



ALeRCE

Automatic Learning for the
Rapid Classification of Events

<http://alerce.science/>



Danilo Alvares, Ismael Alvarez, Javier Arredondo, Nicolás Astorga, Franz Bauer, Guillermo Cabrera-Vives, Rodrigo Carrasco-Davis, Ernesto Castillo, Márcio Catelan, Andrew Connolly, Demetra De Cicco, Cristóbal Donoso, Felipe Elorrieta, Pablo Estévez, Susana Eyheramendy, **Francisco Förster**, Germán García, Matthew Graham, Pablo Huijse, Ashish Mahabal, Giovanni Motta, Rosario Molina, Giuliano Pignata, Pavlos Protopapas, Esteban Reyes, Ignacio Reyes, Diego Rodríguez, Daniela Ruz, Juan Sáez, Paula Sánchez-Sáez, Camilo Valenzuela, Jorge Vergara



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE



CMM
Center for
Mathematical
Modeling



MILLENNIUM
INSTITUTE OF
ASTROPHYSICS

DO
DATA OBSERVATORY



CD
CENTER FOR DATA-DRIVEN DISCOVERY



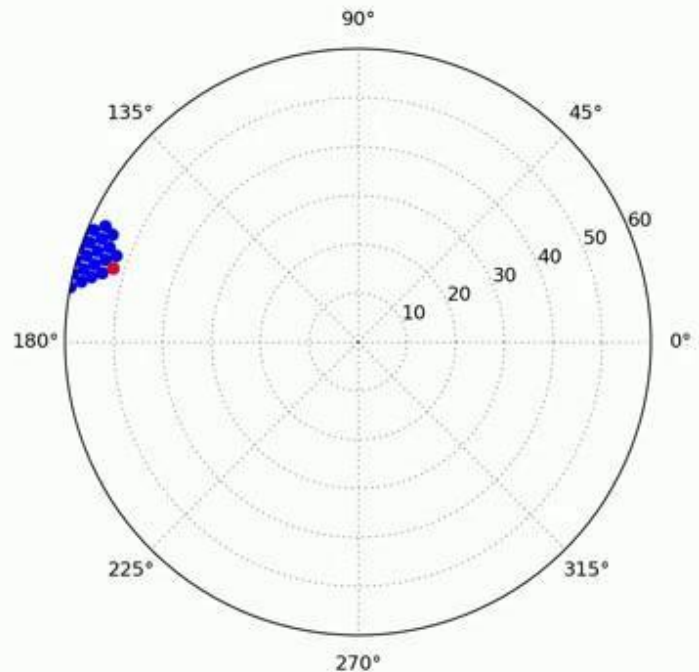
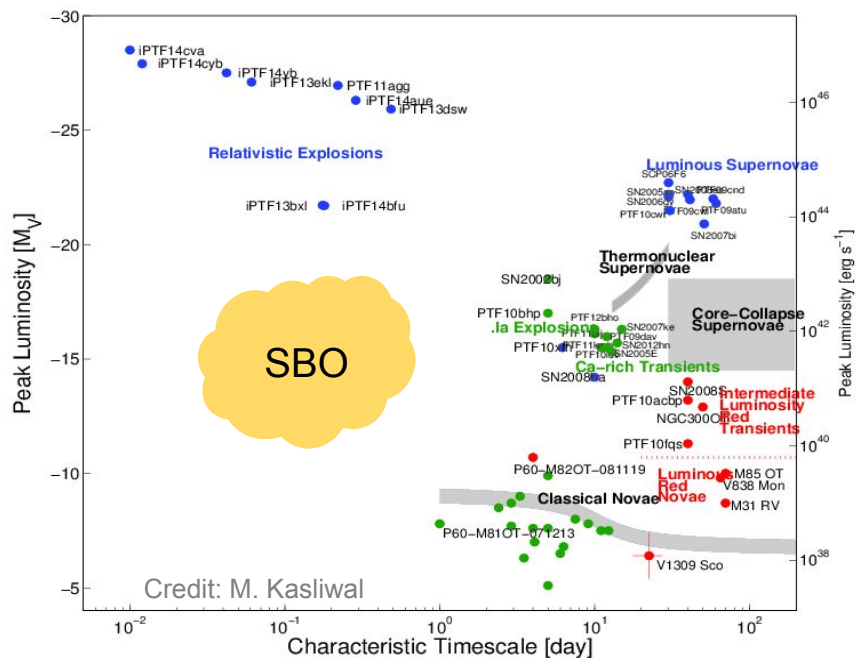
Overview:

1. High Cadence Transient Survey (HiTS)
2. Automatic Learning for the Rapid Classification of Events (ALeRCE)
3. DEMO!

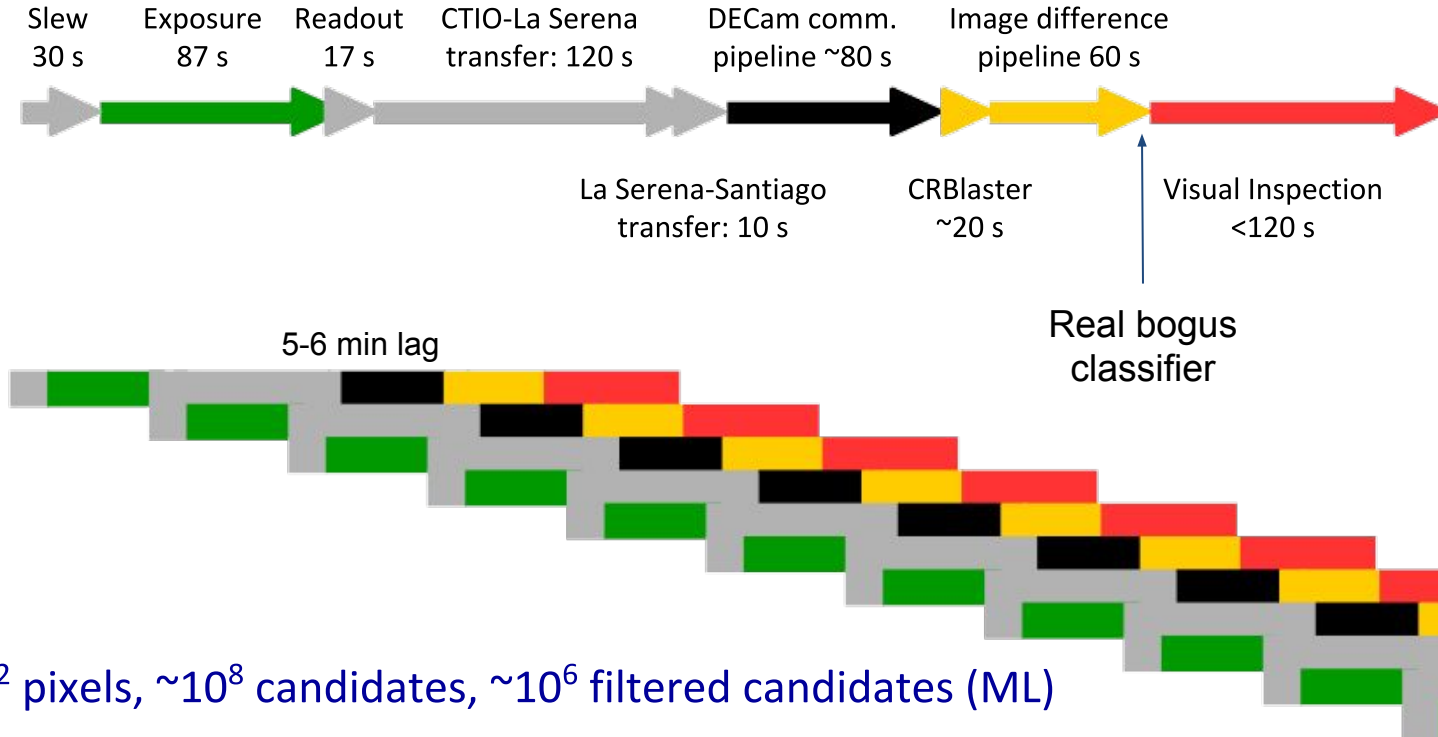


1. High Cadence Transient Survey (HiTS)

The High cadence Transient Survey (HiTS)



