and transferability to light curves with incomplete observations. Similar strategy has been adopted in prior work, such as that of V. G. Shah et al. (2025).

Following augmentation, we reconstruct each sequence to the same length, because the `Mgformer` model requires a unified length for all time sequences. For this purpose, we adopt a two-dimensional Gaussian process (GP) interpolation with `GpyTorch`, (J. R. Gardner et al. 2018) to jointly reconstruct the light curves in both bands. In cases where the GP fails to converge or prone to convergence toward the prior mean, due to data sparsity or noise, we instead employ a multilayer perceptron (MLP) based interpolation. The probability of this phenomenon occurring is approximately 7%, and in most cases, it arises from the small numerical differences between data points.

We adopt GP interpolation as the primary approach because it provides smoother results and is generally more robust to a small number of outliers, thanks to its smoothing priors and kernel-based structure. In contrast, MLPs are more prone to overfitting. In cases where the GP fails to converge or converges to a constant function, we instead choose the MLP, since it remains capable of producing an interpolation under such circumstances.

In the end, we apply min-max normalization again. The distribution of labels in the training and test sets used for the final evaluation of the model performance is summarized in Table 3.

Table 3
Samples in the Training and Test Sets

Class	Training Set	Test Set	Total
TDE	142	77	219
SN Ia	11267	6086	17353
SN Ib/c	725	385	1110
SN II	3906	2106	6012
SLSN	486	283	769
AGN	147	73	220
Total	**16673**	**9010**	**25683**

Note. Sample counts of each transient class used for training and testing the classifier after data augmentation. Totals (in bold) are also directly reflected in Figure 3.

Table 4
Hyperparameters Set for Training

Hyperparameters	Value
multi_group	[1,2]
batch	1
lr	0.0001
nlayers	2
emb_size	32
nhead	4
emb_size_c	32
masking_ratio	0.15
ratio_highest_attention	0.35
dropout	0.01
nhid	32
nhid_c	32
Training epochs	40

Note. The definition and default setting of these hyperparameters are introduced in J. Wen et al. (2024).

3.2.2. Performance of the ML Filter

Given the limited number of available TDE samples, it is crucial to maximize data utilization. Therefore, after obtaining promising results in test data, the previously designated test set is incorporated into the training set, to enhance the model's learning capacity. In this section, we only present the performance on the test set.

The hyperparameters set by the training the model and the training epoch are detailed in Table 4. Based on our experiments, the hyperparameters listed in the table represent the minimum values required to ensure model convergence. We also explored the default settings from J. Wen et al. (2024) and larger parameter settings, including various parameter combinations, and found no significant difference in the convergence outcomes. However, excessively large parameter values substantially increase the training time.

This model was deployed for training and testing on a single NVIDIA A100-PCIE-40 GB. Each training epoch takes about a few minutes.

Our evaluation is based on three key metrics: precision, recall and F1-score. Precision quantifies the accuracy of TDE detection, while recall measures the completeness of TDE identification, The F1-score is the harmonic mean of precision and recall. The calculation of these metrics requires a four-fold classification:

1. *True Positive (TP).* TDE correctly classified as TDE.
2. *False Positive (FP).* Non-TDE incorrectly classified as TDE.
3. *False Negative (FN).* TDE incorrectly classified as non-TDE.
4. *True Negative (TN).* Non-TDE correctly classified as non-TDE.

Precision, recall, and F1-score are defined as

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{5}$$

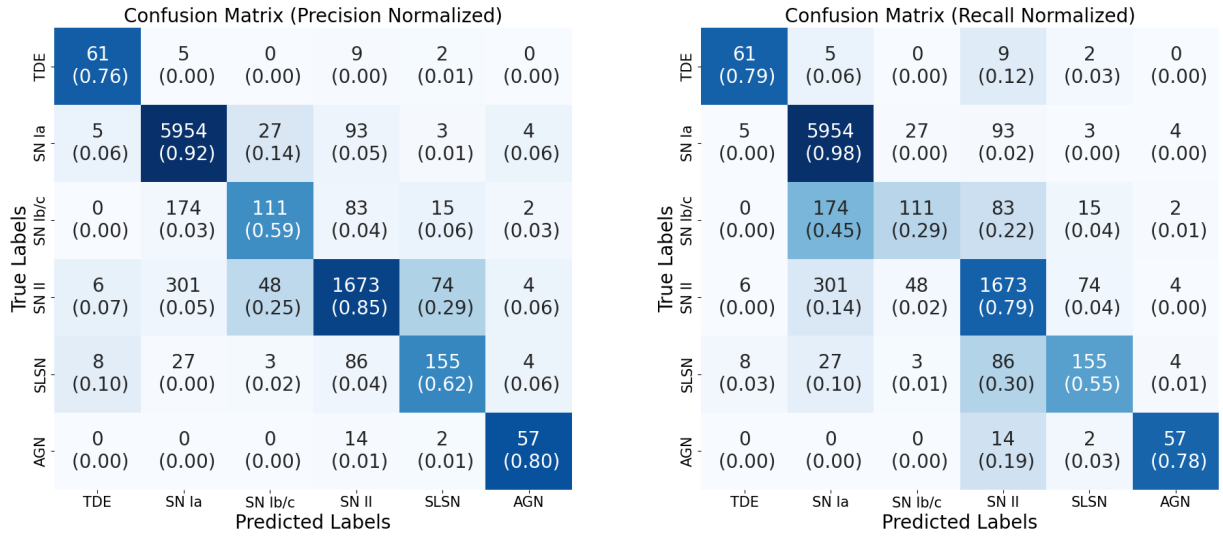

Figure 2. The performance on the test set, the classification achieves a precision of 0.76 and a recall of 0.79. Notably, the classification process relies solely on the information provided by the light curve.

