Transient Classifiers for Fink

Benchmarks for LSST

B.M.O. Fraga[1]⋆, C.R. Bom[1,2], A. Santos[1], E. Russeil[3], M. Leoni[4], J. Peloton[4], E.E.O. Ishida[3], A. Möller[5], and S. Blondin[6,7]

[1] Centro Brasileiro de Pesquisas Físicas, Rua Dr, Xavier Sigaud 150, Rio de Janeiro, Brazil
[2] Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rodovia Márcio Covas, lote J2, Itaguaí, Brazil
[3] Université Clermont-Auvergne, CNRS, LPCA, 63000, Clermont-Ferrand, France
[4] Université Paris-Saclay, CNRS/IN2P3, IJCLab, 15 rue Georges Clemenceau, 91405 Orsay, France
[5] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Mail Number H29, PO Box 218, 31122 Hawthorn, VIC, Australia
[6] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[7] European Southern Observatory, Karl-Schwarzschild-Straße 2, Garching, D-85748, Germany

Yuan FANG

Introduction

- Legacy Survey of Space and Time (LSST) at the Vera C. Rubin Observatory will be a wide-field ground-based system. The telescope will have an 8.4 m (6.5 m effective) primary mirror, a 9.6 deg2 field of view, and a 3.2 Gigapixel camera.
- The survey area will be imaged multiple times in six bands, ugrizy, covering the wavelength range 320--1050 nm. The typical 5σ point-source depth in a single visit in r will be ~24.5 (AB). The project is in the construction phase and will begin regular survey operations by 2022.
- The standard observing sequence will consist of pairs of 15-second exposures in a given field, with two such visits in each pointing in a given night. With these repeats, the LSST system is capable of imaging about 10,000 square degrees of sky in a single filter in three nights.

Introduction

- LSST is expected to detect a few million transients per night, which will generate a live alert stream during the entire 10 years of the survey.
- For time-domain astronomy, the ability to quickly process the data and obtain meaningful results has become critical due to current and upcoming projects such as the Zwicky Transient Facility (ZTF, Bellm et al. 2019) and the LSST, respectively.
- These projects employ a difference imaging analysis pipeline which stream to community brokers, in the form of alerts, every detection above a given signal to noise threshold.Brokers are subsequently tasked with filtering and analysing the data in detail, selecting the most promising objects for different science cases and redirecting them to different research communities.

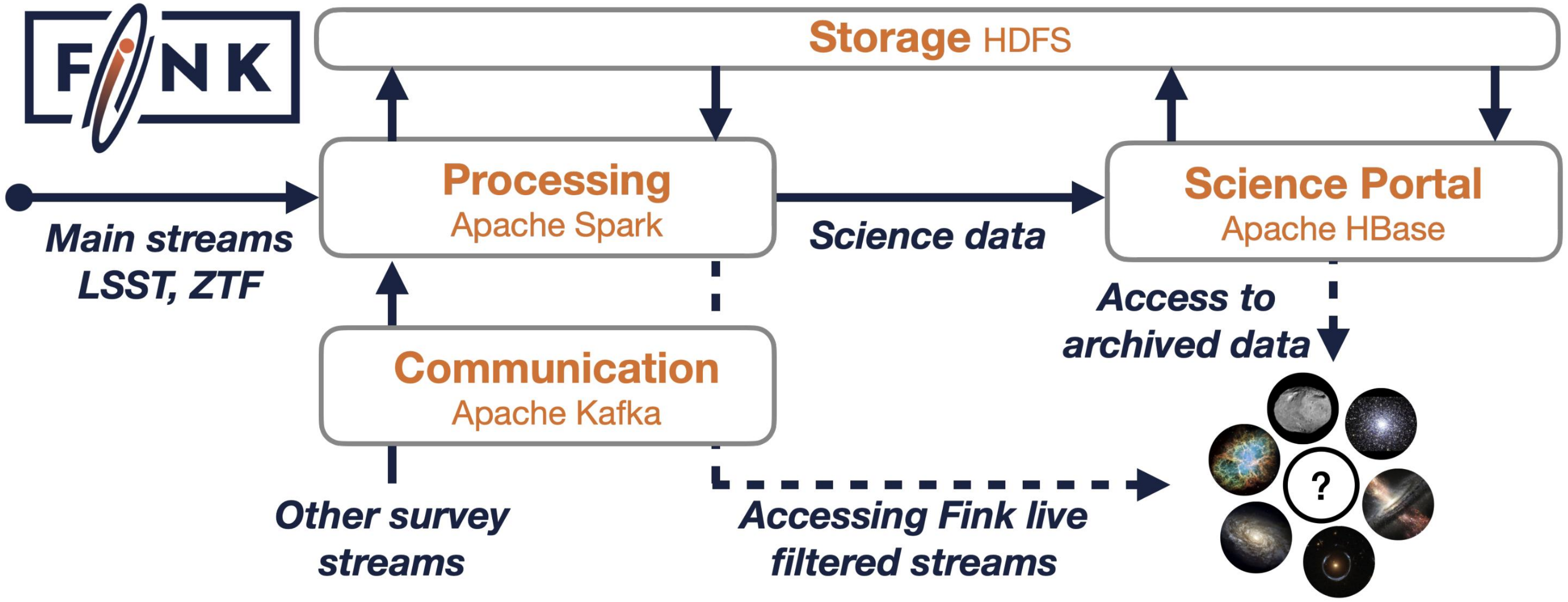
Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **501**, 3272–3288 (2021) doi:10.1093/mnras/staa3602
Advance Access publication 2020 November 19

FINK, a new generation of broker for the LSST community

Anais Möller,[1★] Julien Peloton,[2★] Emille E. O. Ishida,[1★] Chris Arnault,[2] Etienne Bachelet,[3] Tristan Blaineau,[2] Dominique Boutigny,[4] Abhishek Chauhan,[5] Emmanuel Gangler,[1] Fabio Hernandez,[6] Julius Hrivnac,[2] Marco Leoni,[2,7] Nicolas Leroy,[2] Marc Moniez,[2] Sacha Pateyron,[2] Adrien Ramparison,[2]

Fink

- Fink (Möller et al. 2021) is one of the official LSST brokers, selected to receive the raw alert stream from the beginning of LSST operations,expected for 2025. In the meantime, brokers systems are operating, and being tested, with alerts from ZTF.
- Based on established technologies for fast and efficient analysis of big data, Fink provides traditional broker features such as catalogue and survey cross-matches, but also uses machine learning techniques to generate classification scores for a variety of time-domain phenomena.



Deployment platform and prototyping

- Processing (APACHE SPARK cluster)

  A cluster of 11 machines (162 cores total for the computation, 2 GB RAM/core) with APACHE SPARK deployed and managed by APACHE MESOS , the associated data store is a HDFS cluster of 11 machines with 35 TB of storage total

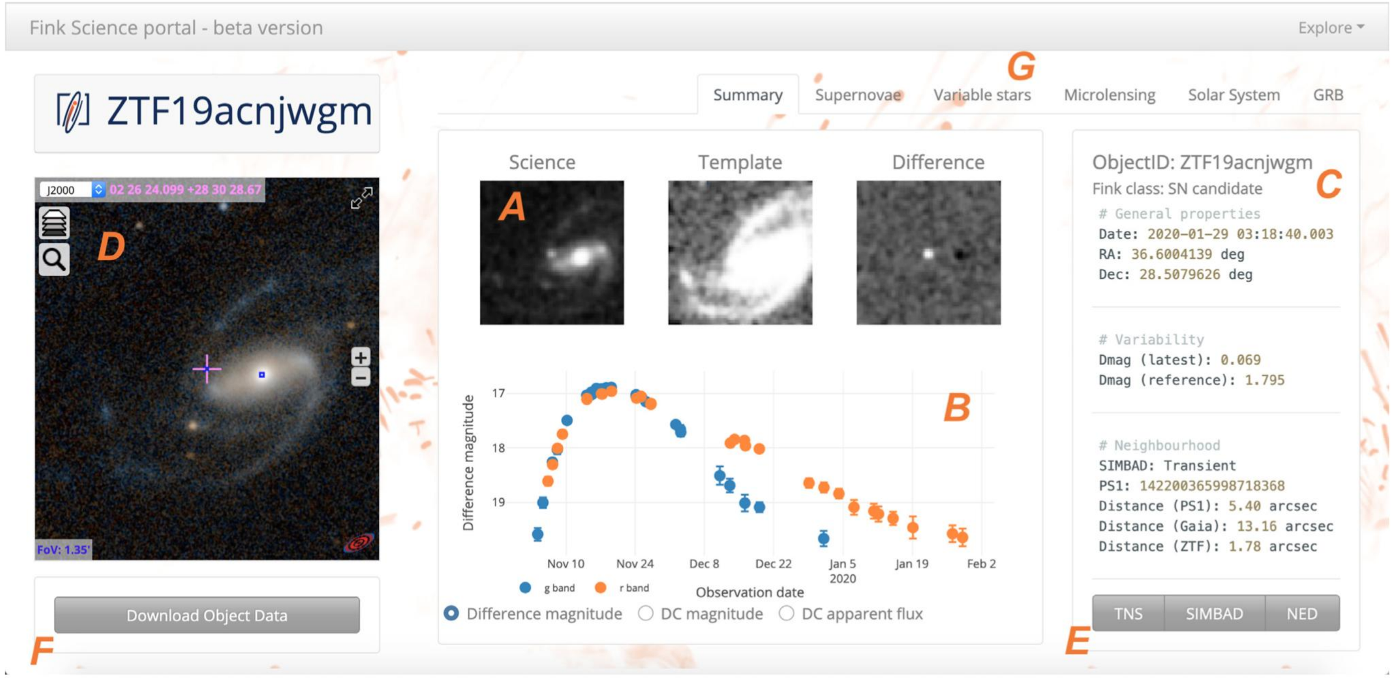
- Communication (APACHE KAFKA cluster)

  The substream data are sent to users via a cluster of five machines (20 cores total, 2 GB RAM/core) where APACHE KAFKA and ZOOKEEPER are deployed.