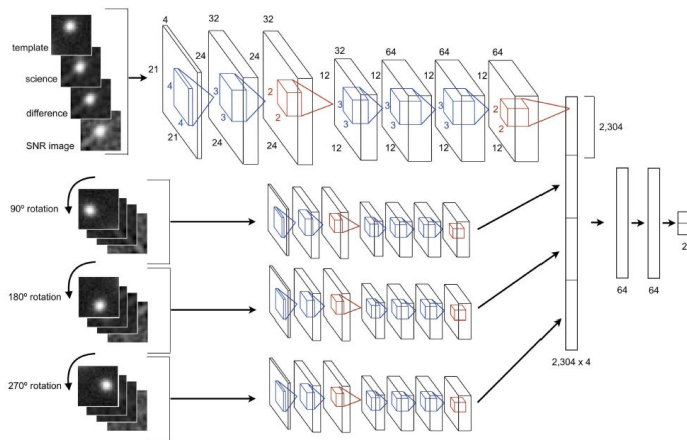
Pipeline flow outline



$\sim 10^{12}$ pixels, $\sim 10^8$ candidates, $\sim 10^6$ filtered candidates (ML)

$\sim 10^4$ visual inspections, 125 SNe

Deep learning for real-bogus classification



Cabrera-Vives+2016 IJCNN, 2017 ApJ

Deep-HiTS: Rotation Invariant Convolutional Neural Network for Transient Detection

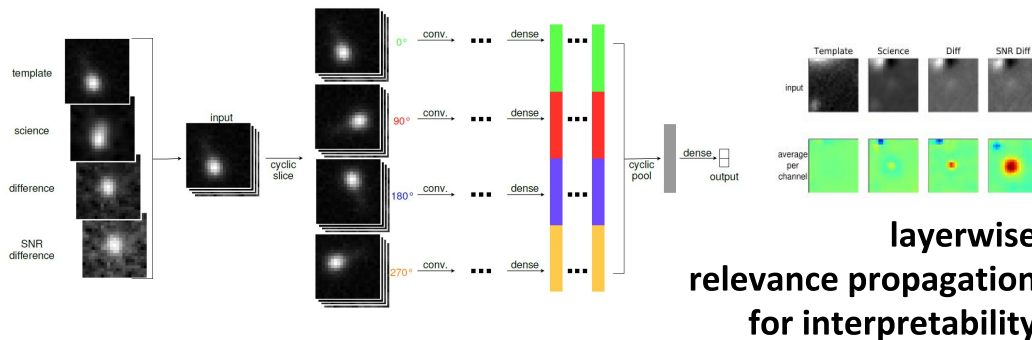
<https://github.com/guille-c/Deep-HiTS>

99.45% vs 98.96% (Random Forest)
F1-score



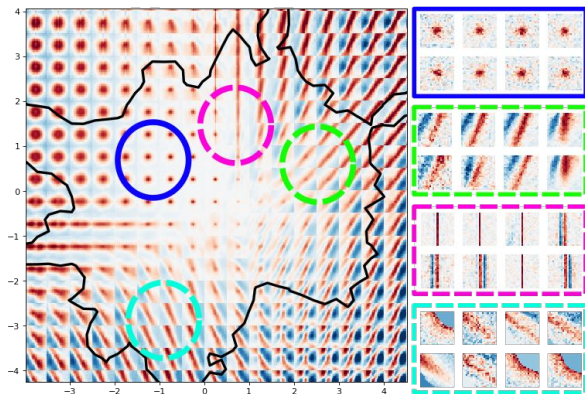
Reyes+2018 IJCNN
Enhanced Rotational Invariant
Convolutional Neural Network for
Supernovae Detection
arxiv:1808.03626

99.53% vs 99.45% (Deep-HiTS)
F1-score



**layerwise
relevance propagation
for interpretability**

Deep learning for real-bogus classification



Huijse+18, ESANN

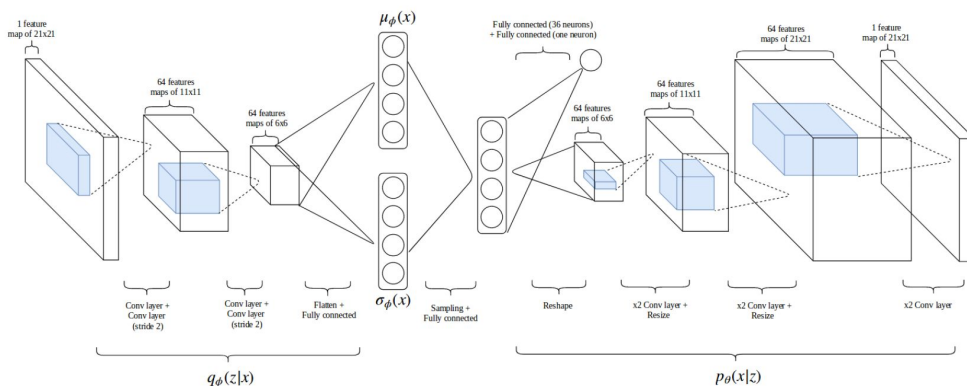
Latent representations of transients from an astronomical image difference pipeline using VAE



Input

Latent dimension

Output

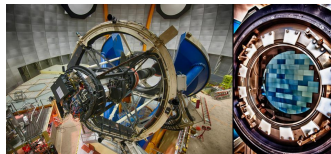


Astorga+18, IJCNN

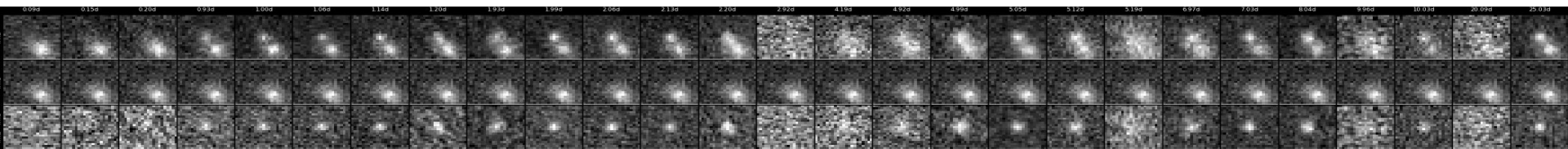
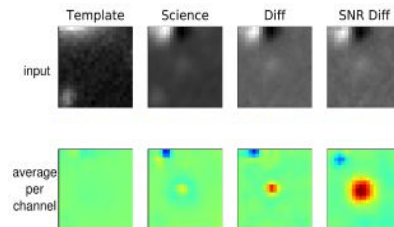
Clustering of Astronomical Transient Candidates Using Deep Variational Embedding



