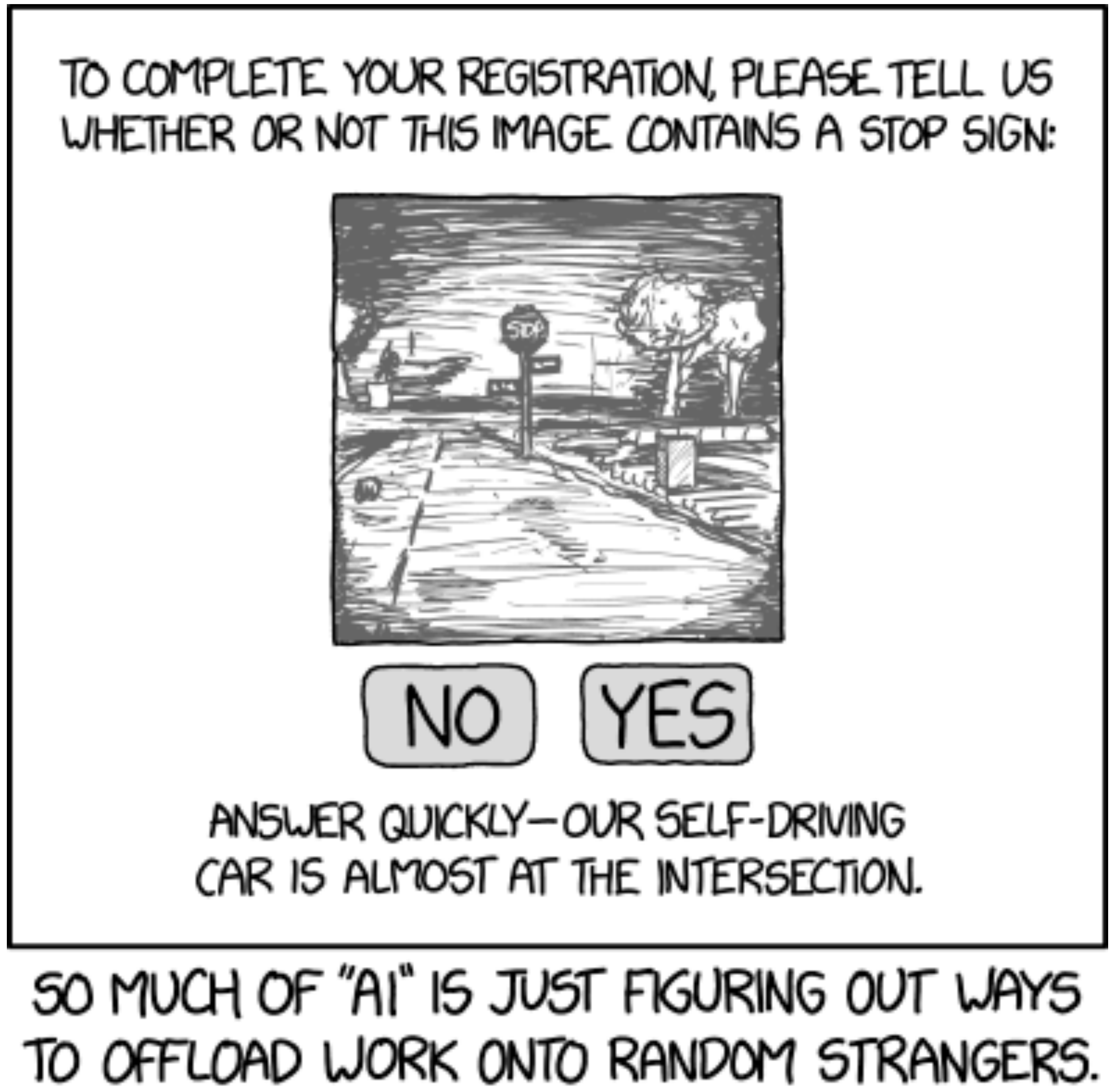
An Overview of Deep Learning in Astronomy



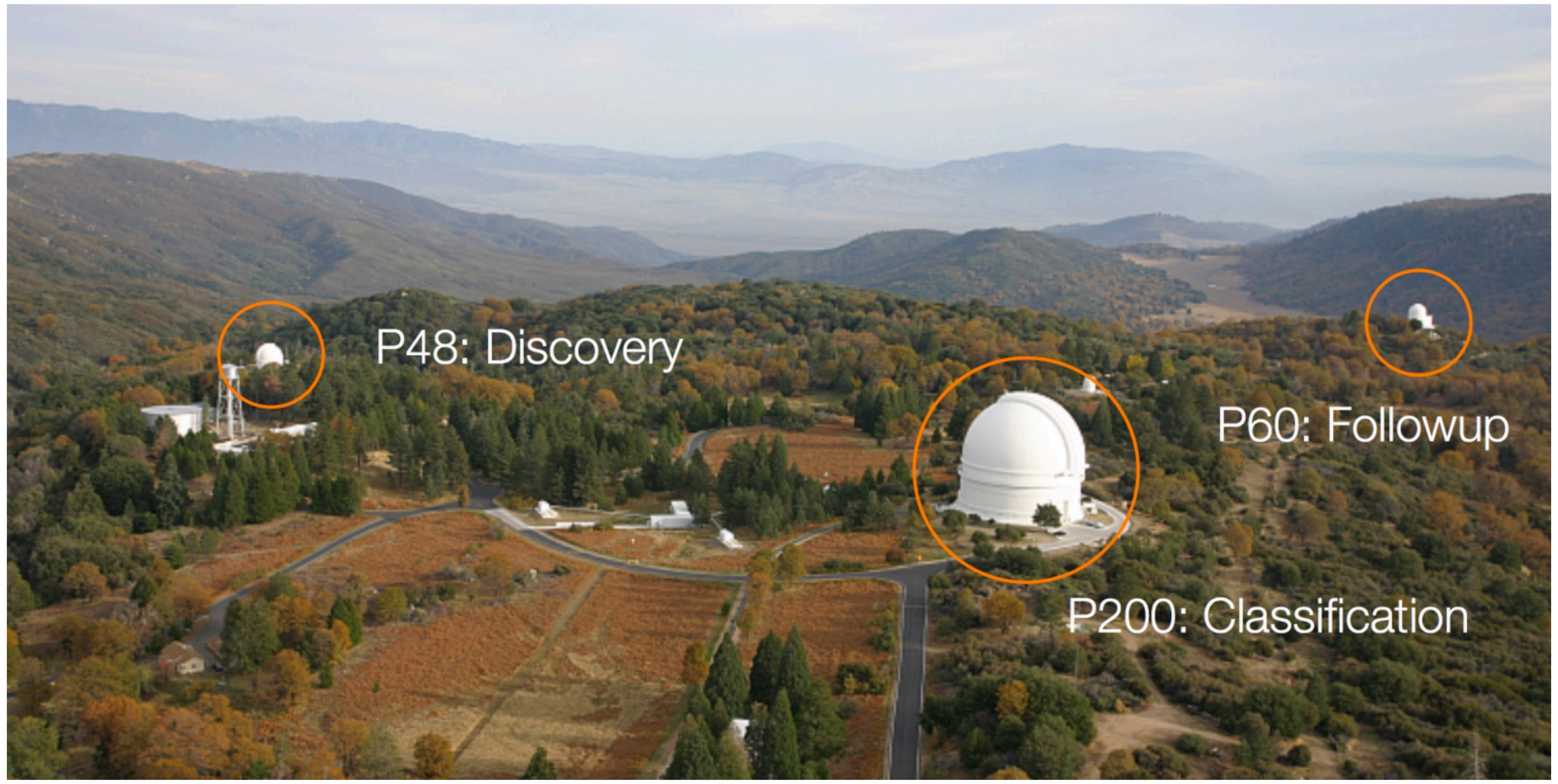
Ashish Mahabal, aam@astro.caltech.edu
Center for Data Driven Discovery, Caltech
ML Lead ZTF
SOMACHINE2020, 26 Nov

Outline

- Intro to ZTF
- ML/DL avenues within ZTF
- Multi Messenger Astronomy

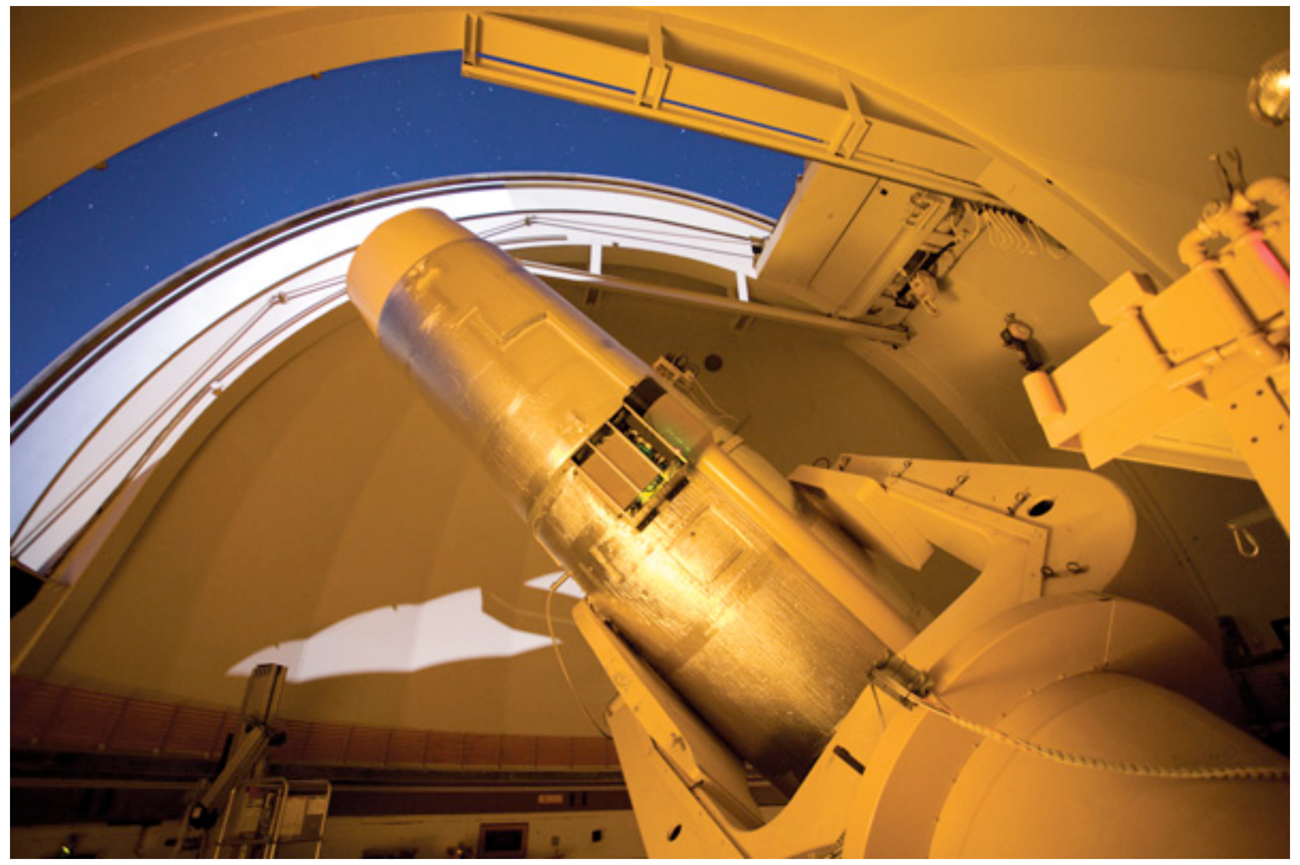
- LSST/Rubin brokers, transfer learning
- Good practices and tricks

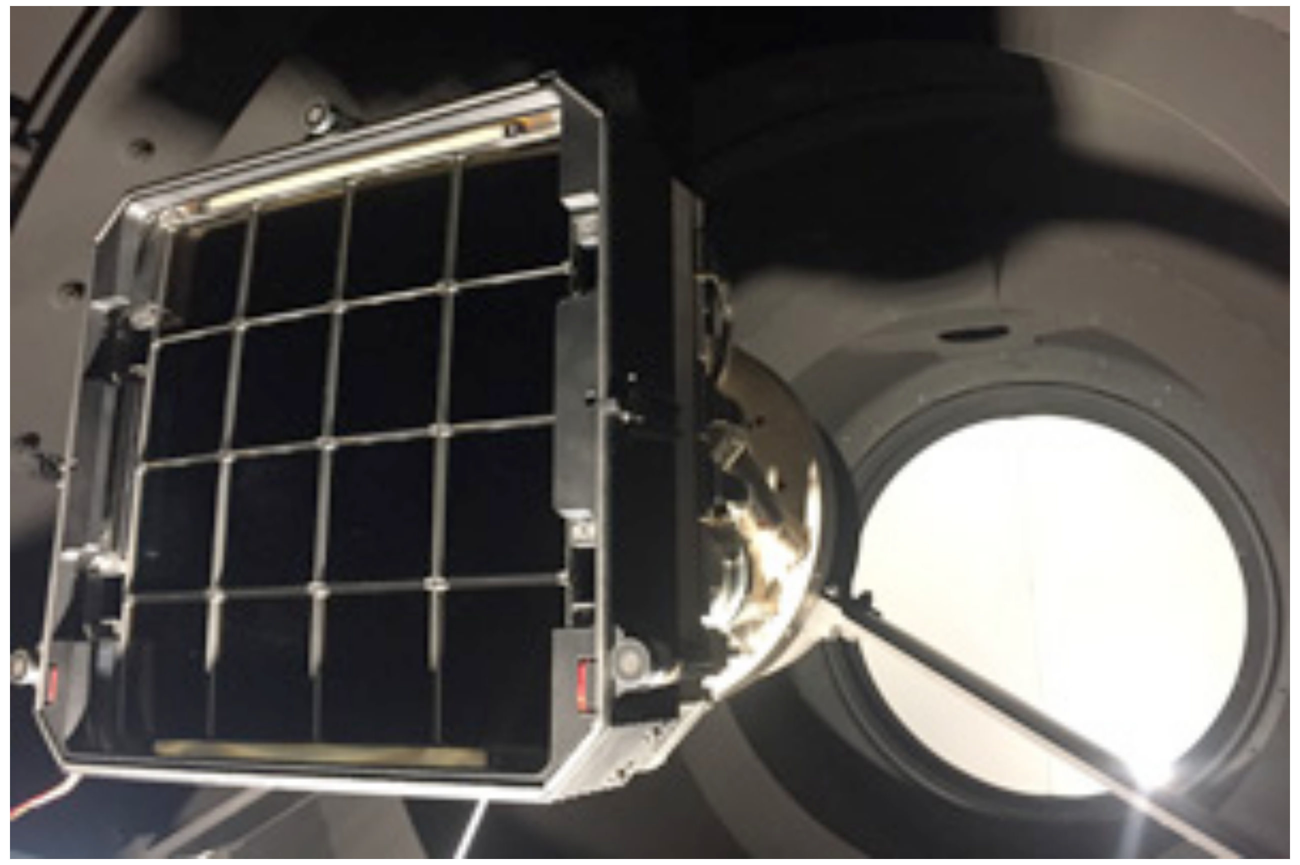
- Exoplanets
- Deep Space Data Fusion



Bryan Penprase

Brief intro to ZTF





47 deg^2 FOV, m_{lim}~20.5 in

30 sec:

- g, r, i filters
- Well-tested subtraction/detection pipeline

~20000 sq deg every 3 nights in two filters (g and r)

ZTF Zwicky Transient Facility

Systematic Exploration of the Dynamic Sky

DR3

https://www.ztf.caltech.edu/page/dr3

Filter(s)	#PSFcat-*sci* sources	#Aperturecat-*sci* sources	#PSFcat-*ref* sources	#Aperturecat-*ref* sources
g	59,554,633,211	37,972,012,960	2,066,428,459	659,475,989
r	171,101,170,546	105,544,929,811	2,981,010,701	1,022,102,891
i	2,759,684,746	1,688,197,984	561,942,811	164,121,447
$g+r+i$	233,415,488,503	145,205,140,755	5,609,381,971	1,845,700,327

Table 3: Number of sources in CCD-quadrant-based catalog files in DR3, according to extraction and image type

ZTF Zwicky Transient Facility
Systematic Exploration of the Dynamic Sky

DR3

https://www.ztf.caltech.edu/page/dr3

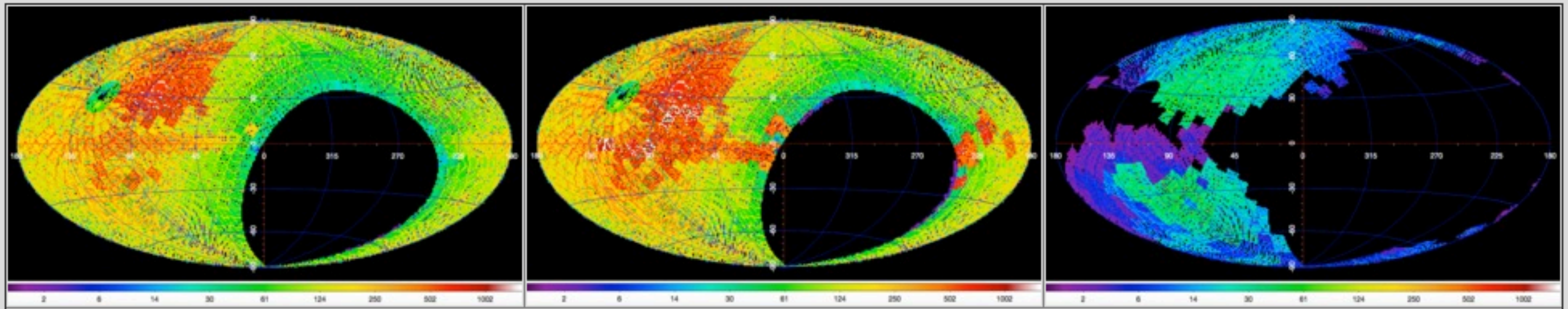
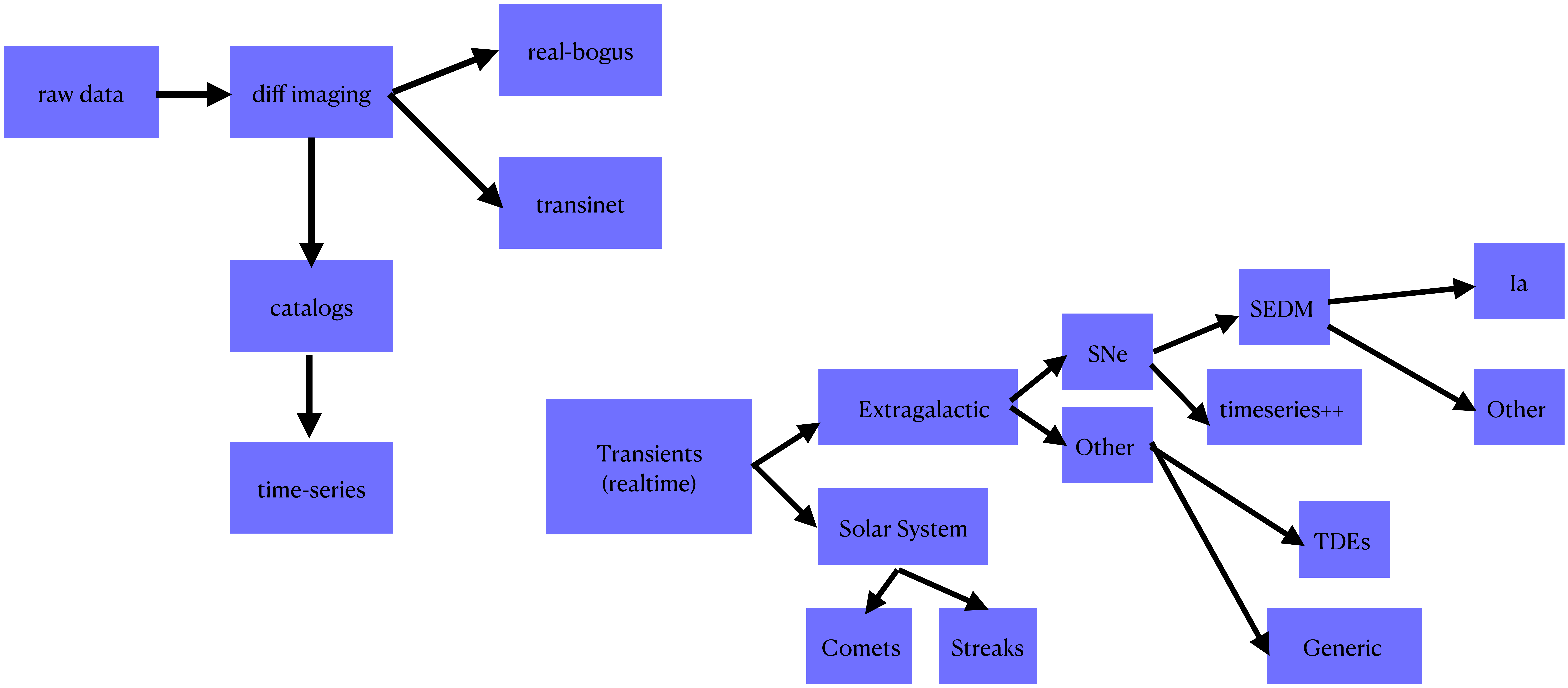