- Science portal (APACHE HBASE cluster)
- Data store (HDFS cluster)
- GRAFANA and GANGLIA.

Fink performance on ZTF

- The experience accumulated in the last few years in Fink with ZTF has been paramount for the design, development and fine tune of the broker services according to the needs of difference scientific communities.
- Nevertheless, given the volume and complexity of the expected data, restructuring algorithms to transition from ZTF to LSST is a non-trivial task.What is deployed for ZTF typical rates (of the order of 200,000 alerts per night can be easily scaled to ELAsTiCC rates (of the order of 1,000,000 alerts in a few hours every night) by adding more machines.



Currently implemented science modules

- Cross-matching modules:
  - (i) Catalogues: Simbad catalogue with a matching radius of 1 arcsec, using the XMATCH service provided by CDS.
  - (ii) Surveys: LIGO/Virgo, Fermi, and Swift alerts via the COMET broker (live), and survey public catalogues (post-processing).
  - (iii) Other services: Transient Name Server (TNS) for recent classifications.
- Classification modules:
  - (i) Microlensing: Classification of events using LIA based on Godines et al. (2019).
  - (ii) Supernovae partial and complete light-curve classification: recurrent neural network architecture on SUPERNNOVA .
  - (iii) We determine potential Solar system object based on a series of filters.

Classifier and anomaly detection modules

- Algorithms that are able to characterize objects with only a handful of light-curve observations for obtaining rapid and reliable characterization of transient events while they are still active
- Active learning algorithms for improving the training set (such as Bayesian neural networks (BNNs))
- Adaptive machine learning for providing anomaly scores whose accuracy improve with the evolution of the survey. FINK is designed to have a specific AD module, based on contemporary adaptive machine learning techniques that will be specifically designed to optimize the use of domain knowledge.

The ELAsTiCC dataset

- The “Extended LSST Astronomical Time-series Classification Challenge” (ELAsTiCC; Knop & ELAsTiCC Team 2023) was designed to test brokers systems and classification algorithms when applied to a state of the art dataset which mostly resembles LSST alerts.
- Its first objective was to test the brokers infrastructure capability of ingesting and processing a real-time alert stream. The second goal was to enable the evaluation of ML classification algorithms

Test Instance 1:ELAsTiCCv1

- Simulated dataset : alert stream of 3 years of LSST (2023.11.27 -2026.11.27)， the dataset was simulated using SuperNova ANAlysis package and contained 19 classes divided into 5 broad (SN-like, Periodic, Non-periodic, Long and Fast). Light-curves, comprising detections and forced photometry in the LSST broad-band filters {u, g, r, i, z, Y} were provided
- Trainning set: The first year of ELAsTiCCv1 as training sample for all our algorithms (17 233 868 alerts labelled from 2023.11.27 to 2024.11.27 and corresponding to 1 676 431 distinct objects). Test set: The remaining two years of ELAsTiCCv1 (34 872 745 alerts corresponding to 2 865 642distinct objects).

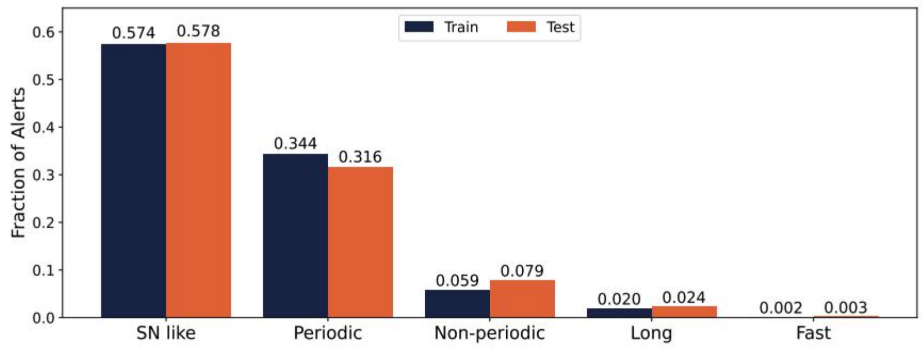


Fig. 2: ELAsTiCC broad class distribution for our training (dark blue) and test (orange) sets.

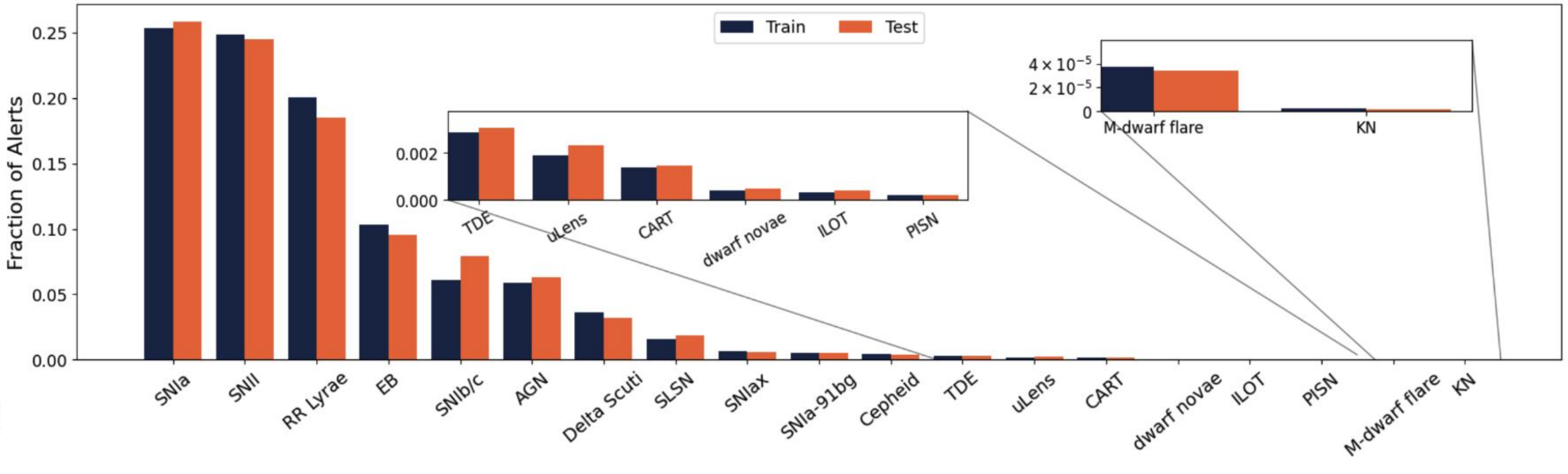
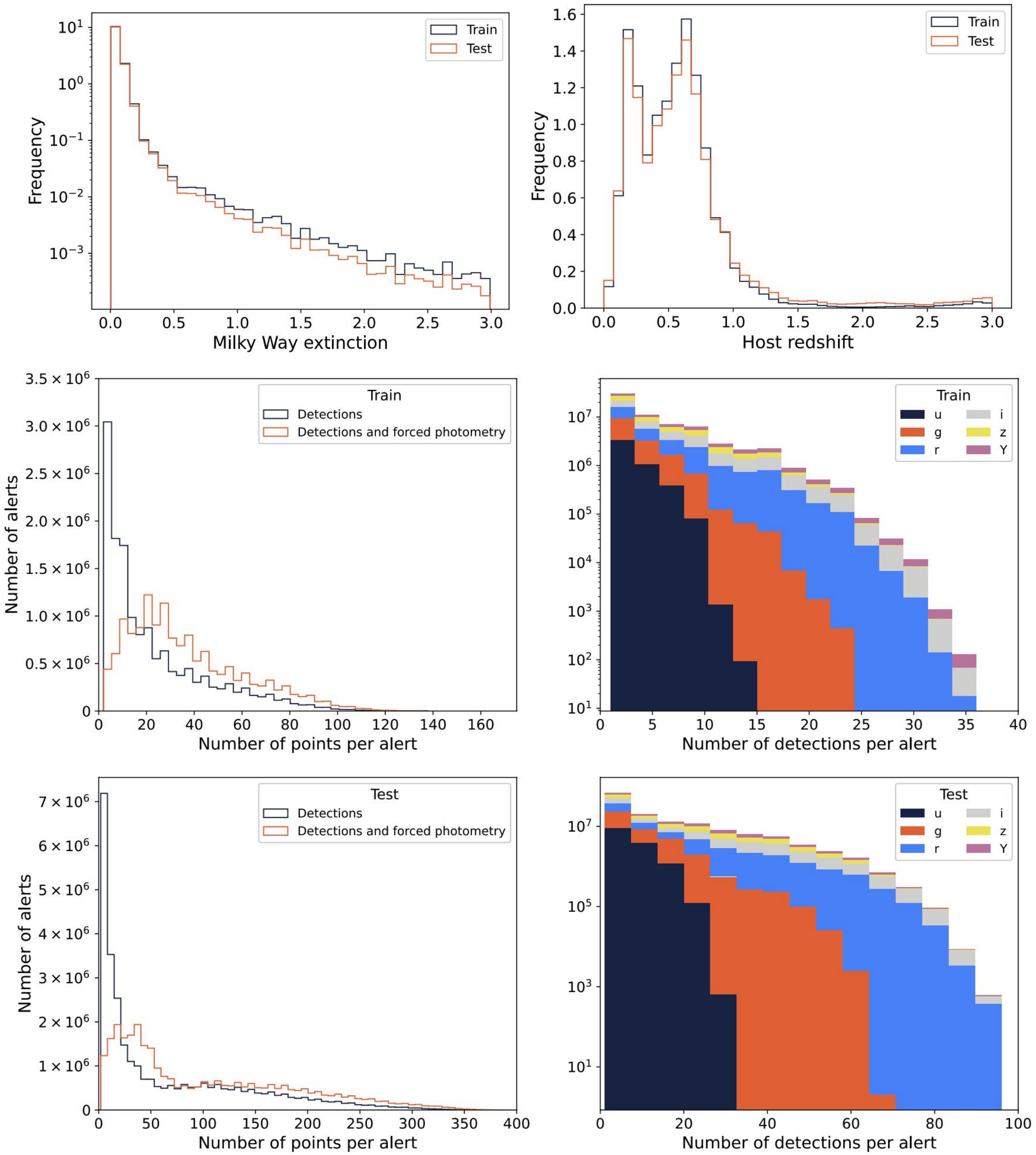


