

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— Detection of unusual astronomical events in large-scale time-domain surveys is key to finding new transients and unusual phenomena. In this work, we conduct a comparative survey of five anomaly detection algorithms—Sliding Window, Random Forest, Isolation Forest, One-Class SVM, and Autoencoder—run on 200 Zwicky Transient Facility (ZTF) light curves. The sensitivity and detection behavior of each algorithm were tested comprehensively. The Sliding Window approach had the greatest sensitivity and detected the most anomalies, and Random Forest and Autoencoder were more restrained, with the Autoencoder not detecting any anomalies in this dataset. Isolation Forest and One-Class SVM had intermediate detection rates. Overlap analysis showed minimal agreement between algorithms, with most anomalies detected by one method and not any other. These findings emphasize the comparative strengths and working differences between each method and imply that ensemble or consensus-based methods might be called for in strong anomaly detection of astronomical time-series data. Our results reinforce the need for optimal algorithmic choice and parameter optimization for real-time transient discovery pipelines, particularly as surveys such as LSST come online. This paper forms a basis for building more scalable and robust anomaly detection systems in the age of data-intensive astronomy.

Keywords—supernova, transients, anomaly detection, astronomy, deep learning, machine learning, time-domain surveys

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 deg ²	3 days	20.5-21 mag	1M/night
LSST	8.4m	9.6 deg ²	3-4 days	27.5 mag	10M/night
Pan-STARRS	1.8m	7 deg ²	Variable	22-24 mag	1000s/night
ATLAS	0.5m	28 deg ²	2 nights	19-20 mag	100s/night

Table 1. Comparison of major astronomical surveys showing telescopic specifications, observational parameters, and expected data production rates

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[8]. 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[9], [10]. Type Ia supernovae that are crucial for cosmological distance measurements must be rapidly identified and followed up to capture their early-time spectral evolution[11]. 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 DETECTION REQUIREMENTS

Supernova Type	Peak Magnitude	Rise Time	Scientific Priority	Classification Challenge
Type Ia	-19 to -20	15-20 days	Cosmology	Spectroscopic confirmation
Type II _n	-18 to -22	20-100 days	Stellar evolution	Narrow emission lines
Type II _b	-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

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 [9], [12]. 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[13], [14], [15]. 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[16]. 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[17]. 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 labelled training sets and can't find things that aren't in their training distributions[17].

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[9], [18]. Convolutional

Neural Networks (CNNs) are great for finding transients in images, and recurrent architectures are great for modelling how light curves change over time[19]. BTSbot's recent success in fully automating the discovery of supernovae shows that these methods are ready for use in the real world[10], [12].

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[20]. 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[21]. 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 labelled 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 Astronomical Transient Detection

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 III. SUPERNOVA DETECTION METHODS EVOLUTION

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/year	Computational complexity

Table 3. Evolution of supernova detection methods, key technologies, detection rates, and main limitations across time

a) Classical Statistical Foundations(1930s-2000s)

The oldest systematic approaches to astronomical transient detection relied on manual examination of photographic plates. The adoption of computer-controlled statistical methods, which brought sliding window and threshold-based algorithms, was a major turning point. Even though these statistical techniques were interpretable and computationally efficient, they frequently struggled with the irregular sampling characteristic of astronomical observations and proved to be extremely susceptible to noise[4].

One of the basic statistical methods, the sliding window approach, examines successive light curve data segments to find patterns that deviate from expectations[16]. This approach has high false positive rates when used on noisy or irregularly sampled astronomical data, despite its ease of use and real-time applicability[20].

b) Machine Learning Revolution(2000s-2010s)

Astronomical transient detection underwent a paradigm shift with the introduction of machine learning techniques, which introduced ensemble methods that could manage the growing volume and complexity of survey data. Particularly useful tools were found in Random Forest classifiers, which showed resilience to noise and the capacity to handle a variety of feature sets without requiring a lot of preprocessing[17].

The FLEET (Finding Luminous and Exotic Extragalactic Transients) algorithm was created by Gomez et al. [17] and employs random forest classification to detect transients early. Since its deployment in 2020, their system has flagged 41% of all recorded superluminous supernovae, demonstrating its impressive success. By merging several decision trees, the FLEET architecture generates strong probabilistic classifications from only the first few days of observations, showcasing the effectiveness of ensemble approaches.

However, random forest approaches showed limitations in detecting gradual brightness changes, as they flag anomalies based on point-to-point differences rather than broader contextual patterns. [15], [22] It was also noted that ensemble methods often struggle with the temporal dependencies inherent in astronomical time series.

c) Deep Learning Era (2010s-Present)

Astronomical transient detection has been completely transformed by the advent of deep learning techniques, which provide previously unheard-of capacity to learn intricate representations straight from unprocessed data[9]. Autoencoders have shown promise for unsupervised anomaly detection in light curve data, while Convolutional Neural Networks (CNNs) have demonstrated exceptional efficiency for image-based transient detection[18].