Figure 1 - Sky coverage and number of observation epochs per approximate CCD-quadrant footprint represented in DR3, in Galactic coordinates centered at *l*,*b* = 0,0. **Left:** *g*-filter; **Center:** *r*-filter; **Right:** *i*-filter. Color bar represents the number of observation epochs. Dark regions on small scales are not holes in coverage, but due to the coarse resampling of CCD-quadrant centers on 1° scales. This resampling also distorts the true number of epochs per pointing (see Figure 3a for the true epoch distribution). **Click on a panel to enlarge.**

Example Query using the APIs

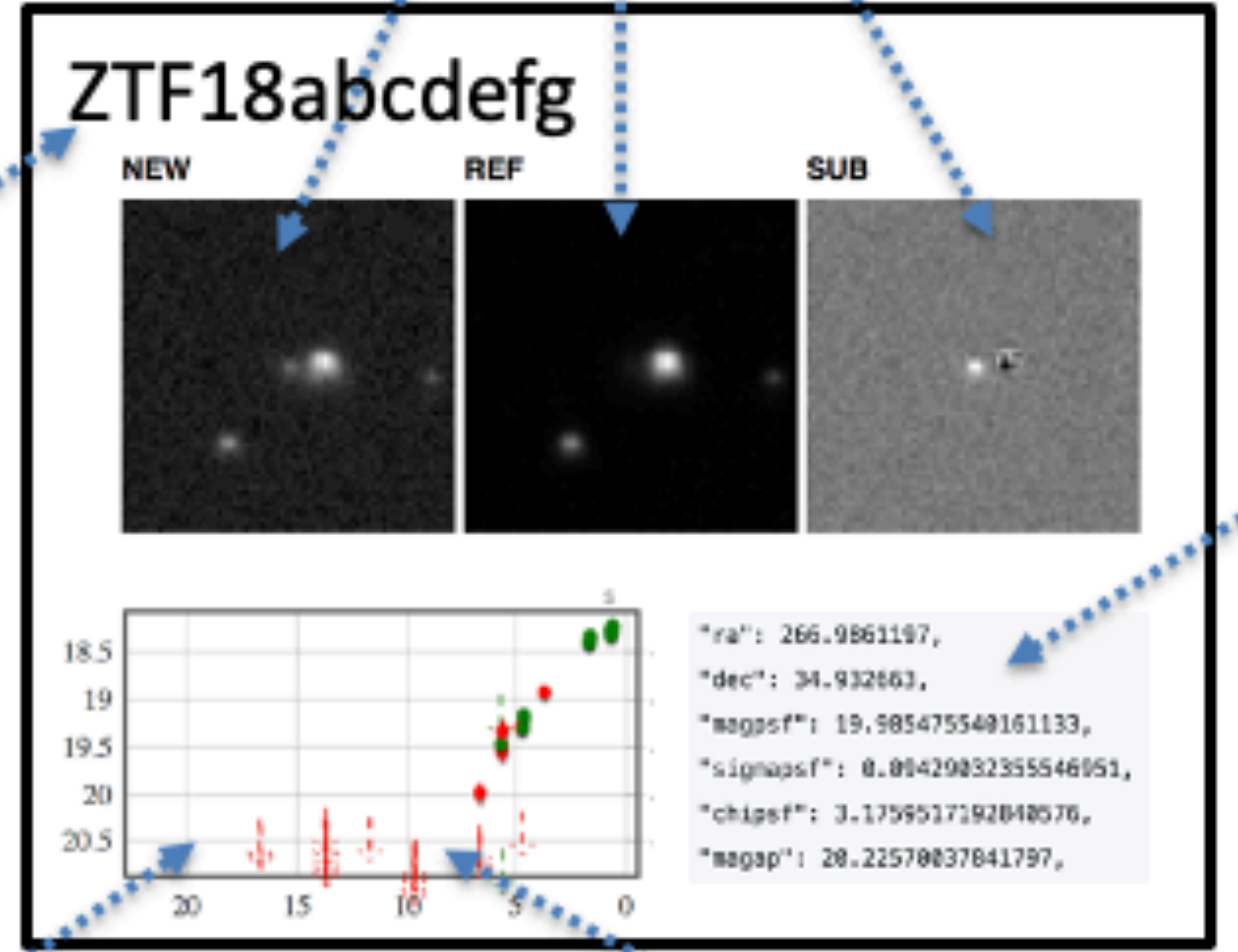
```
wget "https://irsa.ipac.caltech.edu/ibe/search/ztf/products/sci?
POS=255.9302,11.8654&WHERE=obsjd>2458219.9678+AND+obsjd<2458228.8155+
      AND+infobits<33554432" -O out.tbl
```

A few ZTF ML Avenues



ZTF Alert Packet

63 x 63 pixel 32-bit images



Unique spatially matched alert name (1.5” radius)

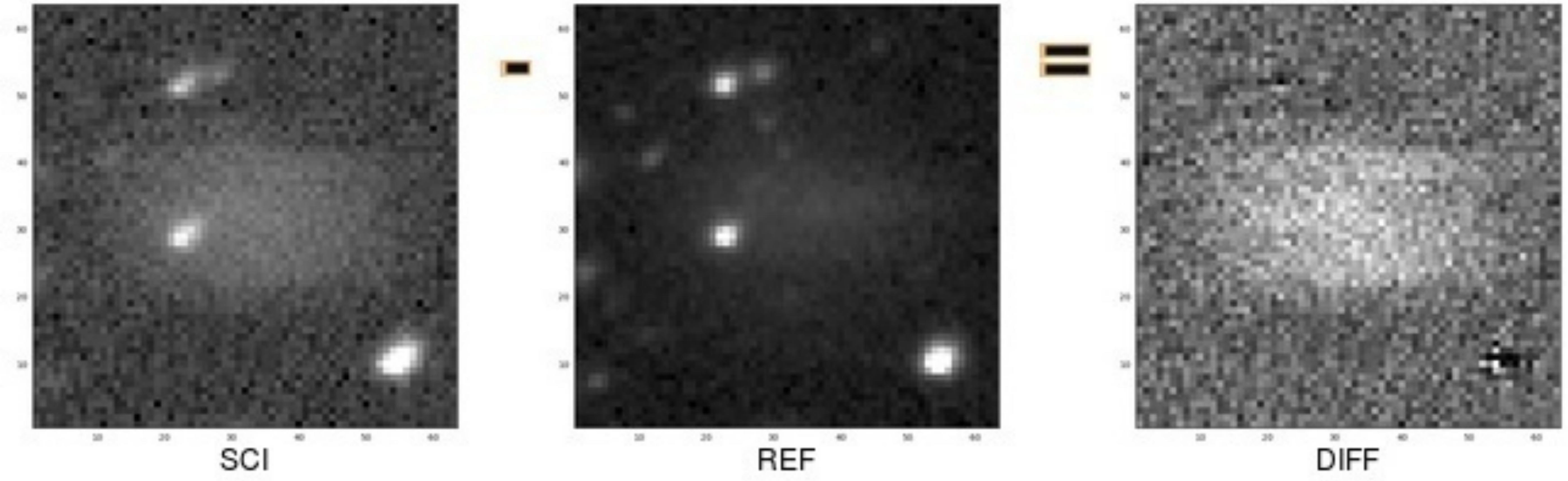
- ZOGY parameters
- Real-bogus score
- Star/galaxy score
- 3 nearest PS1 sources
- Nearest SS object
- Alert history

Rolling 30-day window light curve

Forced photometry instead of upper limits

Smears - bogus

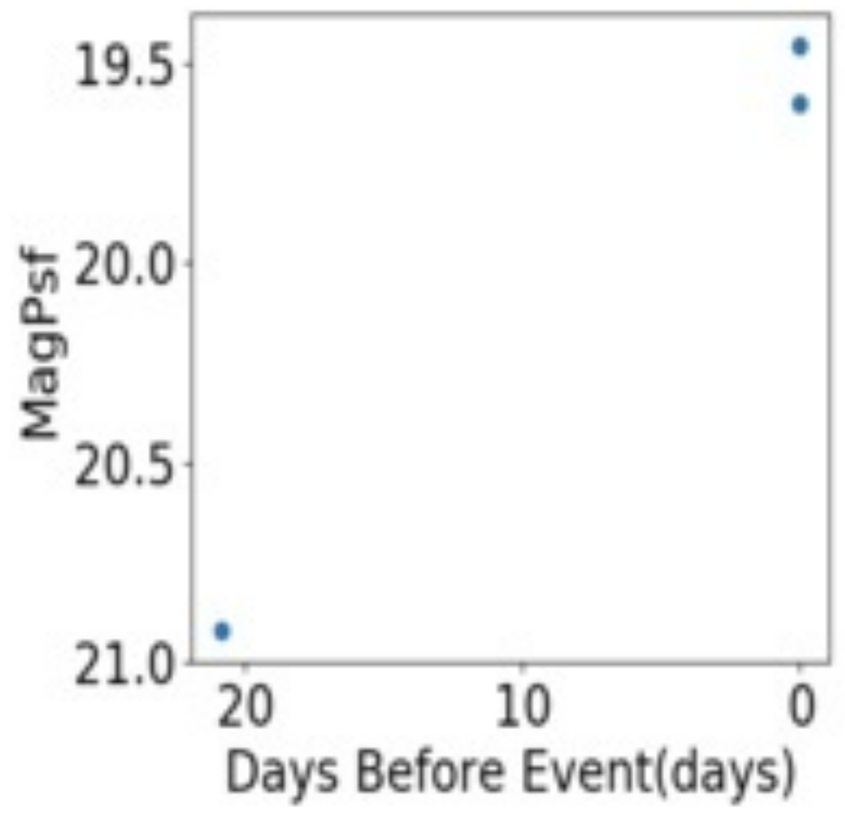
Labeling Objects for training



metadata from ZTF alerts

fwhm: 44.73"
magpsf: 19.46
sigmapsf: 0.07
magdiff: -0.81
elong: 1.42
classtar: 0.0
sgscore: 0.99
isdiffpos: t

Pan-STARRS



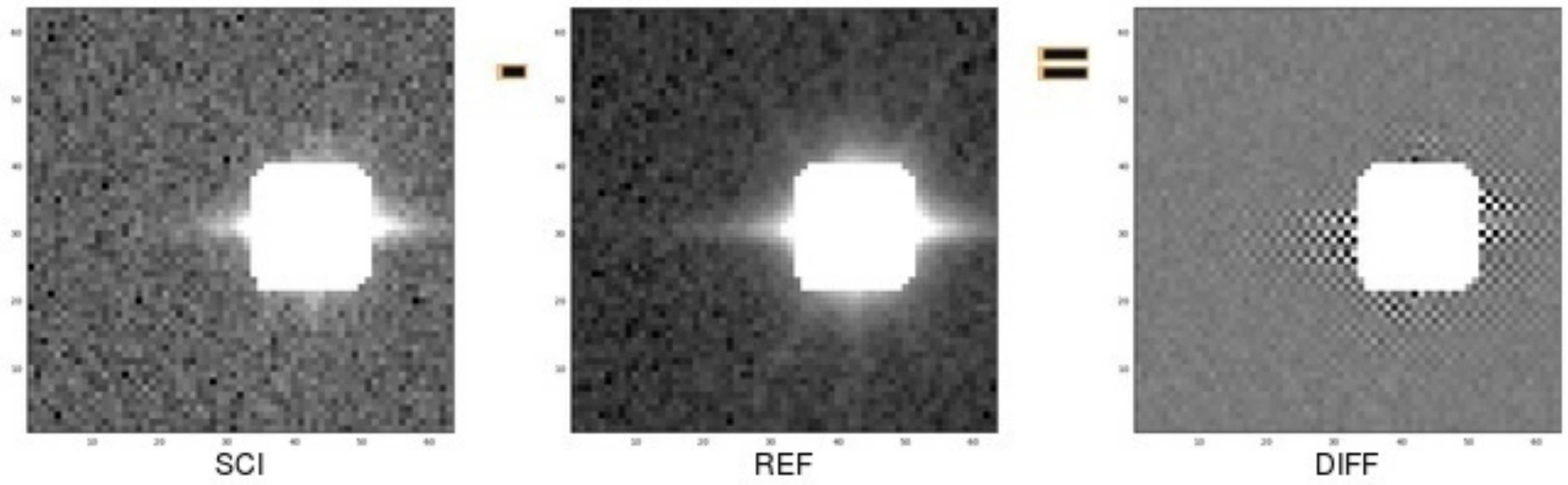
light curve

FWHM
isdiffpos
magpsf
...

archival data

Easy to remove artifact with enough examples

Bogus - streak from masked star



fwhm: 5.37"

