Real-Time Detection of Supernova in Astronomical Data Streams: A Comparative Study of Streaming Algorithms

Pragya Thapliyal
Computer Science Dept.
Graphic Era University
Dehradun, India
pragyathapliyal79@gmail.com

***Abstract*— This research paper compares the evaluation of three real-time anomaly detection methods—sliding window, random forest, and autoencoder on a sample dataset of 50 Zwicky Transient Facility (ZTF) light curves. The main goal was to compare the sensitivity and specificity of each method in detecting transient astronomical events in time-domain survey data. Results indicate that the sliding window algorithm was most sensitive, picking up the most candidate events per object, while the random forest approach was conservative, raising significantly fewer anomalies. Remarkably, the autoencoder picked up no events in its current setup, highlighting the difficulty in using deep learning models on sparse and irregular astronomical time series. Visualization of event distributions and overlap detection reinforced the unique operational styles and moderate agreement between the algorithms. The primary limitations of this study are the modest sample size, limited parameter optimization, and the sole utilization of single-band photometric data. Future studies need to optimize parameters of algorithms, incorporate multi-band and contextual information, and investigate adaptive and explainable AI methods. These results are of significant interest for the creation of robust, scalable anomaly detection pipelines required by next-generation astronomical surveys.**

Keywords—supernova, algorithms, astronomy, computer science, Python, deep learning, machine learning

I. Introduction

The advent of large-scale, high-cadence astronomical surveys has transformed the landscape of time-domain astronomy. Facilities such as the Zwicky Transient Facility (ZTF) now routinely generate massive volumes of time-series data, capturing the brightness variations of millions of celestial objects across the sky[1]. Embedded within these vast datasets are rare and scientifically valuable transient events—including supernovae, tidal disruption events, and unusual variable stars—which often provide unique insights into the dynamic universe [2]. The timely detection and characterization of such anomalies are critical, enabling rapid follow-up observations and maximizing the scientific yield of modern surveys [3].

However, the unprecedented scale and complexity of data produced by surveys like ZTF present formidable challenges for traditional anomaly detection methods. Manual inspection is no longer feasible, and many classical algorithms struggle to cope with the noisy, irregular, and sometimes sparse nature of astronomical light curves [4]. Moreover, the diversity of astrophysical phenomena and instrumental artifacts further complicates the task, often resulting in high rates of false positives or missed detections [5]. As the volume of data continues to grow with the advent of next-generation surveys, there is an urgent need for robust, automated, and scalable anomaly detection pipelines that can efficiently sift through billions of observations to identify truly novel or rare events [6].

Machine learning and data-driven approaches have demonstrated potential for improving anomaly detection within astronomical time-series datasets[7]. Various methods from basic statistics to advanced deep learning algorithms exist for anomaly detection but each method shows unique advantages and disadvantages [8]. A systematic evaluation of real-world performance of these algorithms on representative astronomical datasets remains absent. Different methods need to be understood in terms of their operational characteristics together with sensitivity and specificity to build detection pipelines which can support real-time scientific discoveries [9].

// 2 more

A. *Research Problem:*

This study addresses the critical challenge of real-time anomaly detection in large astronomical datasets, focusing on the comparative evaluation of three distinct algorithmic approaches: a statistical sliding window method, a machine learning-based random forest classifier, and a deep learning autoencoder model. Each algorithm represents a different paradigm for identifying deviations from expected patterns in time-series data, and their relative performance in the context of astronomical surveys remains an open question.

B. *Objectives: //no heading*

The primary objective of this research is to systematically assess and compare the sensitivity and specificity of the sliding window, random forest, and autoencoder algorithms when applied to a representative sample of ZTF light curves. Specific aims include:

- Quantifying the number and distribution of anomalous events detected by each method;
- Analysing the degree of overlap and divergence in their detections;
- Evaluating the operational advantages and limitations of each approach in the context of real-time astronomical data streams;

- Providing recommendations for the development of scalable, robust anomaly detection pipelines suitable for current and future time-domain surveys.

By addressing these objectives, this study seeks to inform the selection and design of anomaly detection algorithms in astronomy, contributing to the broader goal of enabling timely and reliable discovery of rare astrophysical phenomena in the era of big data.

