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 200 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, transients, algorithms, astronomy, computer science, Python, deep learning, machine learning

I. Introduction

A. The Era of Large-Scale Time-Domain Astronomy

With the emergence of extensive, high-cadence sky surveys that produce terabytes of data annually, contemporary astronomy is also changing its paradigms to monitor the sky with efficiency[1], [2]. Facilities such as the Zwicky Transient Facility (ZTF) and the forthcoming Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) represent a fundamental paradigm shift from traditional targeted observations to comprehensive, automated sky surveys that capture the universe in motion [3], [4].This paradigm shift has unlocked extraordinary directions for scientific exploration while simultaneously posing difficulties in data management and analytical methodologies.

Modern time-domain surveys generate data at an unprecedented scale. For example, the Zwicky Transient Facility (ZTF) currently produces up to one million transient alerts per night across two photometric bands, while the Legacy Survey of Space and Time (LSST) is projected to generate over 10 million alerts nightly when it begins operations[5], [6]. These facilities scan thousands of square degrees of sky spanning timescales from fleeting stellar flares to month-long supernova explosions. These surveys create an unprecedented census of astronomical variability through repeated observations of large areas of the sky at intervals of hours to days[7].

Table I. Current and Future Survey Capabilities

Survey	Telescope Aperture	Field of View	Cadence	Limiting Magnitude	Alert Rate
ZTF	1.2m	47 deg²	3 days	20.5-21 mag	1M/night
LSST	8.4m	9.6 deg²	3-4 days	27.5 mag	10M/night
Pan-STARRS	1.8m	7 deg²	Variable	22-24 mag	1000s/night
ATLAS	0.5m	28 deg²	2 nights	19-20 mag	100s/night

Table 1. Comparison of major astronomical surveys showing telescopic specifications, observational parameters, and expected data production rates [5], [6], [8], [9], [10]

B. The Supernova Discovery Challenge

Among the various phenomena captured by these surveys, supernova has its own identity and detection complexity. These stellar explosions serve as cosmic laboratories for understanding stellar evolution, nuclear physics, and cosmology, also serving as crucial instruments for determining cosmic distances and understanding the characteristics of dark energy[11]. However, the discovery and classification of supernovae present more challenges in astronomical transient detection.

There are different types of supernovae based on light curve morphologies, peak luminosities, and temporal evolution patterns. Their classification requires algorithms capable of distinguishing between genuine astrophysical events and instrumental artifacts[12], [13]. Type Ia supernovae that are crucial for cosmological distance measurements must be rapidly identified and followed up to capture their early-time spectral evolution[14]. Information on massive star evolution and explosive nucleosynthesis is provided by Core-collapse supernovae. While exotic types such as superluminous supernovae and pair-instability supernovae offer glimpses into extreme astrophysical conditions.

Table II. Supernova Classification Requirements

Supernova Type	Peak Magnitude	Rise Time	Scientific Priority	Classification Challenge
Type Ia	-19 to -20	15-20 days	Cosmology	Spectroscopic confirmation
Type IIn	-18 to -22	20-100 days	Stellar evolution	Narrow emission lines
Type IIb	-17 to -18	10-15 days	Progenitor systems	Hydrogen loss mechanism
Superluminous	-21 to -23	30-70 days	Exotic physics	Extreme luminosity

Table 2. Scientific characteristics and observational requirements for major supernova classes, demonstrating the complexity of automated classification in astronomical transient detection[15], [16], [17], [18]

Due to the temporal constraints imposed by supernova evolution, early detection and classification are crucial. The best scientific results are obtained by early identification, as many spectroscopic features evolve rapidly within the first weeks of the explosion [12], [19]. Moreover, the limited resources available for follow-up observations necessitate the implementation of intelligent prioritization systems. These systems should be capable of differentiating the most scientifically valuable targets from the extensive array of transient candidates.

C. *Algorithmic Approaches to Real-Time Detection*

To address these problems, sophisticated algorithmic approaches for automated transient detection and classification are being adopted[20], [21], [22]. These methods range from classical statistical techniques to machine learning and deep learning architectures, each having its distinct limitations and advantages for different aspects of the detection problem.