magpsf: 20.03

sigmapsf: 0.12

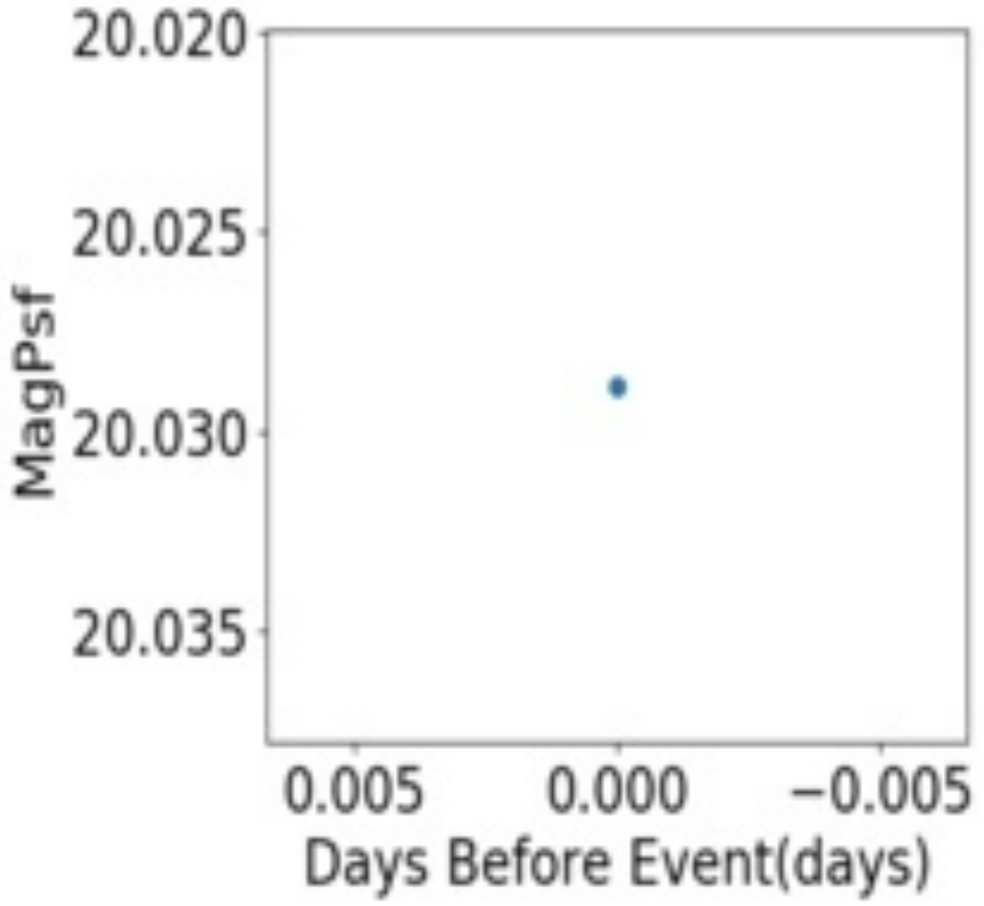
magdiff: -0.42

elong: 1.33

classtar: 0.99

sgscore: 0.5

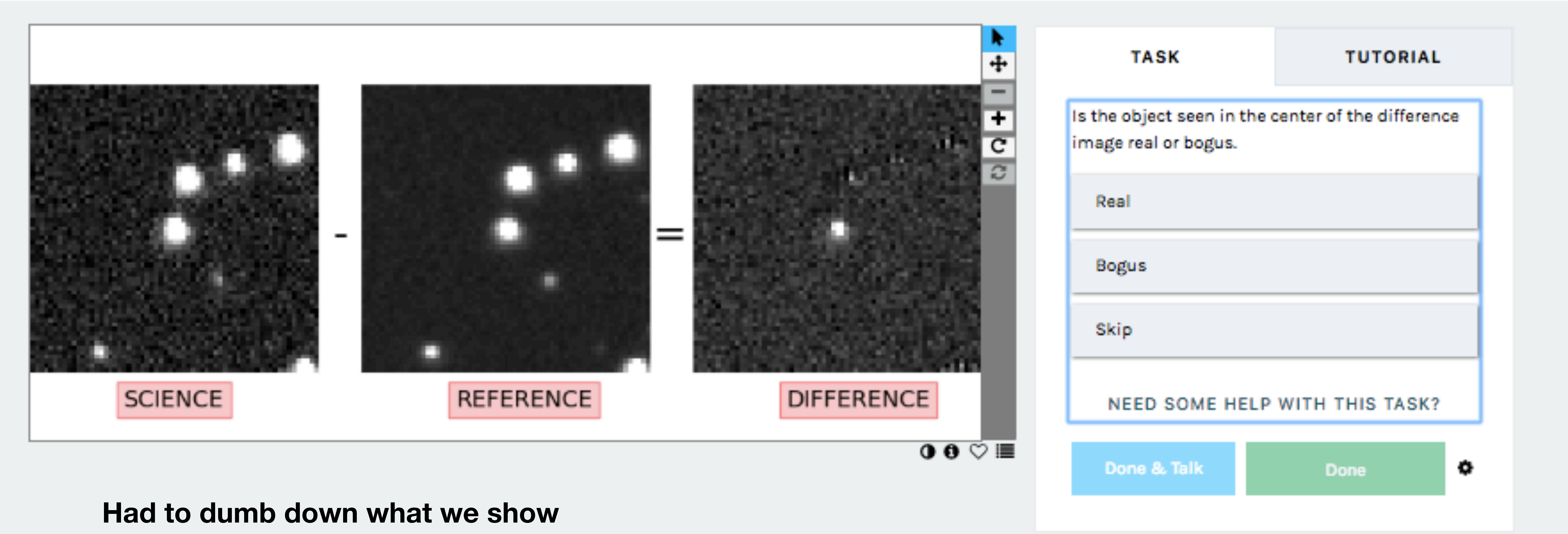
isdiffpos: t



Easy to remove artifact with enough examples

Zooniverse campaigns

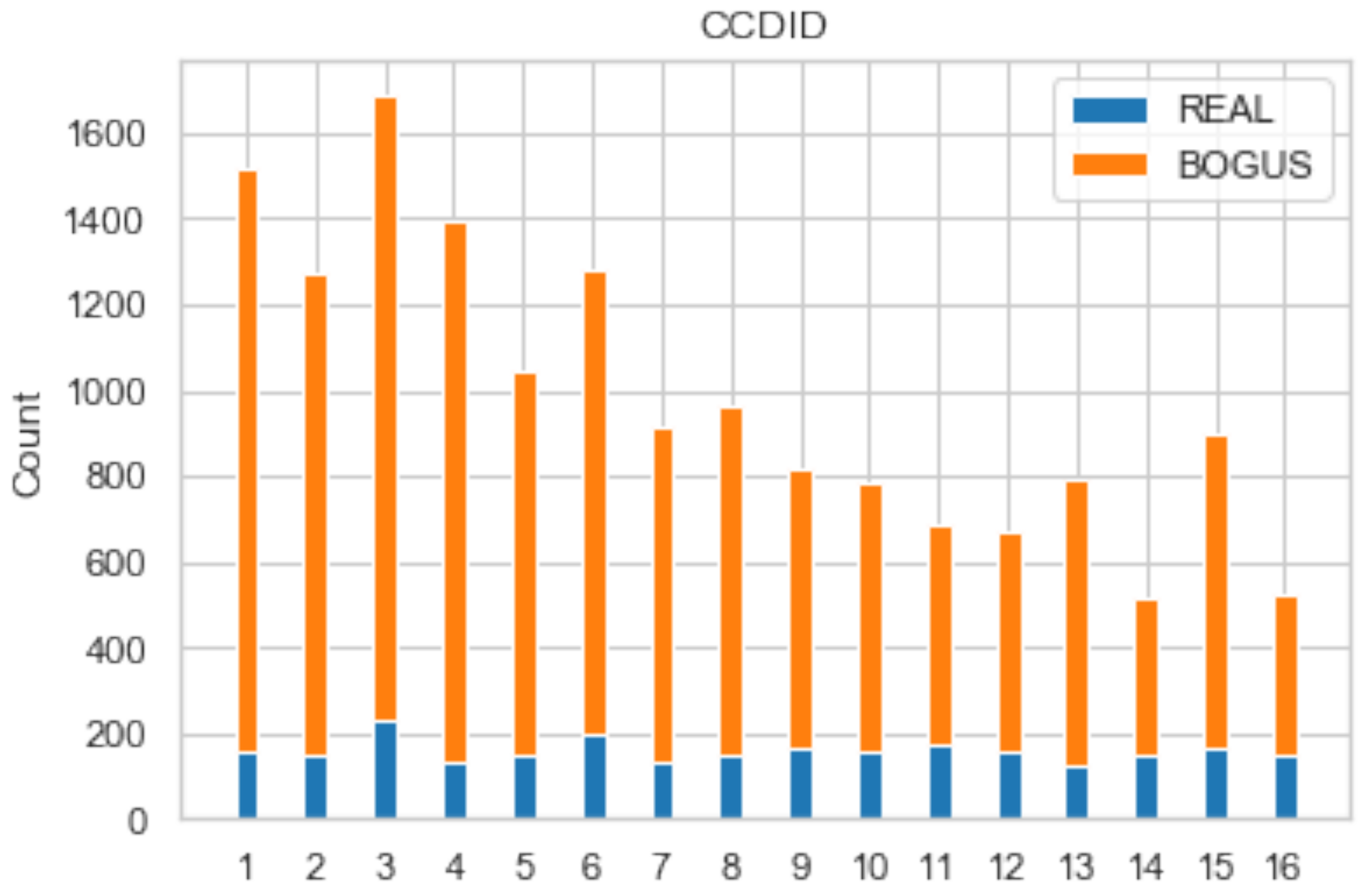
Richard Walters++



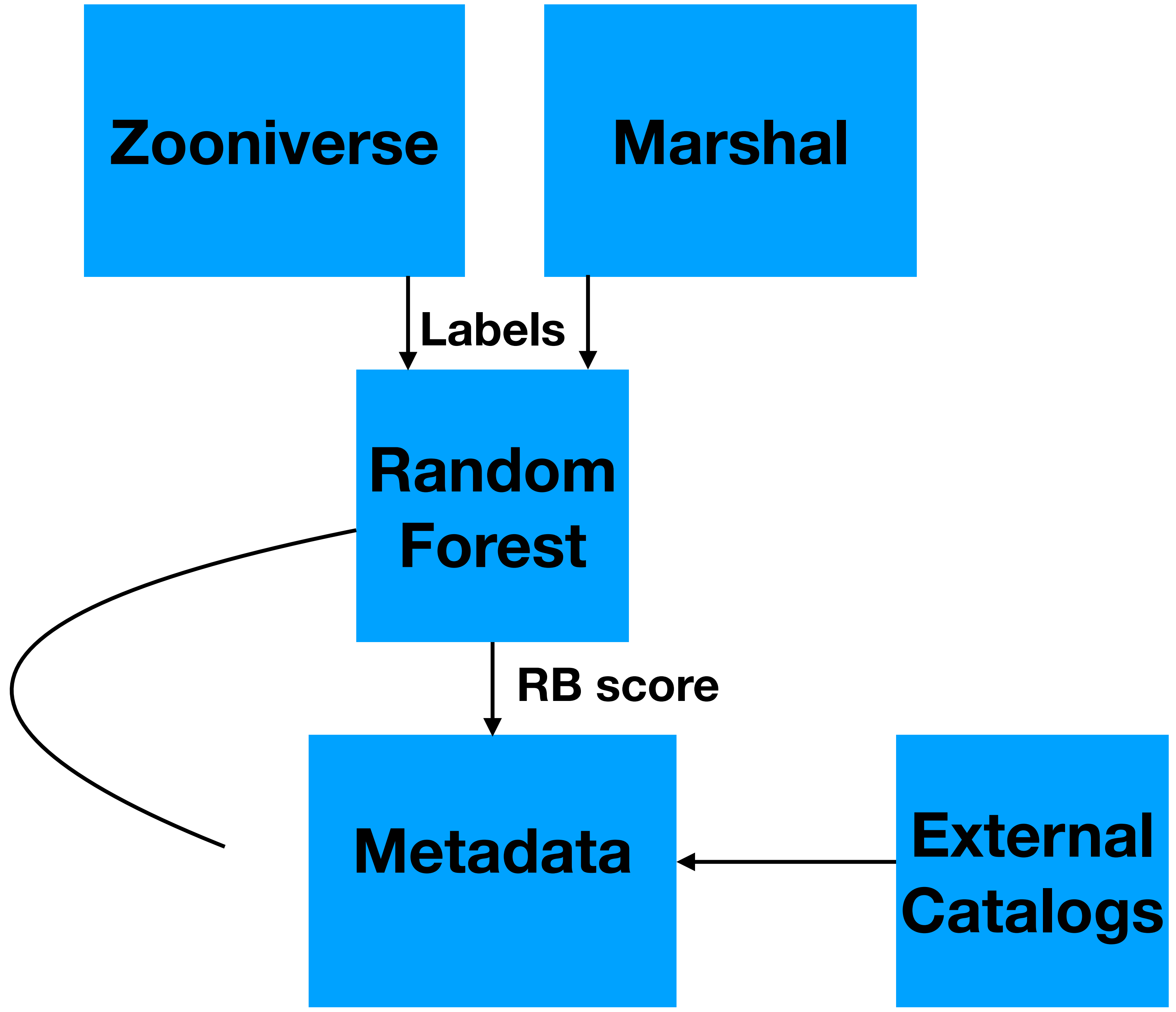
Had to dumb down what we show

b > .4 (No upper limit cutoff)
programid = 1 = MSIP
ssdnr > 8
isdiffpos = true
n > 1

Sources of labels



By CCD



PS1
Gaia

Zwickyverse Home Projects Docs

Low Galactic rb edit remove

id: 5c58cd4c67b6e500be33b6f3

Training data-sets for low Galactic real/bogus classification in ZTF

Project metadata

Classes

Users

Datasets

20181115_0.2<rb<0.4 classify inspect edit remove

id: 5c5aa5b367b6e500be33b6f4

299 random alerts from 20181115 with 0.2<rb<0.4

Number of objects: 299
Number of classifications: 0
Number of classifications by all users: 297
Dataset
Classifications
Classifications from all users

20181126_0.2<rb<0.4 classify inspect edit remove

id: 5c5acbb667b6e500c117d63a

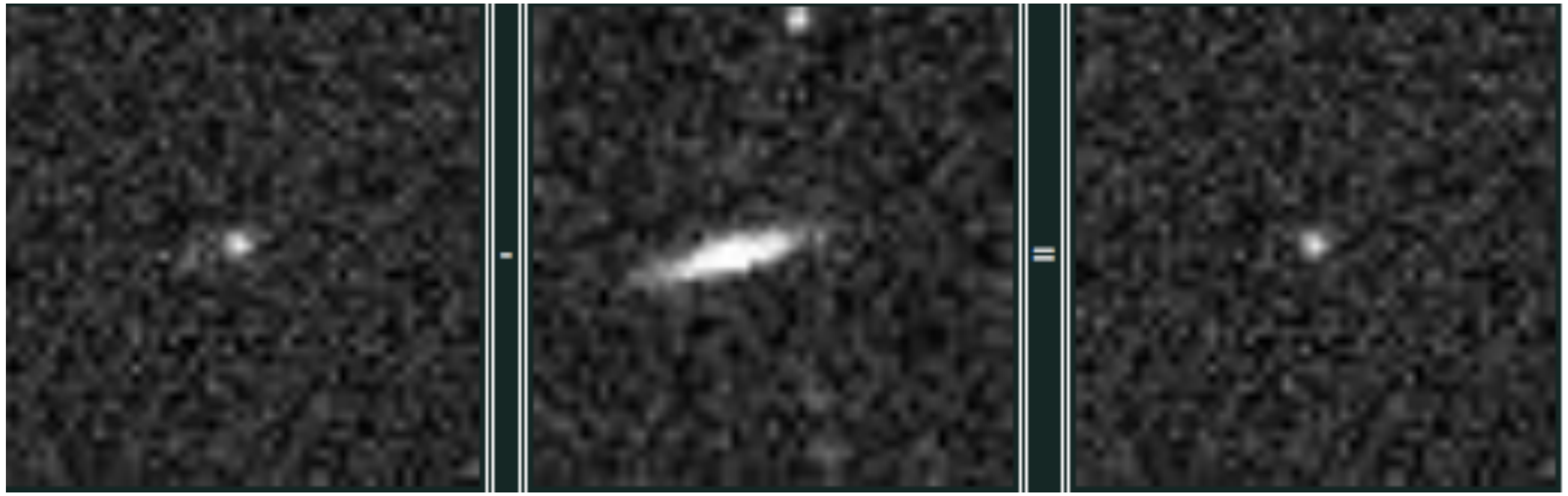
304 random alerts from 20181126 with 0.2<rb<0.4

Number of objects: 304
Number of classifications: 0
Number of classifications by all users: 304

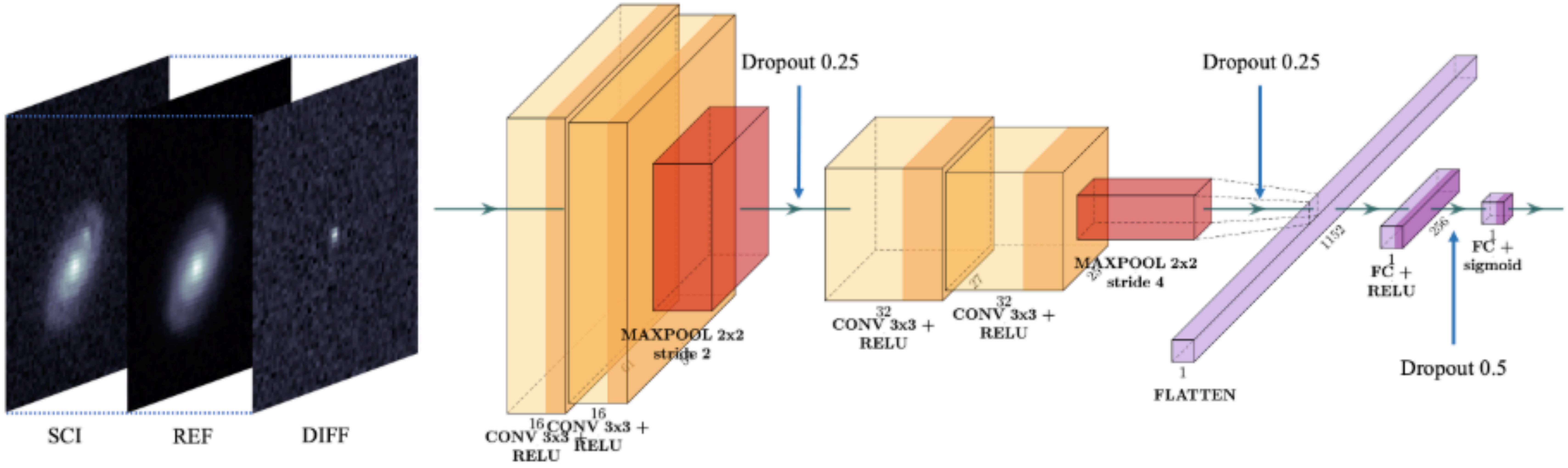
private.caltech.edu

Dmitry Duev

Real-Bogus separation ('braai')

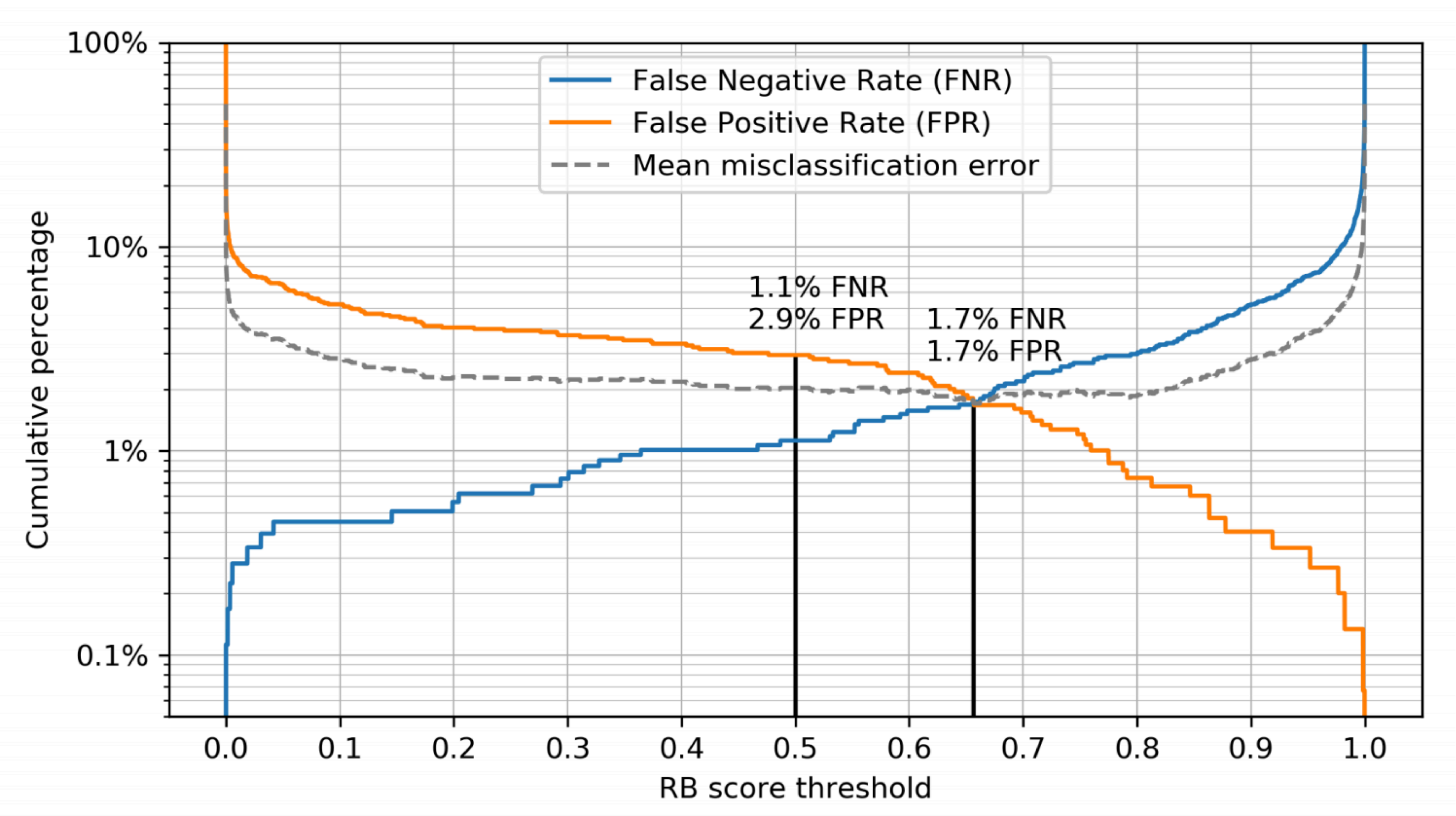


Sci Ref Diff



https://github.com/dmitryduev/braai

Duev, Mahabal, … arXiv 1907.11259



$$A = \frac{TP + TN}{TP + FP + TN + FN}$$

Total predictions

How often is the model correct?