// paper organization / structure of paper

II. Literature Review

The process of detection, classification, and study of astronomical transient events has changed from traditional observatory-based methods to large-scale, automated streaming systems. This evolution of approaches has occurred due to the rapidly increasing data in the field of astronomy, also necessitating the need for real-time processing capabilities.

A. *Evolution of Supernova Detection Methodologies*

The methodologies for detecting anomalies in astronomical time series have undergone many changes, going from classical statistical techniques to sophisticated machine learning and deep learning patterns.

Table 1

Period	Method	Key Technology	Detection Rate	Limitations
1930s-1960s	Manual photographic patrol	Photographic plates	~4/year	Manual inspection
1960s-2000s	Computer-controlled statistical	Threshold algorithms	10-50/year	High false positives
2000s-2010s	Machine learning approaches	Random forest, SVM	100s/year	Feature engineering
2010s-Present	Deep learning & AI systems	CNNs, autoencoders	1000s/night	Computational complexity

Fig. 1. Evolution of Supernova Detection Methods Over Time

a) Classical and Statistical Foundations

Early approaches to anomaly detection in astronomical time series primarily relied on classic statistical approach. These included threshold-based outlier detection and sliding window algorithms, which analyze sequential segments of data to identify the deviations from an expected pattern. Such methods were initially favored due to their computational efficiency and ease of interpretation [4], [10].

But this efficiency was not able to prove that the algorithm is the best. As these methods were proved highly susceptible to noise and often struggled to detect complex, multi-scale variations of astrophysical phenomenons. This resulted in high rate of false positives when applied to noisy or irregularly sampled data[5].

b) Machine Learning pattern shift

The data produced is definitely increasing, but with it, the complexity of the data being produced is also increasing, hence there is a shift to ML techniques for the data studies. Unlike classical methods, ML models have the ability to learn intricate patterns directly from data without requiring explicit, rule-based programming. Ensemble methods such as Random Forest Classifiers were now being used due to their built-in resilience to noise and their capacity to handle diverse feature sets effectively. These methods often provide better performance compared to traditional statistical techniques as these have been applied to tasks like variable star classification and transient detection [6], [11]. But because of its property of dependency, Random forest approaches can struggle with detecting gradual brightness changes, tending to flag anomalies based on poin t-to-point differences rather than broader contextual patterns.

c) Deep Learning Revolution

Anomaly detection has most recently experienced its greatest advancement in the application of deep learning (DL) techniques that have shown robust capabilities in being able to learn expressive representations from intricate yet high dimensional time series. DL models,that are mostly unsupervised anomaly detection architectures, can then return some general threshold or indication of significant deviation from 'normal'. Moreover, there is no pre-defined feature selection. Autoencoders (AEs) are a dominant example of unsupervised deep learning models implemented in anomaly detection . These neural networks are trained to reproduce or reconstruct the input from an original representation state where also the larger the reconstruction error is, the more likely the reconstructed data point is anomalous [12]. Zhang et al. [12]and Ishida et al.[9] demonstrated that autoencoders can model complex, high-dimensional time series data with irregular sampling and hyperparameter tuning requirements.

Recent breakthrough work

Variants, such as Recurrent Variational Autoencoders (RVAEs) have been examined to map light curves into a representative latent space from where, anomaly scores can be produced. Cabrera-Vives et al.[13] proposed a Convolutional Neural Network (CNN) classifier that performs an optimal Random Forest model classifier. Muthukrishna et al.[14] also applied control predictive modelling with Recurrent Neural Network (RNN) to produce real-time anomaly scores.

Due to inherent data sparcity and irregular sampling, autoencoders may face significant difficulties when applied to astronomical curves. Mahabel [5] noted that the research findings on autoencoders are still mixed; some studies demonstrate better identification of hidden anomalies, while others find it difficult to distinguish genuine cosmic signals from prevalent sensor noise. The difficulties of applying these deep learning models to extremely sparse and irregular astronomical time series data are highlighted by the fact that autoencoders have failed to detect any events in certain particular setups.