Statistical methods, such as sliding window algorithms and threshold-based detection, are fast and can work in real time with little delay[23]. These methods work best when the noise properties are well-known and the observing conditions are stable. However, they have trouble with the irregular sampling and complicated systematic effects that are common in modern astronomical surveys.

Machine learning methods, especially ensemble methods like Random Forests, proved to be better at dealing with the noisy data and different feature spaces that are common in astronomy[24]. The FLEET (Finding Luminous and Exotic Extragalactic Transients) algorithm is a great example of this method. It has had great success in finding superluminous supernovae through early-time photometric classification. But these supervised methods need a lot of labeled training sets and can't find things that aren't in their training distributions[24].

Deep learning techniques are the latest method to find astronomical transients. They let us learn complex patterns directly from raw observational data in ways that have never been possible before[12], [25]. Convolutional Neural Networks (CNNs) are great for finding transients in images, and recurrent architectures are great for modeling how light curves change over time[26]. BTSbot's recent success in fully automating the discovery of supernovae shows that these methods are ready for use in the real world[13], [19].

D. *Streaming Data Challenges and Real-Time Constraints*

Real-time operation requirements create fundamental constraints that set astronomical transient detection apart from other machine learning applications[27]. In retrospective analysis, all the data is available, whereas real-time systems have to make decisions based on incomplete and changing information streams. This time-based causality limit has an impact on every part of algorithm design, from obtaining features to making predictions and figuring out how certain they are.

The amount and acceleration of modern survey data streams make it even harder for real-time processing systems to work[28]. ZTF's alert distribution system needs to be able to handle and send out millions of alerts within minutes of being observed. This means it needs very efficient algorithms and computing power[4]. The expected data rates from LSST will push these requirements even further, requiring new algorithms that can keep up with scientific performance while meeting strict latency constraints.

E. *Comparative Analysis Framework*

This research addresses the need for a systematic comparison of streaming algorithms for real-time supernova detection through a comprehensive evaluation of three distinct algorithmic paradigms: sliding window methods, machine learning-based Random Forest classifiers, and deep learning autoencoder architectures. Each approach represents a distinct perspective on balancing computational efficiency, detection sensitivity, and discovery potential within the context of extensive astronomical surveys.

The sliding window approach represents the classical statistical paradigm, providing computational simplicity and real-time performance at the expense of limited pattern recognition capabilities. Random Forest methods, which are currently the mainstream in astronomical machine learning, offer robust performance on heterogeneous data, but they need substantial training overhead and labeled datasets. Autoencoder architectures offer unsupervised discovery, but they pose challenges related to training stability and interpretability when applied to sparse astronomical time series.

Using representative data from the Zwicky Transient Facility, our comparative framework assesses these methods in different ways such as detection sensitivity, false positive rates, computational demands, and operational limitations. This analysis offers important insights for the design of next-generation transient detection systems that will have to meet the astronomical community's demands for interpretability and reliability while managing the previously unknown data volumes and discovery requirements of future surveys.

The impending commissioning of LSST and other next-generation facilities, which will drastically alter the field of time-domain astronomy, highlights the urgency of this research. The scientific output of these billion-dollar investments will be determined by the algorithmic frameworks developed and validated through studies like these.

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 II

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 [25], [26].

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[31].

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 [32], [33]. 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 [34]. Zhang et al. [34]and Ishida et al.[35] 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.[36] proposed a Convolutional Neural Network (CNN) classifier that performs an optimal Random Forest model classifier. Muthukrishna et al.[37] 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 [31] 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.[38]. 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 [13], [19] 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.[37] 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[39] 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.[31], 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[31] 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 **200 representative light curves** sourced from the ZTF Data Release 23, obtained via the IRSA archive[6]. 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

A. Environment Setup

The analysis was conducted using a Python-based computational environment within Jupyter Notebook, which facilitated interactive data processing, visualization, and algorithm implementation. The primary libraries used were pandas for data manipulation, NumPy for numerical operations, and Matplotlib and Seaborn for visualization. Machine learning and anomaly detection algorithms were implemented using scikit-learn and PyTorch for deep learning models.