HiTS in a nutshell



- 320 deg² deep & high cadence survey, 1st real time analysis of DECam (Feb 2014), 125 SNe!
- Supernova shock breakout model constraints (Förster+16, ApJ)
- **1st deep learning real/bogus classifier (Cabrera-Vives+16,17; Reyes+18, Huijse+18, Astorga+18)**
- Distant RR Lyrae (Medina+17,18, ApJ)
- ~10k new asteroids (Peña+18, AJ)
- ~22M public variable catalog (Martínez+18, AJ)
- Evidence for CSM around most SNe II (Förster+18, Nat. Ast.)
- **1st RCNN image sequence classifier (Carrasco-Davis+19, PASP)**
- New population of intermediate mass black holes (Martínez-Palomera+19, ApJ)



2. Automatic Learning for the Rapid Classification of Events (ALeRCE)



ALeRCE: from HiTS to LSST



HiTS

2013-2015 (~3 weeks)

0.2 TB per night

~20 million objects

~100 million measurements

~0.1 million alerts per night

→
10x



ZTF

2018-2020

1.4 TB per night

~1 billion objects

~1 trillion measurements

~1 million alerts per night

→
10x



LSST

2022-2032

15 TB per night

~37 billion objects

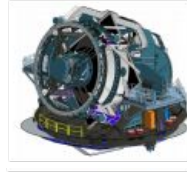
~7 trillion measurements

~10 million alerts per night



Future time domain astronomy ecosystem

Survey telescopes



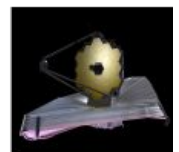
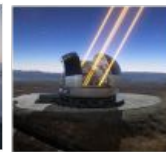
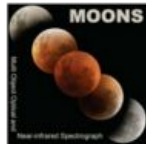
Alert brokers/TOMs



+ *other brokers/TOMs*

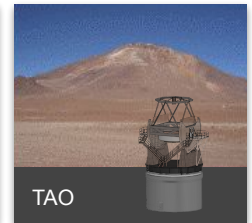
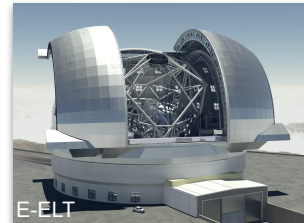
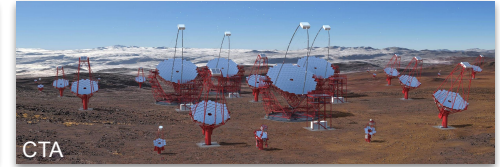
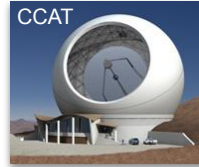


Follow up telescopes





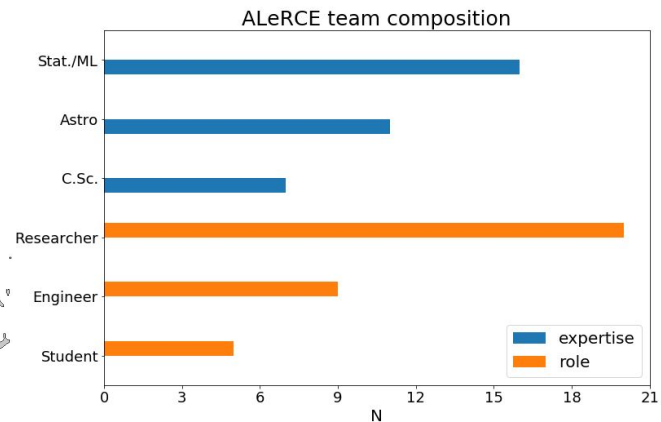
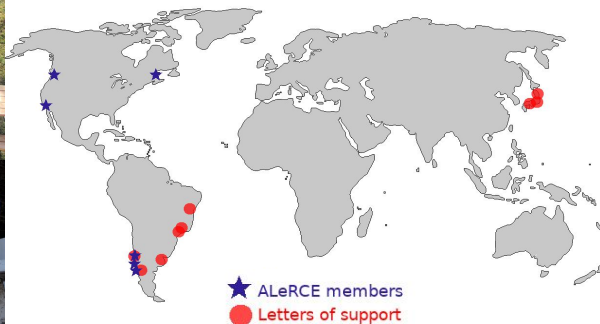
Astronomical infrastructure in Chile



Chilean institutions: access to ~10% observing time



ALeRCE



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

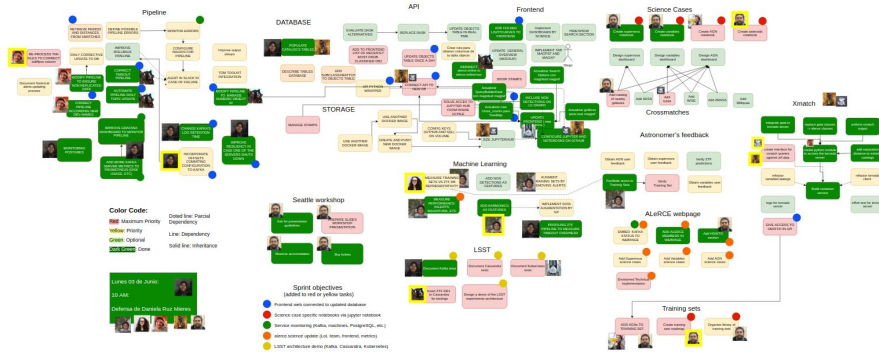


CMM
Center for
Mathematical
Modeling

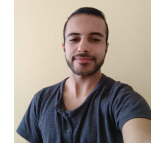




Agile Methodology



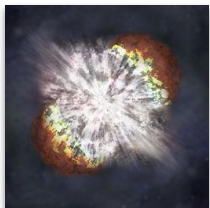
- Daily meetings + 2 weeks sprints
- Product owner
- Project wall
- Weekly ML/training set meetings
- Hackathons



Javier Arredondo, Ernesto Castillo, FF, Ismael Álvarez, Diego Rodríguez, Daniela Ruz, Juan Sáez, Ignacio Reyes, Camilo Valenzuela



Scientific Questions



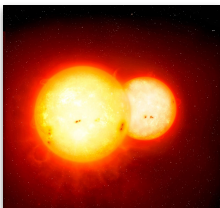
Transients

Rapid photometric & spectroscopic follow-up: nature of the progenitors (outermost layers) & **explosion physics** (ejecta structure).

Short-lived transients (GWs, GRBs)

Cosmological **distance rulers**.

Rare populations of events?



Variable stars

Rapid photometric & spectroscopic follow up: low mass microlensing events, changing mode stellar pulsators, rapid reaction to eclipsing events, eruptive events

Analysis of large populations of events: study **Milky Way structure & formation**.