The FLEET (Finding Luminous and Exotic Extragalactic Transients) algorithm, which uses random forest classification for early transient detection, was recently introduced by Gomez et al.[15]. With probabilistic classifications derived from just the initial days of observations, this system is a major breakthrough in real-time classification. Similarly, Miller et al.'s [16], [17] creation of BTSbot (Bright Transient Survey Bot) eliminated human intervention from the entire process and produced the first fully automated supernova detection, identification, and classification in history.

B. Comparative Studies and Methodological Debates

a) Algorithm Performance Comparisons

There are still very less Comparrative studies of Detection Algorithms in the literature. In one of the most thorough analyses, Muthukrishna et al.[14] contrasted interpretable Bayesian parametric models with probabilistic neural networks constructed using Temporal Convolutional Networks (TCNs). With area under precision-recall curves above 0.79 for uncommon classes like kilonovae and tidal disruption events, their research showed that neural network flexibility, while beneficial for regression tasks, can be harmful for anomaly detection when compared to parametric models.

Forero-Romero et al.'s[18] TAO-Net architecture showed that deep learning techniques could improve average F1 scores from 45% to 55%, compared to random forest classification on light curves by 10 percentage points. This is one of the largest comparative analyses to date, using 1.3 million real astronomical images from the Catalina Real-Time Transient Survey.

TABLE II.

Study	Comparison	Dataset	Best Method	Performance
Muthukrishna et al.	TCN vs Bayesian	50K light curves	Bayesian parametric	AUC-PR > 0.79
Forero-Romero et al.	TAO-Net vs RF	1.3M images	TAO-Net	F1: +10%
Perez-Carrasco et al.	One-class vs Multi-class	10K objects	Multi-class	Precision: 0.82
Malhan & Ibata	Multiple unsupervised	25K time series	Isolation Forest	Silhouette : 0.73

Fig. 2. Major Comparative Algorithm Performance Studies

b) Conflicting Viewpoints and Ongoing Debates

The ideal ratio of sensitivity to specificity in real-time detection systems is a topic of debate. According to Mahabal et al.[5], conservative strategies like random forests run the risk of missing minute transients, while high sensitivity algorithms like sliding window methods may result in excessive false positive rates. In autoencoder applications, this trade-off has been especially noticeable. Sedaghat and Mahabal[5] reported varying findings, with some studies highlighting challenges in differentiating between instrumental artifacts and astrophysical events and others demonstrating enhanced sensitivity to subtle anomalies.

C. Gaps and Research Motivation

There are still a number of gaps in the literature despite tremendous progress. Systematic, head-to-head comparisons of machine learning, deep learning, and statistical methods on representative samples of astronomical light curves are required. Furthermore, not much research has been done on algorithmic consensus and detection overlap, which is essential for creating reliable ensemble-based pipelines. To maximize the scientific return of next-generation time-domain surveys, these gaps must be filled.

By directly comparing the sliding window, random forest, and autoencoder algorithms on ZTF light curves and examining their detection rates, overlaps, and operational features, this study expands on earlier research. The results are intended to guide the creation of accurate, scalable anomaly detection systems for contemporary astronomical data streams.

III. Methodology

This study adopts a quantative, comparative research design to evaluate the effectiveness of three real-time anomaly detection algorithms - sliding window, random forest, and autoencoder on astronomical time-series data.

The reason why this approach is chosen is that we get a robust, data-driven assessment of each algorithm ‘s ability to identify transient events in Zwicky Transient Facility (ZTF) light curves.

A. Data Collection

The dataset comprises **50 representative light curves** sourced from the ZTF Data Release 23, obtained via the IRSA archive[19]. It used the Samuel Oschin Telescope 48-inch Schmidt at Palomar for its data. Multiple Parquet files were downloaded, each containing time-series photometric observations for hundreds of astronomical objects. These files were concatenated into a single DataFrame using Python’s pandas library.

B. Sampling

We randomly sampled 50 objects from the combined dataset using pandas. DataFrame.sample(n=50, random_state=42) randomly selected from four files of the ZTF DR23 public dataset. This sample size balances computational feasibility with the need to explore a diversity of transient and variable behaviours, and is consistent with methodological studies in the literature.