The dataset comprised time-series photometric observations from the ZTF, accessed and processed using custom Python scripts. The computational experiments were performed on a system equipped with enough storage and processing power to handle multi-algorithm comparative study on 200 sampled objects.

This environment setup ensured efficient iterative experimentation and reproducibility of results, enabling seamless integration of data processing, feature engineering, and application of five distinct anomaly detection algorithms.

B. Summary of Anomaly Detection Results

The five anomaly detection algorithms-Sliding Window, Random Forest, Isolation Forest, One-Class SVM, and Autoencoder were applied to a sample of 200 objects from the ZTF time-series dataset. The total number of objects flagged as anomalous is summarized in Table 7.

Table III. NO. OF OBJECTS FLAGGED AS ANOMALOUS BY EACH ALGORITHM

Algorithm	**Objects Flagged**
Sliding Window	76
Random Forest	0
Isolation Forest	10
One-Class SVM	12
Autoencoder	0

Table 7. Flagged objects by each algorithm

- The Sliding Window algorithm detected the most number of anomalies, reflecting its sensitivity to local fluctuations in the light curves.

- Random Forest and Isolation Forest flagged fewer objects, indicating a more conservative detection approach.

- The One-class SVM and Autoencoder identified zero number of anomalies, consistent with their modelling assumptions and sensitivity to data sparsity.

C. Overlap and Consensus Among Algorithms

The consensus analysis reveals substantial divergence in the objects identified as anomalous by the five algorithms.

Table 8 shows that a majority of objects (121 out of 200) were not flagged by any algorithm, indicating that most of the sample was considered normal by all the five algorithms. 68 objects were flagged anomalous by only a single algorithm, highlighting that most detections were unique to individual methods. Only a small number of objects were flagged by two (3 objects) or three (8 objects) algorithms, and none were flagged by four or all five algorithms. This low consensus underscores the distinct detection strategies and sensitivities of the algorithms, with little overlap in their anomaly selections.

Table IV. OVERLAP OF ANOMALIES DETECTED (CONSENSUS TABLE)

Number of Algorithms Flagging	**Number of Objects**
0	121
1	68
2	3
3	8

Table 8. Distribution of objects flagged by varying numbers of algorithms

Table 3 further illustrates the limited agreement between methods. The Sliding Window algorithm showed some overlap with Isolation Forest and One-Class SVM (9 objects in common with each), but no overlap with Random Forest or Autoencoder. Similarly, Isolation Forest and One-Class SVM shared 9 anomalous objects, but neither overlapped with Random Forest or Autoencoder. Notably, Random Forest and Autoencoder did not share any anomalous objects with any other algorithm, nor with each other.

These results indicate that each algorithm is sensitive to different types of anomalies. That ensemble or consensus-based approaches may be necessary to capture a broader range of unusual behavior in the data.

TABLE V. PAIRWISE OVERLAP BETWEEN ALGORITHMS

Algorithm 1	Algorithm 2	Objects in Both
Sliding Window	Random Forest	0
Sliding Window	Isolation Forest	9
Sliding Window	One-Class SVM	9
Sliding Window	Autoencoder	0
Random Forest	Isolation Forest	0
Random Forest	One-Class SVM	0
Random Forest	Autoencoder	0
Isolation Forest	One-Class SVM	9
Isolation Forest	Autoencoder	0
One-Class SVM	Autoencoder	0

Table 9. Number of objects jointly flagged by each pair of algorithms

- These findings demonstrate that while some algorithms (Isolation Forest and One-Class SVM) have much more number of flagged anomalous objects in common, the overall agreement is low. This highlights the complementary nature of different anomaly detection strategies and suggests that a multi-algorithm or ensemble approach may be more effective for comprehensive anomaly identification in astronomical time-series data.

D. *Visualization of Anomaly Detection Results*

Figure 1 illustrates the number of objects flagged as anomalous by each of the five algorithms. The bar chart highlights the relative sensitivity of each method. The Sliding Window detects the largest number of anomalies, and Random Forest and Autoencoder are the most conservative.