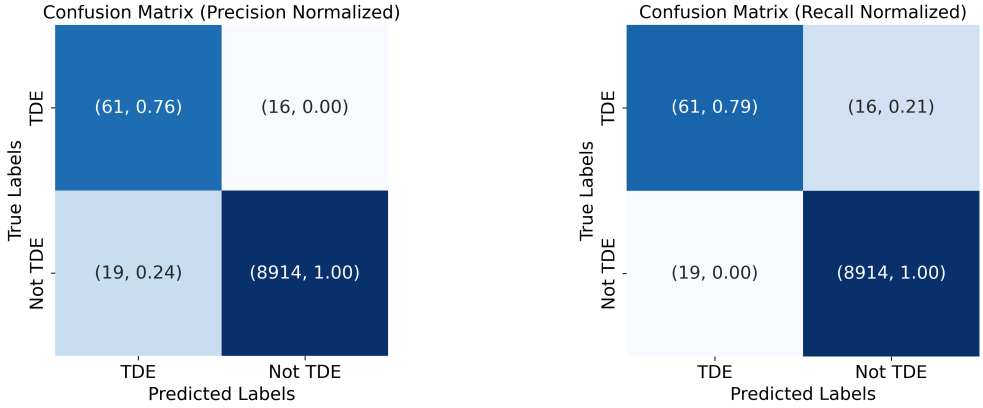

Figure 3. The confusion matrix shown here is analogous to that in Figure 2, with the distinction that all non-TDE classes have been merged into a single "Not TDE" category.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7}$$

We adopt the `Adam` optimizer (D. Kingma & J. Ba 2014) and evaluate the model performance after each epoch using two metrics: the F1 score for TDE classification on the test set (F1_{TDE}) and the global F1 score across all classes (F1_{total}). The epoch yielding the maximum sum of $\text{F1}_{\text{TDE}} + \text{F1}_{\text{total}}$ is selected as the final model checkpoint. Despite the substantial class imbalance in the dataset, we employed the standard (unweighted) cross-entropy loss and obtained satisfactory performance, which we attribute primarily to the effectiveness of our data augmentation strategy.

The evaluated performance of the ML filter is demonstrated as follows. We present the confusion matrix on the test set in Figure 2. This model achieves a precision of 0.76 and a recall of 0.79 on the test set. For simplicity, the weights produced by the classification layer of the machine learning algorithm are hereafter referred to as the "TDE score" throughout this study. Given our primary interest in TDE classification, we represent the classification performance for TDEs using a simplified 2×2 confusion matrix, as shown in Figure 3.

We further explore the relationship between the TDE score threshold and TDE recall, precision, and F1-score in Figure 4. Based on this relationship, we assign three strategies based on TDE score thresholds: 0.03, 0.5, and 0.8, that yield recall $\approx$0.9 (recall-oriented strategy), precision$\approx$0.9 (precision-oriented strategy), and recall$\approx$precision (balanced strategy), respectively.

3.3. Application of the Integrated Filter on the Latest ZTF Data

The final routine of our filter is described as follows. First, the parametric fitting method is employed to evaluate the shape and color trends of the light curves. Subsequently, the ML model is utilized to assign scores. In the Table A1, we present the scores for TDEs included in the training set.

To assess the reliability of our model on new data, we retrieved the latest light curves from the ZTF by using the Lasair API. The selection criteria for light curves are straightforward: (1) the source must have been observed at least once within 10 days prior to the date of our test, 2025

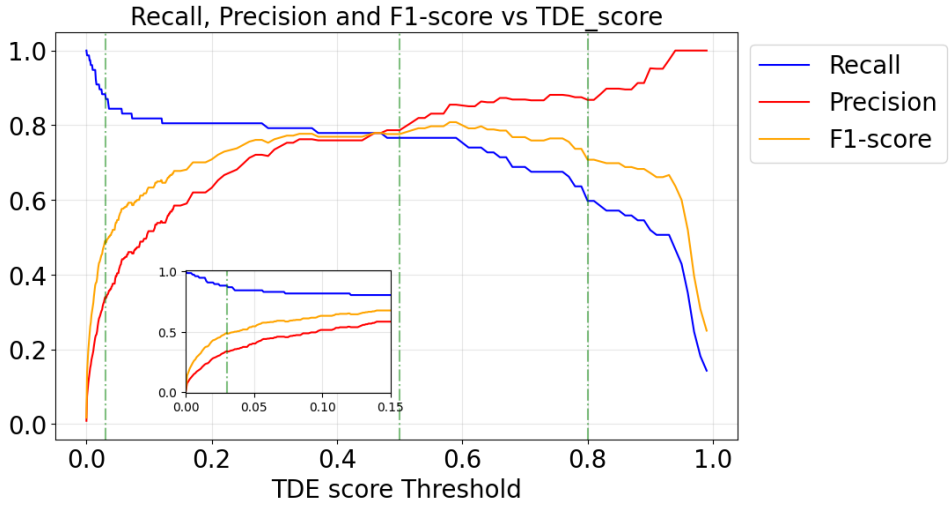

Figure 4. The relationship between the TDE score and precision, recall, and F1-score is depicted, with the blue solid line representing recall and the red solid line representing precision. The intervals (0, 0.15) is magnified in the left subplots. The three vertical green dashed–dotted lines represent the thresholds of the three strategies.

Table 5
ZTF TDE Candidates Selected by Model

ZTF Source	IAU Name	Transient Type (TNS)	Discovery Date	TDE Score
ZTF25aafofcs	TDE 2025bri	TDE-He	2025-02-15	0.59
ZTF25aafwfzz	TDE 2025cyh	TDE	2025-02-15	0.86
ZTF25aagevje	TDE 2025chm	TDE	2025-02-20	0.94
ZTF25aajjeon	TDE 2025hbw	TDE	2025-03-19	0.98

Note. This table shows TDE candidates that we have selected from 129 transients that have names beginning with "ZTF25" and at least one detection within 10 days prior to 2025 April 28. ZTF25aajjeon was classified on 2025 May 2.

April 28, and (2) the light curve must contain more than 10 data points. No constraints were imposed regarding angular separation from the host galaxy center or the host galaxy properties. (3) The source's name must begin with "ZTF25," which means this source is officially discovered in 2025.

Applying these criteria, we collected 129 light curves in total. Following the methodology described in previous sections, we selected candidates based on a cut of TDE score$>$0.5. This selection cut finds all four spectroscopically confirmed TDEs with 100% precision. The names, discovery dates, and TDE scores of these sources are summarized in Table 5. In particular, ZTF25aajjeon was selected as a highly promising candidate 4 days before it was classified as a featureless TDE at $z = 0.6278$,[12] demonstrating the capability of finding distant TDEs.