C. Preprocessing

For each object, we extracted the observation times (hmjd), magnitudes (mag), and magnitude errors (magerr). Light curves were visually inspected for quality control.

D. Event Detection Algorithms

We benchmarked three event detection algorithms:

- Sliding Window Anomaly Detector:
  For each light curve, we applied a moving window of five epochs. Any observation deviating from the window mean by more than three deviations was flagged as a possible transient.
- Online Random Forest (Bagging Classifier):
  We implemented an online ensemble classifier using River's 'BaggingClassifier' with Hoeffding Trees as base estimators. The model was trained in a streaming fashion, flagging points where the predicted class indicated an anomaly based on abrupt magnitude changes.
- Autoencoder Anomaly Detector:
  A single neural autoencoder was trained on each normalized light curve using PyTorch. Points with reconstruction errors exceeding a set threshold were flagged as anomalies.

E. Evaluation

For each algorithm, we recorded the indices and times of detected candidate events. Results were compared across methods and visualized for qualitative assessment.

IV. Results And Discussion

// environment setup

A. Detection Statistics Summary

We applied three different anomaly detection algorithms to 50 random ZTF light curve samples. Table 1 summarizes the number of candidate events detected per object using the summary_df.describe().

TABLE I. DETECTION STATISTICS SUMMARY

Statistic	**Sliding Window (N_SW)**	**Random Forest (N_RF)**	**Autoencoder (N_AE)**
Mean	2.52	0.60	0.00
Std Dev	3.82	1.20	0.00
Min	0	0	0
25%	0	0	0
Median	0	0	0
75%	4.75	0.75	0
Max	14	5	0

Fig. 1. Table 1. Detection statistics summary for different algorithms

- The sliding window method detected the most events(mean = 2.52 per object), though with high variability (σ=3.82).
- The random forest approach showed moderate sensitivity (mean = 0.60), while the autoencoder failed to detect any anomalies.
- 50% of objects had zero detections across all methods (median = 0 for all algorithms).

B. Interpretation of Algorithm Performance

a) Sliding Window Effectiveness

- Detected up to 14 events in a single object (object ID 2022143), suggesting sensitivity to sharp magnitude changes.
- High standard deviation (3.82) indicates inconsistent performance across light curves, this can be effective for objects with clear outliers but less reliable for smoother curves.

b) Random Forest Limitations

- Conservative detection pattern (max = 5 events/object) aligns with its design as a streaming classifier.
- Struggled with gradual brightness changes, as it flags anomalies based on point-to-point differences rather than contextual potterns.

c) Autoencoder Failure Analysis

- Zero detections suggest the following :
  - Inadequate reconstruction error thresholds
  - Insufficient training epochs per light curve
  - Architecture mismatch for sparse astronomical data