Active Galactic Nuclei

Rapid photometric & spectroscopic follow-up: changing look AGNs, reverberation mapping studies

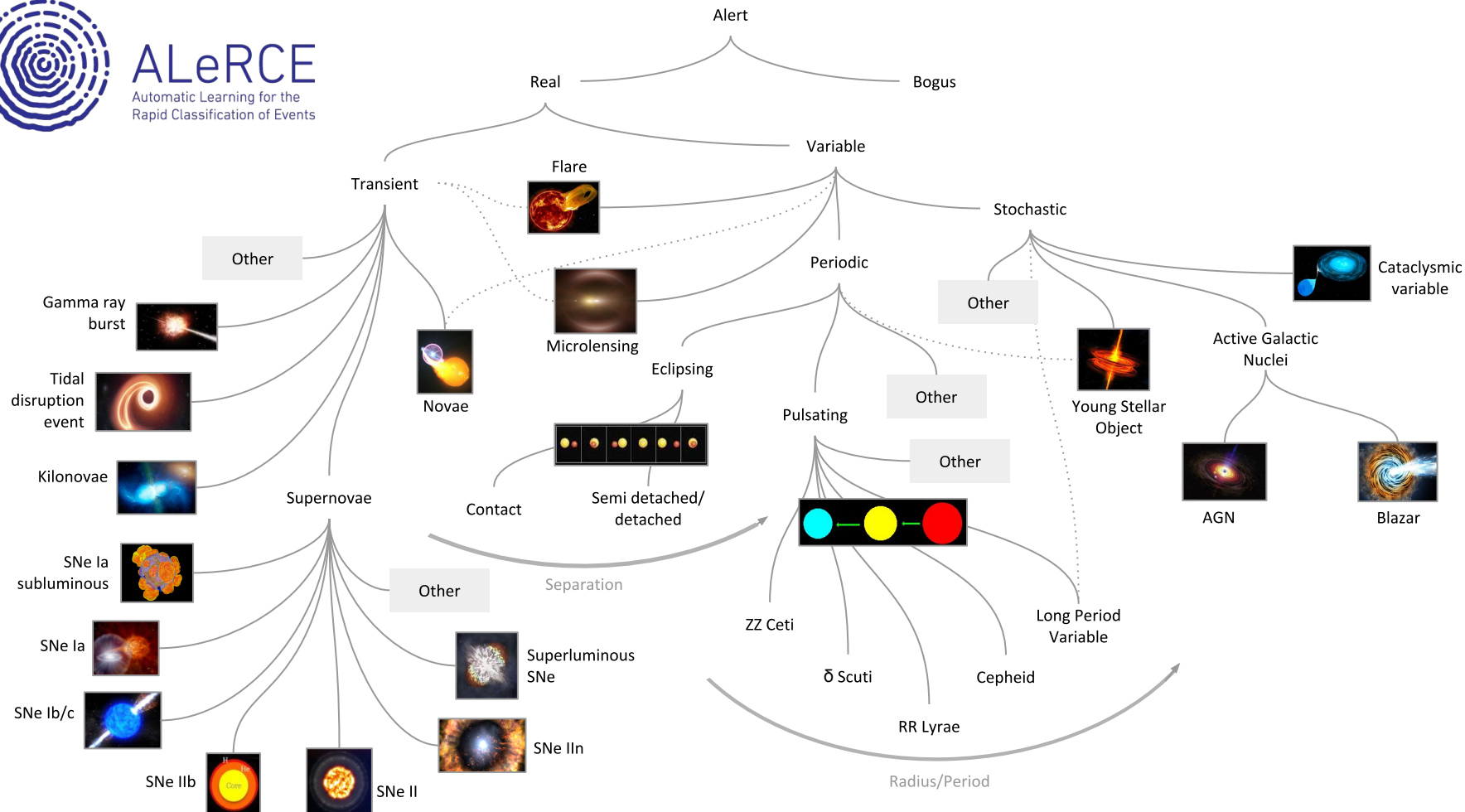
New populations of **faster, redder, dimmer AGNs**

Detection of **intermediate mass black holes, tidal disruption events**



ALeRCE

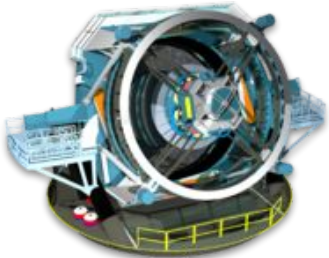
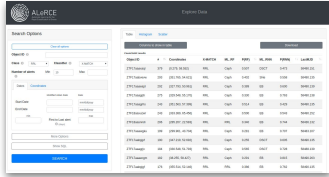
Automatic Learning for the Rapid Classification of Events



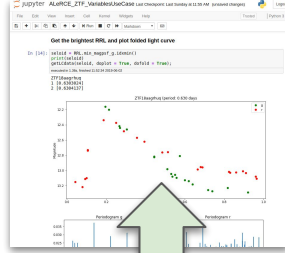


Frontend

(web based filtering, visualization, download)



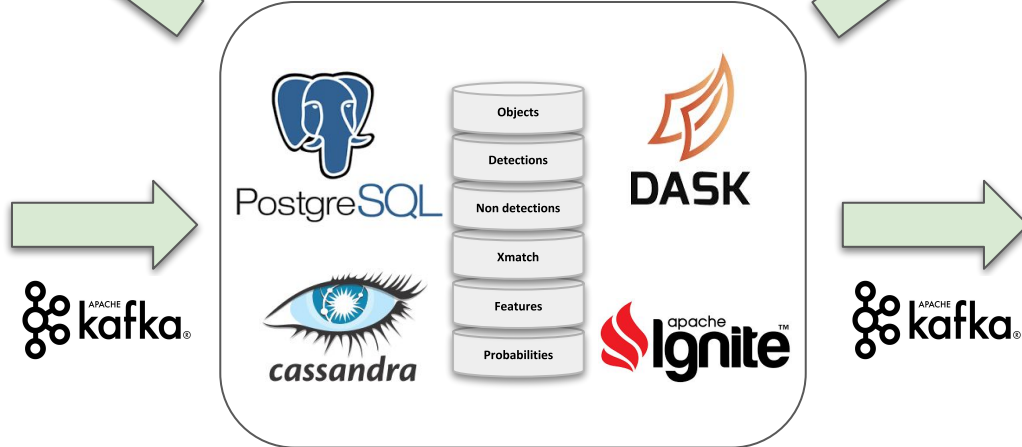
Jupyter notebook (local processing)



Jupyter hub (cloud processing)

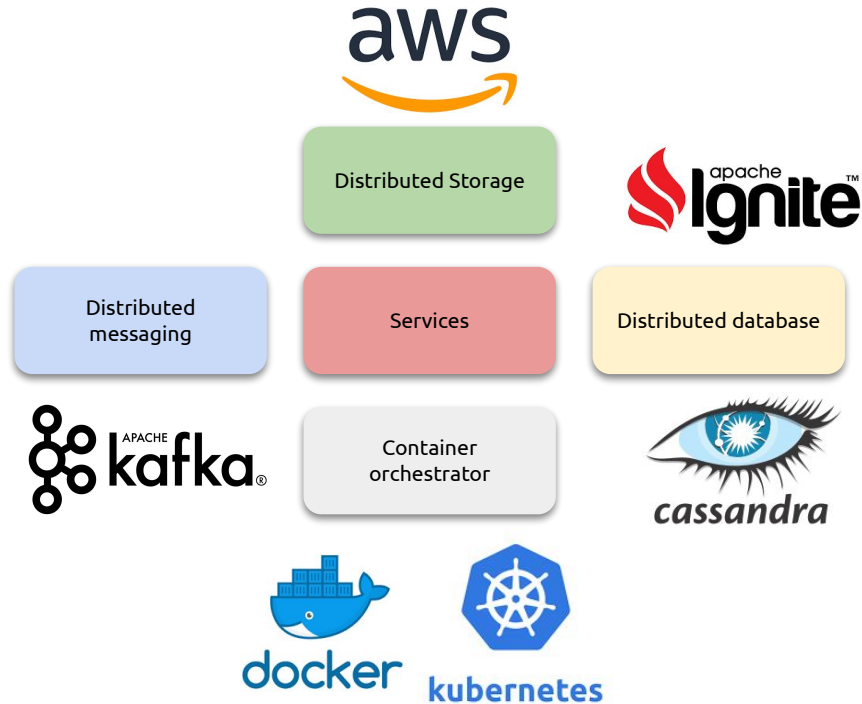


Output stream (real-time follow-up)