4. Application on WFST Survey Data

The sky survey strategy of WFST consists of two main components: the Wide-Field Survey (WFS) and the Deep High-Cadence u-band Survey (DHS) program (T. Wang et al. 2023). The cadence of these surveys is designed to align with their respective scientific objectives, with each occupying approximately 45% of the total observation time. During the operation of WFST, we generate a vast dataset. The raw data are processed using a standardized data pipeline (M. Cai et al. 2025).

The WFS program will cover $\sim$8000 deg^2 of the northern sky, utilizing u, g, r, i bands with a single-exposure time of 30 s. In addition to WFS, the DHS program is designed to leverage WFST's superior u-band imaging performance for time-domain surveys. DHS will conduct regular monitoring of two $\sim$360 deg^2 sky regions: the "Spring" and "Autumn" fields near the celestial equator. Following the completion of reference-image preparation, the DHS of WFST has been systematically conducting planned science survey since March 2024. Reference images for the WFS are currently under preparation, with construction having begun in late 2024. In the meantime, we have initially applied our method to the deep field component of the survey.

We apply the same min-max normalization to the WFST light curves, followed by a standardization step using the mean and variance derived from the training set. This ensures that the input scale remains consistent with that of the training samples, thereby maintaining compatibility and stability during inference.

Since the ZTF dataset used for training contains only the g and r bands, our classifier is currently limited to using these two bands when applied to WFST data. In the future, as WFST observations accumulate, we plan to retrain and evaluate the model using the WFST dataset itself, which includes additional photometric bands. We expect that incorporating more bands will further improve the classification performance by providing richer color information.

[12] https://www.wis-tns.org/astronotes/astronote/2025-131

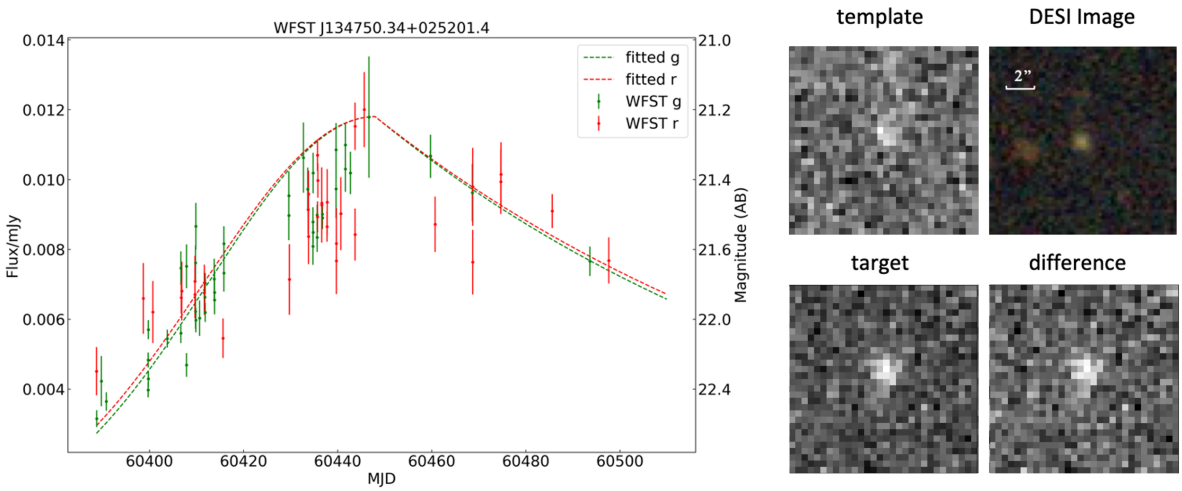

Figure 5. Left panel: The differential light curve of WFST J134750.34+025201.4. The green and red data points represent the *g*- and *r*-band differential flux respectively, while the green and red dotted lines represent the best fits to the rise-and-fall function (Equation (1)). Right panel: WFST *r*-band cutouts of WFST J134750.34+025201.4 on MJD = 60444: "target," "template," and "difference" mark the reference, science, and difference cutouts, respectively. The cutout size is $16\farcs5 \times 16\farcs5$. The upper right corner shows the *griz*-colored cutout in DESI Legacy Surveys DR10.

4.1. WFST TDE Candidates

During each quarter of the deep field survey, WFST is capable of finding ∼8 TDEs (Z. Lin et al. 2022a). As of 2025 May 1, we have found a total of about 20 candidates via the method mentioned in Section 3, but unfortunately all of them have been fainter than 21 mag in both the *g* and *r* bands since their discoveries, which is too faint for our spectroscopic follow-up observations. We expect the situation will be fundamentally improved when the WFS differential database becomes ready, and we anticipate that the discovery rate should increase by an order of magnitude.

For reference, we hereby present two promising TDE candidates selected by our method: WFST J134750.34 +025201.4 and WFST J015306.85+065606.7 (alert name WFST0467evz).[13]

For the former source, we demonstrate the potential of WFST to detect faint TDEs. The latter source is currently among the brightest WFST TDE candidates, featuring the best *u*-band sampling and a relatively high classification score.

WFST J134750.34+025201.4 (TDE score = 0.0592), is associated with a host galaxy in the HSC-Wide catalog (H. Aihara et al. 2019), with a photometric redshift of $z_{\rm phot} = 0.840^{+0.093}_{-0.350}$ (M. Tanaka et al. 2018) within $0\farcs082$. Its *g*-, *r*-, *i*-, and *z*-band magnitudes, as recorded in the DESI Legacy Imaging Surveys (A. Dey et al. 2019) DR10 catalog, are 23.64 ± 0.20, 22.97 ± 0.16, 22.34 ± 0.09, and 21.97 ± 0.14, respectively. The observed rise timescale and color evolution near the peak are consistent with a TDE classification. However, we cannot entirely rule out alternative explanations, such as a SLSN or a flare from an AGN, based solely on the available light curves. Because some SLSNe may also exhibit bluer colors and slower color evolution (Z. H. Chen et al. 2023), the possibility of such confusion cannot be fully excluded. Moreover, the possibility of an AGN flare cannot be entirely excluded as alternative interpretations, due to the lack of information to confirm whether there is AGN activity in this source. Relying solely on the light curve in the *g* and *r* bands is insufficient to conclusively rule out the presence of an AGN, as AGN activity can exhibit similar photometric variability (S. van Velzen et al. 2021). Additional multiwavelength data are needed to enable a more definitive interpretation. The absence of pre-outburst photometric coverage further complicates the identification and characterization of the transient's origin.