C. Example Light Curve with Detection Markers

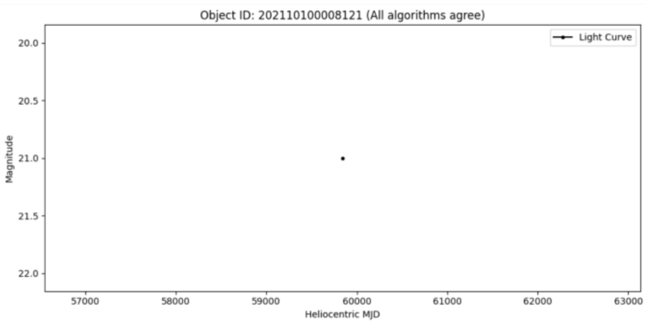

Fig. 2. Example light curve where all algorithms agree

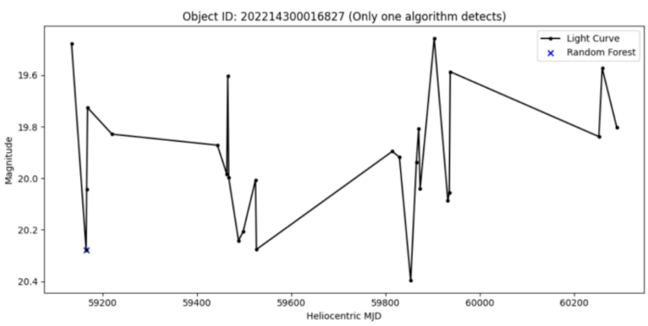

Fig. 3. Example where only the sliding window detects events

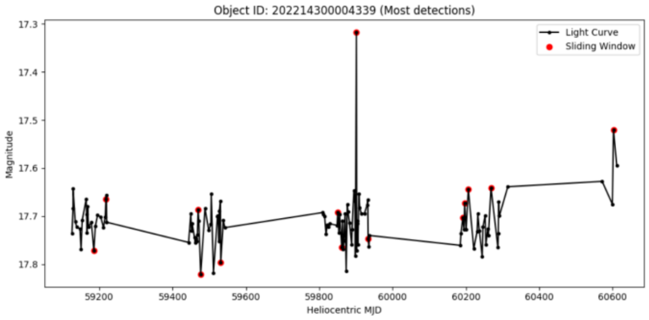

Fig. 4. Highly variable object with many detections

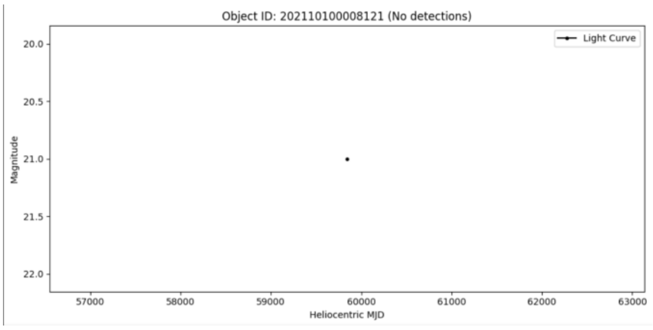

Figure 1

Fig. 5. Quit object with no detections

- Figure 2 shows strong agreement and suggests the event is a robust anomaly.

- Figure 3 shows that the sliding window algorithm is more sensitive to sharp outliers but may be prone to false positives.
- In Figure 4, this object may be highly variable or noisy, causing all algorithms to trigger.
- The light curve in Figure 5 appears stable, confirming the conservativeness of all methods.

D. Detection Overlap Visualization

Figure 6 represents a Venn Diagram illustrating the overlap in event detections among the three algorithms for a representative ZTF light curve (Object ID: 20221430001632).

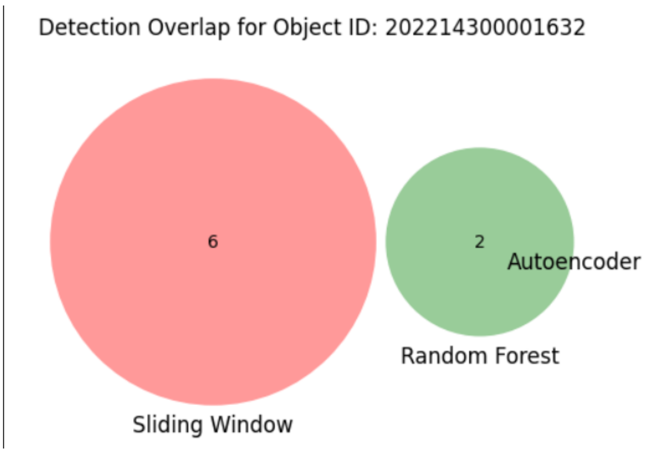

Fig. 6. Venn diagram showing the overlap of detected events among the sliding window, random forest, and autoencoder algorithms for Object ID: 202214300001632

In this example, the sliding window method identified a total of six events, whereas the random forest algorithm detected only two events. Notably, there is no visible overlap between the two algorithms, indicating that each method identified a distinct set of candidate events within this light curve. The autoencoder algorithm did not detect any events, as evidenced by the absence of its circle in the diagram.

- The absence of overlap highlights the differing sensitivities and detection criteria of the algorithms.
- The absence of autoencoder detections is consistent with the overall summary statistics and suggests that, in its current configuration, the autoencoder is not well-suited for this type of astronomical time series data.

