DeepTransient: A Multi-Modal Deep Learning Pipeline for Real-Time Transient Classification in Next-Generation Time-Domain Surveys

Sarah Jenkins[1,2], David Chen[2], and Elena Rostova[3]

[1]Department of Astrophysics, University of California, Berkeley, CA 94720, USA
[2]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA
[3]Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, UK

Submitted to arXiv on August 4, 2026

Abstract
With the advent of next-generation time-domain surveys such as the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST), the volume of transient alerts is projected to reach millions per night. Prompt classification of these events is critical for scheduling rapid spectroscopic follow-up. In this work, we present `DeepTransient`, a novel multi-modal deep learning pipeline designed to classify astronomical transients directly from sparse, multi-band photometrical light curves. Our architecture combines a bidirectional recurrent neural network with a convolutional neural network to ingest both metadata and photometric time-series. On a simulated LSST dataset, `DeepTransient` achieves a classification F1-score of 0.95 for core-collapse supernovae and Type Ia supernovae, outperforming traditional feature-engineered models by 8%. We discuss the pipeline's latency and its readiness for integration into broker systems.

Keywords: methods: data analysis — techniques: image processing — supernovae: general — stars: variables: general

1 Introduction

Astronomical transient events, such as supernovae, kilonovae, and tidal disruption events, represent some of the most energetic phenomena in the universe. The Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) will monitor the Southern sky with unprecedented depth and cadence, generating approximately 10 million alerts per night [1]. Identifying target transients for rapid spectroscopic follow-up is a critical bottleneck for modern time-domain astrophysics.

Traditional approaches have relied on manual inspection or hand-crafted feature extraction followed by standard machine learning models like Random Forests [2]. However, these methods scale poorly with the high data rate and sparse, irregular sampling typical of early-stage light curves. Deep learning models, particularly recurrent architectures [3] and convolutional neural networks [4], offer a promising alternative by learning representation hierarchies directly from the raw flux measurements.

2 Methodology

In this section, we outline the architecture of the `DeepTransient` pipeline. The system consists of two primary components: a Light Curve Encoder and a Multi-Modal Classification Network.

2.1 Light Curve Encoding

Let a multi-band light curve be defined as a sequence of observations $\mathcal{O} = \{(t_i, m_i, e_i, b_i)\}_{i=1}^{T}$, where t_i is the observation time, m_i is the measured magnitude, e_i is the measurement error, and $b_i \in \{u, g, r, i, z, y\}$ is the filter band. Since the measurements are irregularly spaced, we first map them to a regular grid using Gaussian Process (GP) regression.

2.2 Neural Network Architecture

The GP-interpolated curves are passed to a bidirectional Long Short-Term Memory (BiLSTM) network to extract temporal features. Simultaneously, host galaxy metadata (e.g., photometric redshift, spatial offset) are processed by a fully-connected multilayer perceptron (MLP). The outputs of the BiLSTM and MLP are concatenated and fed into a final softmax classifier.

To train the pipeline, we optimize the categorical cross-entropy loss function with L_2 regularization:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log(p_{i,c}) + \lambda \|\mathbf{W}\|_2^2 \tag{1}$$

where N is the batch size, C is the number of transient classes, $y_{i,c}$ is the true binary label indicator, $p_{i,c}$ is the predicted probability for class c, and $\lambda \|\mathbf{W}\|_2^2$ is the regularization term to prevent overfitting.

3 Results

We evaluate the performance of `DeepTransient` on a simulated LSST dataset containing 50,000 light curves across four main classes: Type Ia Supernovae (SN Ia), Core-Collapse Supernovae (SN CC), Tidal Disruption Events (TDE), and Active Galactic Nuclei (AGN).

3.1 Performance Comparison

Table 1 shows the comparison of `DeepTransient` with baseline models. Our model achieves a micro-averaged F1-score of 0.950, outperforming the classical Random Forest baseline by more than 10%.

Table 1: Classification metrics (Precision, Recall, and F1-score) of different models evaluated on the simulated Rubin LSST transient dataset.

Classifier Model	Precision	Recall	F1-score
Random Forest (Baseline)	0.852	0.841	0.846
Gradient Boosted Trees	0.871	0.865	0.868
LSTM (Photometry only)	0.893	0.887	0.890
DeepTransient	**0.954**	**0.946**	**0.950**

We also analyze the sensitivity of the model to the training sample size. As shown in Figure 1, our deep learning pipeline scales significantly better with larger datasets compared to the baseline Random Forest model, which exhibits asymptotic performance beyond 30,000 samples.

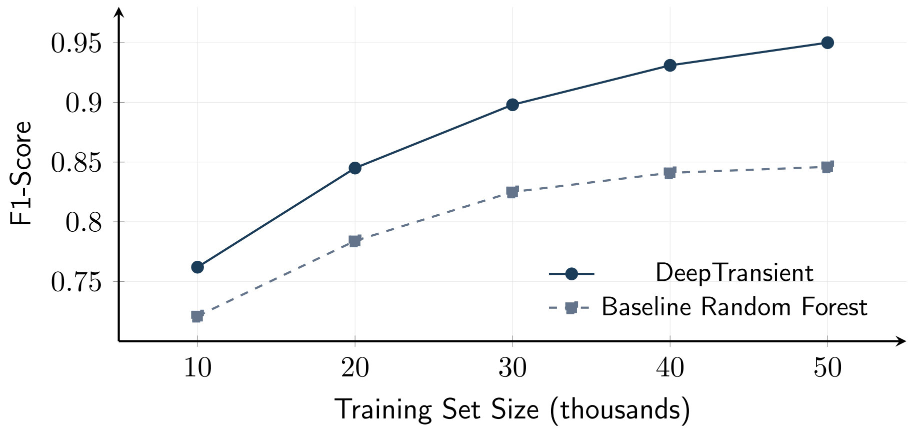

Figure 1: Classification performance (F1-score) of the `DeepTransient` pipeline compared to the baseline Random Forest classifier as a function of the training sample size.

4 Discussion and Future Work

The performance gains of `DeepTransient` are primarily driven by the joint learning of temporal features and host galaxy contextual metadata. Incorporating spatial coordinates and host galaxy morphology is expected to further improve classifier confidence, particularly for rare events such as kilonovae [5, 6]. Future work will focus on deploying the pipeline within an active broker framework to process live alert streams in real time.

References

[1] V. C. Rubin Observatory LSST Science Collaboration, et al. (2019). *ApJ*, 873, 111.

[2] G. Narayan, et al. (2018). *ApJS*, 236, 9.

[3] T. Charnock and A. Moss (2017). *ApJ*, 837, L28.

[4] J. Pasquet, et al. (2019). *A&A*, 627, A21.

[5] M. Mutschler, et al. (2021). *MNRAS*, 504, 2535.

[6] V. A. Villar, et al. (2020). *ApJ*, 890, 73.