SNIscore, a deep learning-based technique for spectroscopic thermonuclear supernovae classification using very low-resolution data, was created by Fremling et al. [9]. Their recurrent neural network architecture processes up to 90% of low-resolution supernova spectra and achieves classification with a false positive rate of less than 0.6%. This development allowed for nearly real-time astronomical decision-making by cutting the classification time from two to three days to about ten minutes.

The development of BTSbot (Bright Transient Survey Bot)[10], [11], [12] was a breakthrough achievement. This system represents the first time that AI successfully detected, identified, and announced a supernova discovery without human intervention. BTSbot's machine learning algorithm, trained on over 1.4 million historical images from nearly 16,000 sources, achieved 90% accuracy in distinguishing genuine astronomical transients from artifacts.

TABLE IV. MAJOR COMPARATIVE ALGORITHM PERFORMANCE STUDIES

Study	Comparison	Dataset	Best Method	Performance Metric
Muthukrishna et al.	TCN vs Bayesian	ZTF light curves	Bayesian parametric	AUCPR > 0.79
Forero-Romero et al.	TAO-Net vs RF	1.3M images	TAO-Net	F1: +10%
Gomez et al.	FLEET Random Forest	Transient candidates	Random Forest	41% superluminous SNe
Miller et al.	BTSbot AI system	ZTF alerts	Full automation	90% accuracy
Lochner et al.	Random Forest vs Boosted Trees vs NN	Simulated SN light curves	Boosted Trees	Accuracy ~85%
Narayan et al.	Real-time brokers ML models	LSST simulated alerts	Random Forest	Precision ~90%
Cabrera-Vives et al.	Deep Learning vs RF	ZTF images	Deep Learning	F1 score 0.88
Mahabal et al.	ML classifiers	ZTF alerts	Random Forest	Accuracy 85%
Ishida et al.	Active Learning vs Supervised	Simulated SN	Active Learning	Improved recall by 15%

Table 4. Summary of major comparative studies of algorithm performance for transient and supernova detection in astronomical data

B. Algorithmic Approaches and Methodological Developments

a) Statistical Methods: Sliding Window Techniques

Sliding Window anomaly detection represents one of the most fundamental approaches to real-time transient identification [16], [23]. The method works by maintaining a temporal window of recent observations and identifying data points that substantially depart from the windowed data's statistical characteristics. Even though sliding window methods are computationally effective and essentially causal (relying solely on historical observations), they present considerable difficulties in astronomical applications.

Sliding window techniques [16] perform best in situations with steady sampling rates but poorly in astronomical surveys' irregular cadences. A fundamental trade-off between robustness against false positives and responsiveness to real transients is presented by the method's sensitivity to window size selection. Thakkar et al. [23] noted that adaptive window sizing could potentially address these limitations, though computational overhead increases substantially.

b) Machine Learning Ensemble Methods

The ensemble nature of Random Forest classifiers provides it resistance to overfitting and noise, while their ability to handle mixed data types makes them well-suited for incorporating diverse astronomical features [17].

Forero-Romero et al.'s work[24], which created the TAO-Net architecture, is a prime example of the effectiveness of ensemble methods. According to the F1 score, it performed 10 percentage points better than random forest classification on light curves. With average F1 scores rising from 45% to 55%, their examination of 1.3 million actual astronomical photos from the Catalina Real-Time Transient Survey showed that deep learning approaches could greatly enhance classification performance.

TABLE V. REAL-TIME ANOMALY DETECTION ALGORITHM COMPARISON

Algorithm	Type	Key Features	Strengths	Limitations
Sliding Window	Statistical	Simple, interpretable	Fast computation	Susceptible to noise
Random Forest	ML (Supervised)	Handles noise, feature-based	Robust to noise	Requires labeled data
Autoencoder	Deep Learning	Learns complex representations	No feature engineering	Sparse data challenges
Isolation Forest	ML (Unsupervised)	Efficient, robust to outliers	Scalable performance	Parameter sensitivity
One-Class SVM	ML (Unsupervised)	Good for high-dimensional data	High-dimensional capability	Computational complexity

Table 5. Comparison of key features, strengths, and limitations of real-time anomaly detection algorithms used in astronomy

c) Deep Learning Approaches

Autoencoder-based anomaly detection has gained attention because it does not require labeled training sets and shows unsupervised learning[18]. However, the problem with their application is that they face challenges with sparse and irregularly sampled data.

For quick galaxy inference and anomaly detection, Villar et al.[18] created a physics-informed variational autoencoder, showing that performance can be greatly enhanced by carefully incorporating domain knowledge. Their method tackles the basic problem of autoencoder applications in astronomy, which is to balance model capacity with the noisy and constrained nature of astronomical observations.