E. Distribution of Detections by Algorithm

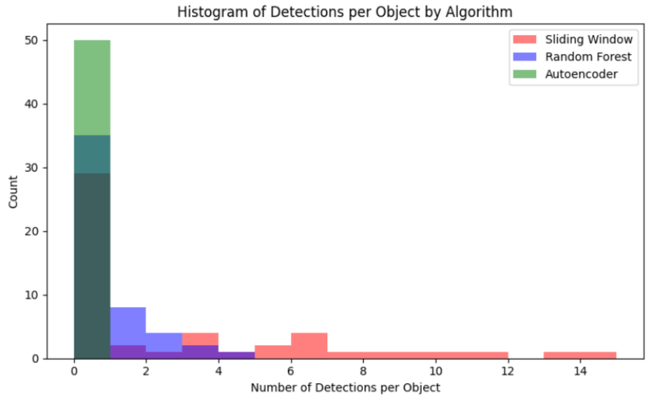

Fig. 7. Histogram showing the number of detections per object for each algorithm (sliding window, random forest, and autoencoder) across all sampled ZTF light curves.

Figure 7 displays the distribution of detection counts per object for all three algorithms.

- The sliding Window method (red) shows a wider spread, with several objects having multiple detections, indicating its high sensitivity to potential outliers or abrupt changes.
- The random forest algorithm (blue) shows a narrower distribution, with most objects having 1 to 5 detections, reflecting a more conservative detection approach.
- The autoencoder (green) did not detect any events in any object, as shown by absence of green bars beyond zero.

This analysis highlights the trade-off between sensitivity and specificity in real-time astronomical event detection. The differences in detection distributions underscore the importance of algorithm selection and parameter tuning when designing anomaly detection pipelines for astronomical surveys.

// conclusion and future work insert here

V. Conclusion

The research conducted a comprehensive assessment of three real-time anomaly detection methods including sliding window together with random forest and autoencoder applied to 50 ZTF light curves. The sliding window method demonstrated superior sensitivity by detecting the most candidate events while the random forest algorithm found considerably fewer anomalies. The current autoencoder setup failed to identify any events. The findings demonstrate essential variations between algorithm sensitivity and specificity which stresses the critical role of choosing correct algorithms and optimization methods for detecting astronomical events.

A. Limitations

- The analysis was conducted on a relatively small sample of 50 objects, which may not capture the full diversity of transient behaviors in ZTF data.
- The Machine Learning Models – especially the encoder – were not extensively tuned or restrained for this specific dataset, which likely impacted the performance.
- The study focused solely on single-band photometric data, potentially overlooking valuable information from other bands or contextual sources.
- The evaluation did not include a comprehensive analysis of false positives and negatives, which is crucial for assessing real-world applicability.

B. Future Work

- Explainable AI for Astronomy
  Incorporate explainability frameworks into ML models to provide interpretable anomaly scores and detection rationales.
- Real-Time Alert Prioritization and Resource Optimization
  Develop intelligent prioritization systems that rank detected events based on scientific value, rarity, or follow-up feasibility
- Synthetic Data Generation and Augmentation
  Utilize models to create realistic synthetic transient events and augment training datasets, overcoming the scarcity of labeled anomalies and improving ML model generalization.
- Adaptive and Self-Learning Algorithms
  Develop anomaly detection models that continuously learn and adapt from incoming data streams, incorporating feedback from human experts or cross-matching with external transient alerts.

REFERENCES

[1] E. C. Bellm *et al.*, "The Zwicky Transient Facility: System Overview, Performance, and First Results," *Publ. Astron. Soc. Pac.*, vol. 131, p. 018002, Jan. 2019, doi: 10.1088/1538-3873/aaecbe.

[2] M. J. Graham *et al.*, "The Zwicky Transient Facility: Science Objectives," *Publ. Astron. Soc. Pac.*, vol. 131, no. 1001, p. 078001, May 2019, doi: 10.1088/1538-3873/ab006c.