$$P = \frac{TP}{TP + FP}$$

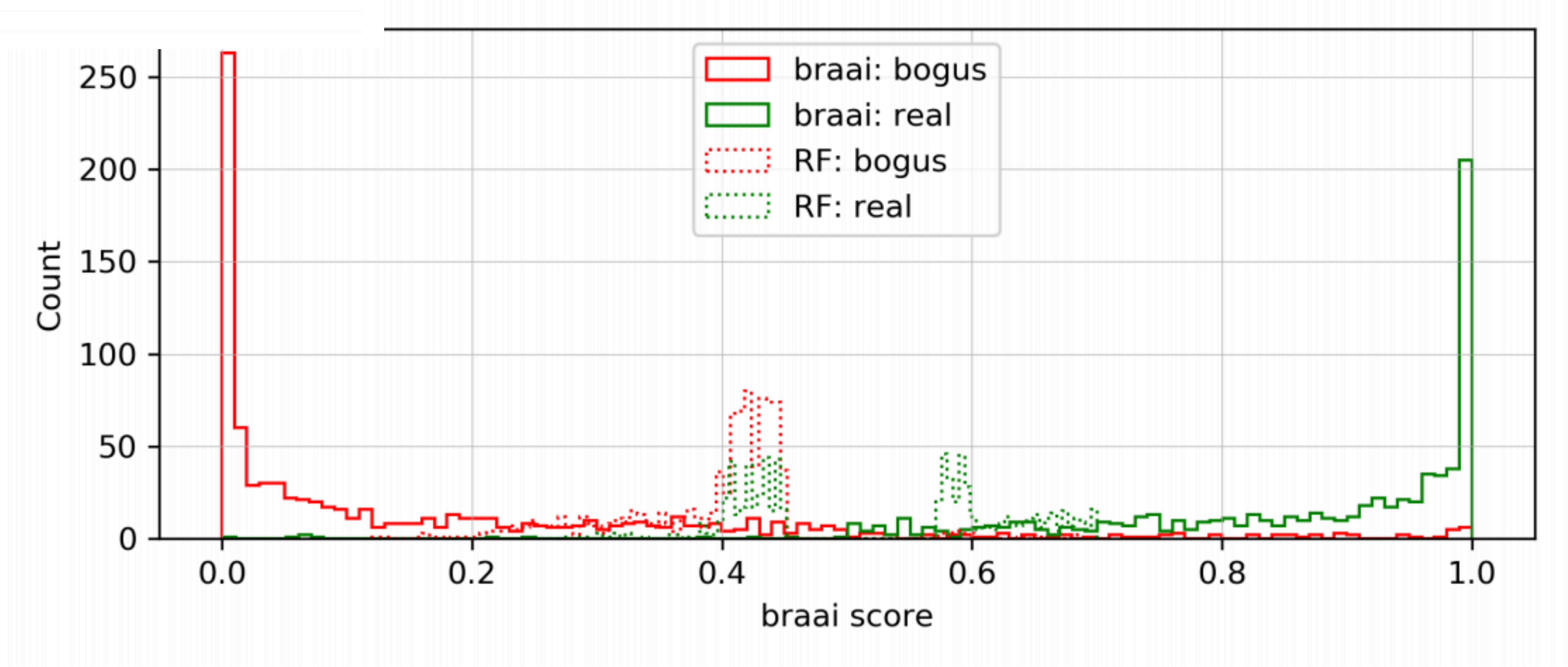
Predicted positives

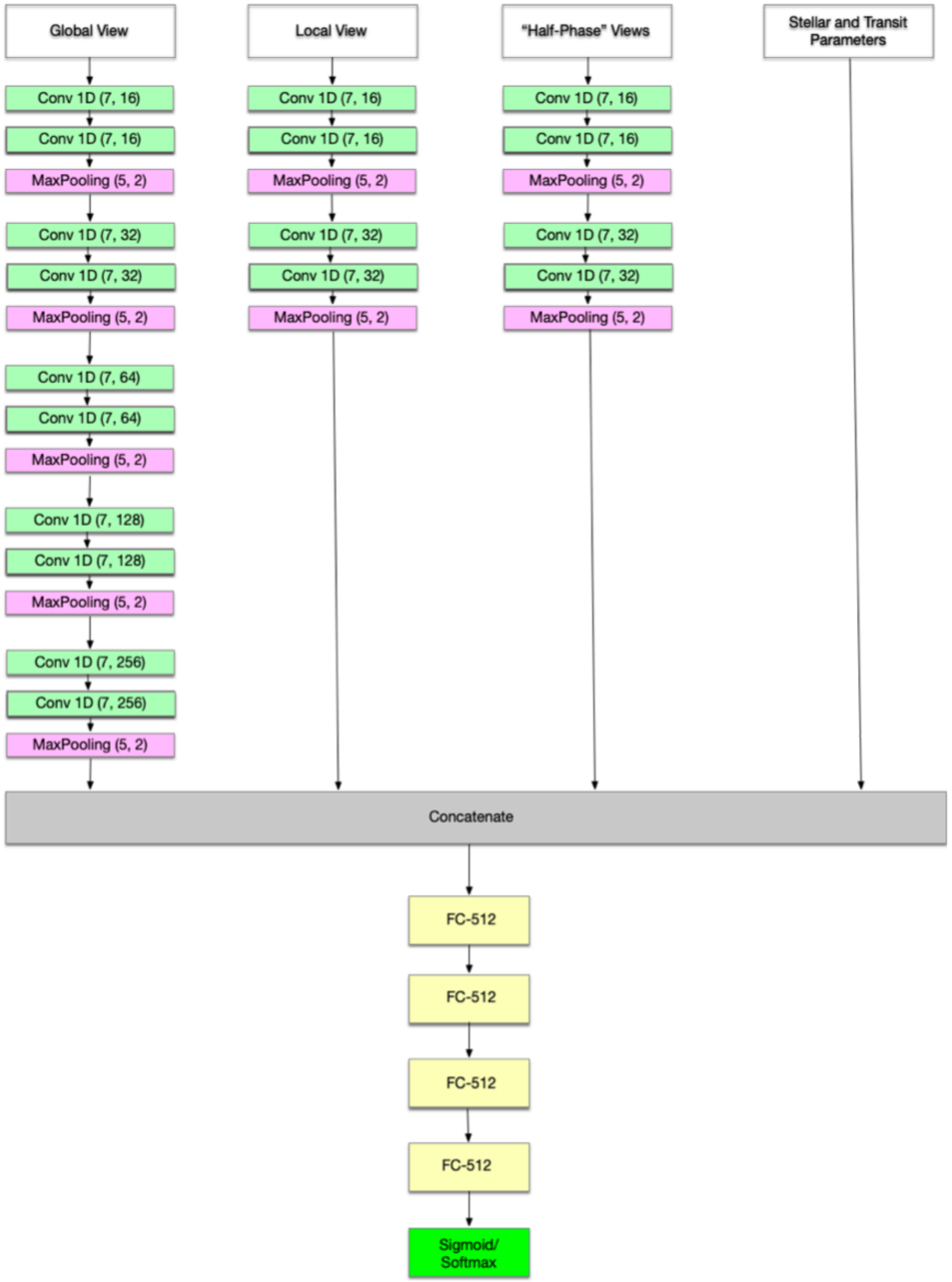
How often is the model correct when it predicts that the candidate is real?

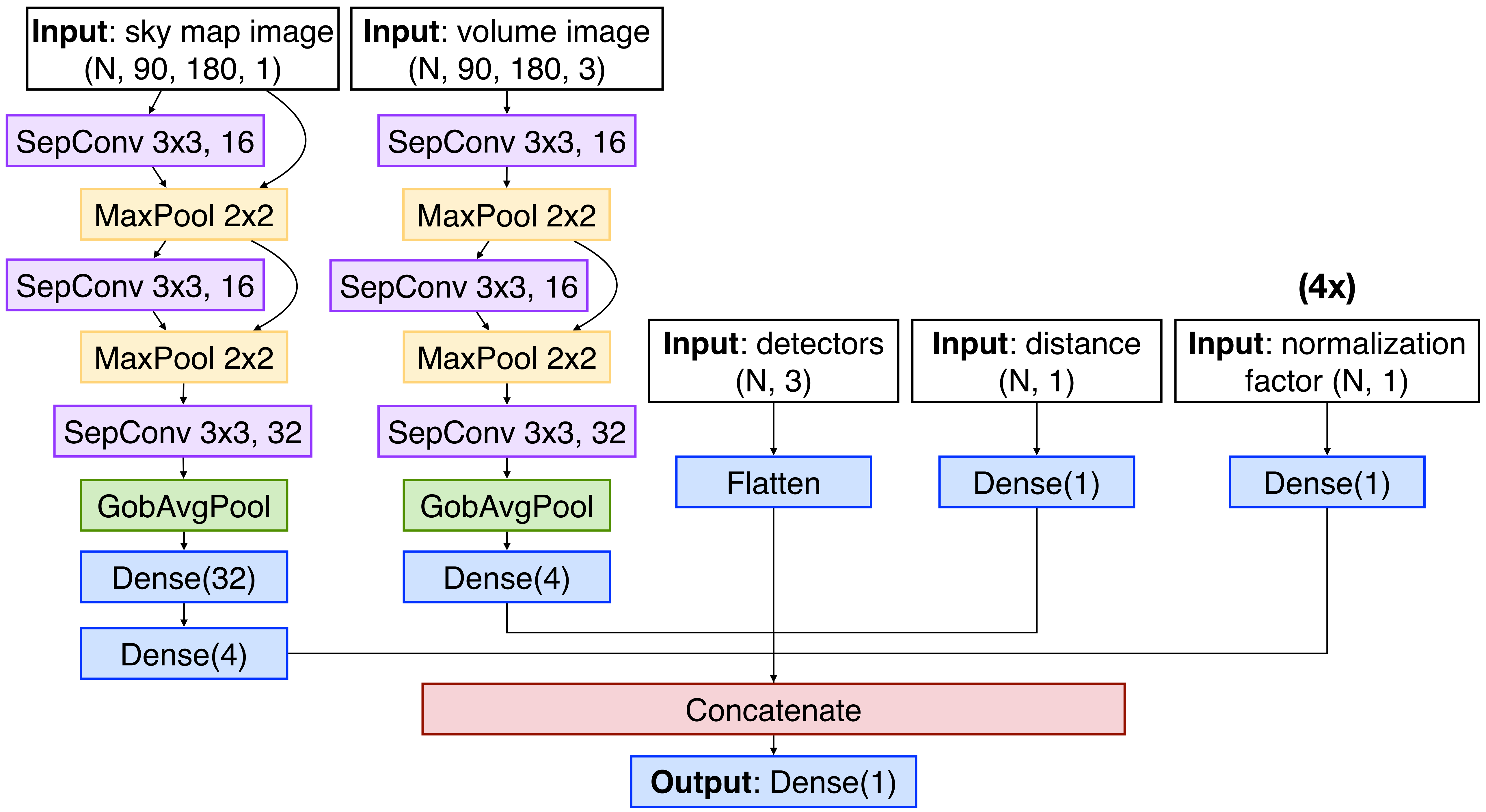
$$R = \frac{TP}{TP + FN}$$

Actual positives

How many real candidates are predicted correctly?



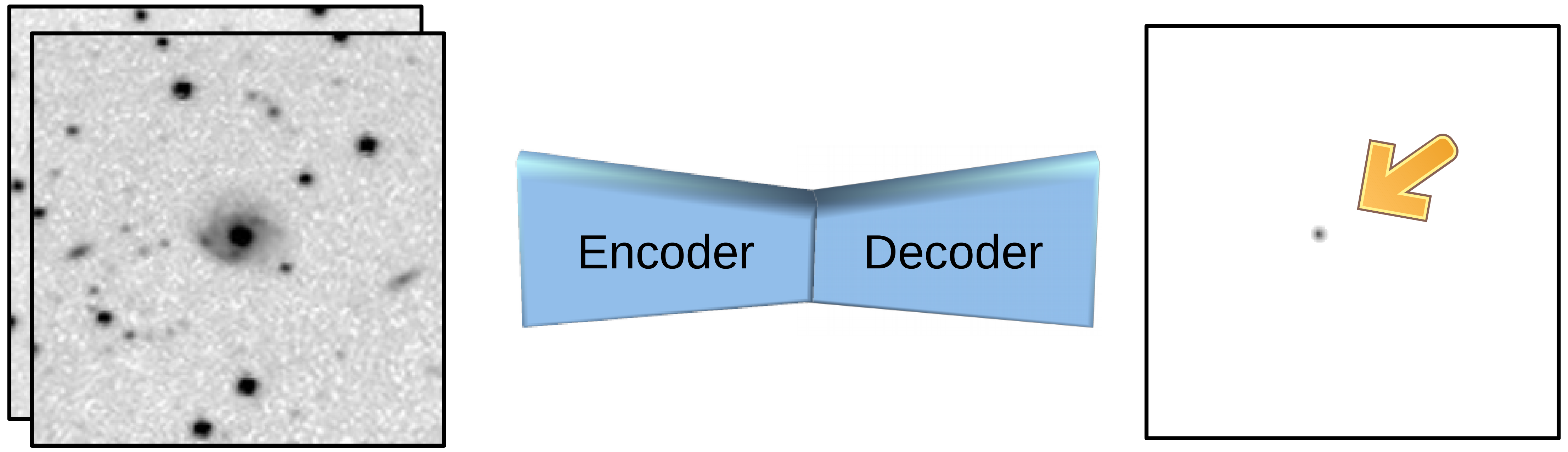


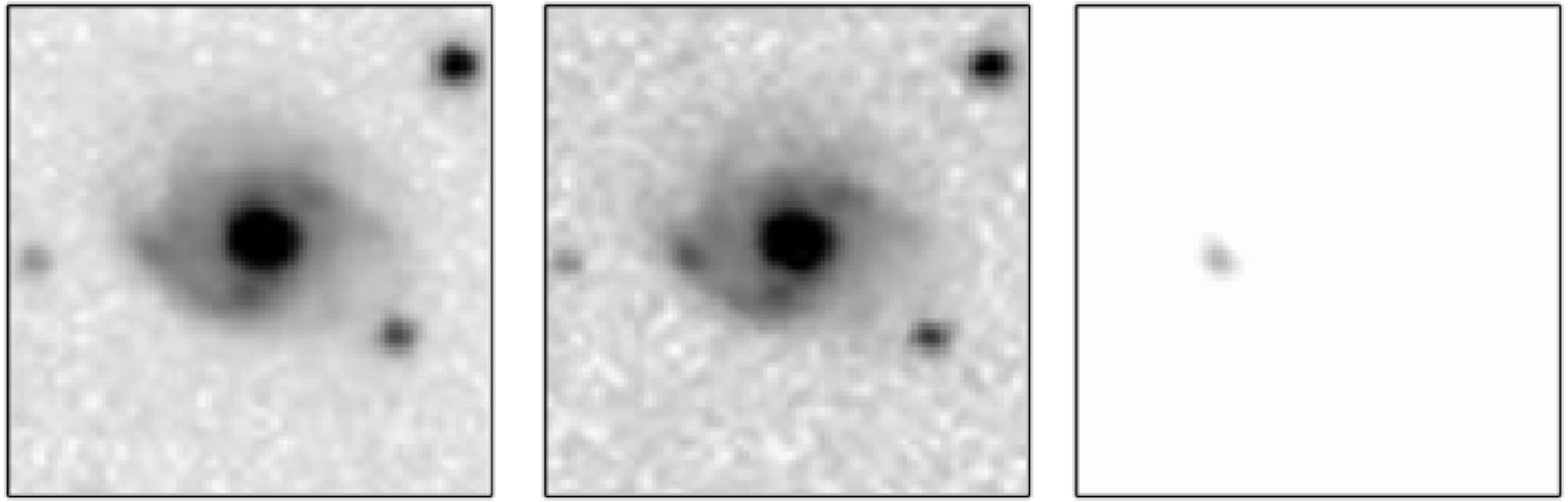


Examples with multple inputs

Similarly one can build models with multiple outputs
e.g. classification and regression