Fig. 1: ELAsTiCC class distribution for our training (dark blue) and test (orange) sets.

Test Instance 1:ELAsTiCCv1

- Each alert package included :light curve data (mjd,fluxcal, fluxcal_err, filter) , metadata includes position, milky way extinction and estimated photometric redshift
- Photometric redshift available:Trainning set: 81% , Test set is 91%,the distributions: displaying a double peaked structure
- The distribution of number of detection points per alert with and without forced photometry.The detections is strongly peaked around 10 detections, dropping heavily after that.
- The distribution of detections per passband is similar for the training and test sets: the redder the band, the larger the maximum number of detections. The exception to this is the z band, which do not appear in longer lightcurves. This is especially important for classifiers that rely on colours or use only specific passbands. Nevertheless, this feature is a direct consequence of the chosen survey strategy, and it is reasonable to expect that the real data will also hold differences in number of detections on each band, thus it is paramount to access the robustness of classifiers in this scenario.
-



Classifiers in Fink:Metric

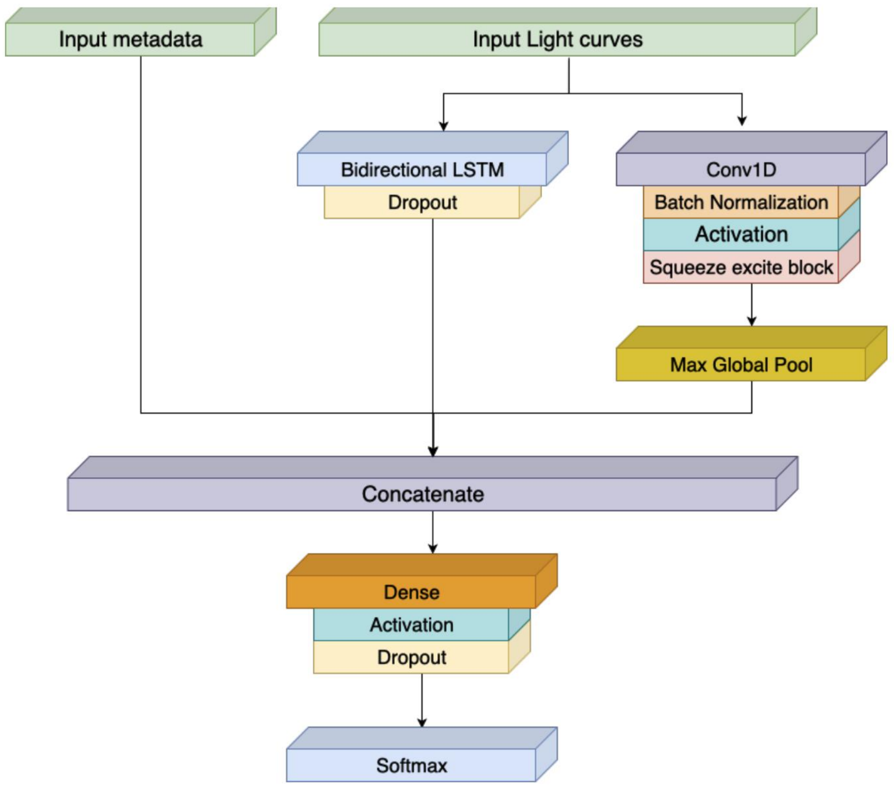
- In a classification task, several metrics can be used to assess the performance of the classifier, such as the Receiver Operating Characteristic (ROC) and Precision-Recall curves, and the Confusion Matrix. These are built from the Precision (P), Recall (R, also called the True Positive Rate,TPR) and False Positive Rate (FPR), which in a binary classification are defined as:

$$\mathrm{P} = \frac{TP}{TP + FP},$$
$$\mathrm{R} = \frac{TP}{TP + FN}$$
$$\mathrm{FPR} = \frac{FP}{FP + TN},$$

With TP(N) the number of true positives (negatives) and FP(N) the number of false positives (negatives). Precision can be understood as the purity of the predictions, while Recall is its completeness or efficiency, and the FPR is the ratio of wrongly classified objects of the negative class (also known as the false alarm rate).

Classifiers in Fink

- The CBPF Alert Transient Search (CATS)



	ROC AUC	PR AUC	Precision	Recall
SN-like	0.99 (0.0002)	0.99 (0.0003)	0.97 (0.002)	0.99 (0.001)
Fast	0.99 (0.0009)	0.82 (0.017)	0.89 (0.012)	0.71 (0.03)
Long	0.96 (0.0025)	0.65 (0.014)	0.79 (0.031)	0.47 (0.03)
Periodic	1.00 (0.00001)	1.00 (0.00003)	1.00 (0.0002)	1.00 (0.0001)
Non-Periodic	1.00 (0.00003)	1.00 (0.0005)	0.97 (0.003)	0.96 (0.005)

Classifiers in Fink

- SuperNNova (SNN)

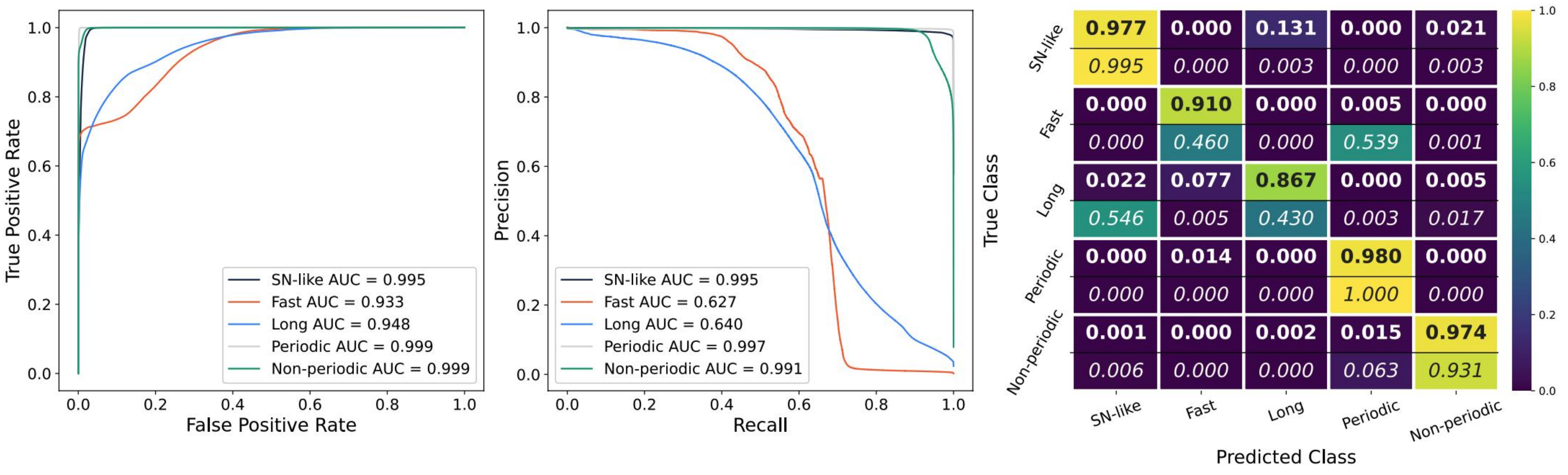
SuperNNova (SNN; Möller & de Boissière 2019) is a deep learning light-curve classification framework based on Recurrent Neural Networks. SuperNNova makes use of fluxes over different band-passes and their measurement uncertainties over time for classification of time-domain candidates in different classes. Additional information such as host-galaxy redshifts and Milky Way extinction and their errors can be included to improve performance.