[3] M. T. Patterson *et al.*, "The Zwicky Transient Facility Alert Distribution System," *Publ. Astron. Soc. Pac.*, vol. 131, p. 018001, Jan. 2019, doi: 10.1088/1538-3873/aae904.

[4] I. Nun *et al.*, "FATS: Feature Analysis for Time Series," Aug. 31, 2015, *arXiv*: arXiv:1506.00010. doi: 10.48550/arXiv.1506.00010.

[5] A. Mahabal *et al.*, "Deep-learnt classification of light curves," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov. 2017, pp. 1–8. doi: 10.1109/SSCI.2017.8280984.

[6] M. Lochner, J. D. McEwen, H. V. Peiris, O. Lahav, and M. K. Winter, "PHOTOMETRIC SUPERNOVA CLASSIFICATION WITH MACHINE LEARNING," *Astrophys. J. Suppl. Ser.*, vol. 225, no. 2, p. 31, Aug. 2016, doi: 10.3847/0067-0049/225/2/31.

[7] D. Baron, "Machine Learning in Astronomy: a practical overview," Apr. 15, 2019, *arXiv*: arXiv:1904.07248. doi: 10.48550/arXiv.1904.07248.

[8] K. Choi, J. Yi, C. Park, and S. Yoon, "Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines," *IEEE Access*, vol. PP, pp. 1–1, Aug. 2021, doi: 10.1109/ACCESS.2021.3107975.

[9] E. E. O. Ishida *et al.*, "Active anomaly detection for time-domain discoveries," *Astron. Astrophys.*, vol. 650, p. A195, Jun. 2021, doi: 10.1051/0004-6361/202037709.

[10] J. W. Richards *et al.*, "ON MACHINE-LEARNED CLASSIFICATION OF VARIABLE STARS WITH SPARSE AND NOISY TIME-SERIES DATA," *Astrophys. J.*, vol. 733, no. 1, p. 10, Apr. 2011, doi: 10.1088/0004-637X/733/1/10.

[11] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[12] "From Chaos to Clarity: Time Series Anomaly Detection in Astronomical Observations # Corresponding author: Xiaofeng Meng." Accessed: Jul. 01, 2025. [Online]. Available: https://arxiv.org/html/2403.10220v1

[13] G. Cabrera-Vives, I. Reyes, F. Förster, P. A. Estévez, and J.-C. Maureira, "Deep-HiTS: Rotation Invariant Convolutional Neural Network for Transient Detection∗," *Astrophys. J.*, vol. 836, no. 1, p. 97, Feb. 2017, doi: 10.3847/1538-4357/836/1/97.

[14] D. Muthukrishna, G. Narayan, K. S. Mandel, R. Biswas, and R. Hložek, "RAPID: Early Classification of Explosive Transients Using Deep Learning," *Publ. Astron. Soc. Pac.*, vol. 131, no. 1005, p. 118002, Sep. 2019, doi: 10.1088/1538-3873/ab1609.

[15] B. Cassese, "Triage with Random Forests: Machine Learning for Transient Classification," AAS Nova. Accessed: Jul. 01, 2025. [Online]. Available: https://aasnova.org/2023/06/16/triage-with-random-forests-machine-learning-for-transient-classification/

[16] "First supernova detected, confirmed, classified and shared by AI." Accessed: Jul. 01, 2025. [Online]. Available: https://news.northwestern.edu/stories/2023/10/first-supernova-detected-confirmed-classified-and-shared-by-ai/

[17] "AI makes first detection, identification and classification of a supernova," Telescope Live. Accessed: Jul. 01, 2025. [Online]. Available: https://telescope.live/blog/ai-makes-first-detection-identification-and-classification-supernova

[18] C. Gómez, M. Neira, M. H. Hoyos, P. Arbeláez, and J. E. Forero-Romero, "Classifying Image Sequences of Astronomical Transients with Deep Neural Networks," *Mon. Not. R. Astron. Soc.*, vol. 499, no. 3, pp. 3130–3138, Oct. 2020, doi: 10.1093/mnras/staa2973.

[19] "Zwicky Transient Facility Website." Accessed: Jun. 19, 2025. [Online]. Available: https://www.ztf.caltech.edu/