The source's light curve, along with the reference image at the peak position (template), the scientific image (target), the difference image, and the host galaxy image, are comprehensively presented in Figure 5.

WFST J015306.85+065606.7 (TDE score = 0.9135) was identified during the observations in late 2024 December. Since its first detection on 2024 December 27 (MJD = 60671), it was monitored well by WFST until 2025 March 2 (MJD = 60736). Its rise-and-fall timescale and constant $g - r$ color of < 0 are consistent with those of TDEs. In addition, the well-sampled *u*-band light curve of WFST shows an evolution trend synchronous with those of the *g* and *r* bands. If the *u*-band flux is slightly underestimated, the corrected color will be typical for a TDE. The *g*-, *r*-, and *z*-band magnitudes of its host galaxy in the DESI Legacy Imaging Surveys DR10 catalog are 22.20 ± 0.09, 20.91 ± 0.04, and 20.09 ± 0.05, respectively. An image similar to that of WFST J134750.34+025201.4 is presented in Figure 6.

Despite the unfortunate lack of spectroscopic confirmation, these two candidates demonstrate the capabilities of WFST in detecting potential TDEs in faint and distant galaxies through our TDE classifier.

5. Discussion

5.1. Ability to Capture Early TDE Candidates

As introduced in Section 3.2.1, data augmentation has been employed to enhance the model's capability to handle early

[13] After the formal survey begins, all sources will be assigned a unique alert name in the format WFST (date since first WFST exposure)+abc or WFST-PS +YYMMDDabc (where PS denotes "pilot survey"). However not all sources have their own alert names, e.g., the first source. In parallel, an internal naming convention using the format WFST+J2000 will be applied to all sources, based on their celestial coordinates. Beginning 2025 July 1, the official WFST alert naming system will be adopted. The standardized format will be WFSTYYMMDDabc, where YYMMDD indicates the discovery date, and abc represents the order of detection on that day. The suffix abc follows a base 26 sequence: a ∼ z, followed by aa ∼ zz, in a manner consistent with the TNS naming convention.

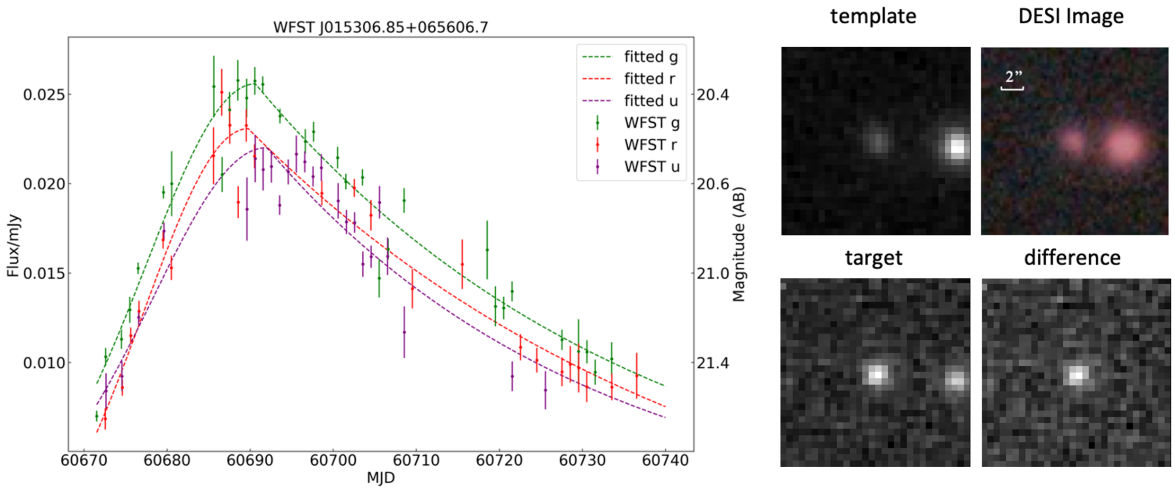

Figure 6. Same as Figure 5, but for WFST J015306.85+065606.7. It has a well-sampled *u*-band light curve. The WFST cutout images were taken in the *r* band and created on MJD = 60686.

light curves. To more rigorously assess the model's classification performance across different phases of transient evolution, we hereby evaluate the recall and precision as functions of time since the first epoch. Based on the results shown in Figure 4 and strategies mentioned in Section 3.2.2, we calculate these metrics at 10 day intervals, plotting them against the number of days elapsed since the initial observation. The results are presented in Figure 7.

Specifically, the balanced strategy (TDE score threshold = 0.5) works out best in finding TDEs at early phases, as it reaches a precision of 0.57 using the light curves within the first 20 days, and the recall quickly rises when given longer light curves. Regarding the deeper flux limit of WFST, this strategy should enable WFST to discover TDEs earlier than ZTF. In comparison, the precision-oriented strategy (threshold = 0.8) and the recall-oriented strategy (threshold = 0.03) are more suitable for the cases of limited and abundant follow-up resources, respectively.

5.2. Trade-off between Performance and Time

The parametric fitting and `Mgformer`-based method are both able to handle the classification task independently. A brief comparison of the strengths and limitations of these two methods is provided in Table 6.

In detail, regarding the individual performance, the `Mgformer`-based method is superior in metrics (with a recall of 0.79 versus 0.72, and a precision of 0.76 versus 0.40), flexibility (it can assign scores that allow for different strategies) and versatility (it is capable of multiclass classification, e.g., of SNe, as detailed in Appendix C).

Despite its performance, the parametric fitting method should consume less time, as it only requires one curve fitting step. To quantify the speed, we run a single-threaded test on a personal computer (Apple M2 Pro, 16 GB RAM). This test uses the same ZTF light-curve dataset introduced in Section 2 and Table 1.

In the parametric fitting stage, the total processing time for the test is 179 s, yielding an average processing time of 0.024 s per source. However, the subsequent light-curve reconstruction and `Mgformer`-based classification cost a substantially longer time of 3557 s, or 0.42 s per source. Therefore, the average running speed of the parametric fitting method is >10 times faster than that of the `Mgformer`-based method. In practical applications, a trade-off between performance and time must be achieved in light of the available computational resources and the amount of processed data each day.