The differences in detection rates affect the distinct operational principles and thresholds of each algorithm.

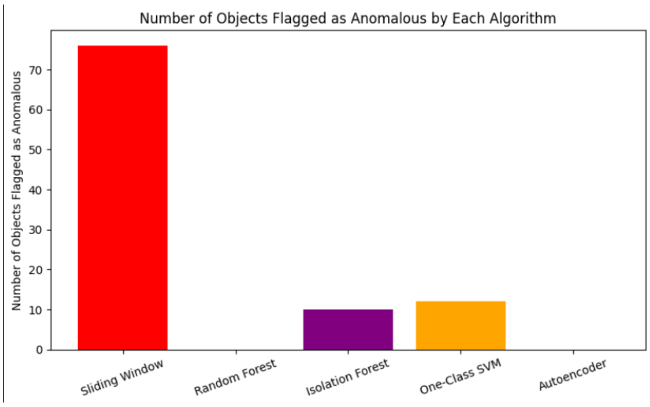


Fig. 1. Bar chart of anomalies detected per algorithm

Figure 2 presents the distribution of how many algorithms flagged each object as anomalous. This histogram provides insight into the consensus levels across the sample, showing how many objects were uniquely flagged by a single algorithm versus those flagged by multiple algorithms.

Most objects were flagged by zero or one algorithm, indicating limited agreement among the methods, while a smaller number of objects were flagged by two or more algorithms, representing higher-confidence anomalies.

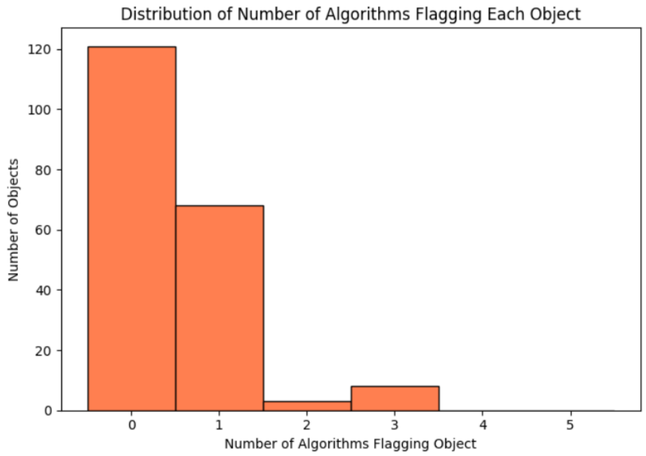


Fig. 2. Histogram showing the distribution of the number of algorithms flagging each object as anomalous

Figure 3 helps to further illustrate the degree of alignment and uniqueness among these algorithms with the help of a venn diagram. It summarizes the overlap in flagged objects by three algorithms - sliding window, Isolation Forest, and One-Class SVM.

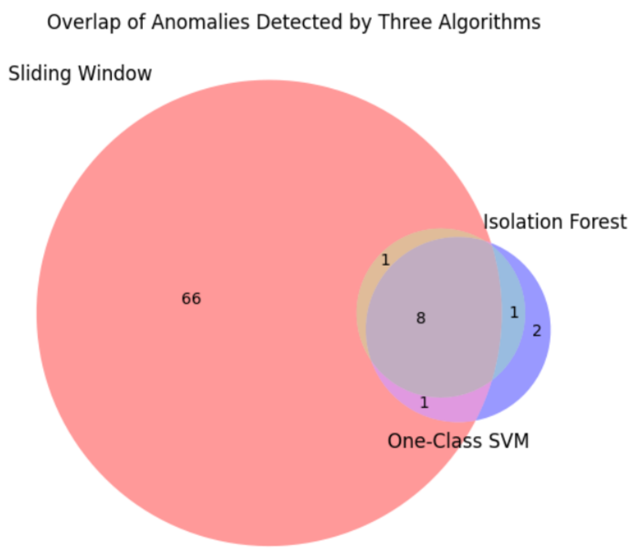
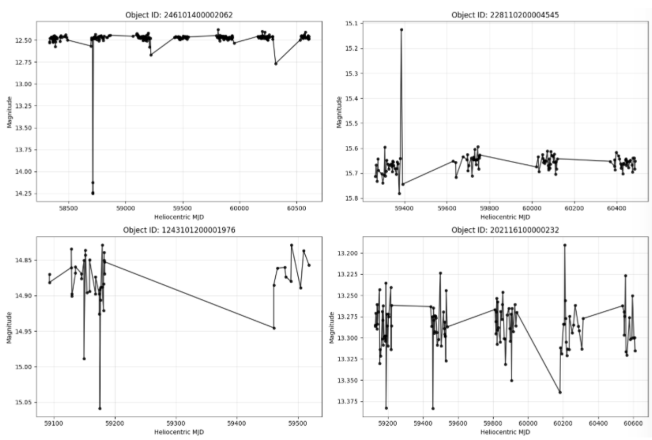