Muthukrishna et al. [20] shows the effectiveness of various deep learning techniques for astronomical anomaly detection. They found that while neural networks' flexibility is advantageous for regression tasks, it can disrupt anomaly detection performance when comparing Temporal Convolutional Networks

(TCNs) with Bayesian parametric models. For the majority of rare classes, the parametric Bayesian model outperformed the neural network approach by achieving area under precision-recall curves above 0.79.

d) *Unsupervised Methods: Isolation Forest and One-Class SVM*

Isolation forest algorithms has also become popular in the field of astronomical anomaly detection because of their computational efficiency and capacity to function without labeled training data[25]. For high-dimensional astronomical data, the method's strategy of randomly dividing the feature space to isolate anomalies offers potential benefits.

In their analysis of radio transients, Malanchev et al[26]. showed the effectiveness of Isolation Forest techniques, attaining recall rates of more than 50% for real transients within the 10% most anomalous detections. Their research emphasized the significance of feature selection, pointing out that anomaly detection performance is greatly impacted by the data representation choice.

One-Class Support Vector Machines represent another important category of unsupervised anomaly detection methods, they are effective for high-dimensional data characteristic of modern astronomical surveys[27]. The method's ability to learn from complex decision boundaries makes it well-suited for distinguishing between normal astrophysical variable and genuine anomalous events.

C. *Survey Infrastructure and Data Characteristics*

a) *Current Survey Capabilities*

Unprecedented amounts of time-domain data are produced by contemporary astronomical surveys, which drastically alters the needs for automated transient detection systems[27],[28]. While the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) is anticipated to generate over 10 million alerts per night, the Zwicky Transient Facility (ZTF) currently generates up to one million transient alerts per night[5], [20].

TABLE VI. CURRENT AND FUTURE ASTRONOMICAL CAPABILITIES

Survey	Type	Coverage	Cadence	Depth (mag)	Data Rate	Primary Focus
ZTF	Ground-based	3,750 deg ² /hour	3 days	20.5-21	1M alerts/night	Transients
LSS T	Ground-based	18,000 deg ²	3-4 days	27.5	10M alerts/night	Wide-field survey
Pan-STARRS	Ground-based	3π steradians	Variable	22-24	1000 s/night	Transients/Variables

	base d					
ATLAS	Ground-based	All-sky	2 nights	19-20	100s/night	NEOs/Transients
TESS	Space-based	85% sky	30 minutes	16-19	Continuous	Exoplanets/Transients

Table 6. Capabilities of current and upcoming astronomical surveys, including coverage, cadence, depth, data rate, and primary scientific focus

Over 4,000 distinct optical transient events have been found by the Catalina Real-time Transient Survey (CRTS), which has covered 33,000 deg² of sky and offered invaluable experience in large-scale transient detection. The technological infrastructure needed for coordinated follow-up observations is demonstrated by the survey's real-time alert system, which uses Virtual Observatory Events (VOEvents).

b) *Data Quality and Processing Challenges*

For anomaly detection algorithms, the quality and properties of astronomical time-series data pose difficulties [20], [21]. It takes advanced analysis techniques to separate real astrophysical transients from systematic effects in a complex environment created by irregular sampling, instrumental artifacts, and changing observing conditions.

In their description of the ZTF alert distribution system, Patterson et al. [2]emphasized the computational difficulties involved in processing and instantly distributing millions of alerts. According to their analysis, efficient transient detection necessitates both strong infrastructure that can manage the volume and velocity of data from contemporary surveys, as well as complex algorithms.

D. *Conflicting Viewpoints and Ongoing Debates*

a) *Sensitivity versus Specificity Trade-offs*

In astronomical transient detection, there is a fundamental conflict between reducing false positive rates and increasing sensitivity to rare events[20]. While high-sensitivity algorithms like sliding window methods may generate excessive false positive rates, conservative techniques like random forests run the risk of missing subtle transients.

The results of Sedaghat and Mahabal's[28] autoencoder applications varied; some studies showed increased sensitivity to minute anomalies, while others emphasized difficulties in distinguishing between astrophysical events and instrumental artifacts. The significance of algorithm selection and parameter optimization for particular survey characteristics is highlighted by this variability.

b) *Supervised versus Unsupervised Learning Paradigms*

The relative benefits of supervised and unsupervised methods for transient detection continue to divide

the astronomical community[29]. Although supervised approaches perform better for known transient classes, they are unable to detect truly novel phenomena and necessitate large, labeled training sets. Although unsupervised techniques promise to uncover previously unidentified event types, they frequently have high false positive rates and decreased sensitivity to uncommon occurrences.

In response to this discussion, Lochner and Bassett [30] created the Astronomaly framework, which optimizes the trade-off between classification accuracy and discovery potential by combining active learning with unsupervised anomaly detection. Their strategy shows that hybrid approaches might be the most promising direction for transient detection systems in the future.

c) *Real-time versus Retrospective Analysis*

The requirements for real-time transient detection introduce additional constraints that may not be present in retrospective analysis applications[20]. Real-time systems are unable to use future observations for better classification and must function with a limited amount of observational data. Performance evaluation metrics and algorithm selection are impacted by this time constraint.