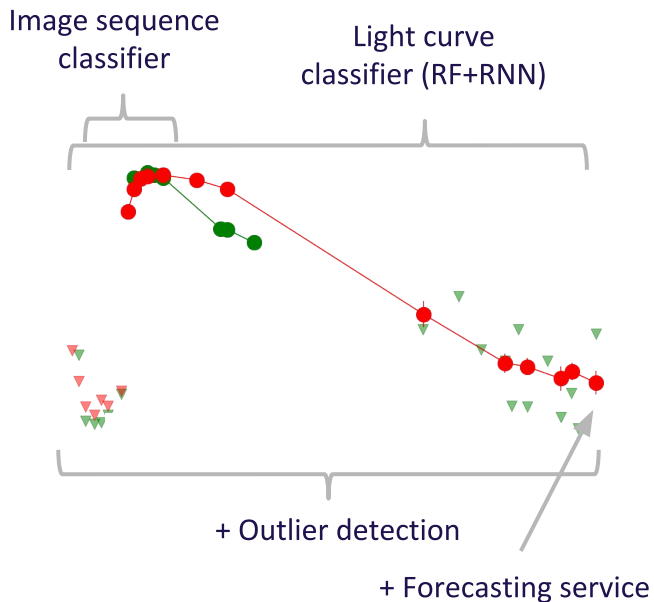
ALeRCE infrastructure





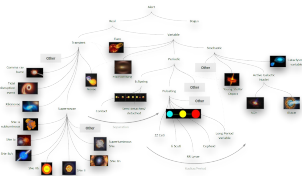
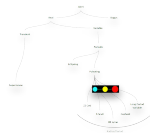
Classification

Classification model



Taxonomy

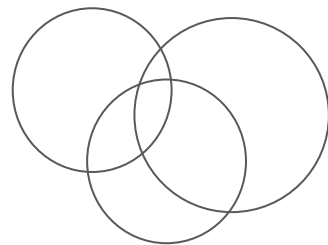
Simplified



Complex & growing

Crossmatch

Label propagation (CRTS, ASAS-SN, LINEAR, TNS, GAIA)

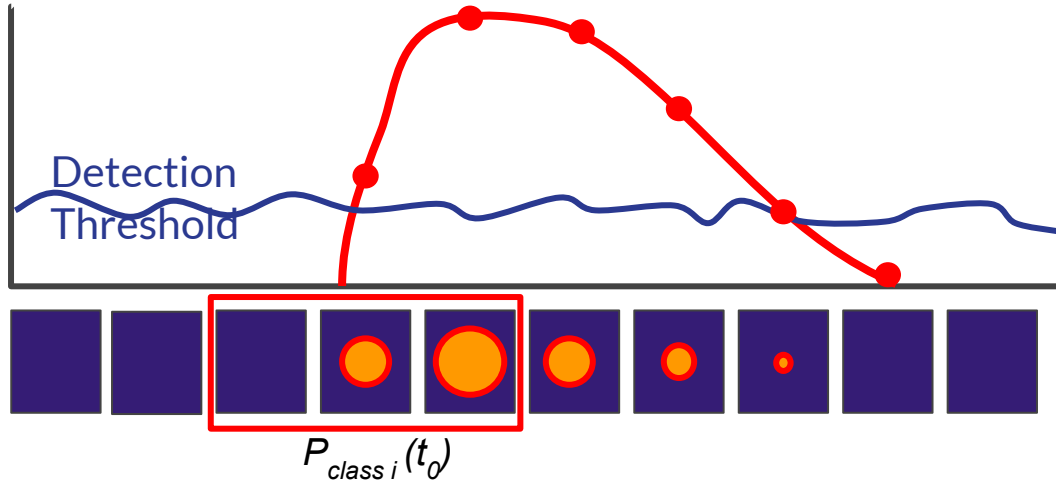


Features input for classifier (GAIA, WISE, eRosita, etc.)

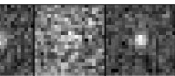
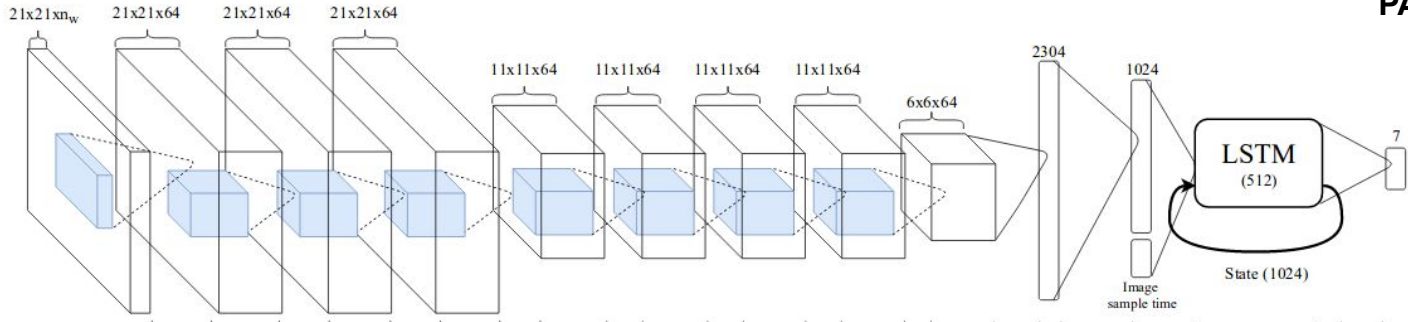