Fig. 3. Venn diagram showing the overlap of objects flagged as anomalous by Sliding Window, Isolation Forest, and One-Class SVM



//

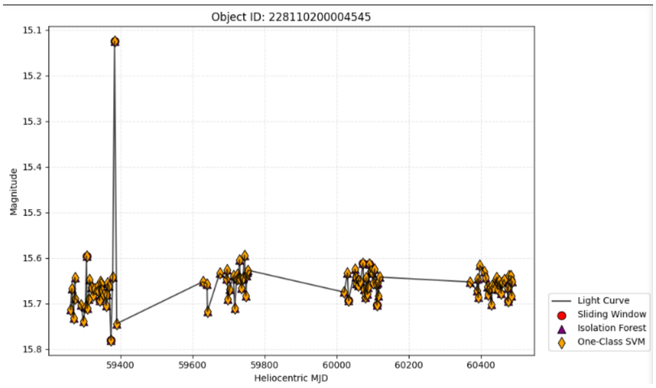


Fig. 4.

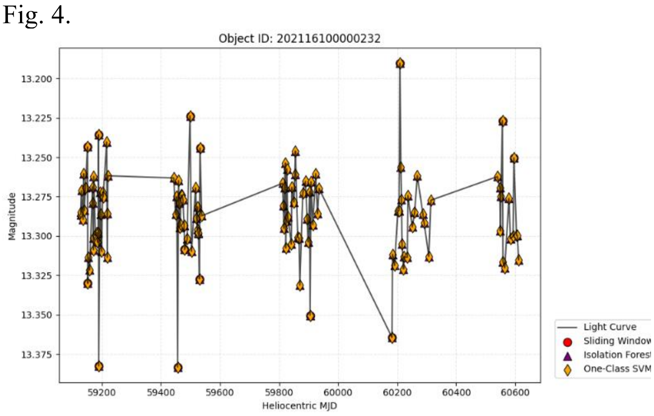


Fig. 5

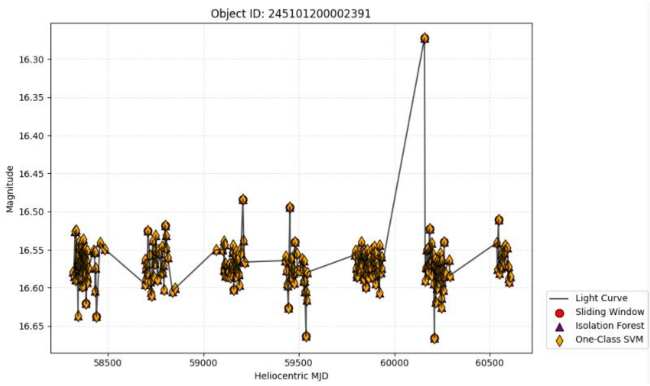


Fig. 6

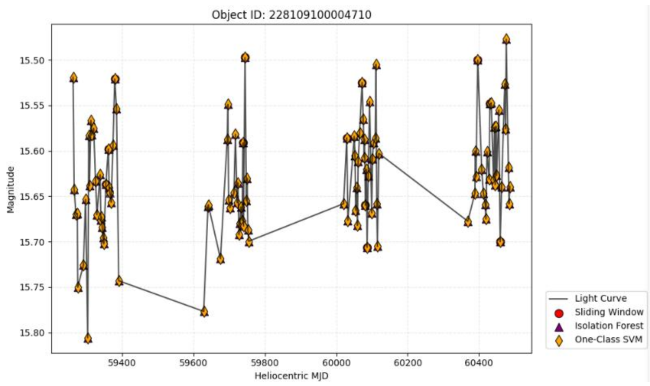


Fig. 7

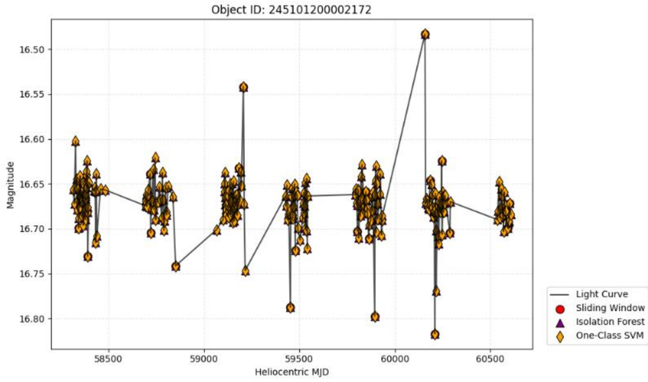


Fig. 8

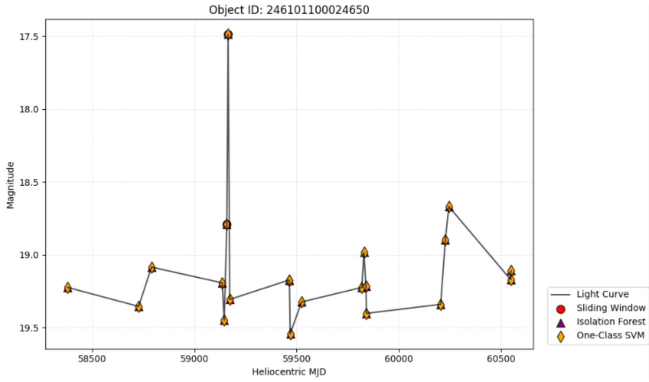


Fig. 9

//

// 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] M. S. P. Kelley *et al.*, "Community Challenges in the Era of Petabyte-Scale Sky Surveys," Nov. 06, 2020, *arXiv*: arXiv:2011.03584. doi: 10.48550/arXiv.2011.03584.

[2] "Revisiting Streaming Anomaly Detection: Benchmark and Evaluation." Accessed: Jul. 03, 2025. [Online]. Available: https://arxiv.org/html/2405.00704v2

[3] S. G. Djorgovski, A. A. Mahabal, M. J. Graham, K. Polsterer, and A. Krone-Martins, "Applications of AI in Astronomy," Dec. 03, 2022, *arXiv*: arXiv:2212.01493. doi: 10.48550/arXiv.2212.01493.

[4] "Sky survey classifies more than 10,000 supernovae." Accessed: Jul. 03, 2025. [Online]. Available: https://news.northwestern.edu/stories/2024/12/sky-survey-classifies/

[5] "Rubin Observatory." Accessed: Jun. 19, 2025. [Online]. Available: https://rubinobservatory.org

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

[7] F. J. Masci *et al.*, "The Zwicky Transient Facility: Data Processing, Products, and Archive," *Publ. Astron. Soc. Pac.*, vol. 131, no. 995, p. 018003, Dec. 2018, doi: 10.1088/1538-3873/aae8ac.

[8] "Panoramic Survey Telescope & Rapid Response System (Pan-STARRS)." Accessed: Jul. 03, 2025. [Online]. Available: https://catalogs.mast.stsci.edu/panstarrs/

[9] "ATLAS - The ATLAS Project." Accessed: Jul. 03, 2025. [Online]. Available: https://atlas.fallingstar.com/

[10] "Mission Overview." Accessed: Jul. 03, 2025. [Online]. Available: https://tess.mit.edu/science/

[11] A. Burrows, "Supernova explosions in the Universe," *Nature*, vol. 403, no. 6771, pp. 727–733, Feb. 2000, doi: 10.1038/35001501.