5.3. Special TDE Types for Future Updates

Our model training heavily relies on current spectroscopically identified transients. This ensures that our filter can effectively capture TDE candidates that share light curves similar to those of the typical TDEs. From a contrary view, this also implies that atypical TDEs, such as TDEs occurring in AGNs (e.g., P. K. Blanchard et al. 2017; C.-H. Chan et al. 2020) and optical-bright repeating partial TDEs (e.g., A. V. Payne et al. 2021; Z. Lin et al. 2024; J. J. Somalwar et al. 2025), will not be selected by our current filter. In future updates, we will explore involving these TDE types or building specialized filters for them.

6. Conclusion

Motivated by the prospect and scientific goal of finding TDEs with WFST, we have developed an automated classifier (`TTC`) that can select TDE candidates from multiband light-curve data with high recall and precision. It is a combination of two modules: light-curve parametric fitting and a transformer (`Mgformer`) based classification network. These two modules can also work independently. We evaluate their individual performance on a ZTF light-curve dataset of 7413 transients found between 2019 and 2024. The metrics are listed below:

1. The parametric fitting method achieves a recall of 0.72 and a precision of 0.40.
2. The `Mgformer`-based method achieves a recall of 0.79 and a precision of 0.76.

The `Mgformer`-based method is superior in terms of performance, yet much slower than the parametric fitting method (Section 5.2). The setup of the modules allows a flexible trade-off between performance and time, depending on the available computational resources and the amount of processed data each day.

To enhance the ability to find TDEs at early stages, data augmentation has been employed (Section 3.2.1). We have

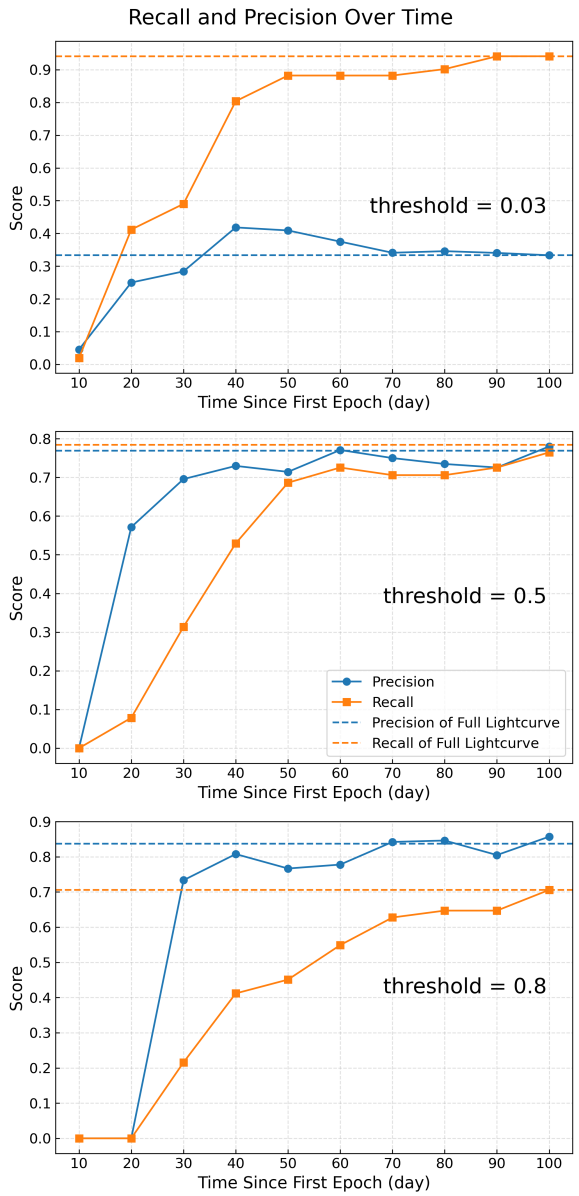

Figure 7. Recall and precision as a function of time since the first observational epoch, corresponding to the three strategies described in Section 3.2.2.

proved TTC's ability in finding TDE candidates within ∼30 days of the first discovery (Figure 7). Regarding the depth of WFST and ZTF ($\gtrsim$22 mag versus $\gtrsim$20 mag), this enables WFST to capture a TDE candidate earlier than ZTF.

To test TTC's capability to find TDE candidates in real time, we apply it to a set of 129 ZTF transients that were first detected in 2025 and have at least one detection in the most recent 10 days, and accurately find four TDEs among them, including ZTF25aajjeon (TDE 2025hbw) four days before it was classified as a featureless TDE at $z = 0.6278$. This result demonstrates TTC's capability in finding distant TDEs.

Table 6
Parametric Fitting versus Mgformer-based Method

	Parametric Fitting	Mgformer
Run independently	✓	✓
Assign quantified scores	✗	✓
Multiclass classification	✗	✓
Speed per source per core	0.024 s	0.42 s
Execution sequence	Former	Latter

Applying this method to the WFST survey dataset, we identify about 20 TDE candidates, including two highly confident candidates: WFST J134750.34+025201.4 and WFST J015306.85+065606.7. Unfortunately, neither of them have been bright enough for our spectroscopic follow-ups since they were discovered. Nonetheless, these two candidates demonstrate the capabilities of WFST in detecting potential TDEs in faint and distant galaxies through TTC. We expect the situation to be greatly improved once the WFS differential database is ready, as it covers a ∼10 times larger sky area, leading to a ∼10 times higher anticipated discovery rate.

In the future, following the commencement of the full-scale sky survey, a substantial number of TDE candidates with well-sampled photometry in the *u* band are expected to be collected. The *u* band plays a particularly important role in the photometric identification of TDEs, due to its sensitivity to the blue continuum. Incorporating high-quality *u*-band data is anticipated to significantly enhance the model's screening capability and improve the reliability of TDE classification.

The code developed for this study is open source and available in Github[14] and in Zenodo (R. Zheng 2025).

Acknowledgments

We sincerely thank the anonymous referee for providing useful comments that have significantly improved the quality of this article. This work is supported by the National Science Foundation of China (grant No. 12233008), the National Key R&D Program of China (2023YFA1608100), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant No. XDB0550200), the Cyrus Chun Ying Tang Foundations, and the 111 Project for "Observational and Theoretical Research on Dark Matter and Dark Energy" (B23042).