Recurrent convolutional neural network

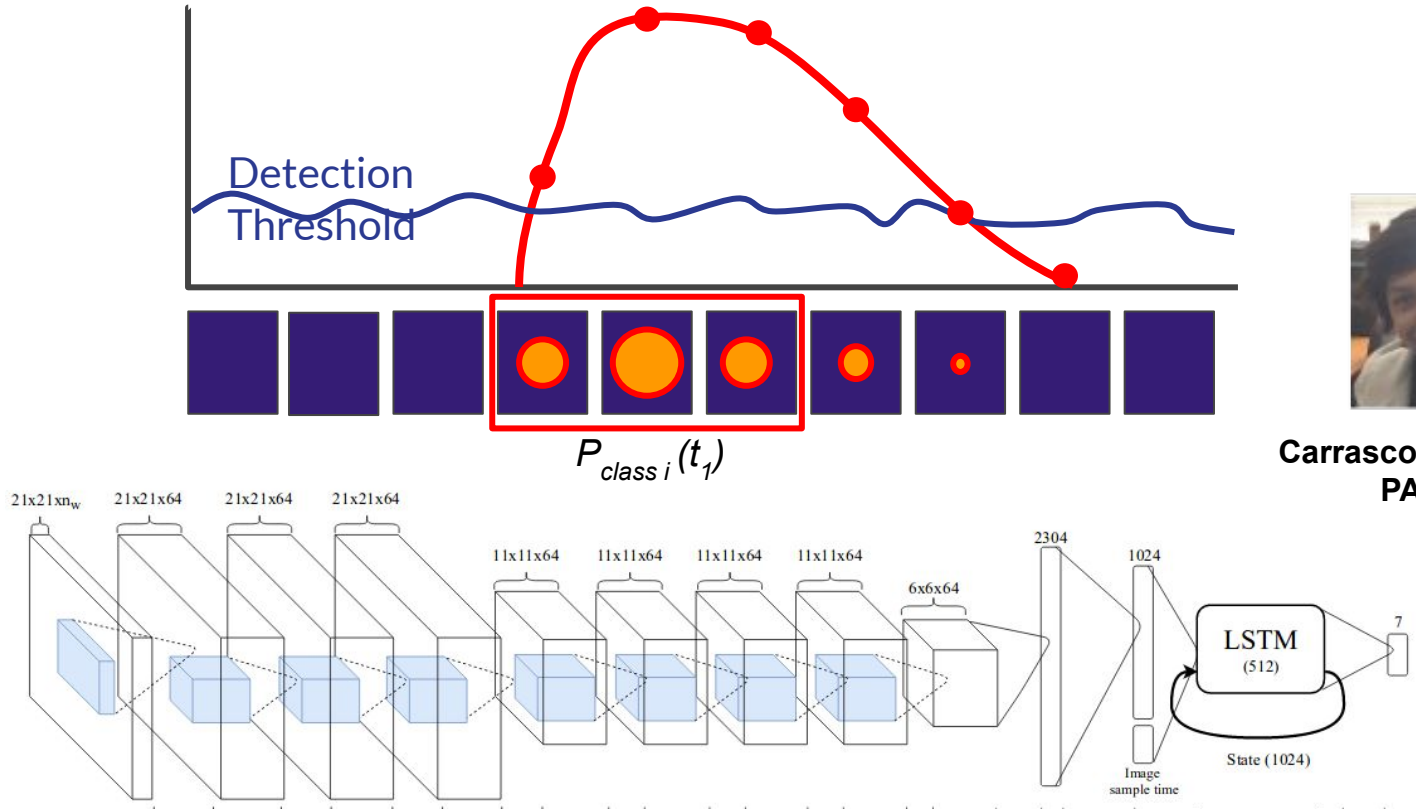


Carrasco-Davis+18
PASP





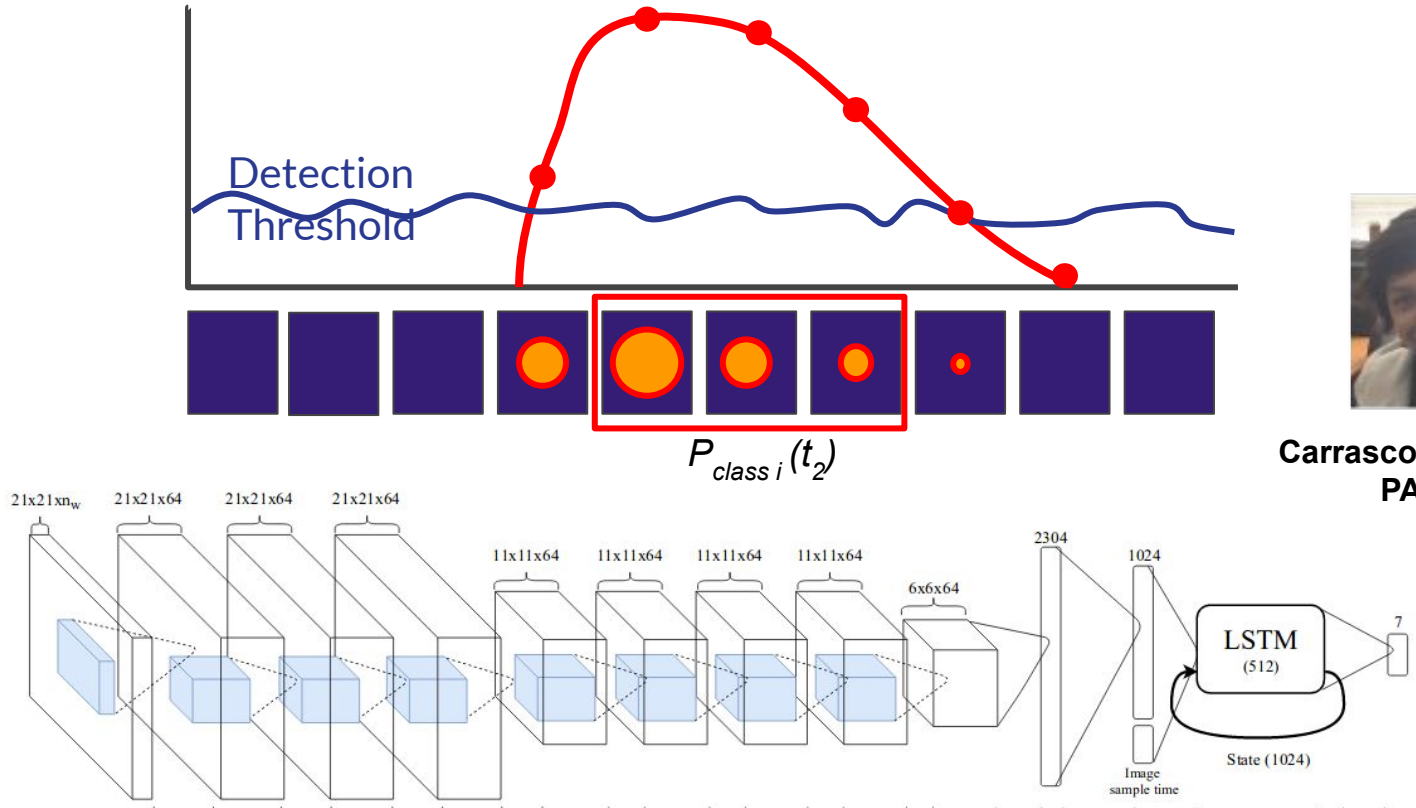
Recurrent convolutional neural network



Carrasco-Davis+18
PASP



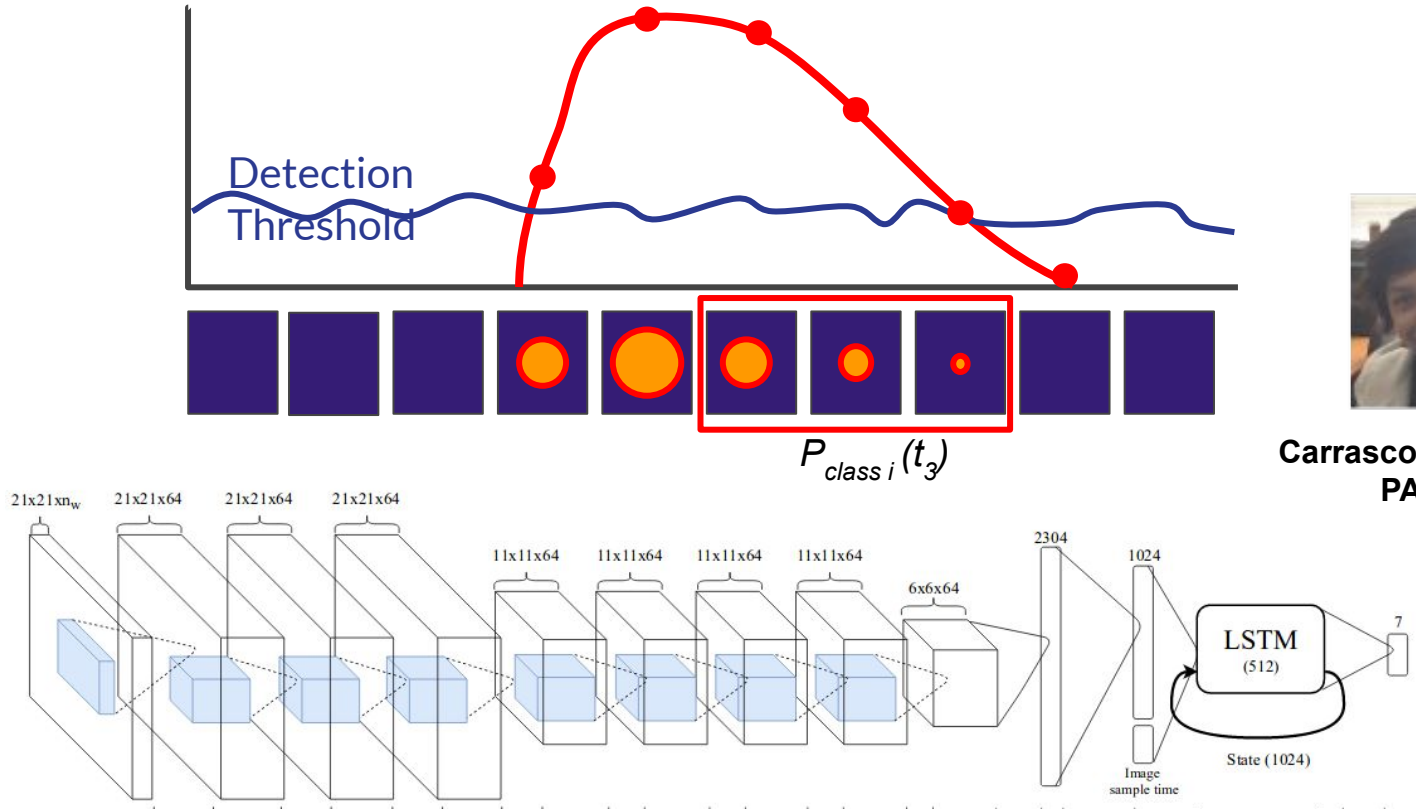
Recurrent convolutional neural network



**Carrasco-Davis+18
PASP**

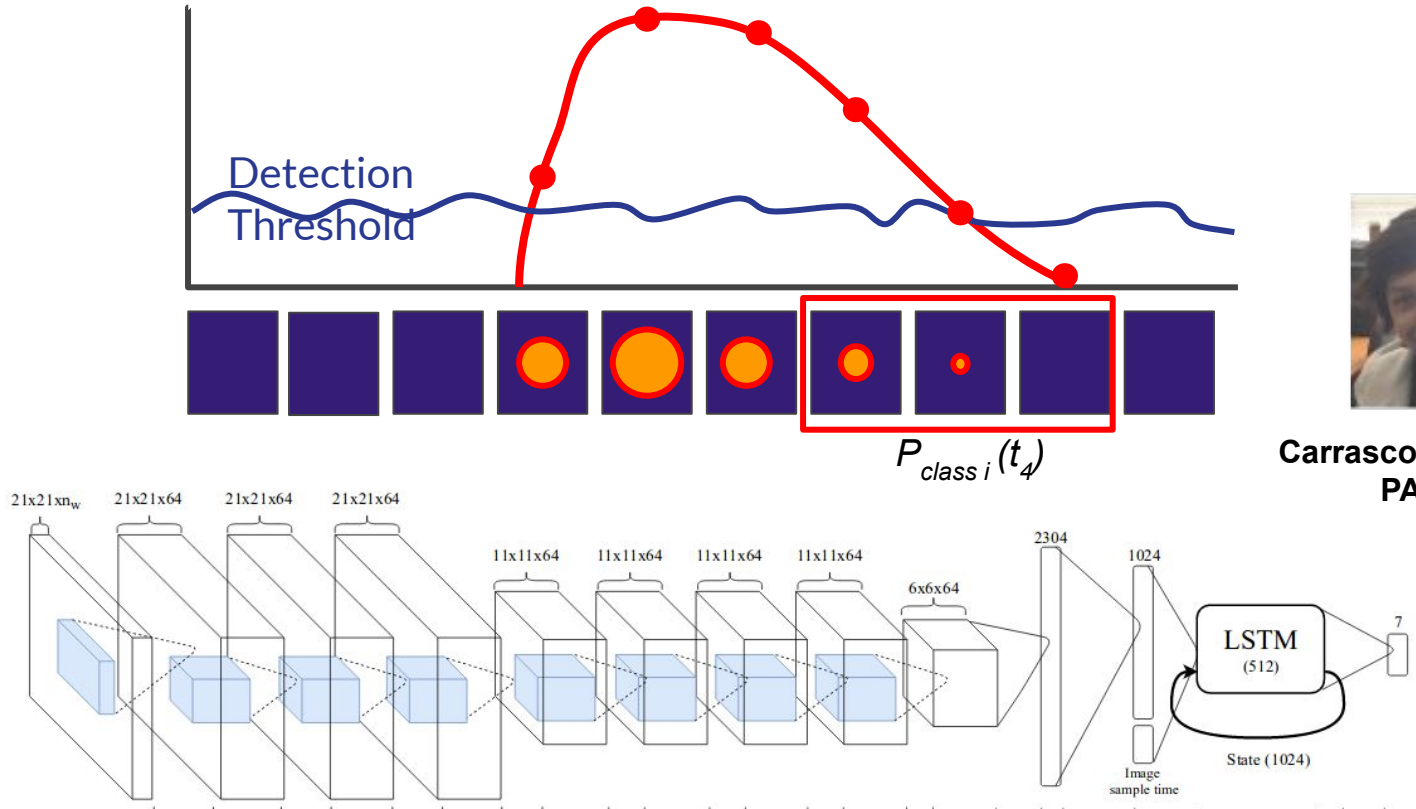


Recurrent convolutional neural network





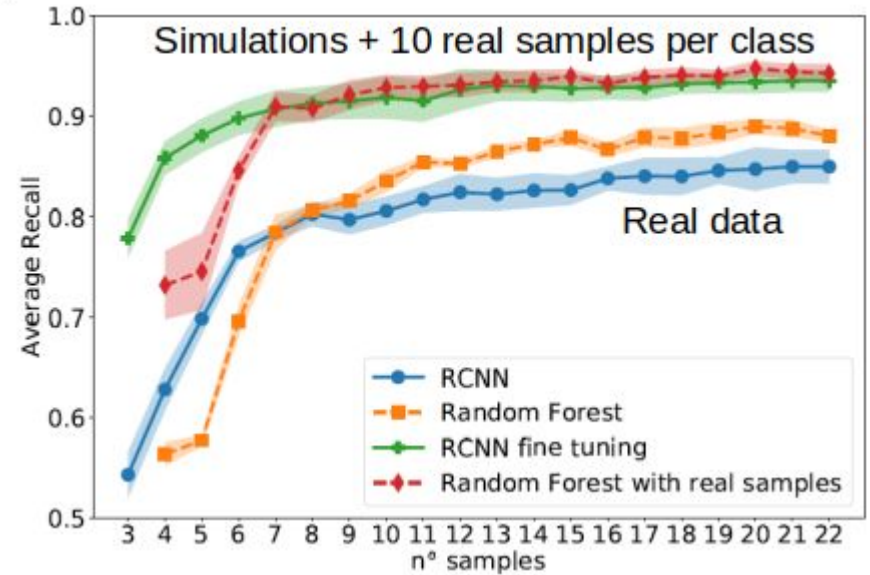