Class	Accuracy	ROC AUC	Precision	Recall
SN-like	97.18	0.9937	95.68	97.70
Fast	99.04	0.9976	99.52	98.57
Long	83.79	0.9198	89.04	77.64
Periodic	99.59	0.9999	99.48	99.76
Non-Periodic	99.59	0.9999	99.48	99.76
Broad	87.96	-	77.61	67.53

Table 2: SuperNNova performance for complete light-curves using an independent test set from the first year of alerts. All rows except the last one show the metrics for a binary target vs. other types.

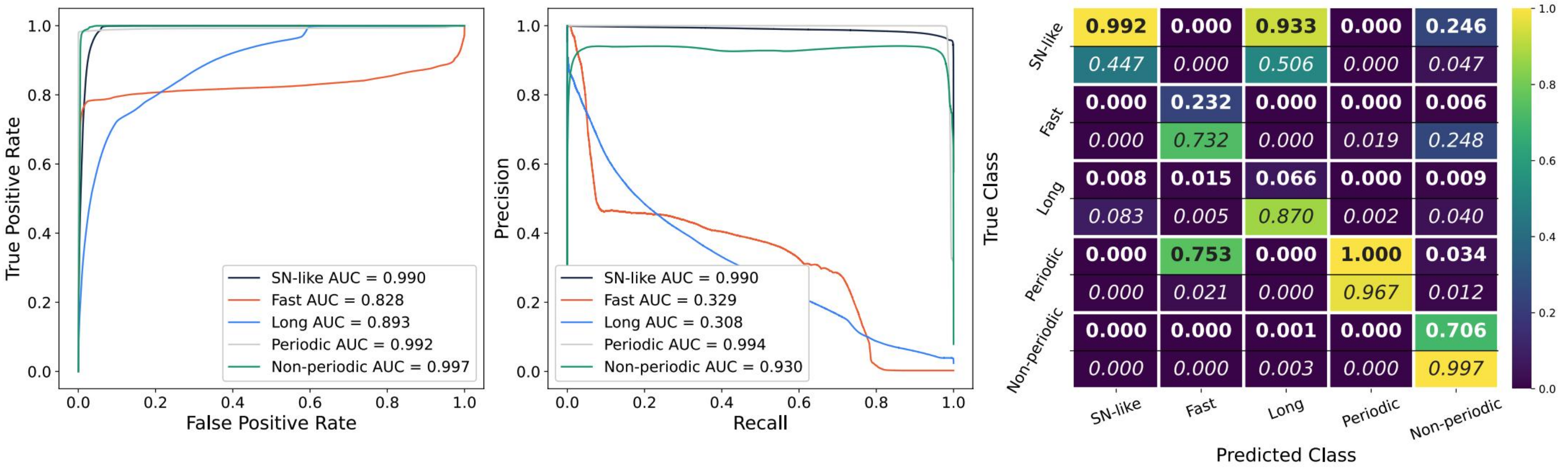
- Binary and broad class models.
  - Superluminous Supernovaee (SLSN) classifer: Superluminous supernovae (SLSN) are SNe whose peak optical luminosity exceeds −21 mag . Their rise times can vary between ~ 20 days to more than 100 days for some events.
  - Early Supernova Ia classifer : a random forest-based classifier

Classifiers in Fink



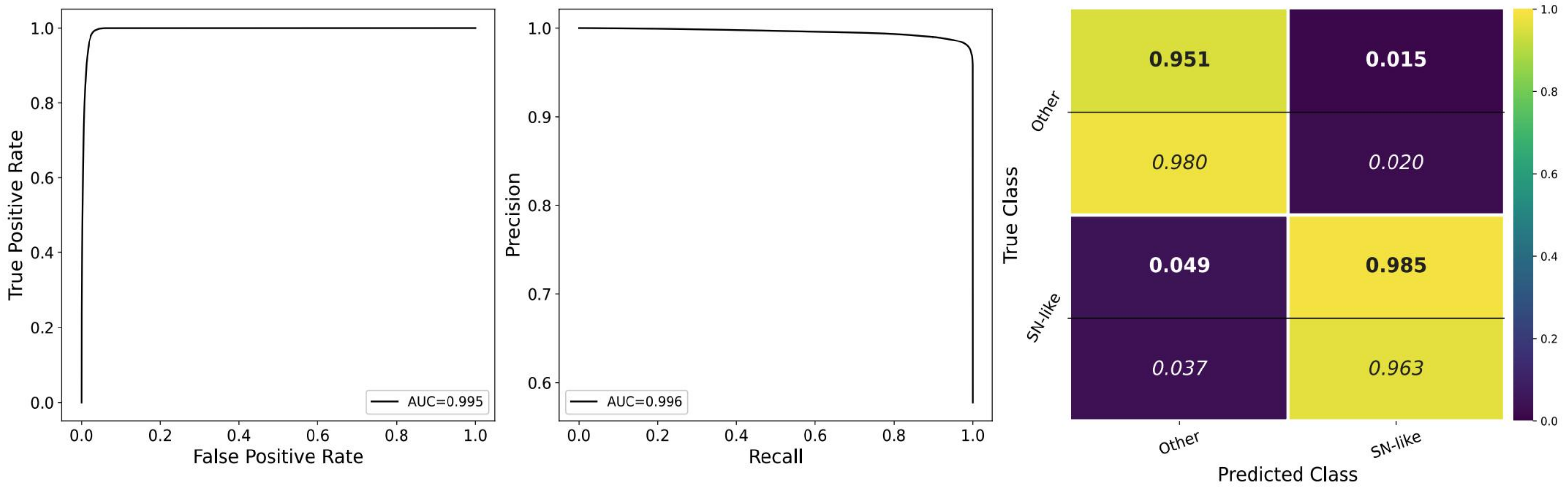
(a) CATS

The model performs excellently with the SN-like, Periodic and Non-Periodic classes, having AUCs above 0.95
However, it had problems with the Long class: more than half of the alerts were classified as SN-like, which is a result of the choice of training set.

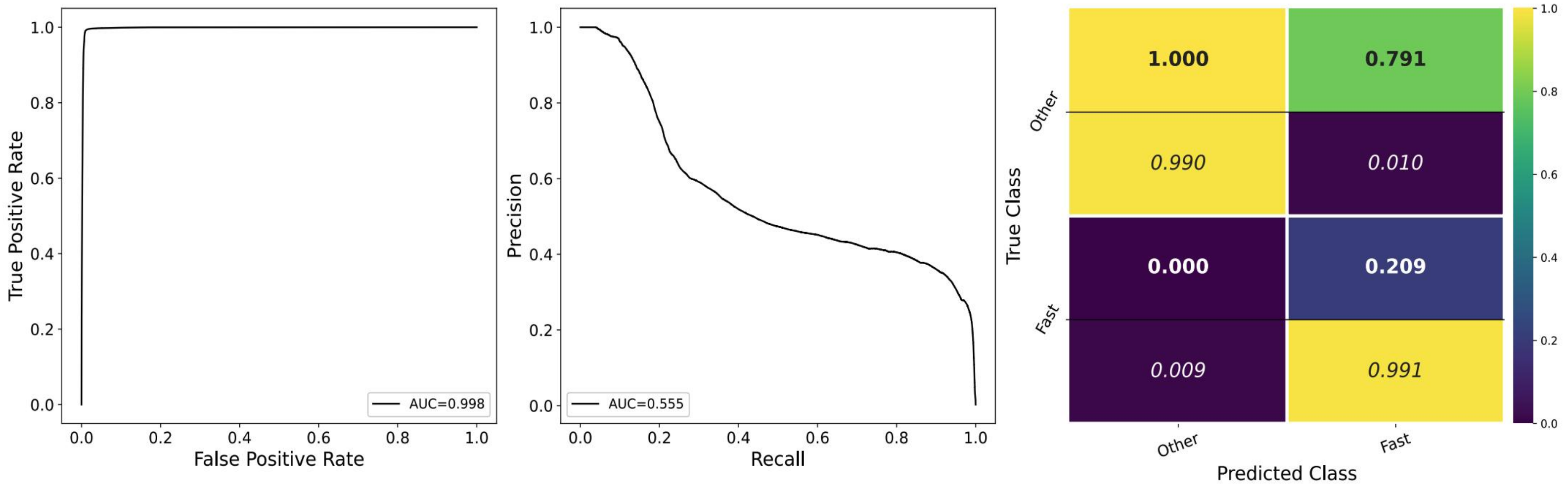


(b) SuperNNova broad

We show the performance metrics and confusion matrix for SNN broad classifier. As found when validating the model, both the SN and Long classes have large classification confusion.