Facilities: WFST(2.5m), P48(1.2m).

Software: Astropy (Astropy Collaboration et al. 2013), Scipy (P. Virtanen et al. 2020), Lasair (R. D. Williams et al. 2024), Pytorch (A. Paszke et al. 2019), Matplotlib (J. D. Hunter 2007), Mgformer (J. Wen et al. 2024), GpyTorch (J. R. Gardner et al. 2018).

[14] https://github.com/Tico-Astro/TTC

Appendix A
TDE Samples Used for Training

In Table A1, we list all TDE samples used for training along with their performance scores in the final model. Overall, the majority of TDEs are correctly classified.

Table A1
Fifty TDEs and Their ML Scores

E. Hammerstein et al. (2023)	Y. Yao et al. (2023)	Samples from TNS
ZTF18abxftqm | 0.9736	ZTF19aakswrb | 0.9323	ZTF22aaabovl | 0.9482
ZTF18acaqdaa | 0.9782	ZTF19aaniqrr | 0.9833	ZTF22aaaedas | 0.9503
ZTF18acnbpmd | 0.9941	ZTF20abgwfek | 0.4769	ZTF22aaahtqz | 0.9201
ZTF19aabbnzo | 0.9903	ZTF20achpcvt | 0.9741	ZTF22aacgcwv | 0.0700
ZTF19aakiwze | 0.9433	ZTF20acnznms | 0.8265	ZTF22aadesap | 0.0787
ZTF19aapreis | 0.9893	ZTF20acwytxn | 0.9512	ZTF22aagvrlq | 0.8859
ZTF19aarioci | 0.7062	ZTF21aaaokyp | 0.6749	ZTF22aavvqyh | 0.3056
ZTF19abhhjcc | 0.9344	ZTF21aakfqwq | 0.8968	ZTF22abajudi | 0.9419
ZTF19abidbya | 0.1680	ZTF21aanxhjv | 0.8643	ZTF22abegjtx | 0.3682
ZTF19abzrhgq | 0.0101	ZTF21abaxaqq | 0.6000	ZTF22abkfhua | 0.9638
ZTF19accmaxo | 0.9855	ZTF21abcgnqn | 0.7749	ZTF23aadcbay | 0.4717
ZTF19acspeuw | 0.8991	ZTF21abhrchb | 0.0720	ZTF23abcvbqq | 0.9733
ZTF20aabqihu | 0.8410	ZTF21abjrysr | 0.6255	ZTF24abmybnp | 0.6760
ZTF20aamqmfk | 0.9944	ZTF21abqhkjd | 0.9129	⋯
ZTF20abefeab | 0.9875	ZTF21abxngcz | 0.9949	⋯
ZTF20abfcszi | 0.8230	ZTF21acafvhf | 0.9777	⋯
ZTF20abjwvae | 0.9925	⋯	⋯
ZTF20abnorit | 0.9829	⋯	⋯
ZTF20abowque | 0.9901	⋯	⋯
ZTF20acitpfz | 0.0206	⋯	⋯
ZTF20acqoiyt | 0.9453	⋯	⋯

Notes. We present the TDE sources used to train the ML model. Of these, 21 sources are from the ZTF-I TDE catalog (E. Hammerstein et al. 2023), 16 sources are from Y. Yao et al. (2023), and the remaining 13 sources were selected based on spectral classifications from the TNS. Their ML scores are also provided.

Appendix B
Visualization of the Selection Cuts of Parametric Fitting

In Figure B1, we plot the distributions of the fitted parameters for all samples. The shaded contour regions illustrate the parameter distribution of non-TDE sources, while red markers denote the TDE population. It is evident that TDEs typically exhibit rising timescales σ exceeding approximately 10 days and declining timescales τ generally greater than 20 days.

Figure B2 presents the peak colors of TDEs and non-TDEs, together with the color evolution within 30 days after the peak. It is evident that the vast majority of TDEs exhibit bluer peak colors and relatively slower color changes following maximum light.