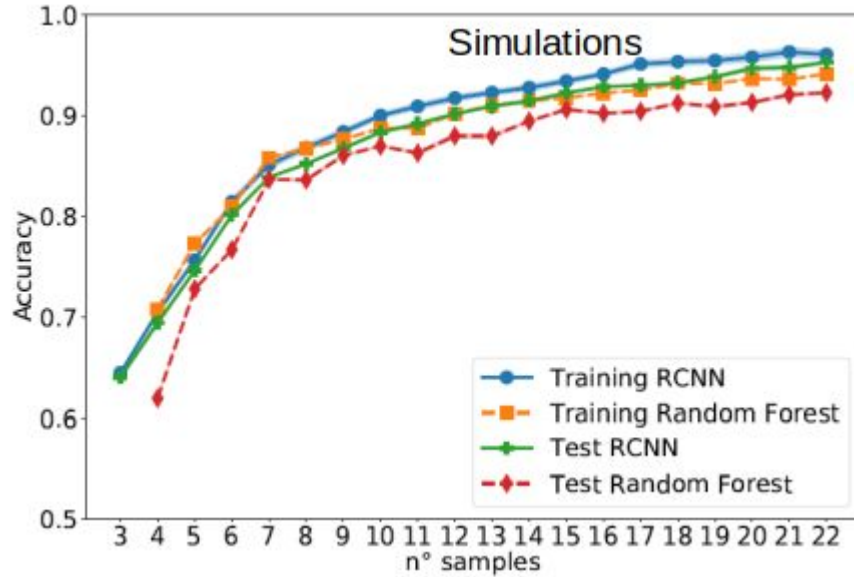
Recurrent convolutional neural network



Carrasco-Davis+18
PASP

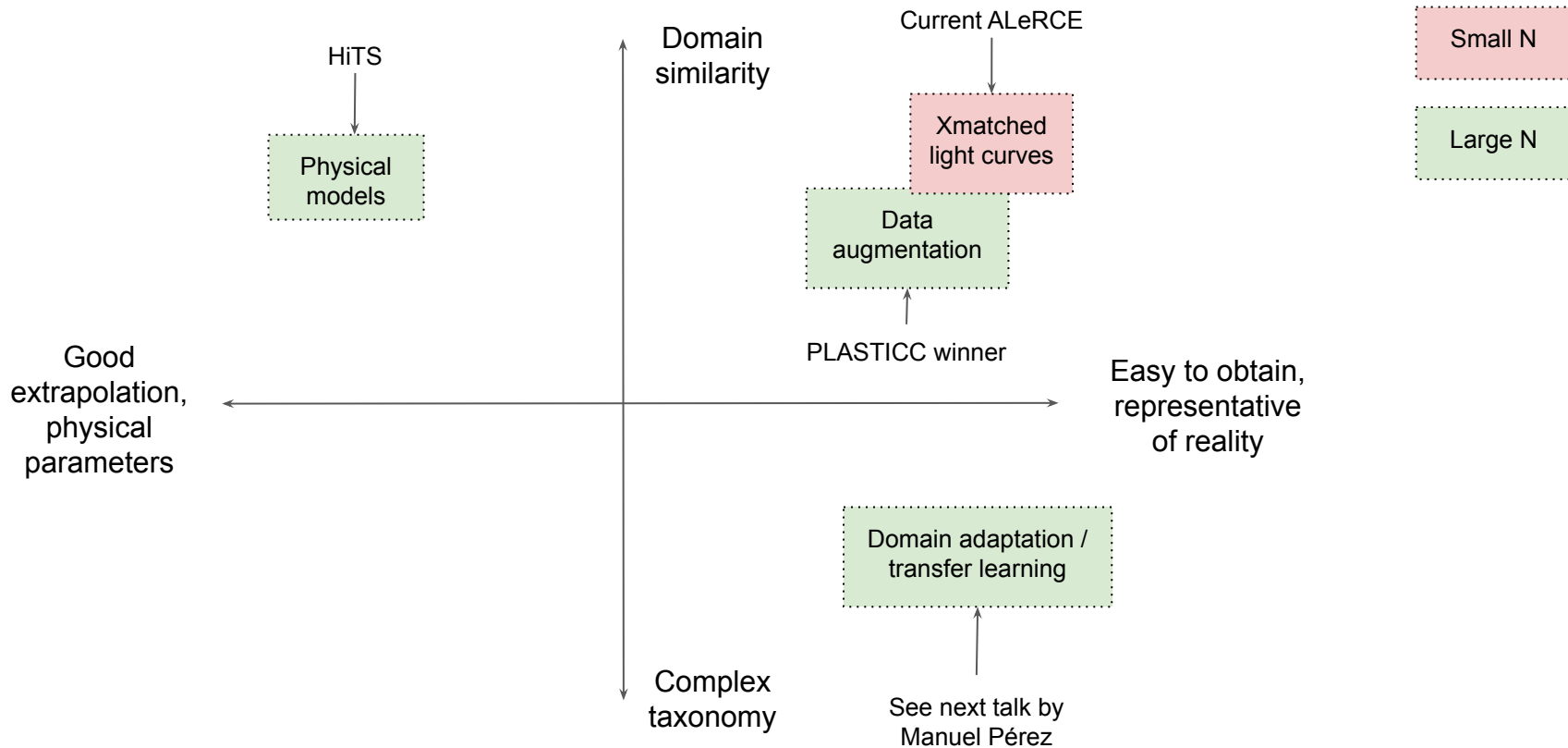


Training vs domain overfitting





Training sets



3. DEMO



Summary

- **ALeRCE**: broker for ZTF, LSST and other large etendue telescopes
- **Interdisciplinary** team born from HiTS survey + young developer team building **distributed** and **scalable** system
- Interaction with the community key to enable effective **follow up**
- Products: living **catalog** of objects, annotated & classified **streams**, **API+frontend**, **jupyter** hub and **batch processing** capabilities
- Learning from **ZTF** to prepare for **LSST**: infrastructure, databases, classification, visualization, transfer learning, forecasting, outlier detection
- Large effort needed to compile **training sets**
- **Exciting future ahead for astrophysics community!**