Image subtraction for hunting transients without subtraction





Sedaghat and Mahabal, 2017
arXiv:1710.01422

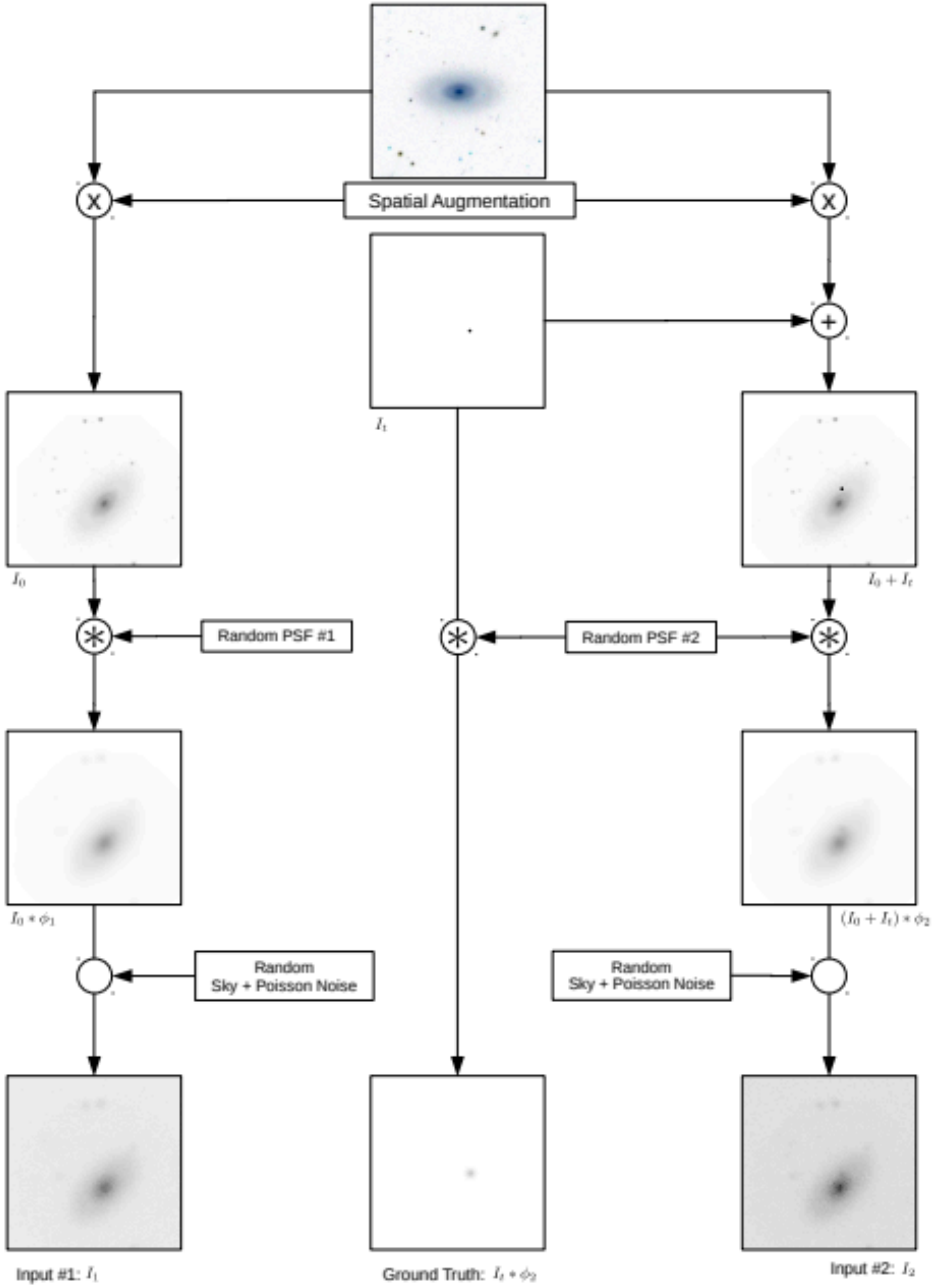
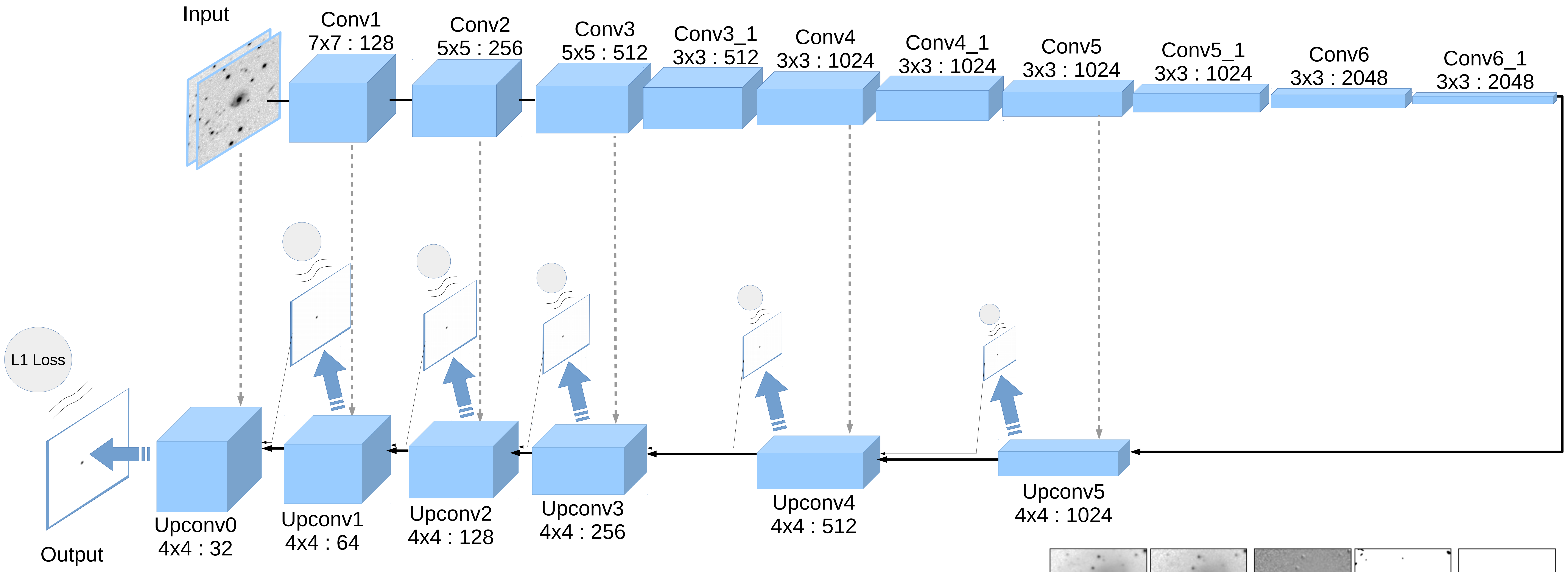
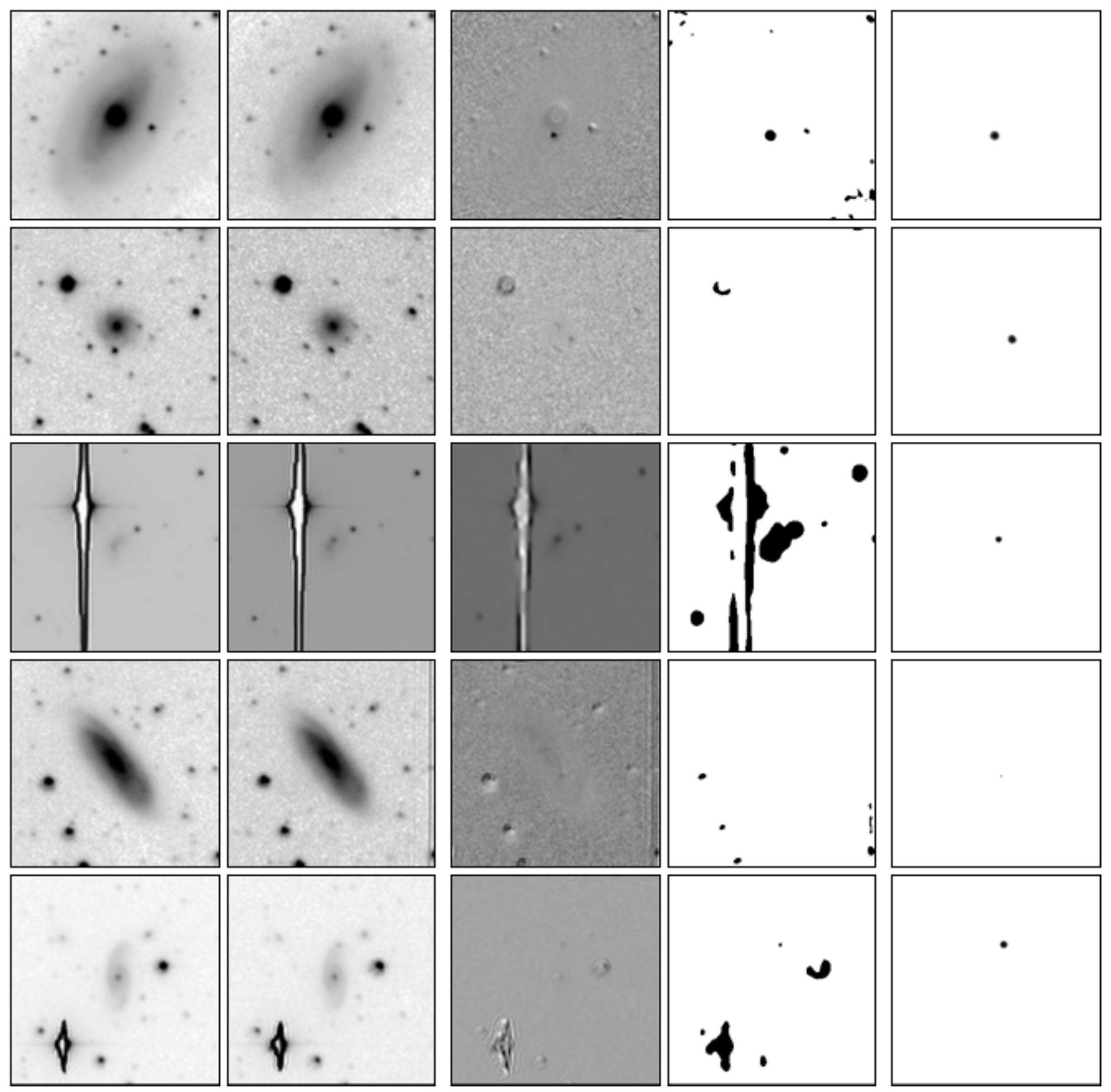


Figure 5. The synthetic sample generation procedure. The notations used here are described in Equations (1) and (2).

Training cycles involving different PSFs

Encoder-decoder network (fully convolutional)



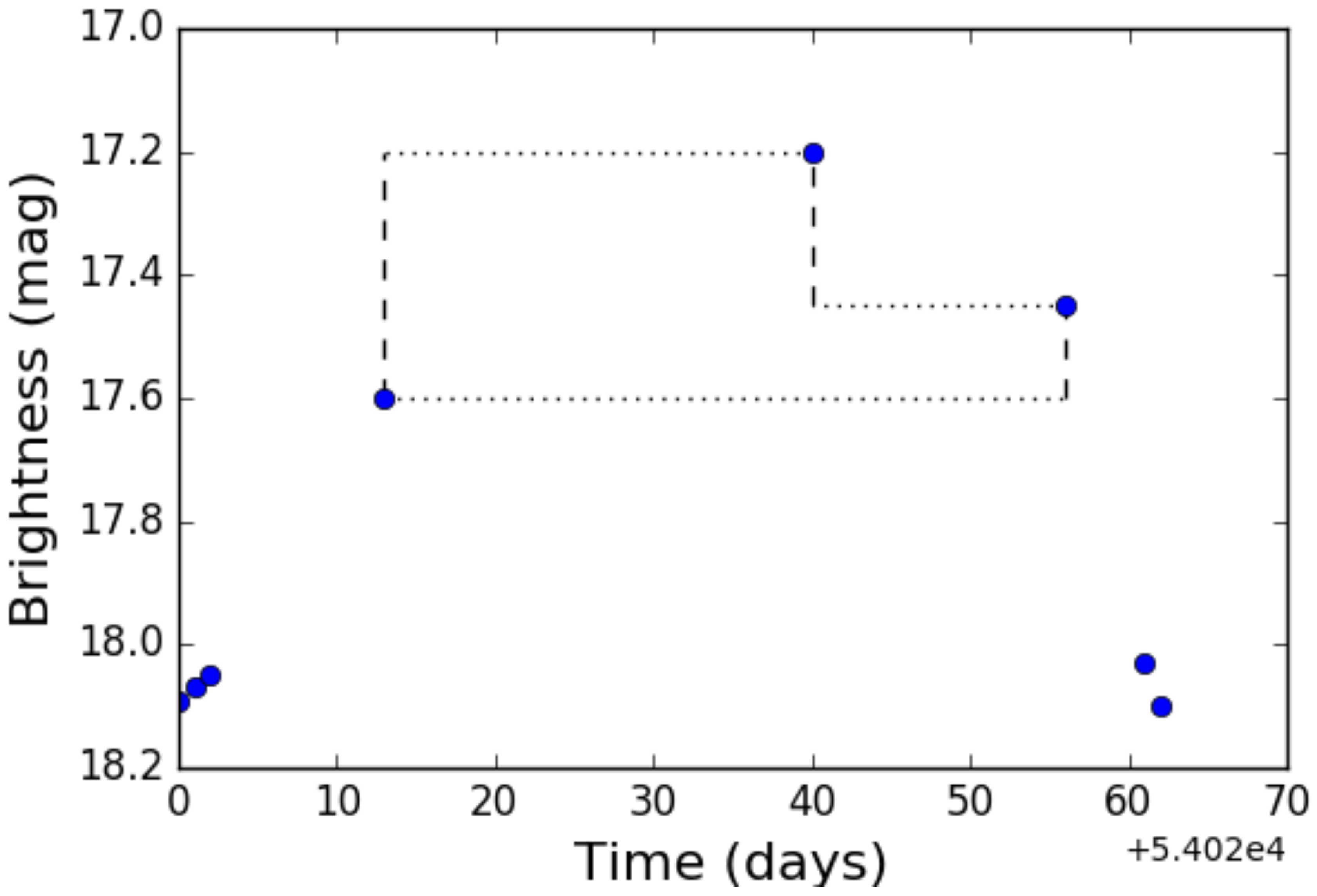


Sedaghat and Mahabal, 2017

arXiv:1710.01422

(dmdt) Image representation

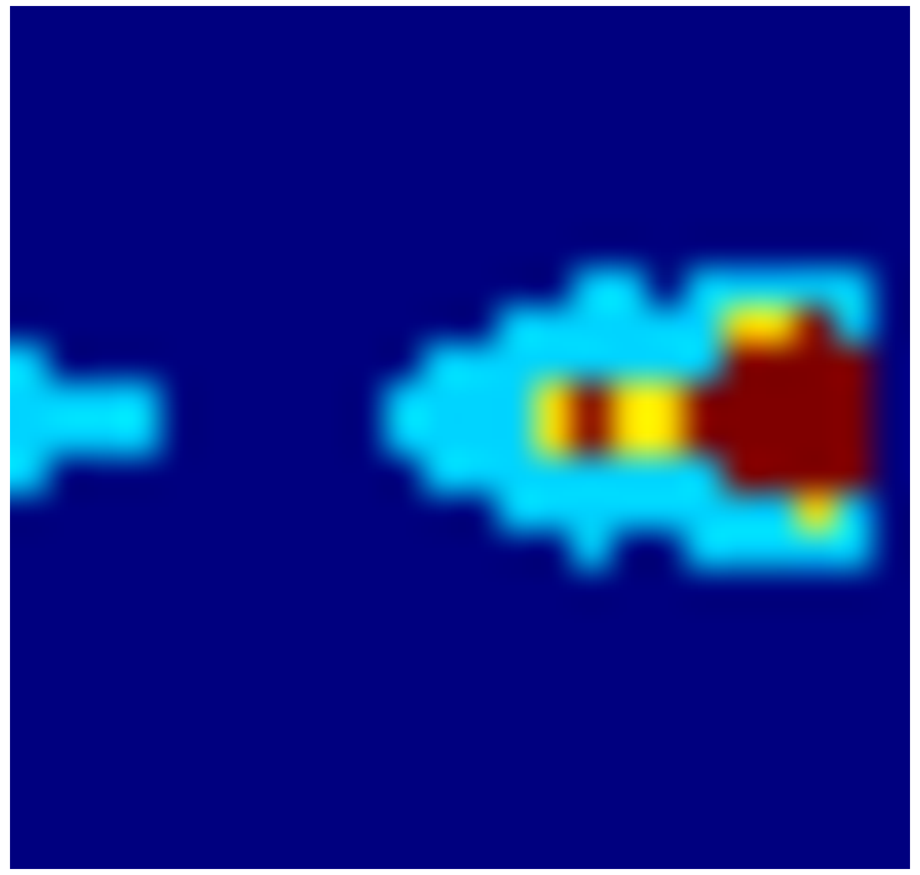
Mahabal et al., 2017

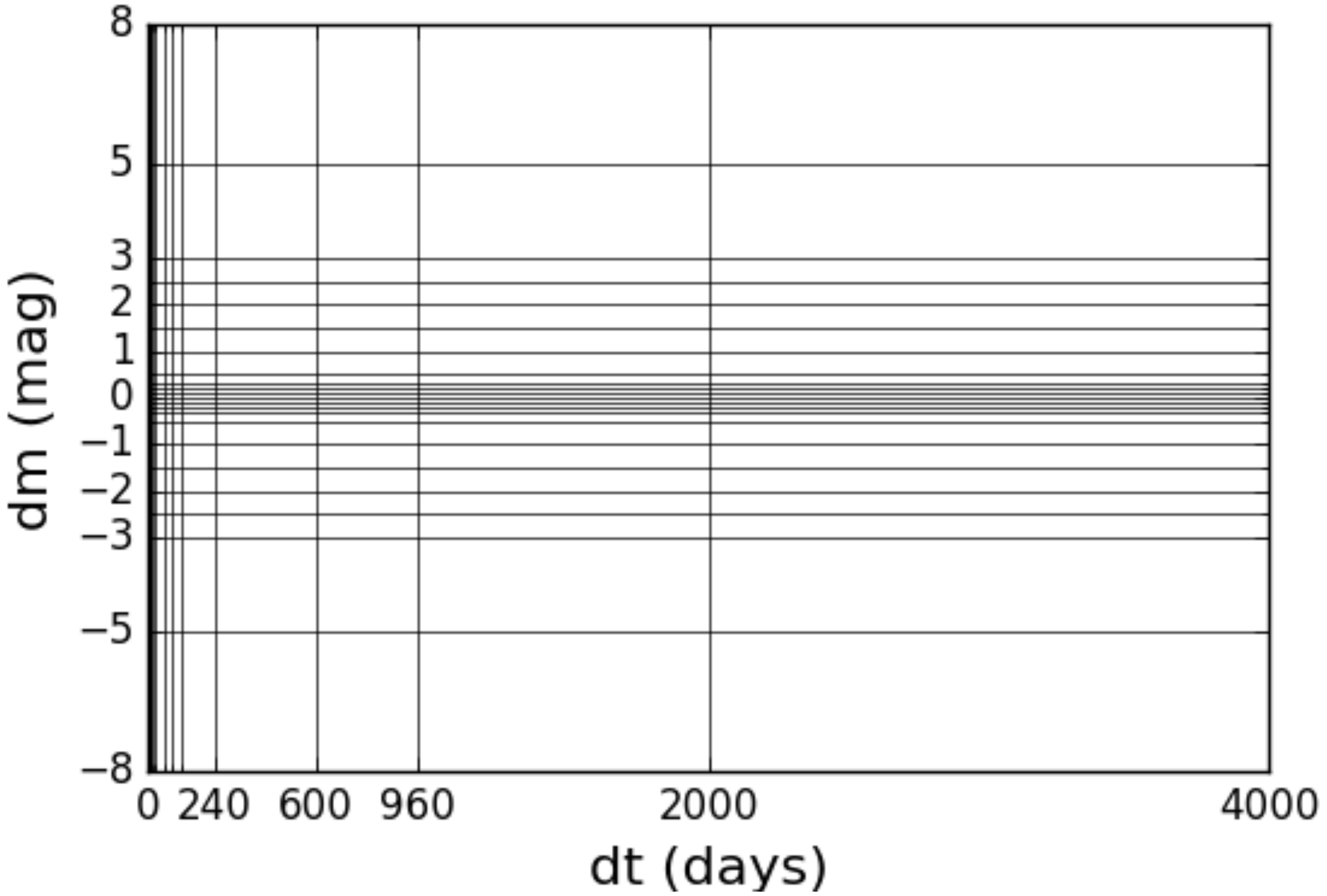


light curve with n points

23 x 24
output grid

n * (n-1)/2 points

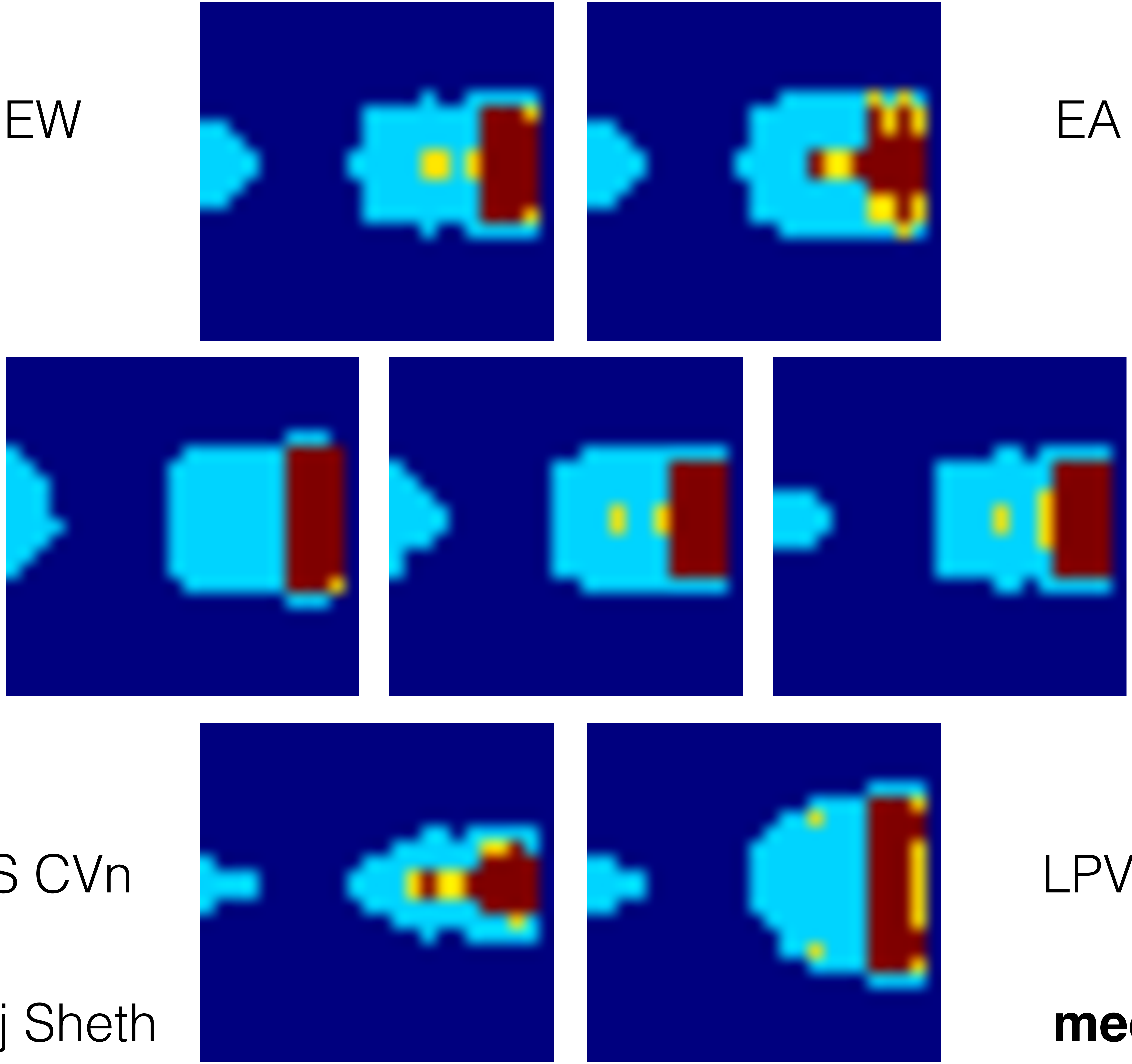




Area equalized pixels

EW

EA



RR

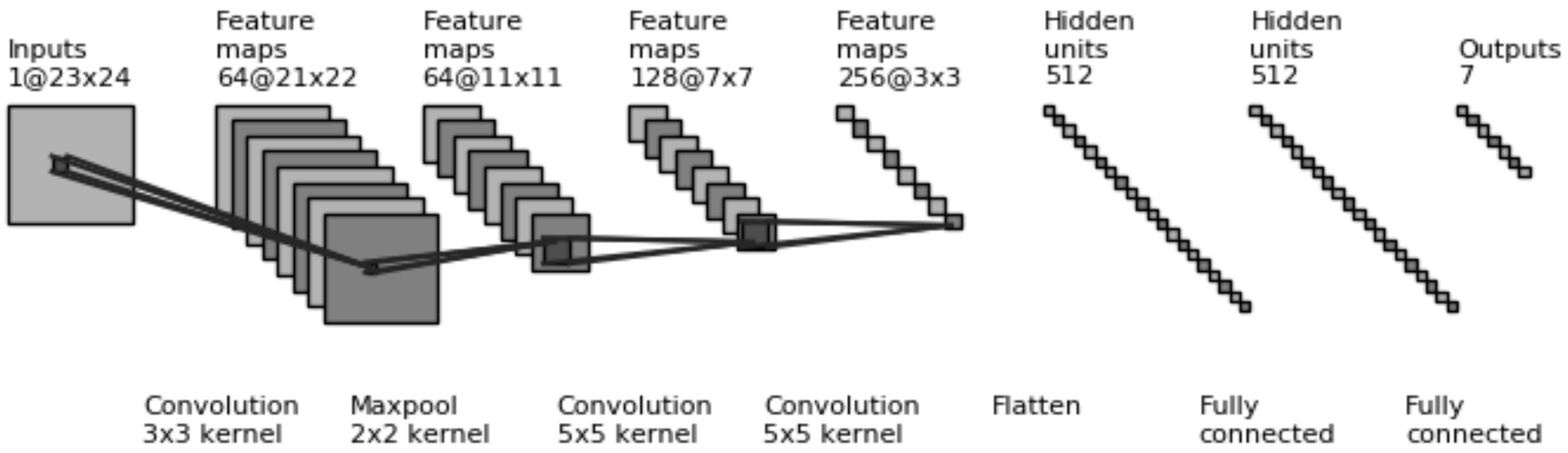
RS CVn

LPV

Kshiteej Sheth

medians

Network architecture



In this case shallow works well too

```
layers = [
    InputLayer,
    Conv2DLayer(32, size:3x3, rectify),
    DropoutLayer(0.1),
    DenseLayer(128),
    DropoutLayer(0.25),
    DenseLayer(128),
    DenseLayer(all, softmax),
```