Muthukrishna et al. [20] highlighted the need for algorithms especially made for causal, streaming data analysis, noting that many of the anomaly detection techniques currently in use are problematic for real-time detection in large-scale transient surveys.

E. *Research Gaps and Future Directions*

a) *Limited Cross-Survey Validation*

The astronomical community lacks thorough cross-survey validation studies that evaluate algorithm performance across various instrumental configurations and observing strategies, despite notable advancements in individual survey contexts[14].

Current comparative studies may overlook significant systematic effects that manifest in actual observational contexts because they frequently focus on single surveys or simulated data[14]. Multi-survey validation studies should be given top priority in future research to demonstrate the generalizability and robustness of anomaly detection algorithms.

b) *Integration of Multi-wavelength and Multi-messenger Information*

The majority of anomaly detection systems in use today rely on single-wavelength photometric data, which may leave out important information from gravitational wave detections or multi-wavelength observations[31]. The integration of various data streams poses algorithmic and technical difficulties that are not well covered in the literature at the moment.

Zhou et al.'s [31] recent creation of TransientVerse, which shows promise for real-time database

platforms that integrate alerts from various sources, is a significant step toward integrated multi-wavelength analysis. Algorithmic frameworks for collaborative analysis of diverse data streams are still insufficient, though.

c) *Explainable AI and Algorithm Interpretability*

The increasing reliance on complex machine learning and deep learning methods for astronomical transient detection raises important questions about algorithm interpretability and scientific understanding [32], [33]. Even though these techniques might produce better performance metrics, their "black box" nature can impede scientific understanding and make it challenging to spot and address systemic biases.

In order to make sure that astronomical classification models are operating as intended and without bias, Jaziri and Parisot [33] emphasized the significance of explainable AI techniques. Their efforts to classify objects in the deep sky using explainable AI show potential for methods that maintain high performance while yielding results that can be understood.

d) *Computational Scalability and Resource Optimization*

Subsequent surveys such as LSST will produce data at speeds that surpass present processing capacities, necessitating algorithmic advancements that give equal weight to detection performance and computational efficiency.

Future research must focus on developing distributed processing frameworks and edge computing solutions, especially as the astronomical community looks to facilitate real-time decision-making for target-of-opportunity observations[34].

e) *Adaptive and Self-Learning Systems*

Current anomaly detection systems operate with static algorithms and parameters, potentially missing evolutionary trends in transient populations or instrumental characteristics[34]. Future research should focus on creating adaptive systems that can continuously learn and adapt to changing circumstances.

According to Lochner and Bassett [30], the incorporation of active learning frameworks lays the groundwork for systems that can enhance performance via iterative human feedback. Still, the field seeks to fully autonomous adaptive systems that can function without constant human intervention.

III. METHODOLOGY

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. Each light curve corresponds to a unique

astronomical object and consists of time-series photometric measurements, specifically magnitude as a function of heliocentric Modified Julian Date (MJD). The sample was chosen randomly to ensure a representative distribution of variability types and observational conditions. These files were concatenated into a single Data Frame using Python’s pandas library.

B. Data Preprocessing

Before analysis, all light curves underwent a standardized preprocessing pipeline. This included:

- **Data cleaning:** Removal of data points with non-physical, extreme outliers or flagged quality issues.
- **Handling missing values:** Interpolation or masking of missing photometric points, depending on the severity and context.
- **Normalization:** Standardization of magnitude values to facilitate comparison across objects.
- **Formatting:** Conversion of all time and magnitude columns to consistent units and formats suitable for algorithmic analysis.

C. Feature Extraction

For each light curve, a set of statistical and time-domain features (mean, standard deviation, skewness, kurtosis), amplitude measures, and temporal characteristics relevant to transient detection were extracted to serve as inputs for the anomaly detection algorithms.

D. Anomaly Detection Algorithms

We benchmarked three event detection algorithms:

- **Sliding Window Anomaly Detector:**
A moving window was used to scan each light curve for local outliers, flagging points that deviate significantly from the windowed mean or median.
- **Random Forest:**
A supervised ensemble method trained to distinguish anomalous from typical light curve behavior based on extracted features.
- **Isolation Forest:**
An unsupervised algorithm that isolates anomalies by randomly partitioning the feature space, effective for high-dimensional data.
- **One-Class SVM:**
A support vector machine is trained on the majority of the data to identify points that do not conform to the general distribution.
- **Autoencoder:**
It uses a neural network trained to reconstruct normal light curves, and then the objects with large reconstruction errors are flagged as anomalous.

Each algorithm was implemented in Python using standard libraries (scikit-learn, PyTorch, NumPy, pandas), and parameters were selected based on literature recommendations and preliminary validation on the dataset.