[12] C. Fremling *et al.*, "SNIascore: Deep Learning Classification of Low-Resolution Supernova Spectra," *Astrophys. J. Lett.*, vol. 917, no. 1, p. L2, Aug. 2021, doi: 10.3847/2041-8213/ac116f.

[13] "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

[14] L. Johnson, "New AI System Discovers Supernova Without Human Help," Tomorrow's World Today®. Accessed: Jul. 03, 2025. [Online]. Available: https://www.tomorrowsworldtoday.com/artificial-intelligence/new-ai-system-discovers-supernova-without-human-help/

[15] "Type Ia supernova," *Wikipedia*. Jun. 10, 2025. Accessed: Jul. 03, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Type_Ia_supernova&oldid=1294963777

[16] S. González-Gaitán *et al.*, "The rise-time of Type II supernovae," *Mon. Not. R. Astron. Soc.*, vol. 451, no. 2, pp. 2212–2229, Aug. 2015, doi: 10.1093/mnras/stv1097.

[17] J. S. W. Claeys, S. E. de Mink, O. R. Pols, J. J. Eldridge, and M. Baes, "Binary progenitor models of type IIb supernovae," *Astron. Astrophys.*, vol. 528, p. A131, Apr. 2011, doi: 10.1051/0004-6361/201015410.

[18] "Types of Supernovae | The Schools' Observatory." Accessed: Jul. 03, 2025. [Online]. Available: https://www.schoolsobservatory.org/discover/projects/supernovae/typeI

[19] "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/

[20] S. C. Tan, K. M. Ting, and T. F. Liu, "Fast Anomaly Detection for Streaming Data".

[21] B. Sánchez *et al.*, "Machine Learning on Difference Image Analysis: A comparison of methods for transient detection," *Astron. Comput.*, vol. 28, p. 100284, Jul. 2019, doi: 10.1016/j.ascom.2019.05.002.

[22] S. G. Djorgovski *et al.*, "Flashes in a Star Stream: Automated Classification of Astronomical Transient Events," in *2012 IEEE 8th International Conference on E-Science*, Oct. 2012, pp. 1–8. doi: 10.1109/eScience.2012.6404437.

[23] "Anomaly Detection Using a Sliding Window Technique and Data Imputation with Machine Learning for Hydrological Time Series." Accessed: Jul. 04, 2025. [Online]. Available: https://www.mdpi.com/2073-4441/13/13/1862

[24] B. Cassese, "Triage with Random Forests: Machine Learning for Transient Classification," AAS Nova. Accessed: Jul. 04, 2025. [Online]. Available: https://aasnova.org/2023/06/16/triage-with-random-forests-machine-learning-for-transient-classification/

[25] A. Gagliano and V. A. Villar, "A Physics-Informed Variational Autoencoder for Rapid Galaxy Inference and Anomaly Detection," Dec. 27, 2023, *arXiv*: arXiv:2312.16687. doi: 10.48550/arXiv.2312.16687.

[26] T. Acero-Cuellar, F. Bianco, G. Dobler, M. Sako, H. Qu, and T. L. D. E. S. Collaboration, "What's the Difference? The potential for Convolutional Neural Networks for transient detection without template subtraction," *Astron. J.*, vol. 166, no. 3, p. 115, Sep. 2023, doi: 10.3847/1538-3881/ace9d8.

[27] D. Muthukrishna, K. S. Mandel, M. Lochner, S. Webb, and G. Narayan, "Real-Time Detection of Anomalies in Large-Scale Transient Surveys," *Mon. Not. R. Astron. Soc.*, vol. 517, no. 1, pp. 393–419, Oct. 2022, doi: 10.1093/mnras/stac2582.

[28] G. Broccia, "Real-Time Detection and Classification of Astronomical Transient Events: The State-of-the-Art," Apr. 30, 2021, *arXiv*: arXiv:2105.00089. doi: 10.48550/arXiv.2105.00089.

[29] 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.

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

[31] 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.

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

[33] 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.

[34] "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

[35] 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.

[36] 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.

[37] 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.

[38] 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/

[39] 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.