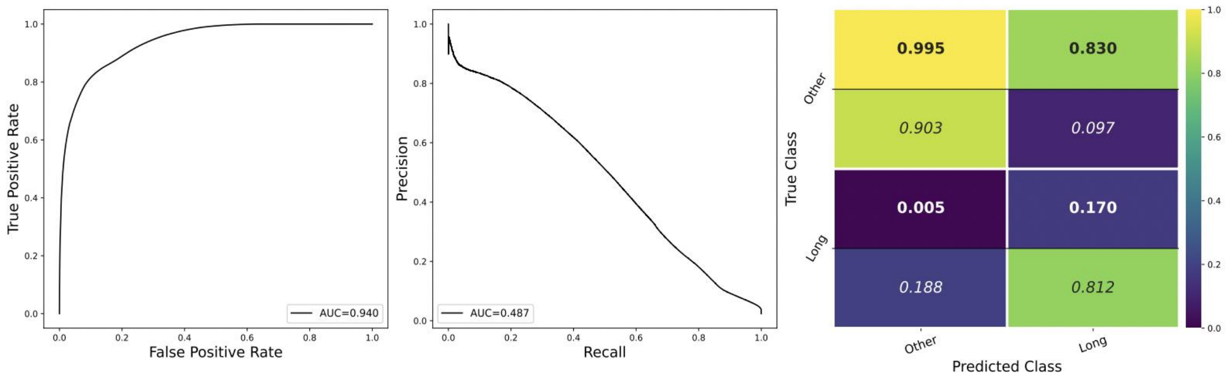
(c) SuperNNova binary for SN-like



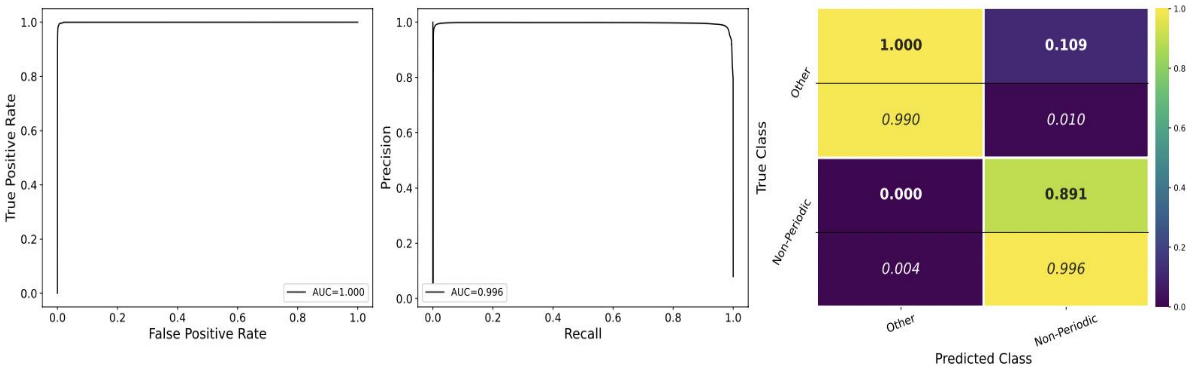
(d) SuperNNova binary for Fast

We find better performance for the SN-like classification with the binary classifier (Figure c) than with the broad one. This suggests, as expected, that the increase of the training set for the target is extremely important for ouralgorithm.