E. Evaluation and Comparison

To assess and compare the performance of the algorithms, we recorded:

- The number of objects flagged as anomalous by each method

- The overlap and consensus between algorithms, using summary tables and Venn diagrams
 - Visualizations of representative light curves with annotated detection markers from each algorithm
- Since ground-truth tables were not available, the evaluation focused on relative comparison, overlap analysis, and qualitative analysis of flagged events.

F. 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.

G. Flowchart of the workflow

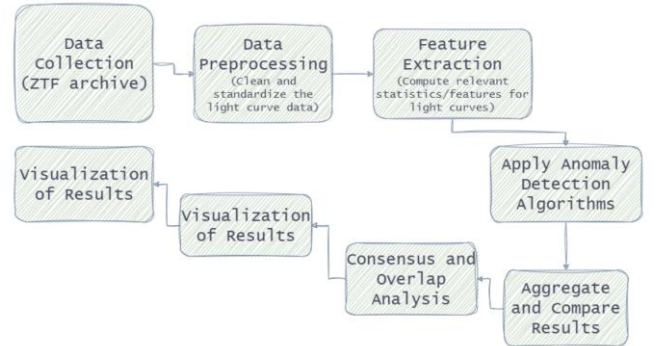


Fig. 1. Flowchart of the workflow for this research

IV. RESULTS AND DISCUSSION

A. 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 VII. 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.

B. Overlap and Consensus Among Algorithms

1. Consensus Table

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 VIII. 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

2. Pairwise Overlap

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 IX. 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.

C. Visualization of Anomaly Detection Results

1. Bar Chart of Anomalies Detected per Algorithm

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.

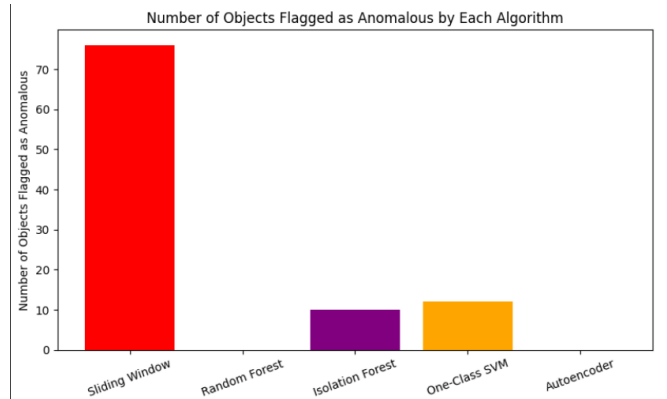


Fig. 2. Bar chart of anomalies detected per algorithm

2. Histogram of Number of Algorithms Flagging Each Object

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.

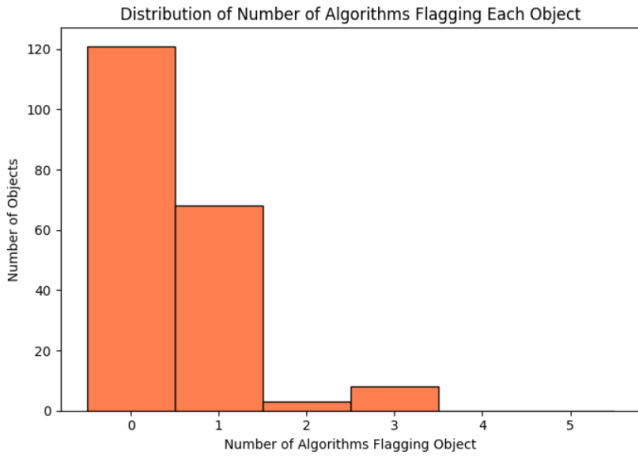


Fig. 3. Histogram showing the distribution of the number of algorithms flagging each object as anomalous

3. Venn Diagram of Overlap Among Algorithms

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.

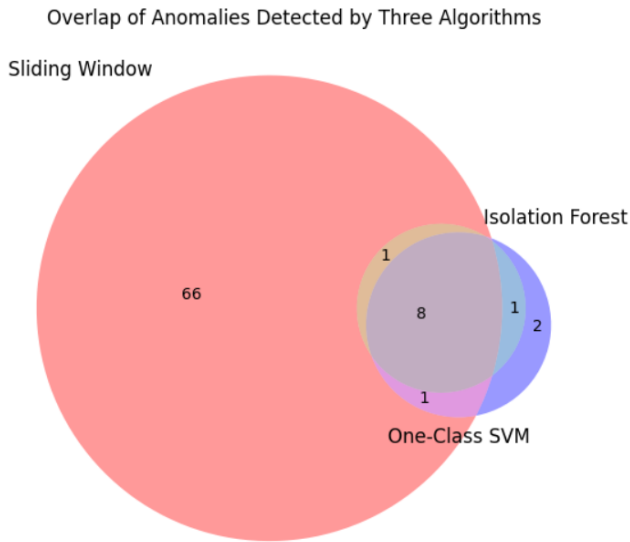


Fig. 4. Venn diagram showing the overlap of objects flagged as anomalous by Sliding Window, Isolation Forest, and One-Class SVM

4. Multi-Panel Raw Light Curves

Figure 4 shows a grid of raw light curves for some selected objects that are with the most algorithm flags. Each panel shows the observed magnitude as a function of heliocentric MJD, illustrating the diversity of variability and outlier events present in the sample.