Random Forest using standard features

True Class \ Prediction	1	2	4	5	6	8	13
1	94	2	0	2	0	0	0
2	15	84	0	0	0	0	0
4	31	0	57	10	0	0	0
5	43	0	2	54	0	0	0
6	43	0	8	38	10	0	0
8	97	0	0	1	0	0	0
13	22	0	5	0	0	0	71

no features
no dimensionality reduction
comparable results

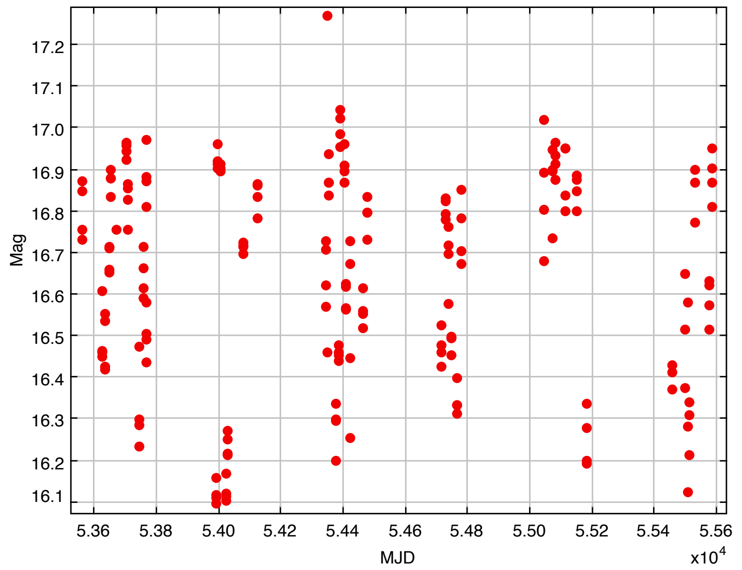
Convolutional Network

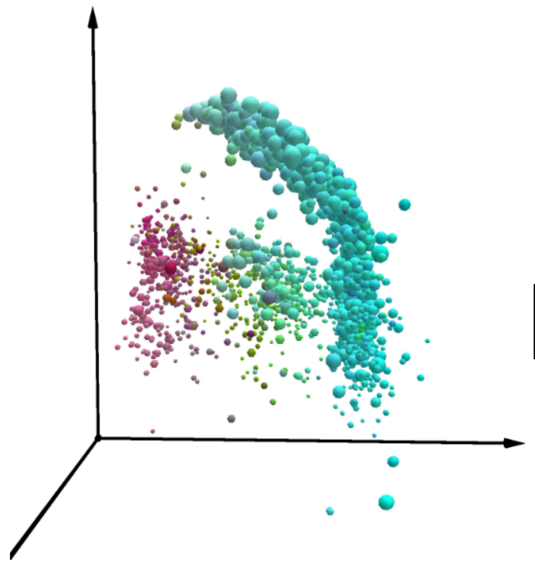
True Class \ Prediction	1	2	4	5	6	8	13
1	94	2	0	2	0	0	0
2	18	81	0	0	0	0	0
4	32	0	53	14	0	0	0
5	32	0	1	65	0	0	0
6	26	0	5	66	0	1	0
8	78	0	0	4	0	13	0
13	1	1	5	1	2	3	83

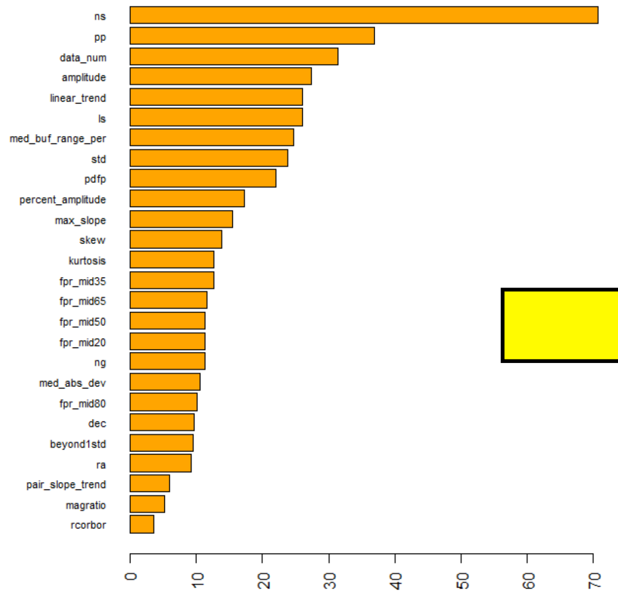
1	EW/EB
2	EA
4	RRab
5	RRc
6	RRd
8	RS CVn
13	LPV

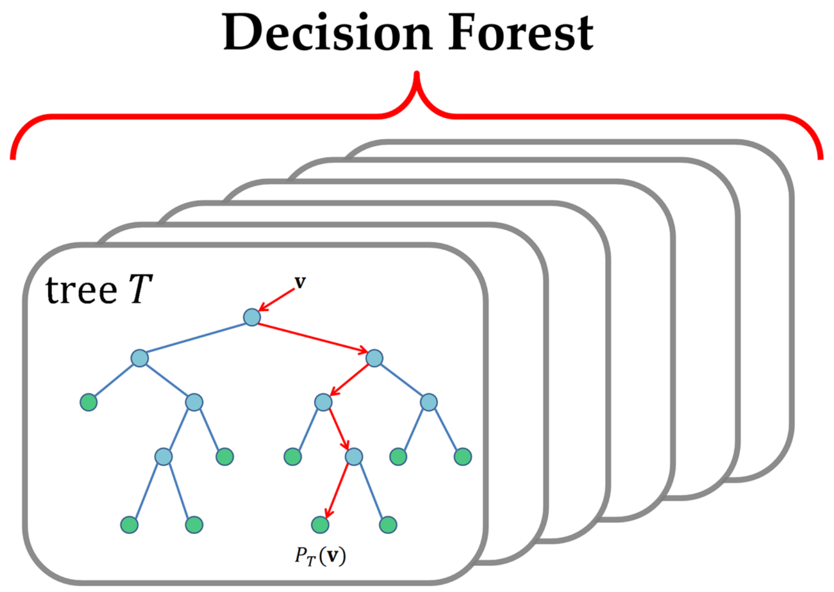
Binary probabilities are better

Classification Workflow









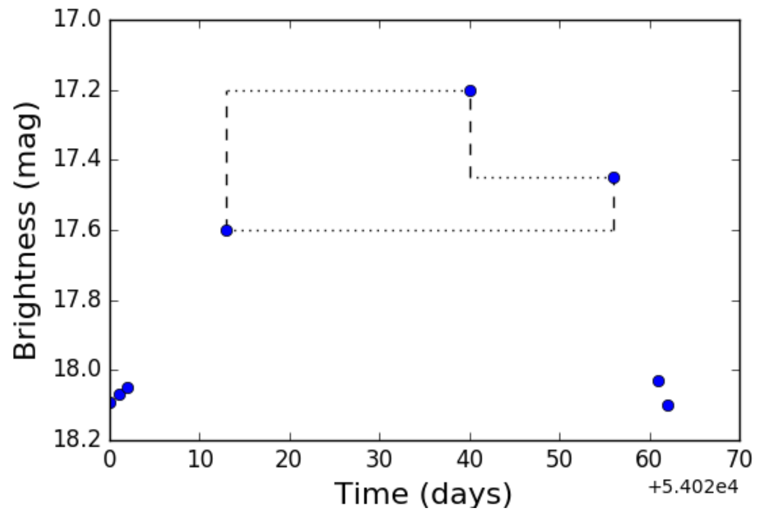
Light curves

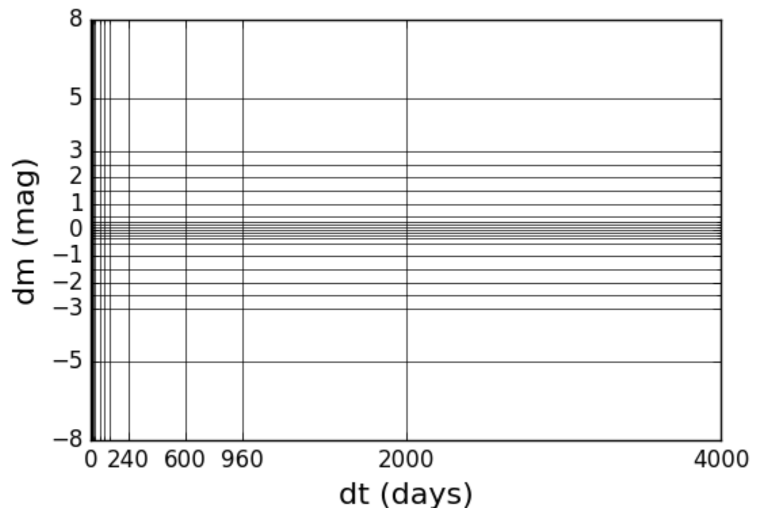
Feature vectors

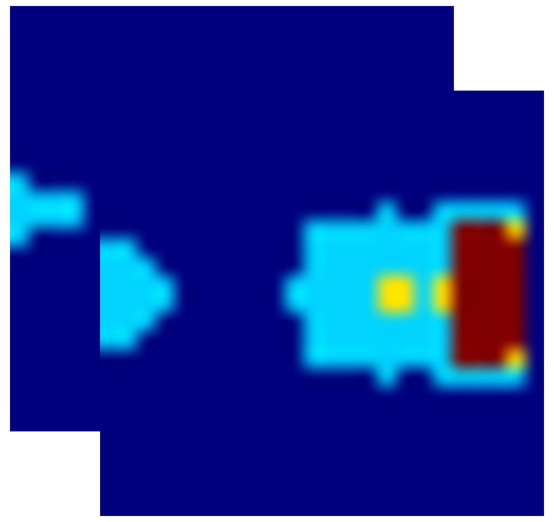
Dimensionality Reduction

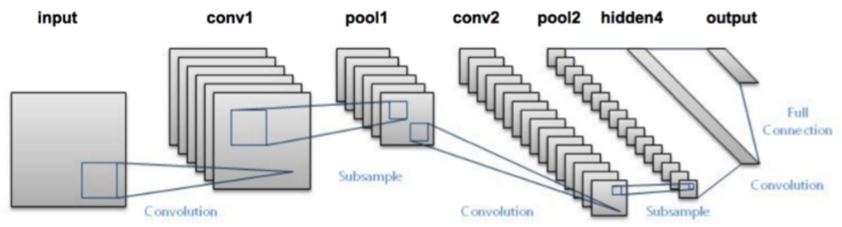
Classification

Domain knowledge/subjectivity









Light curves

Density representation

Equi-area images

Convolutional Neural Network

Improving LIGO detector duty cycle

New methods to assess and improve LIGO detector duty cycle

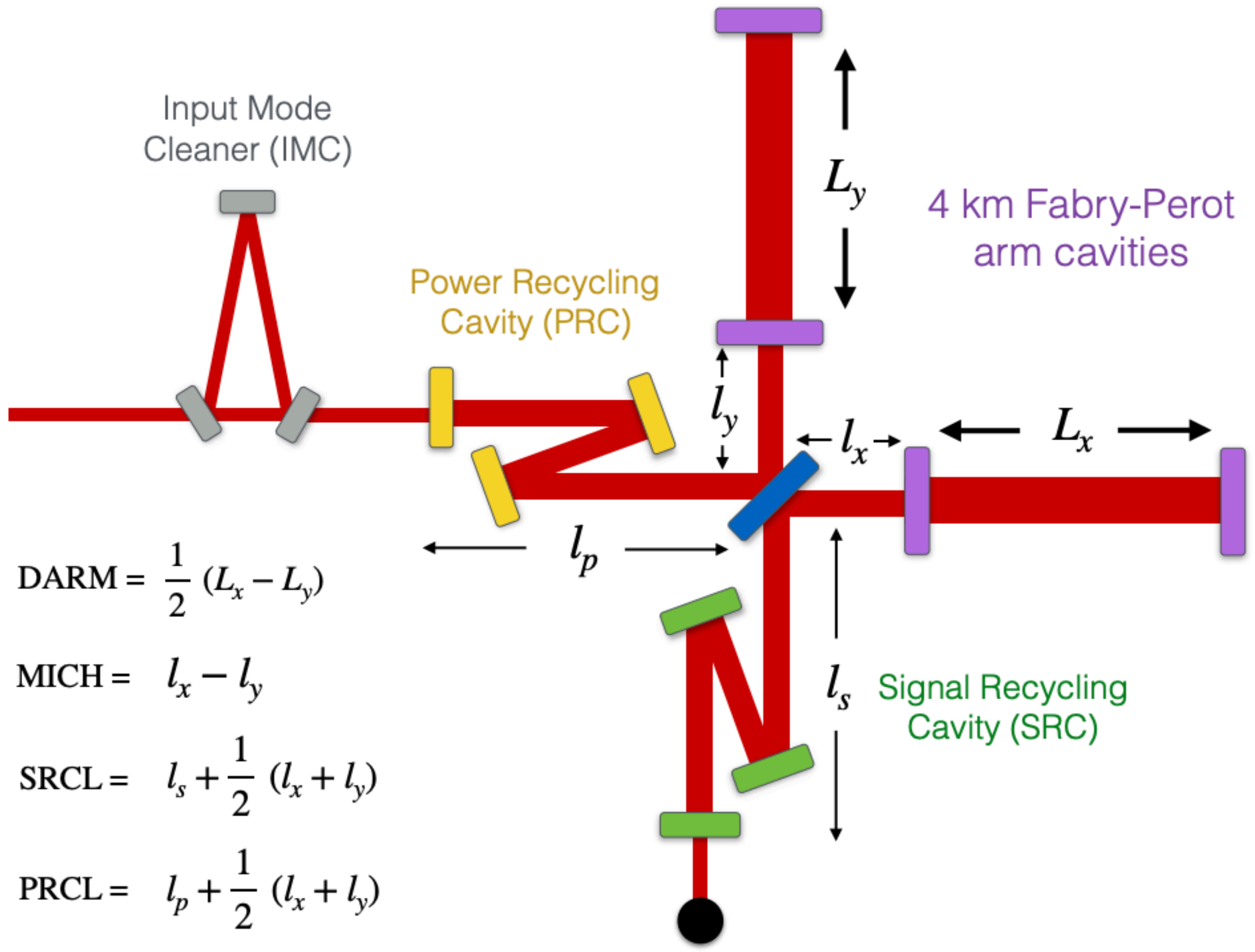
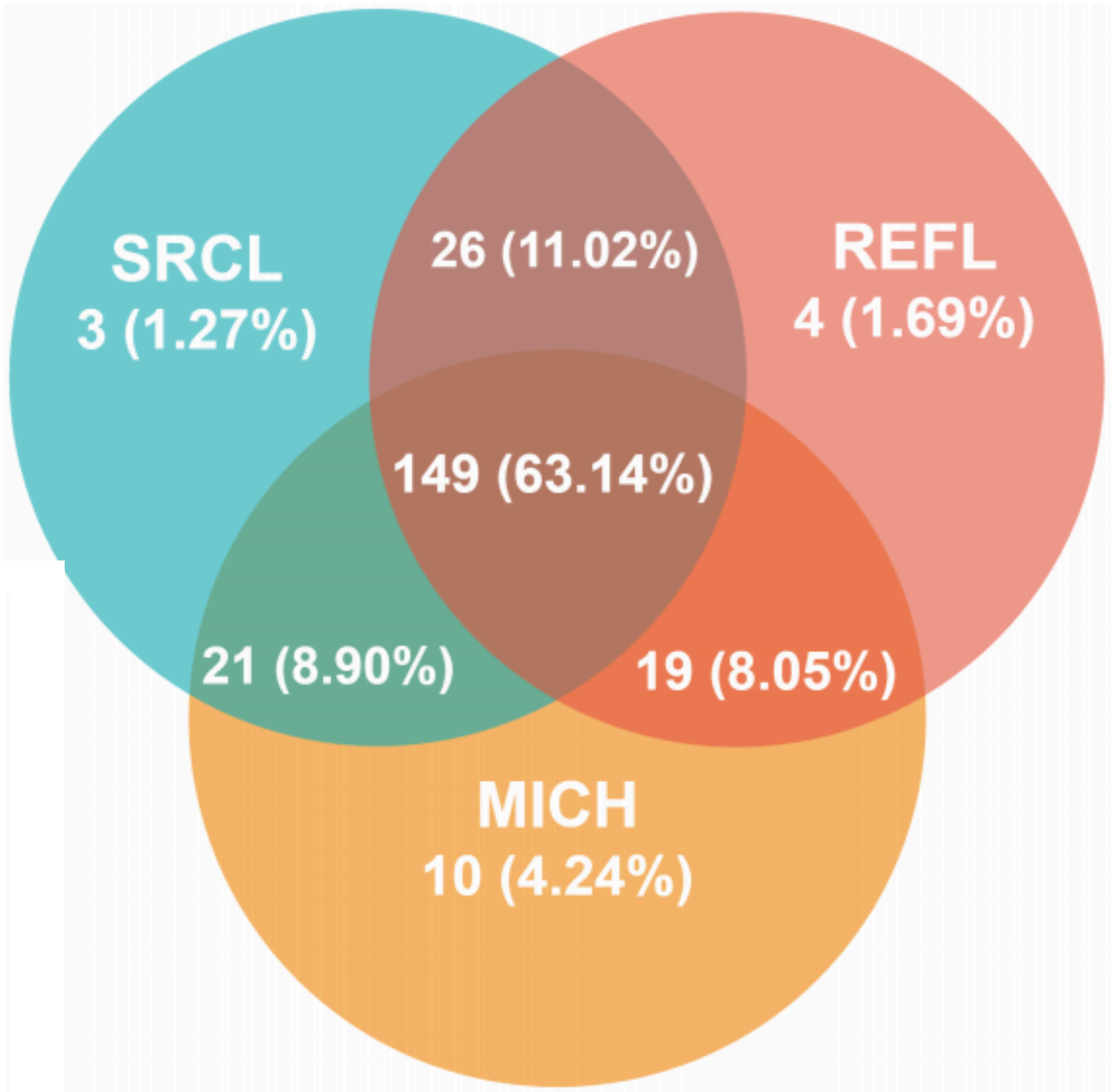
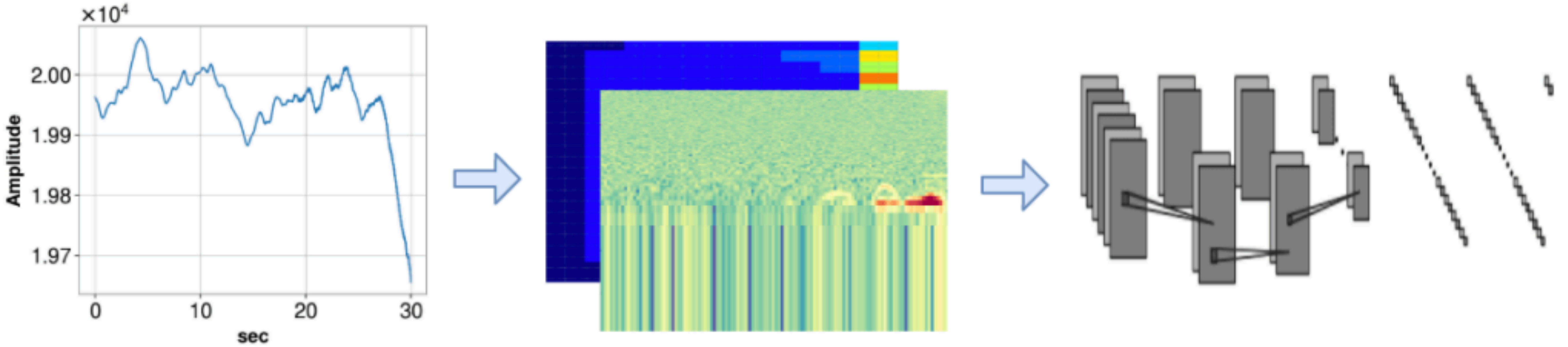