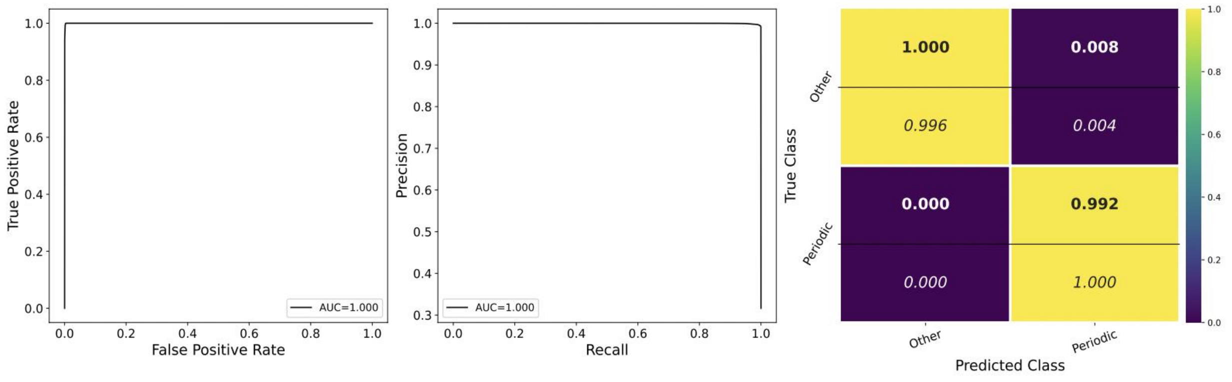
Classifiers in Fink



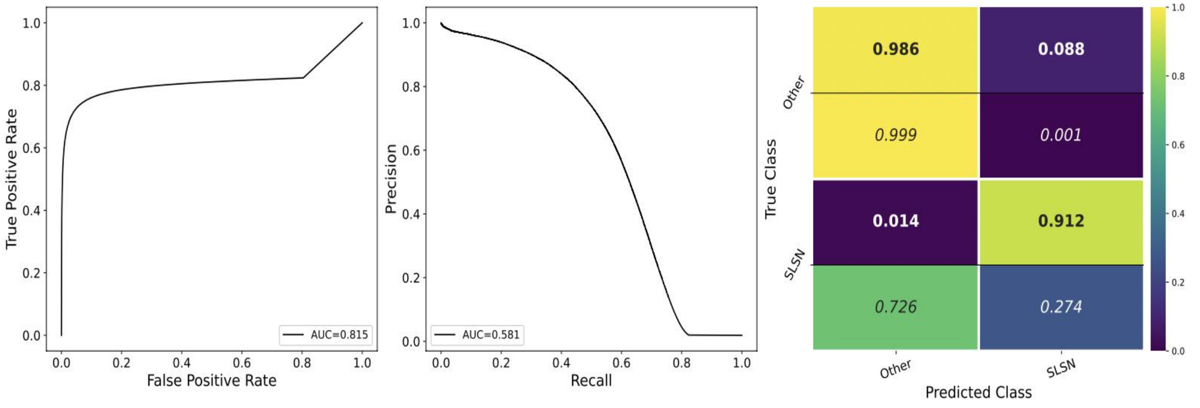
(e) SuperNNova binary for Long



(g) SuperNNova binary for non-Periodic



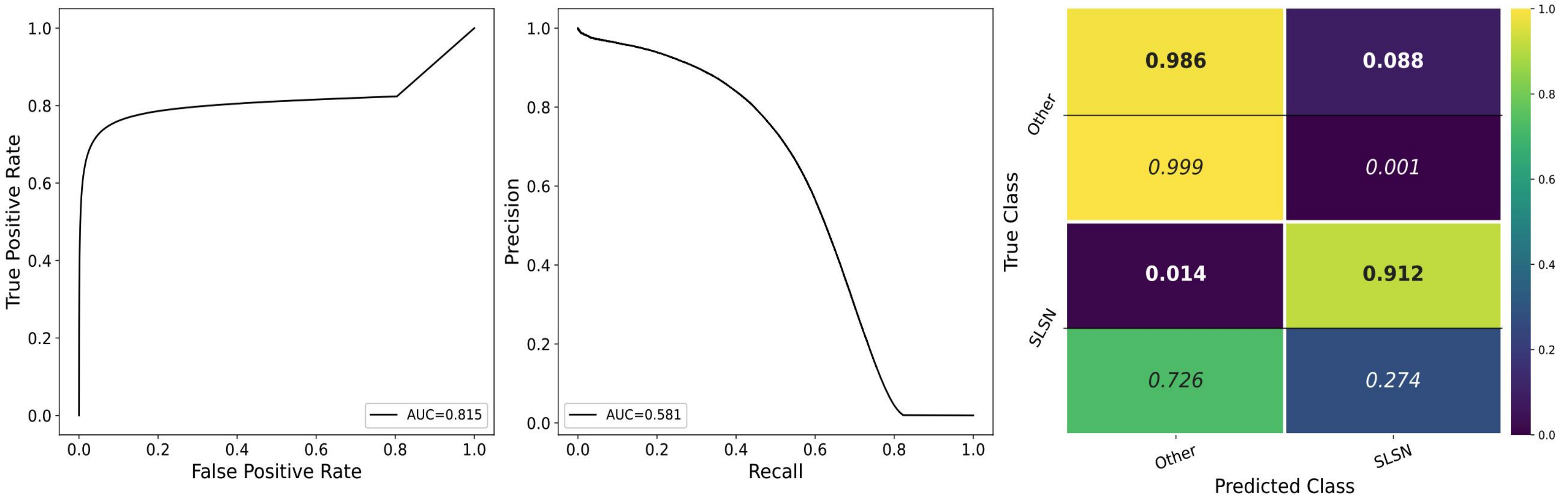
(f) SuperNNova binary for Periodic



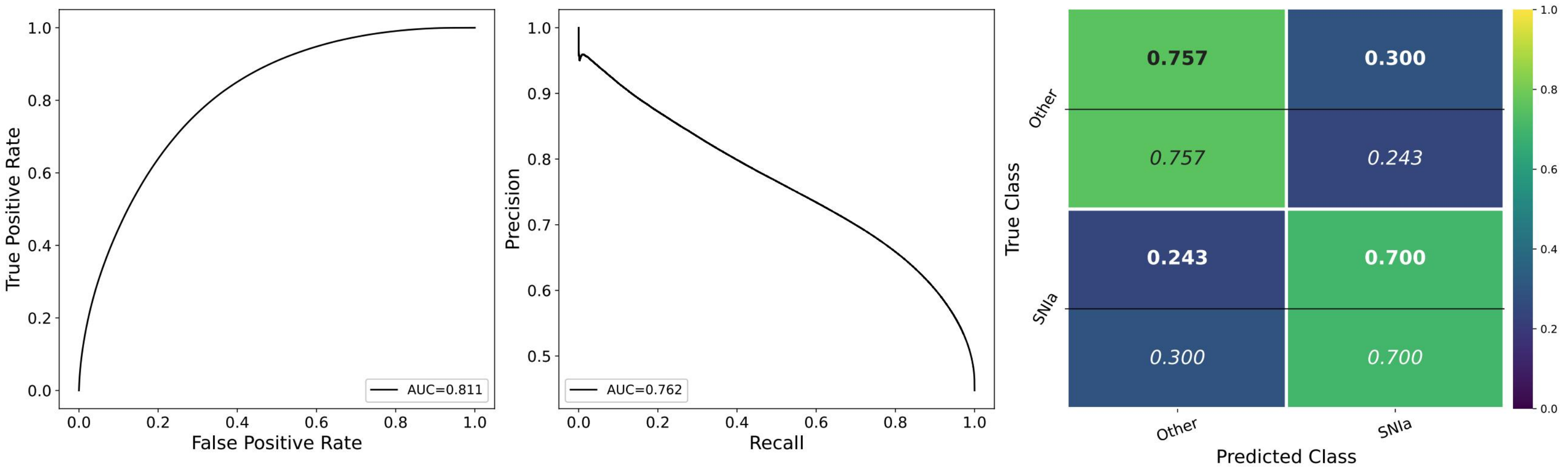
(h) SLSN

We also find an improvement for classes with smaller training sets such as Fast (Figure d) and Long(Figure e). However, these two classes still are challenging go classify.