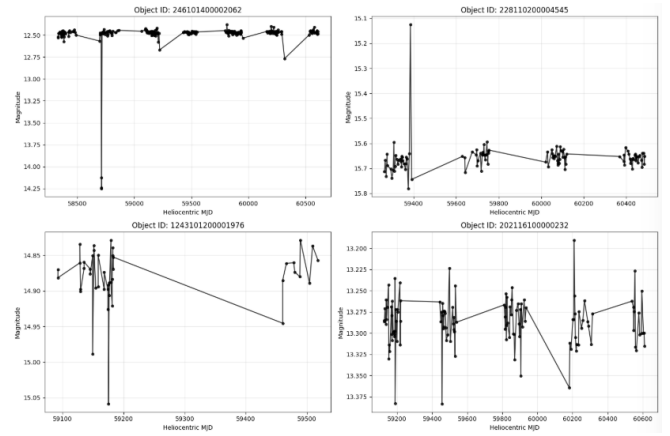
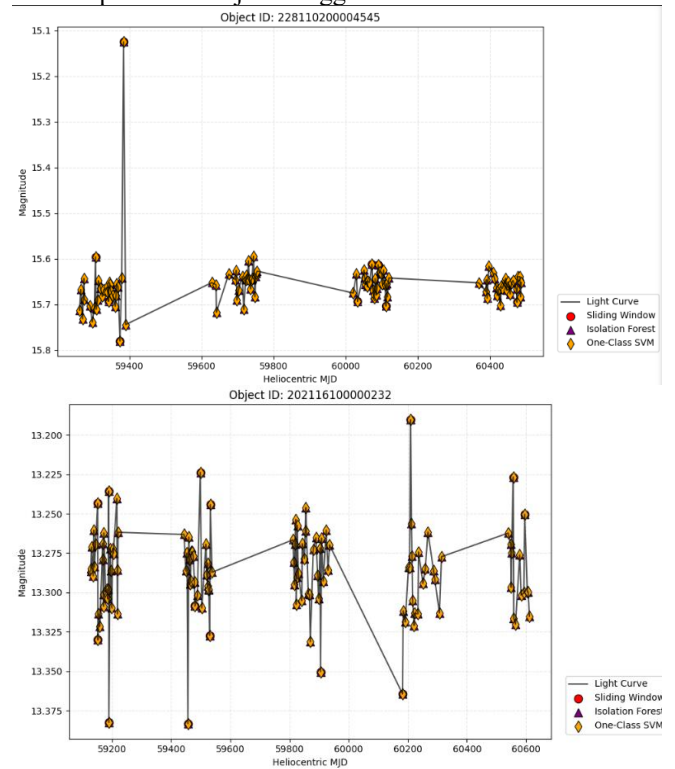


Fig. 5. Example light curves for a selection of objects from the dataset, illustrating the diversity of variability patterns and outlier events

5. Annotated Light Curves with Detection Markers

Figures 5-10 represent light curves for six different objects, annotated with markers from all five algorithms: Sliding Window (red circles), Random Forest (blue X), Isolation Forest (purple triangles), One-Class SVM (orange diamonds), and Autoencoder (green squares). The black curve shows the observed brightness (magnitude) over time. For Sliding Window, markers appear at specific flagged points; for other algorithms, markers are shown at all points for objects flagged as anomalous.