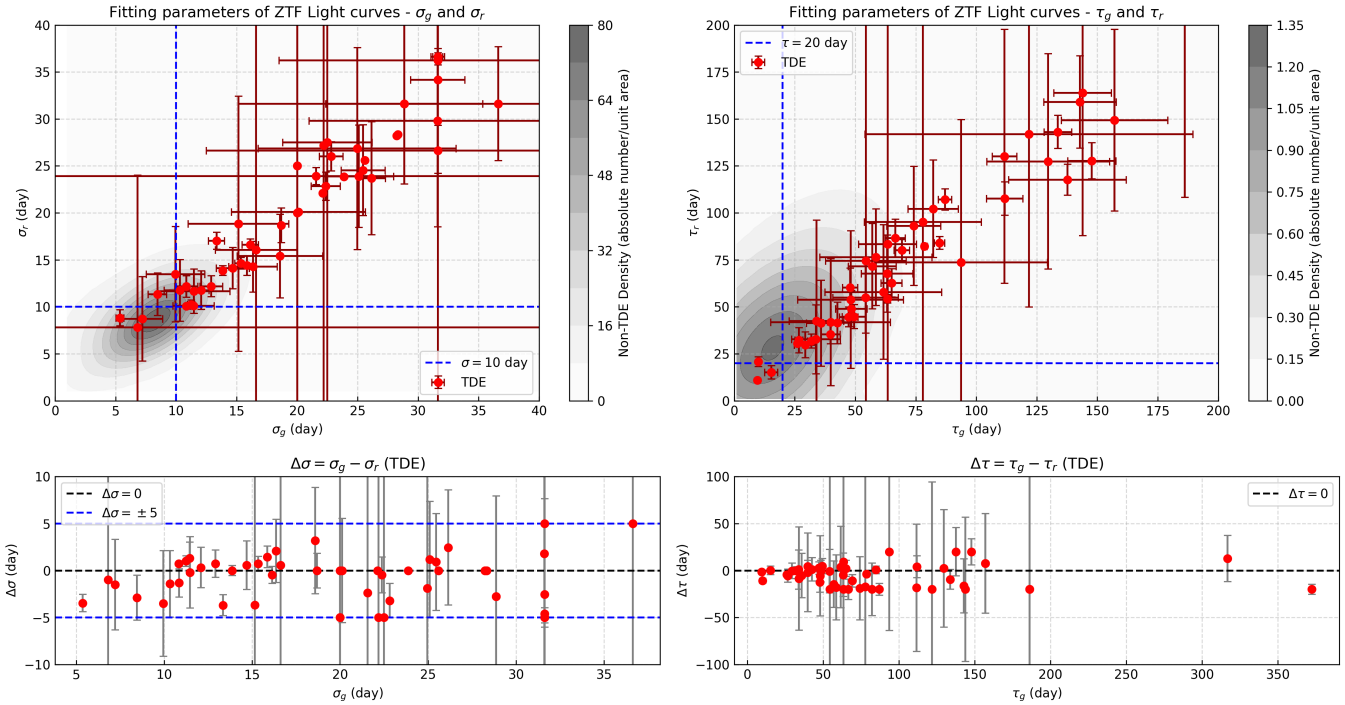

Figure B1. Distribution of the fitted parameters. The parameters σ and τ represent the characteristic rise and decay timescales, respectively, as derived from the adopted fitting function. The gray shaded contour regions represent the distribution of non-TDE sources, while TDEs are indicated by red data points. The upper panels show the correlations between σ and τ for TDE and non-TDE samples, whereas the lower panels display the distributions of $\Delta\sigma = \sigma_g - \sigma_r$ and $\Delta\tau = \tau_g - \tau_r$ for TDEs. The blue dashed lines in the upper panels denote the reference thresholds for σ and τ, and the lower panels mark the limited fitting threshold for $\Delta\sigma$.

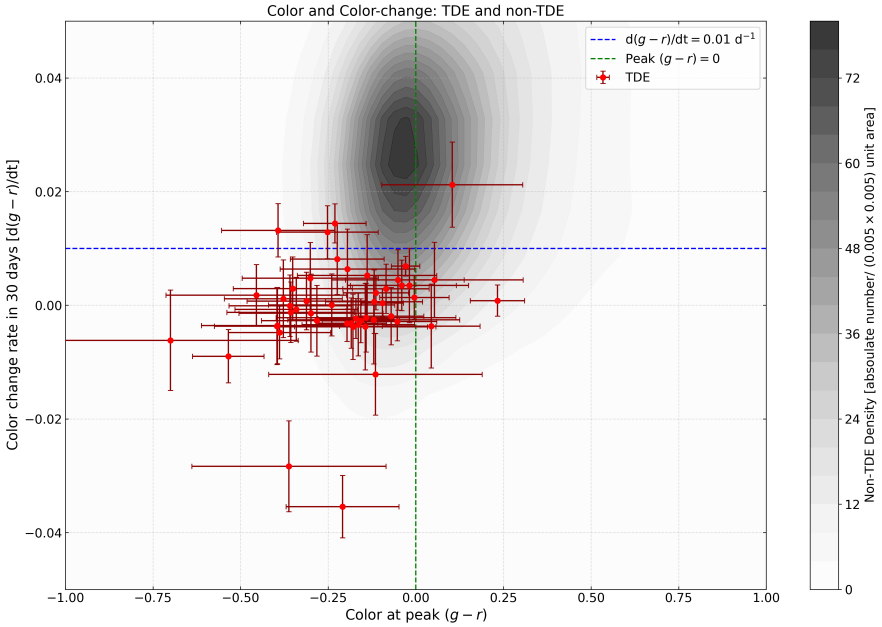

Figure B2. Peak colors $(g - r)$ and color changes within 30 days after the peak for TDEs and non-TDEs. The green dashed line marks $g - r = 0$, and the blue dashed line indicates $\mathrm{d}(g - r)/\mathrm{dt}$. The gray background shows the distribution of non-TDEs, while the red points represent TDEs. The two dashed lines divide the plot into four quadrants, with most TDEs located in the third quadrant.

Appendix C
Bonus Ability—Classify Supernovae

Supernova classification is not the central focus of this study. Nonetheless, we observe that the model performs well when applied to supernova classification in WFST data.

We obtained 137 suspected SN candidates through the function fitting criterion without color and timescale restrictions, which has a peak difference flux >0.05 mJy in both *g*-band and *r*-band data, and then we performed a cross-match with 2″ between the coordinates of supernovae reported in the TNS and the WFST database as of 2025 May 10. This resulted in the identification of 35 supernovae with well-sampled light curves in WFST—defined as having at least five photometric measurements in both the *g* and *r* bands, and confirmed by spectroscopic classification. Among these, 25 are SN Ia, 9 are SN II, and 1 is an SN Ic.

The verification results showed that the correct selection rate for SN Ia is 23/25, while that for SN II is 9/9, and that for SN Ib/c is 1/1. These results are consistent with our expectations. Notably, none of the sources were misclassified as TDE, and all exhibited TDE scores below 0.03. Therefore, we conclude that, although the model is trained on photometric data from ZTF, its applicability can also extend to WFST data.

ORCID iDs

Ranfang Zheng https://orcid.org/0009-0006-4622-1417
Zheyu Lin https://orcid.org/0000-0003-4959-1625
Xu Kong https://orcid.org/0000-0002-7660-2273
Lulu Fan https://orcid.org/0000-0003-4200-4432
Ji-an Jiang https://orcid.org/0000-0002-9092-0593
Ning Jiang https://orcid.org/0000-0002-7152-3621
Jie Lin https://orcid.org/0000-0003-3965-6931
Tinggui Wang https://orcid.org/0000-0002-1517-6792
Qingfeng Zhu https://orcid.org/0000-0003-0694-8946
Wentao Luo https://orcid.org/0000-0003-1297-6142
Jian Wang https://orcid.org/0000-0003-1617-2002
Yongquan Xue https://orcid.org/0000-0002-1935-8104
Hong-fei Zhang https://orcid.org/0000-0002-1463-9070
Wen Zhao https://orcid.org/0000-0002-1330-2329
Xianzhong Zheng https://orcid.org/0000-0003-3728-9912