Classifiers in Fink

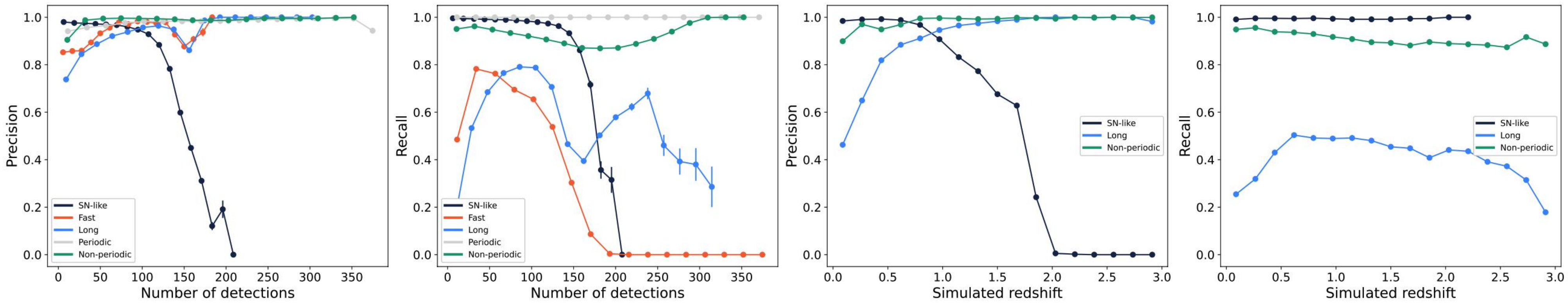


(h) SLSN

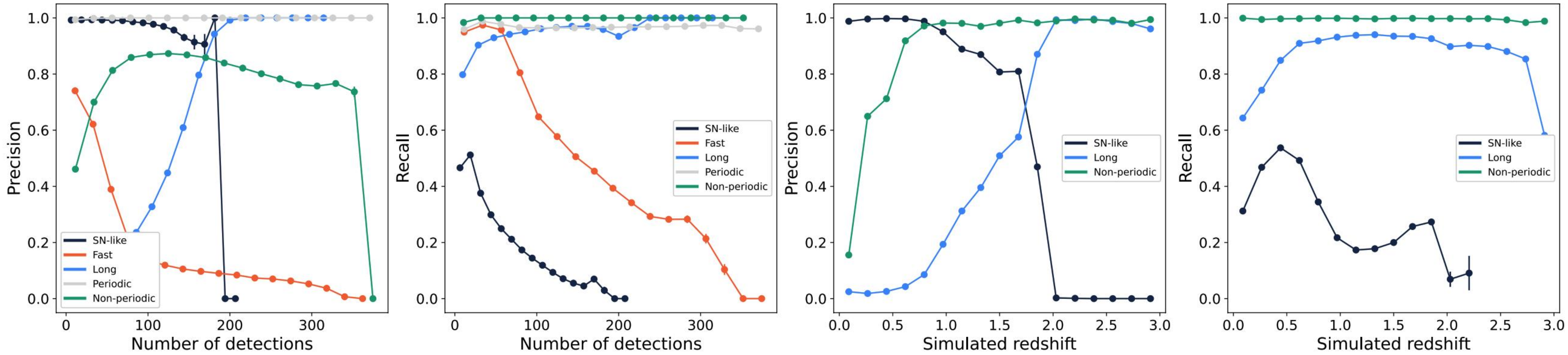


(i) EarlySNIa

Classifiers in Fink

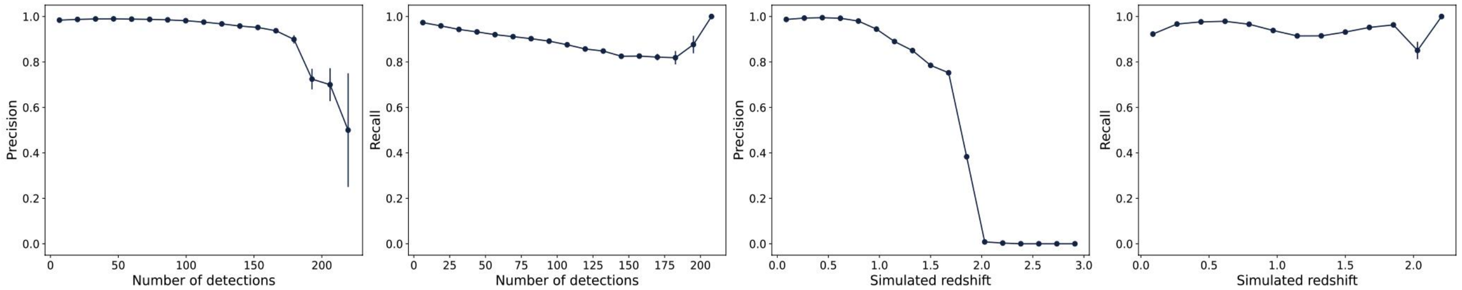


(a) CATS

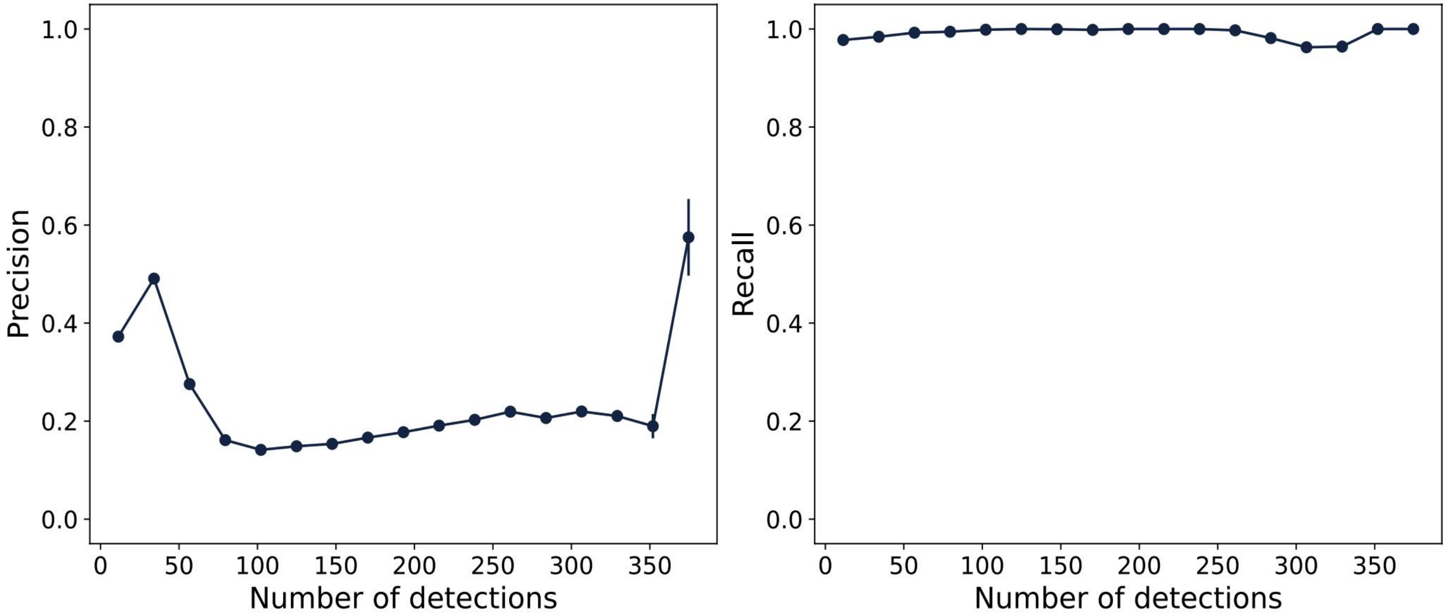


(b) SuperNNova broad

Classifiers in Fink



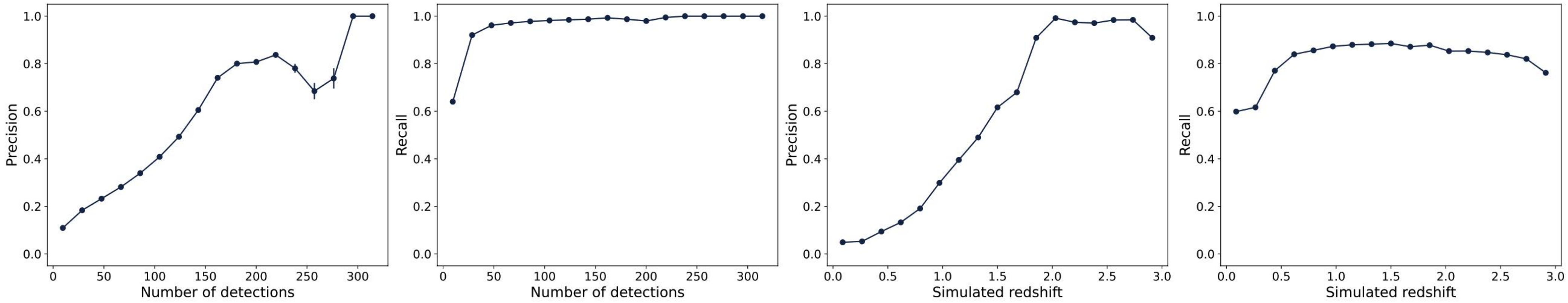
(c) SuperNNova binary for SN-like



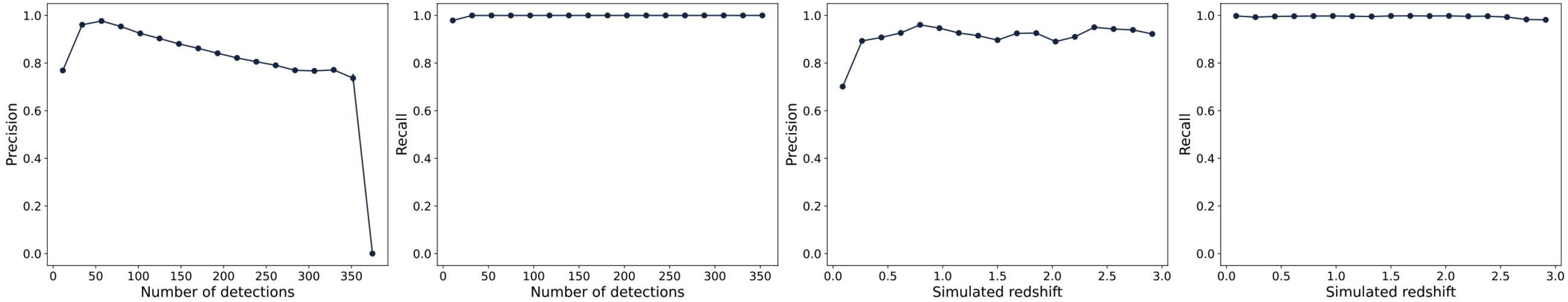
(d) SuperNNova binary for Fast

Evolution of Precision and Recall as a function of number of detections (left two panels) and host galaxy redshift (right two panels). Fast and Periodic alerts have no redshift available and thus have only the first two panels.

Classifiers in Fink



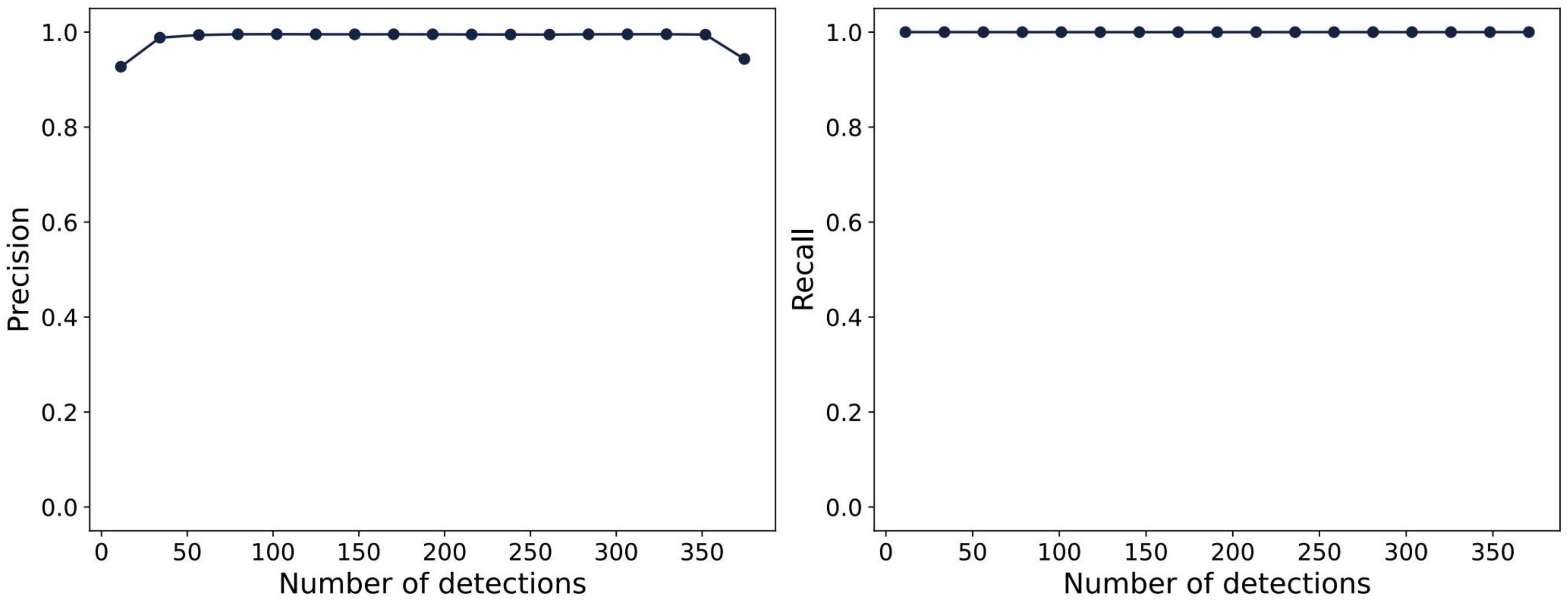
(e) SuperNNova binary for Long



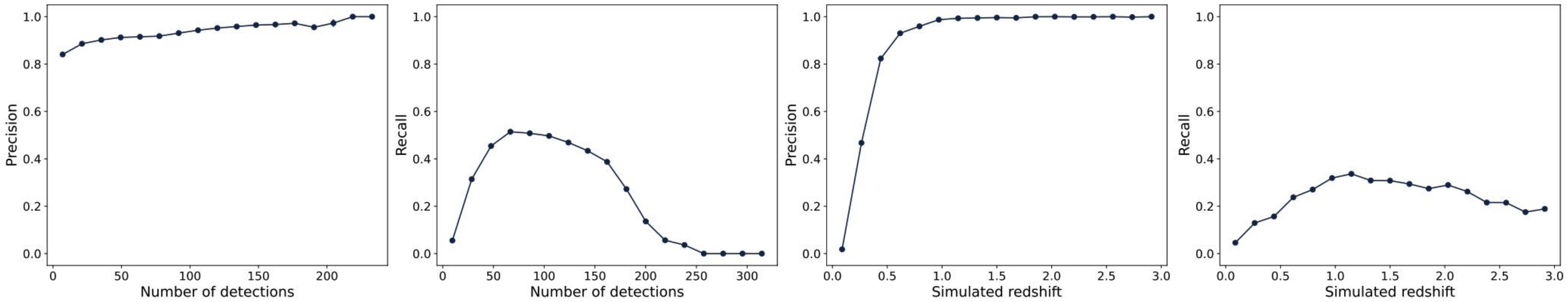
(f) SuperNNova binary for non-Periodic

The evolution of metrics as a function of number of detections and redshift for the SNN binary classifiers. The same trend as in the broad classifier is found, our Long classifier increases its precision as more detections are available (Figure e). Non- Periodic classification (Figure 9f) is found to be very stable with respect of number of detections and redshift provided