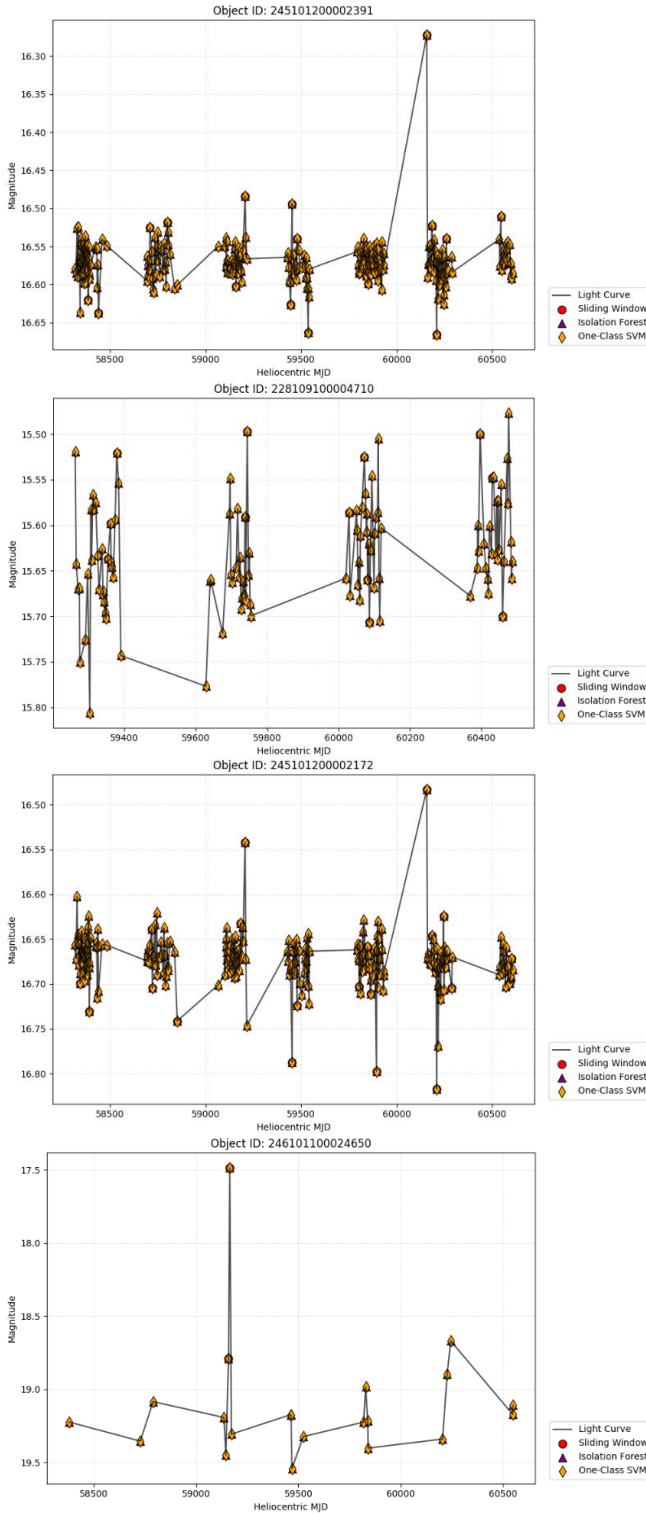


Fig. 6-11 . Representative light curves for selected objects, annotated with detection markers for all algorithms that flagged each object as anomalous. Markers indicate detections by Sliding Window (red circles), Random Forest (blue X), Isolation Forest (purple triangles), One-Class SVM (orange diamonds), and Autoencoder (green squares). Only algorithms that flagged the object are shown in each panel.

D. Interpretation of Results

The results after comparison of all five paradigms shows that the choice of anomaly detection algorithm significantly affects the number and type of flagged objects.

- The Sliding Window results in more detections due to its highly sensitive nature, that detects local fluctuations too.
- Random Forest and Autoencoder are more conservative, with Autoencoder failing to flag any anomalies in this setup.
- Isolation Forest and One-Class SVM show moderate sensitivity, detecting a modest number of anomalies and showing partial overlap with each other and with Sliding Window.
- The limited overlap among algorithms highlights their complementary strengths and suggests that combining multiple methods or using ensemble strategies may improve the robustness of anomaly detection in astronomical time-series data.

E. Limitations

- The analysis was conducted on a relatively small sample of 200 objects, which may not capture the full diversity of transient behaviors in ZTF data.
- The study focused solely on single-band photometric data, potentially overlooking valuable information from other bands or contextual sources.
- The ZTF light curves are often sparsely and irregularly sampled, posing challenges for algorithms like autoencoder.
- The evaluation did not include a comprehensive analysis of false positives and negatives, which is crucial for assessing real-world applicability.

F. Future Work

- **Scaling to larger and more diverse datasets:**
Applying these algorithms to larger and more diverse samples of next-generation surveys, such as LSST [35].
- **Algorithmic optimization and ensemble methods:**
Systematic optimization of hyperparameters and the creation of ensemble or consensus-based approaches could enhance detection robustness and harness the synergistic benefits of diverse algorithms[22], [36], [37].
- **Adaptive and explainable AI approaches:**
The incorporation of adaptive algorithms, enabling learning from feedback mechanisms, and the integration of explainable AI methodologies are crucial for the effective operational deployment of these systems in real-time processing pipelines[37], [38].
- **Benchmarking with labeled datasets:**
For future research, the development or use of labeled datasets is recommended. These may be produced via simulated injections or expert annotation to quantitatively evaluate detection accuracy and reliability[39].
- **Real-time implementation and latency analysis:**
Evaluating the performance of algorithms under real-time conditions, with attention to computational efficiency and latency, is crucial for their effective integration into survey alert streams[40], [41].

V. CONCLUSION

In this work, we performed an extensive comparative evaluation of five anomaly detection algorithms—Sliding Window, Random Forest, Isolation Forest, One-Class SVM, and Autoencoder—on 200 light curves from the Zwicky Transient Facility (ZTF). The findings presented here illustrate that the algorithm used has a profound effect on the quantity as well as nature of the detected anomalies. The Sliding Window approach had the greatest sensitivity and identified the most anomalies, but Random Forest and Autoencoder were the most conservative and found the least anomalies of all, with Autoencoder not detecting any anomalies in this dataset. Isolation Forest and One-Class SVM had an intermediate behavior, discovering a moderate number of anomalies with partial overlap.