References

Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bade, N., Komossa, S., & Dahlem, M. 1996, A&A, 309, L35
Blanchard, P. K., Nicholl, M., Berger, E., et al. 2017, ApJ, 843, 106
Boone, K. 2019, AJ, 158, 257
Bricman, K., & Gomboc, A. 2020, ApJ, 890, 73
Bučar Bricman, K., van Velzen, S., Nicholl, M., & Gomboc, A. 2023, ApJS, 268, 13
Cabrera-Vives, G., Moreno-Cartagena, D., Astorga, N., et al. 2024, A&A, 689, A289
Cai, M., Xu, Z., Fan, L., et al. 2025, arXiv:2501.15018
Carrasco-Davis, R., Reyes, E., Valenzuela, C., et al. 2021, AJ, 162, 231
Chan, C.-H., Piran, T., & Krolik, J. H. 2020, ApJ, 903, 17
Chang, J. N. Y., Dai, L., Pfister, H., Kar Chowdhury, R., & Natarajan, P. 2025, ApJL, 980, L22
Charalampopoulos, P., Kotak, R., Wevers, T., et al. 2024, A&A, 689, A350
Charalampopoulos, P., Pursiainen, M., Leloudas, G., et al. 2023, A&A, 673, A95
Chen, Z. H., Yan, L., Kangas, T., et al. 2023, ApJ, 943, 41
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Förster, F., Cabrera-Vives, G., Castillo-Navarrete, E., et al. 2021, AJ, 161, 242
Fragione, G., Leigh, N. W. C., Ginsburg, I., & Kocsis, B. 2018, ApJ, 867, 119
Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q., & Wilson, A. G. 2018, in NIPS'18: Proc. of the 32nd Int. Conf. on Neural Information Processing System, ed. S. Bengio et al. (Curran Associates Inc.), 7587
Gezari, S. 2021, ARA&A, 59, 21
Gomez, S., Villar, V. A., Berger, E., et al. 2023, ApJ, 949, 113
Greene, J. E., Strader, J., & Ho, L. C. 2020, ARA&A, 58, 257
Hammerstein, E., van Velzen, S., Gezari, S., et al. 2023, ApJ, 942, 9
Hoogendam, W. B., Hinkle, J. T., Shappee, B. J., et al. 2024, MNRAS, 530, 4501
Hunter, J. D. 2007, CSE, 9, 90
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Jiang, N., Zhou, Z., Zhu, J., Wang, Y., & Wang, T. 2023, ApJL, 953, L12
Jin, C. C., Li, D. Y., Jiang, N., et al. 2025, arXiv:2501.09580
Kingma, D., & Ba, J. 2014, arXiv:1412.6980
Komossa, S., & Bade, N. 1999, A&A, 343, 775
Lanza, M. L., Karpov, S., Russeil, E., et al. 2026, A&A, 705, A144
Lin, D., Strader, J., Carrasco, E. R., et al. 2018, NatAs, 2, 656
Lin, Z., Jiang, N., & Kong, X. 2022a, MNRAS, 513, 2422
Lin, Z., Jiang, N., Kong, X., et al. 2022b, ApJL, 939, L33
Lin, Z., Jiang, N., Wang, T., et al. 2024, ApJL, 971, L26
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
Malz, A. I., Hložek, R., Allam, T., et al. 2019, AJ, 158, 171
Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, PASP, 131, 018003
Masterson, M., De, K., Panagiotou, C., et al. 2024, ApJ, 961, 211
Paszke, A., Gross, S., Massa, F., et al. 2019, in NIPS'19: Proc. of the 33rd Int. Conf. on Neural Information Processing Systems, ed. H. Wallach et al. (Curran Associates, Inc.), 8024
Patra, K. C., Foley, R. J., Earl, N., et al. 2025, arXiv:2510.12572
Pavez-Herrera, M., Sánchez-Sáez, P., Hernández-García, L., et al. 2025, A&A, 696, A153
Payne, A. V., Shappee, B. J., Hinkle, J. T., et al. 2021, ApJ, 910, 125
Perez-Carrasco, M., Cabrera-Vives, G., Hernandez-García, L., et al. 2023, AJ, 166, 151
Rees, M. J. 1988, Natur, 333, 523
Ricarte, A., Tremmel, M., Natarajan, P., & Quinn, T. 2021, ApJL, 916, L18
Sánchez-Sáez, P., Reyes, I., Valenzuela, C., et al. 2021, AJ, 161, 141
Shah, V. G., Gagliano, A., Malanchev, K., Narayan, G., & Collaboration, T. L. D. E. S. 2025, ApJ, 995, 4
Sheng, X., Nicholl, M., Smith, K. W., et al. 2024, MNRAS, 531, 2474
Somalwar, J. J., Ravi, V., Yao, Y., et al. 2025, ApJ, 985, 175
Stein, R., Mahabal, A., Reusch, S., et al. 2024, ApJL, 965, L14
Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, PASJ, 70, S9
Tang, Y., Fan, L., Wan, Z., Liu, Y., & Lu, Y. 2025, AJ, 169, 304
Uchiyama, T. 2023, in 2023 IEEE Int. Conf. on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops) (IEEE), 264
van Velzen, S. 2018, ApJ, 852, 72
van Velzen, S., Gezari, S., Cenko, S. B., et al. 2019, ApJ, 872, 198
van Velzen, S., Gezari, S., Hammerstein, E., et al. 2021, ApJ, 908, 4
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wang, T., Liu, G., Cai, Z., et al. 2023, SCPMA, 66, 109512
Wang, Y., Jiang, N., Wang, T., et al. 2022, ApJL, 930, L4
Wen, J., Zhang, N., Lu, X., Hu, Z., & Huang, H. 2024, Engineering Applications of Artificial Intelligence, 133, 108633
Wen, Q., Zhou, T., Zhang, C., et al. 2023, arXiv:2202.07125
Williams, R. D., Francis, G. P., Lawrence, A., et al. 2024, RASTI, 3, 362
Yao, Y., Chornock, R., Ward, C., et al. 2025, ApJL, 985, L48
Yao, Y., Ravi, V., Gezari, S., et al. 2023, ApJL, 955, L6
Zabludoff, A., Arcavi, I., La Massa, S., et al. 2021, SSRv, 217, 54
Zheng, R. 2025, TTC: Transformer-based TDE Classifier for the Wide Field Survey Telescope (WFST), v1, Zenodo, doi:10.5281/zenodo.17775799