Classifiers in Fink



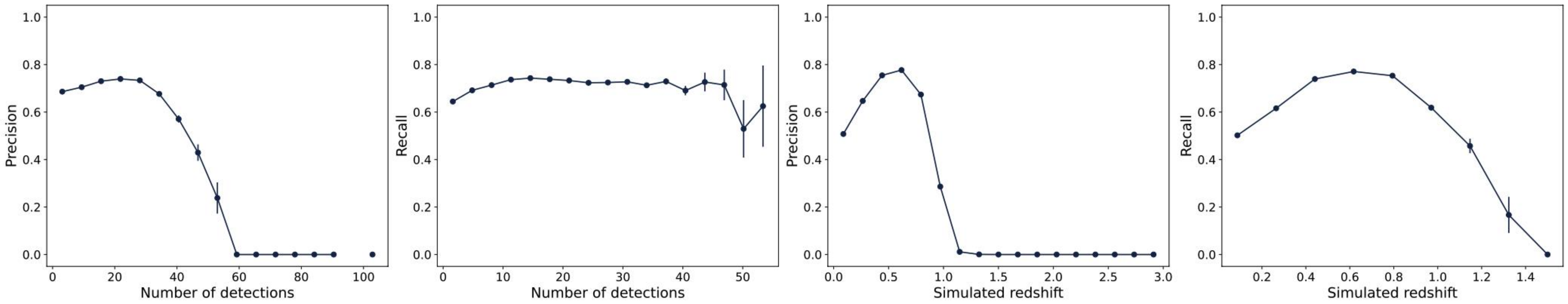
(g) SUPERNNOVA binary for Periodic



(h) SLSN classifier

Periodic and SN-like classification also show stable performance with respect to number of detections (Figure g). Thus, we expect a good performance in the classification of early light-curves for these classes. This is an important feature when scheduling follow-up observations as explored for Rubin SNe Ia in Möller et al. (2024).

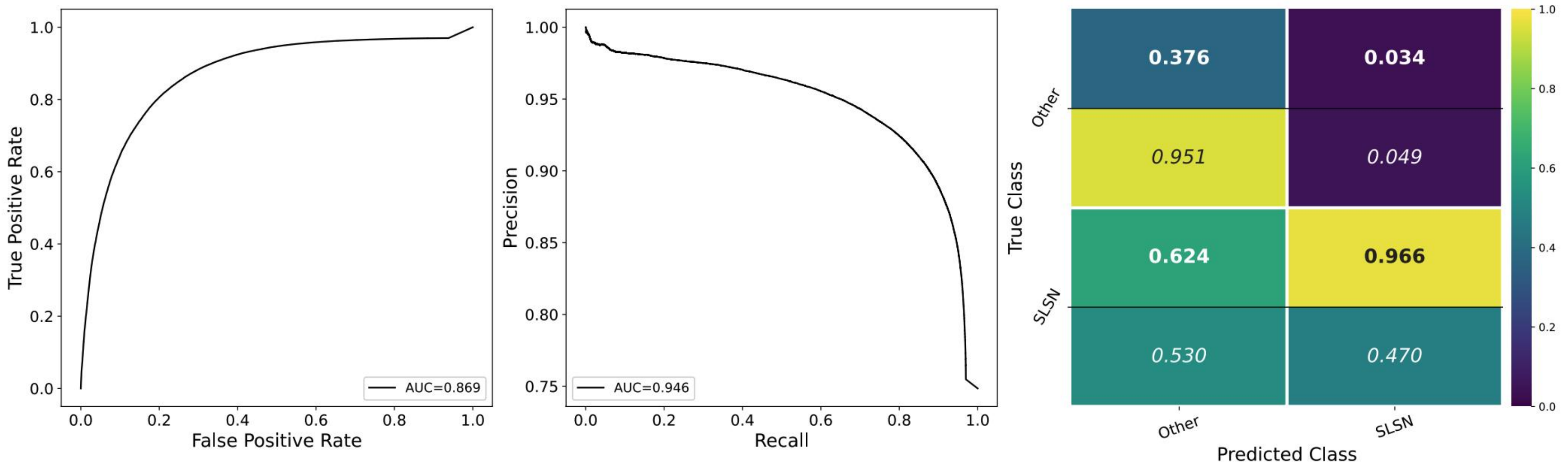
Classifiers in Fink



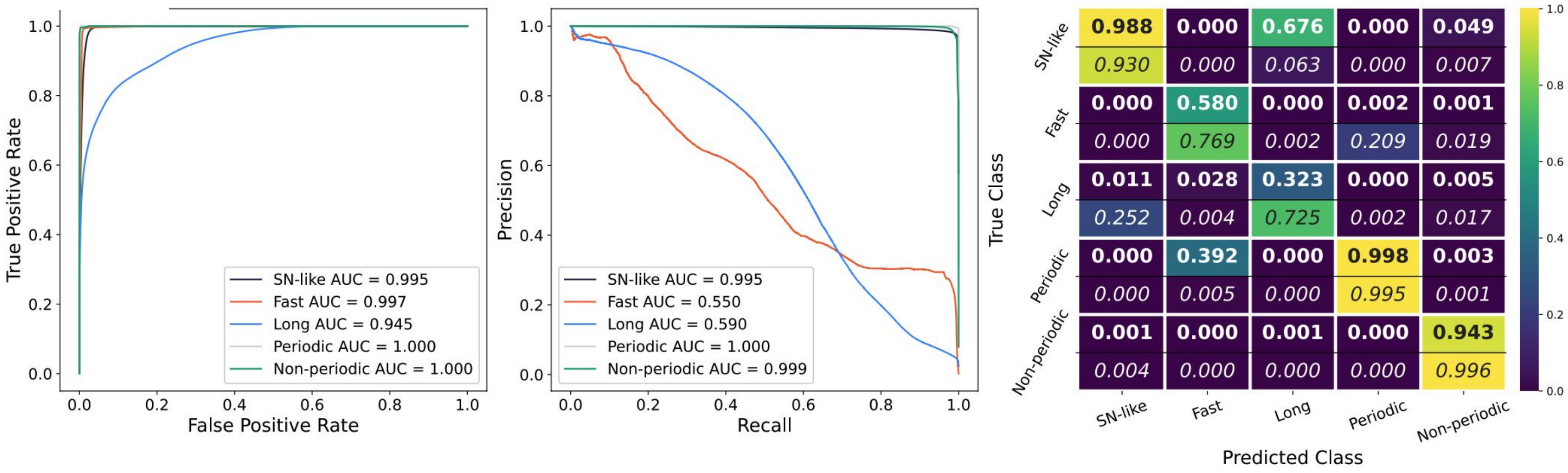
(i) EarlySNIa

Figure i shows how classification results evolve with the number of detections and simulated redshift. We note that precision already starts higher than 0.6 for 7 observed data points (the minimum requirement) and peaks around 20 photometric points, while recall remains almost sta-ble even with more detections. The sample identified as EarlySNIa by the algorithm is highly skewed towards small light curves, with ~ 75% of them having 10 detections or less.

Classifiers in Fink



Performance metrics for the SLSN classifier applied to the sample of alerts classified as Long by CATS.



Performance metrics for the combined SuperNNova binary classifiers.

Discussion

- For a few years now, broker teams have been successfully working with the ZTF alert stream and communication protocols as a test bench for what is to be expected for LSST.This experience has been extremely successful and has alowed the development of an entire broker ecosystem, along with a diverse and interdisciplinary community.
- ELAsTiCC is a kind reminder that, beyond hardware and data format, machine learning models and broker infrastructure will need to change significantly in order to fulfil expectations which rise with the arrival of LSST.
- This includes the design of algorithms themselves, protocols for massive data transfer between geographically disconnected science teams, experiment design for proper evaluation and optimisation of trained models to allow processing of millions of alerts per night. The analysis presented here describes the strategies developed by the Fink team to address these issues.