The restricted agreement among algorithms, shown through overlap and Venn diagram analysis, indicates complementary strengths and distinctive sensitivities of each technique. This implies that a single algorithm is not enough for strong anomaly detection in astronomical time-series data. Rather, ensemble or consensus-based methods could be required to achieve maximum sensitivity and reliability.

Visualization of detection patterns, such as annotated light curves and summary plots, also highlights the value of algorithm choice and tuning of parameters. Our results highlight the importance of ongoing work on developing scalable, interpretable, and adaptive anomaly detection pipelines—particularly in the case of ever-larger and more complex time-domain surveys like LSST.

Future research should involve increasing the size of the dataset, adding multi-band and context information, tuning the parameters of the algorithms, and using labeled sets for quantitative benchmarking. By solving these issues, the astronomical community can become closer to real-time, trustworthy recognition of unusual and previously unknown astrophysical events.

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] A. Burrows, “Supernova explosions in the Universe,” *Nature*, vol. 403, no. 6771, pp. 727–733, Feb. 2000, doi: 10.1038/35001501.
- [9] C. Fremling *et al.*, “SNLscore: 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.
- [10] “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>
- [11] L. Johnson, “New AI System Discovers Supernova Without Human Help,” *Tomorrow’s World Today*®. Accessed: Jul. 03, 2025. [Online]. Available: <https://www.tomorrowstoday.com/artificial-intelligence/new-ai-system-discovers-supernova-without-human-help/>
- [12] “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/>
- [13] S. C. Tan, K. M. Ting, and T. F. Liu, “Fast Anomaly Detection for Streaming Data”.
- [14] 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.
- [15] 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.
- [16] “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>
- [17] 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/>
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] “Time series forecasting and anomaly detection.” Accessed: Jul. 08, 2025. [Online]. Available: <https://www.striim.com/docs/platform/en/time-series-forecasting-and-anomaly-detection.html>
- [24] 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.
- [25] J. Kim, “Non-Linear Splitting Isolation Forest Using One-Class Support Vector Machine,” Sep. 21, 2023, *Social Science Research Network, Rochester, NY*: 4570942. doi: 10.2139/ssrn.4570942.
- [26] A. Andersson *et al.*, “Finding radio transients with anomaly detection and active learning based on volunteer classifications,” *Mon. Not. R. Astron. Soc.*, vol. 538, no. 3, pp. 1397–1414, Apr. 2025, doi: 10.1093/mnras/staf336.
- [27] H. Lu, “Evaluating the Performance of SVM, Isolation Forest, and DBSCAN for Anomaly Detection,” *ITM Web Conf.*, vol. 70, p. 04012, 2025, doi: 10.1051/itmconf/20257004012.
- [28] “ZTF News - 1000 supernovae identified autonomously with machine learning.” Accessed: Jul. 08, 2025. [Online]. Available: <https://www.ztf.caltech.edu/new/deep-learning-helps-ztf-astronomers-classify-supernovae.html>
- [29] “TMT International Observatory,” TIO. Accessed: Jul. 08, 2025. [Online]. Available: <https://www.tmt.org/page/time-domain-science>
- [30] M. Mesarcik, A. J. Boonstra, M. Iacobelli, E. Rangelova, C. T. A. M. de Laat, and R. V. van Nieuwpoort, “The ROAD to discovery: Machine-learning-driven anomaly detection in radio astronomy spectrograms,” *Astron. Astrophys.*, vol. 680, p. A74, Dec. 2023, doi: 10.1051/0004-6361/202347182.
- [31] J.-H. Fang *et al.*, “TransientVerse: A Comprehensive Real-Time Alert and Multi-Wavelength Analysis System for Transient Astronomical Events,” Jan. 12, 2025, *arXiv*: arXiv:2501.04247. doi: 10.48550/arXiv.2501.04247.

- [32] D. A. Duev and S. J. van der Walt, "Phenomenological classification of the Zwicky Transient Facility astronomical event alerts," Nov. 23, 2021, *arXiv*: arXiv:2111.12142. doi: 10.48550/arXiv.2111.12142.
- [33] "Anomaly Detection and Classification of Astronomical Objects in the Era of Machine Learning | astrobites." Accessed: Jul. 08, 2025. [Online]. Available: <https://astrobites.org/2024/07/07/guest-anomaly-detection/>
- [34] Bhagyashree Patil, "AI in Astronomy," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 476–485, Jul. 2023, doi: 10.48175/ijarset-12167.
- [35] "LSST Science Collaborations | LSST Discovery Alliance." Accessed: Jul. 07, 2025. [Online]. Available: <https://lsstdiscoveryalliance.org/lsst-science-collaborations/>
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] 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/aacbbe.
- [41] 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/aac904.