Table 8: Performance metrics for stacked *dmdt* of multiple channels.

channels	Test accuracy	Matthew's coefficient	F1 score	Specificity	Recall	Precision
MICH+POP	0.971	0.941	0.968	0.985	0.953	0.983
MICH+REFL	0.975	0.949	0.972	0.996	0.949	0.996
MICH+SRCL	0.971	0.941	0.968	0.989	0.949	0.987
MICH+POP+SRCL	0.984	0.969	0.983	1.000	0.966	1.000
MICH+PRCL+REFL	0.984	0.969	0.983	0.996	0.970	0.996
MICH+REFL+SRCL	0.986	0.973	0.985	0.996	0.975	0.996
IMC+MICH+POP+SCRL	0.988	0.976	0.987	0.996	0.979	0.996
IMC+MICH+PRCL+REFL	0.986	0.973	0.985	1.000	0.970	1.000
MICH+PRCL+REFL+SRCL	0.986	0.973	0.985	0.996	0.975	0.996
IMC+MICH+POP+PRCL+SCRL	0.990	0.980	0.989	1.000	0.979	1.000



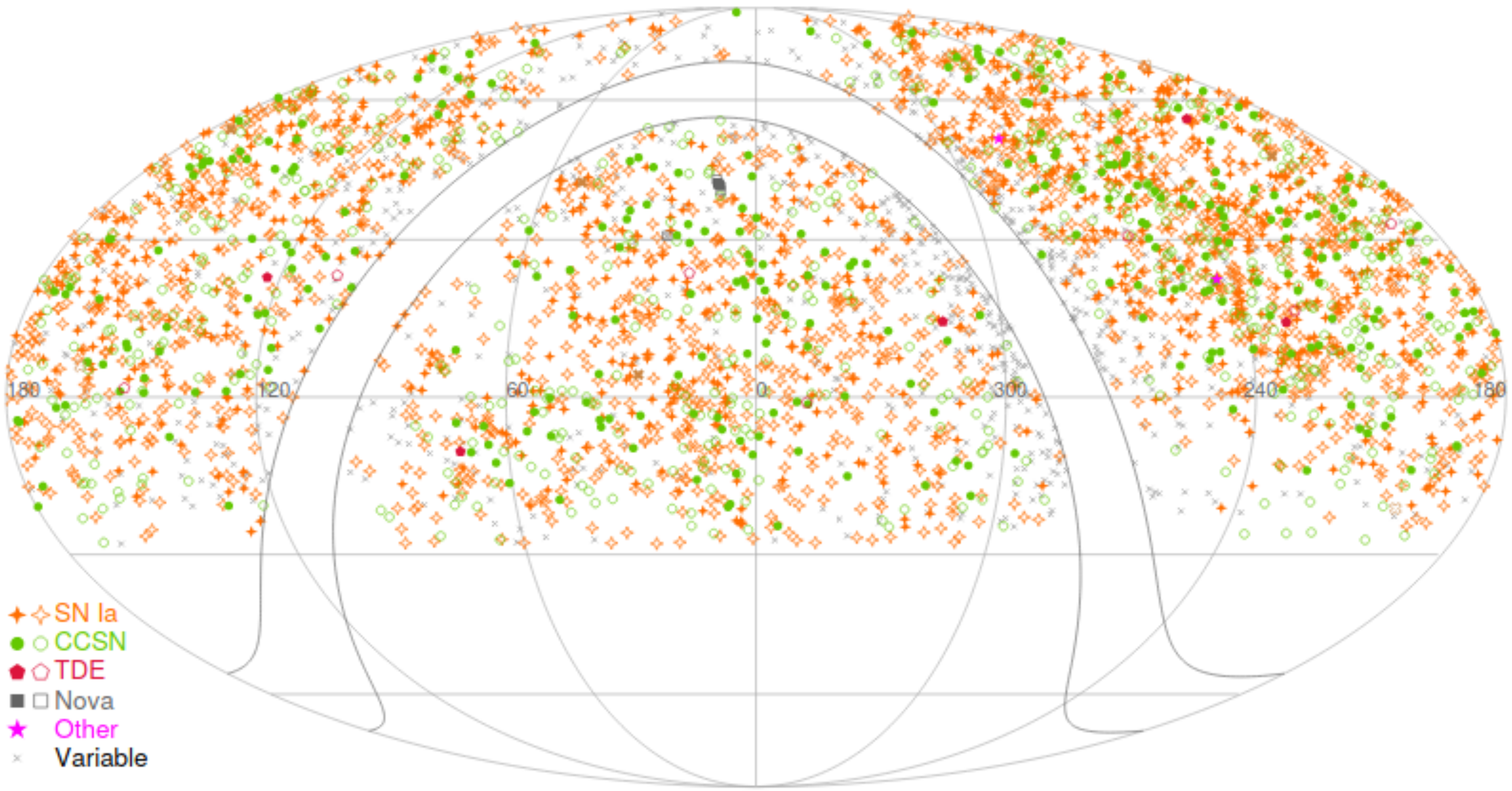

```
Layers = [
InputLayer,
Conv2DLayer(16,3x3,ReLU),
Conv2DLayer(32,3x3,ReLU),
MaxPoolLayer(3x3),
DropoutLayer(0.25),
DenseLayer(16),
DropoutLayer(0.5),
DenseLayer(2)
]
```

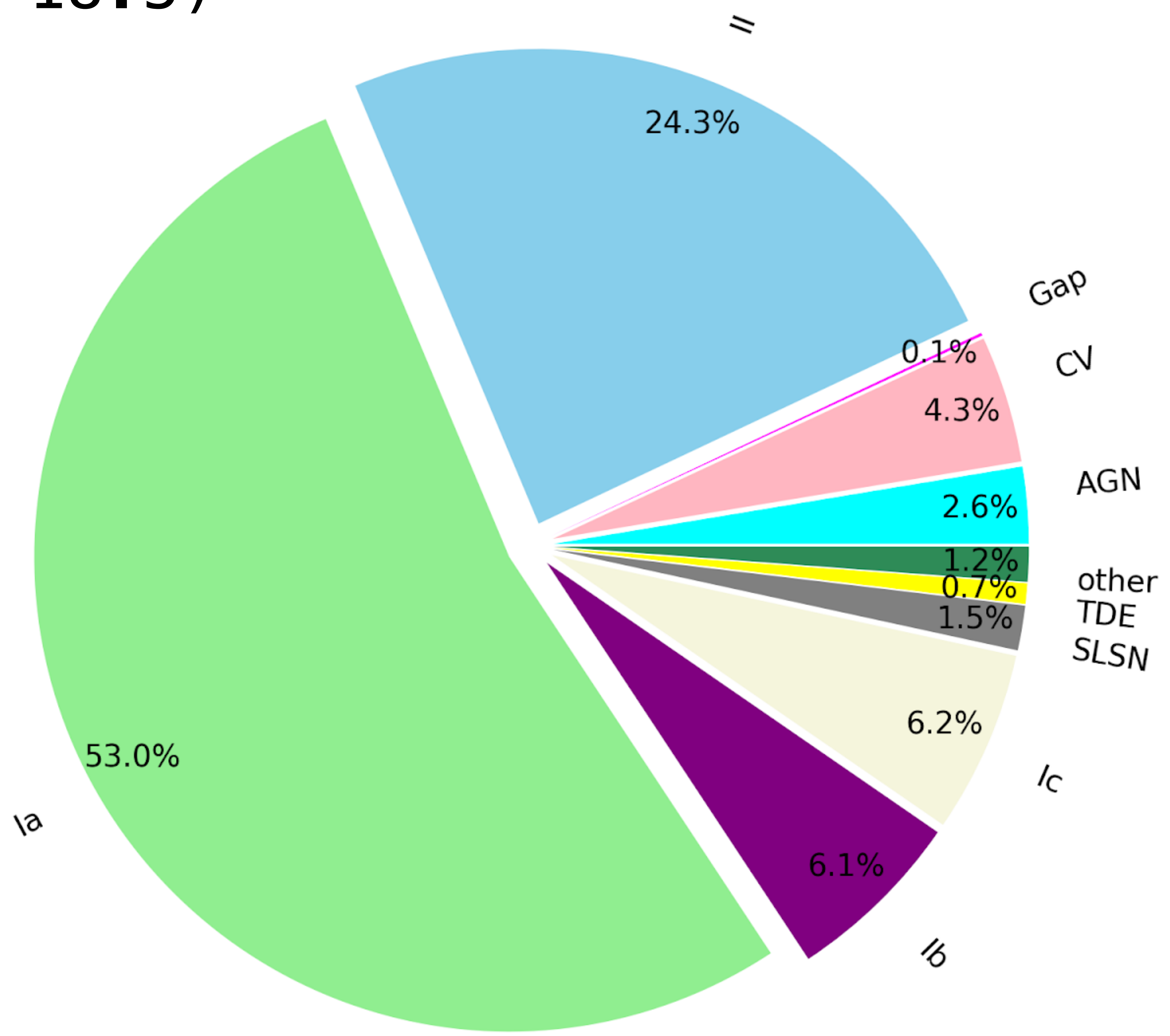


Time Series **2-D features** **Convolutional neural network**

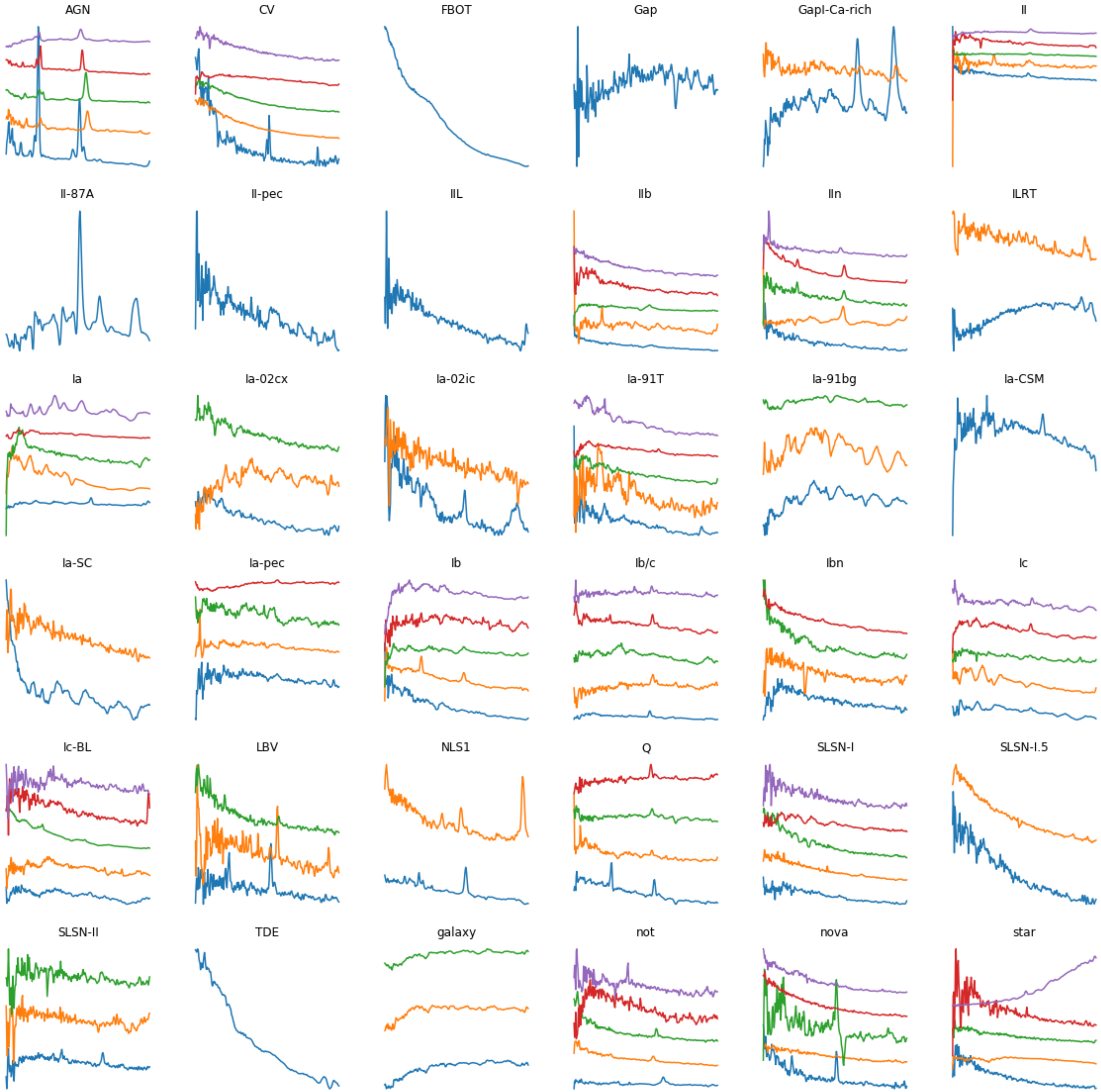
BTS : An unbiased view of transient night sky!

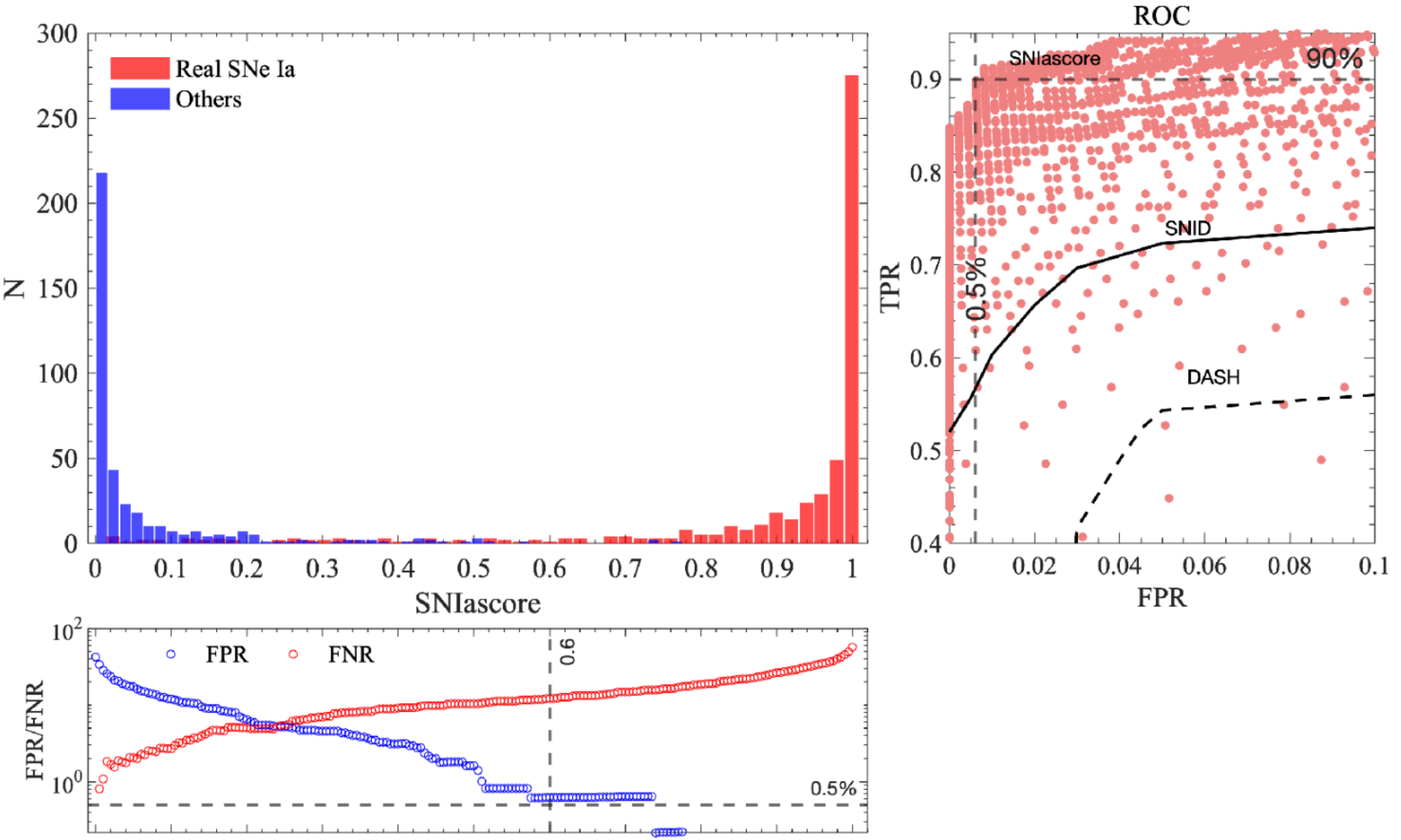
Untargeted and flux limited (m < 19)
No strong variables (like CV, AGN) or moving objects
Spectroscopically complete (m < 18.5)



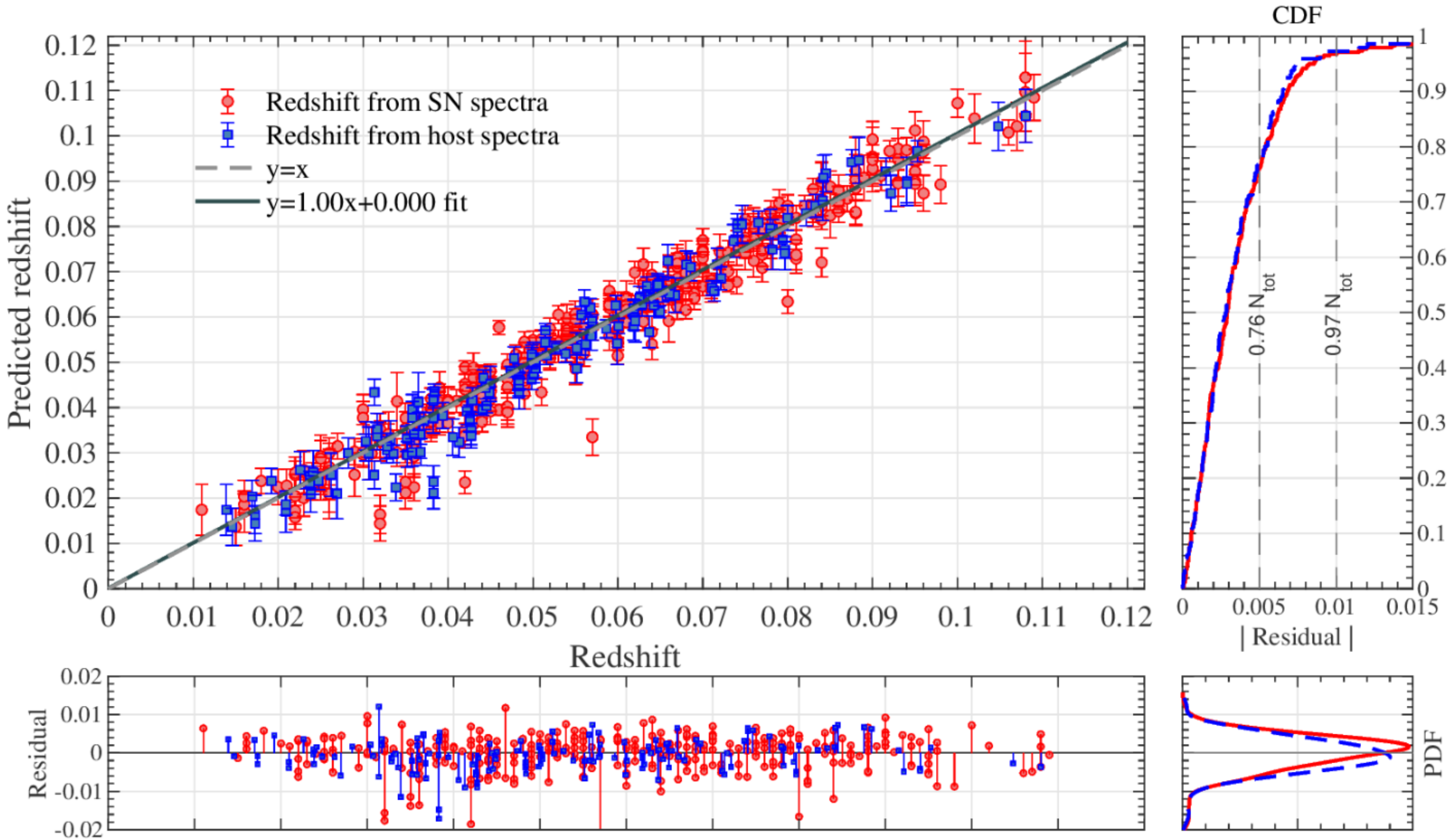


Christoffer Fremling,
Yashvi Sharma, ...


AGN
CV
FBOT
Gap
GapI-Ca-rich
II
II-87A
II-pec
IIL
IIb
IIn
ILRT
Ia
Ia-02cx
Ia-02ic
Ia-91T
Ia-91bg
Ia-CSM
Ia-SC
Ia-pec
Ib
Ib/c
Ibn
Ic
Ic-BL
LBV
NLS1
Q
SLSN-I
SLSN-I.5
SLSN-II
TDE
galaxy
not
nova
star

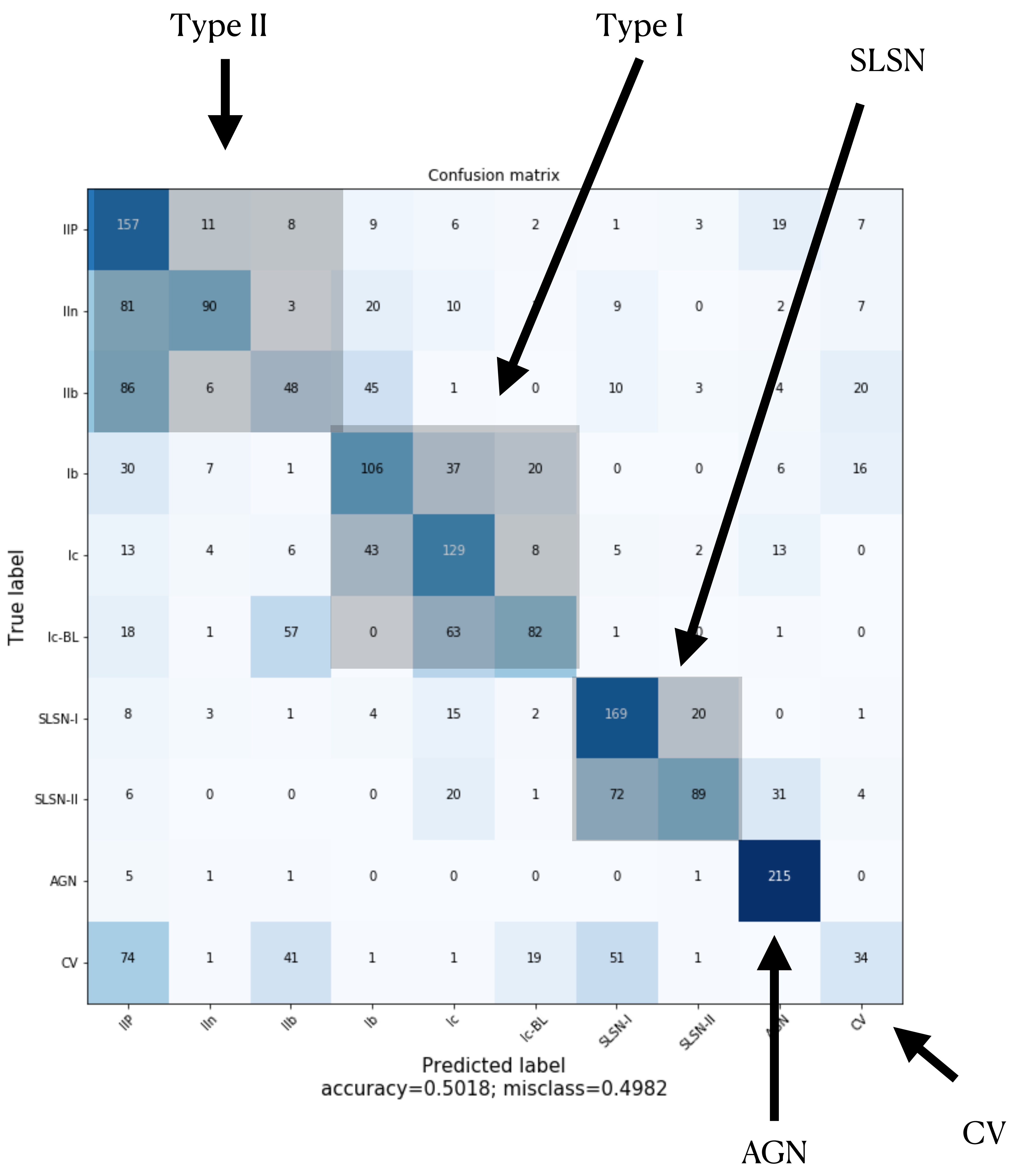

Real SNe Ia
Others
N
SNIascore
FPR
FNR
FPR/FNR
0.6
0.5%
ROC
SNIascore
90%
0.5%
SNID
DASH
TPR
FPR

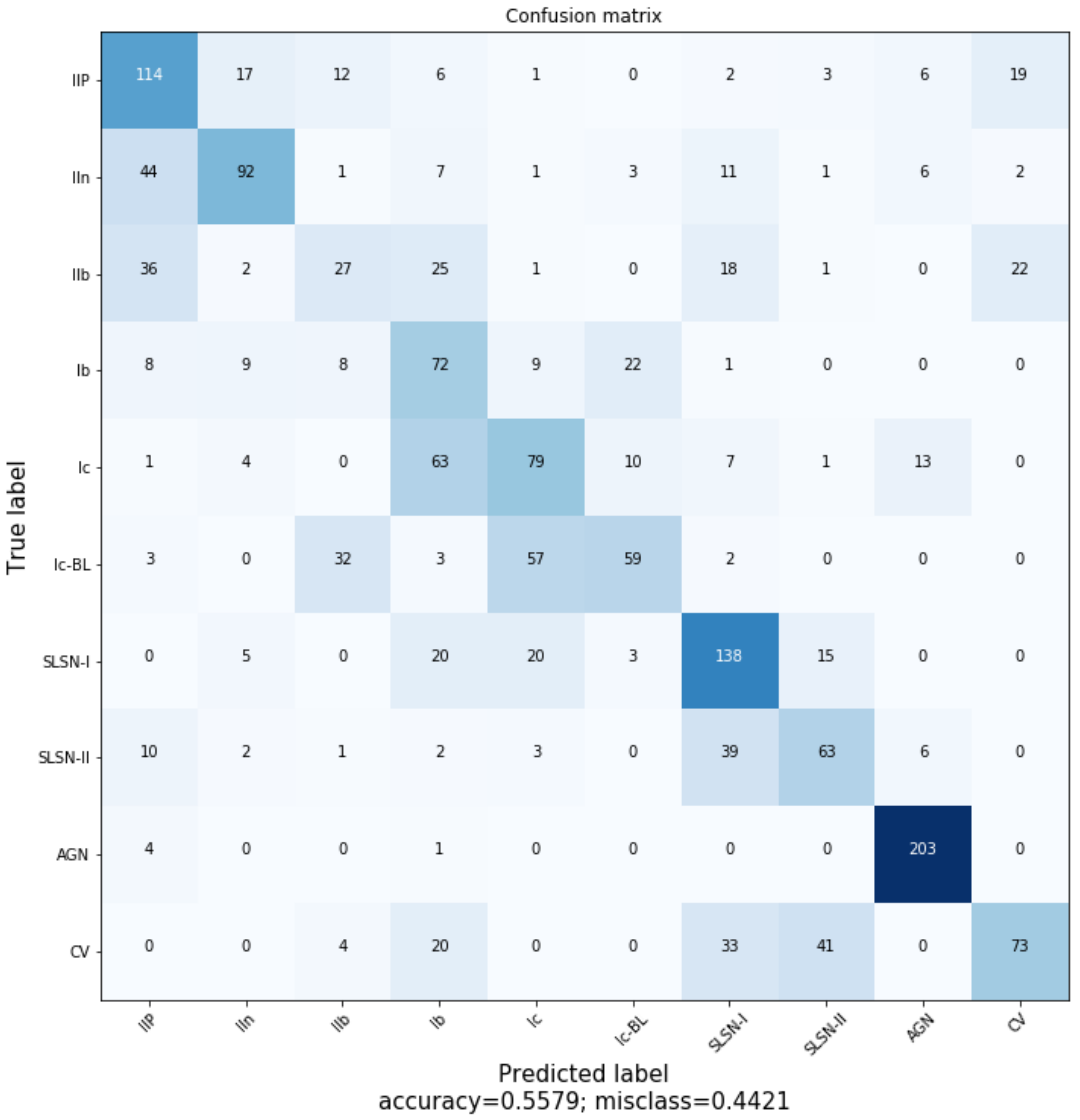
Redshift Regression



Model with two outputs

Multiclass, and binary non-SNIa SEDM classifiers



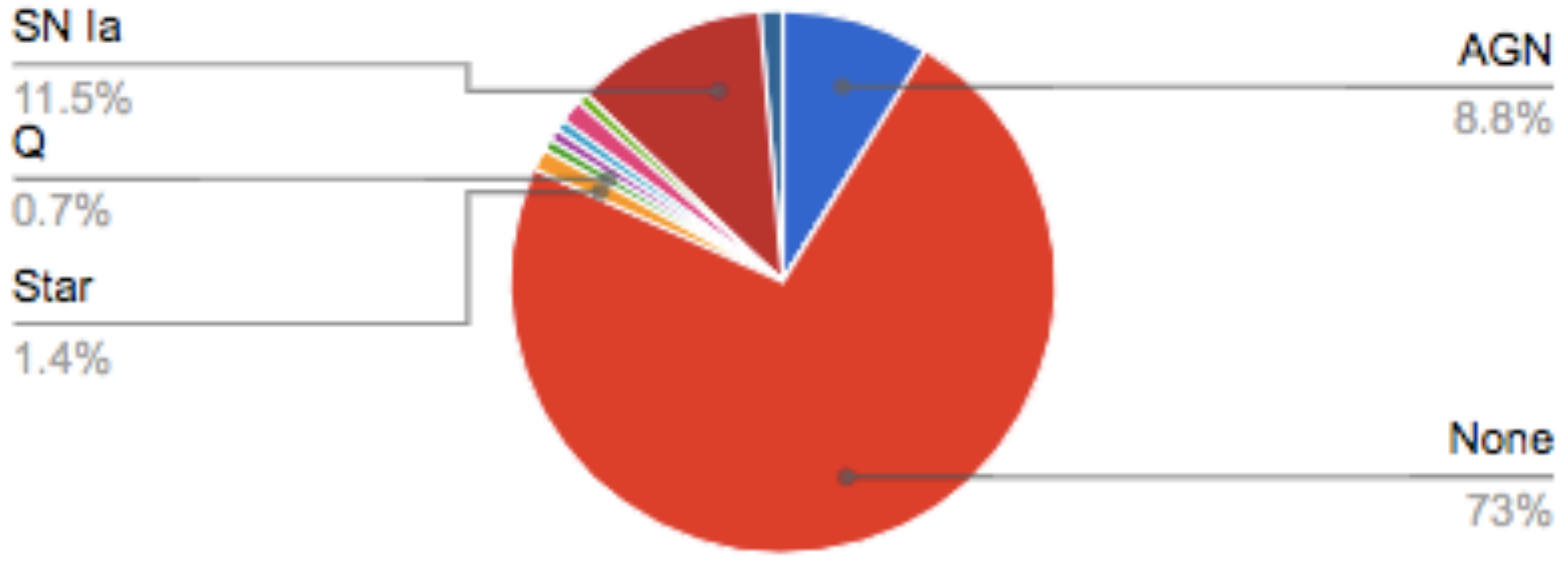


Analysis of Zooniverse data for fainter transients

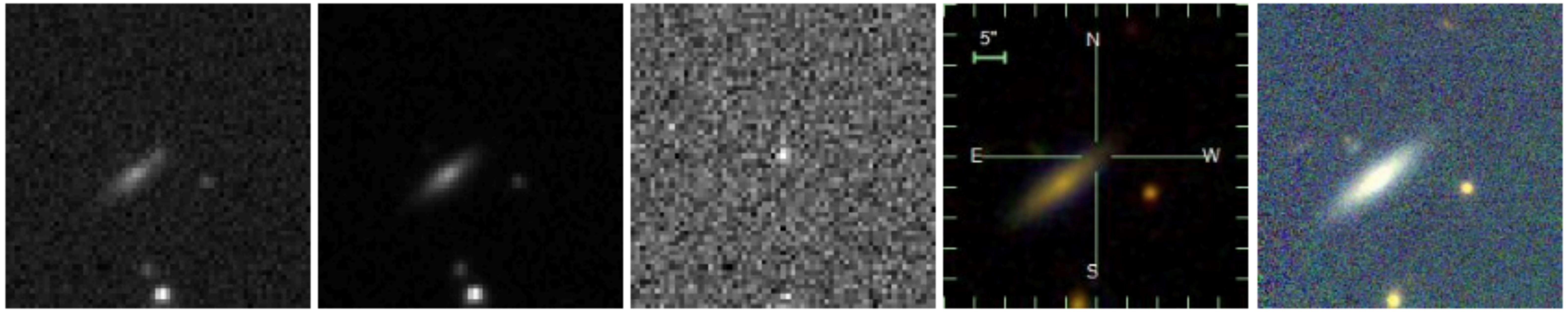
different campaigns
science groups still after
low hanging fruit
help with completeness
but also rare objects
automating analysis

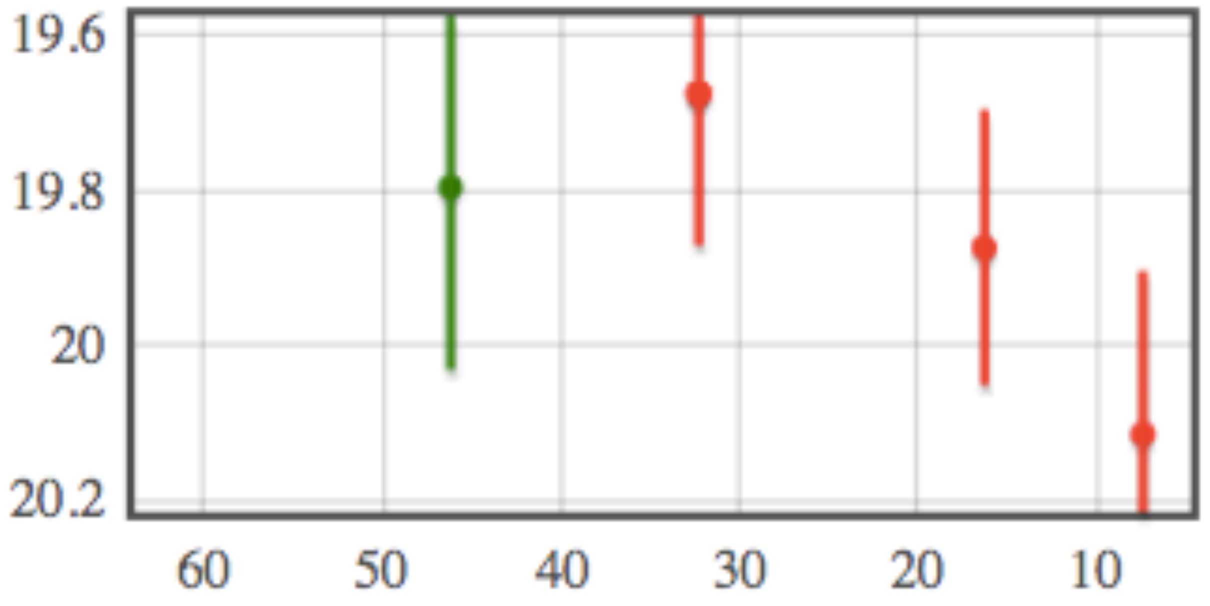
Subtypes:

Any Type: 148
AGN: 13
None: 108
Star: 2
Q: 1
LINER: 1
stellar: 1
SN Ic: 2
SN Ib: 1
SN Ia: 17



NEW **REF** **SUB** **SDSS** **PS1**





r = 20.1 (7.4 d) | Upload New Photometry

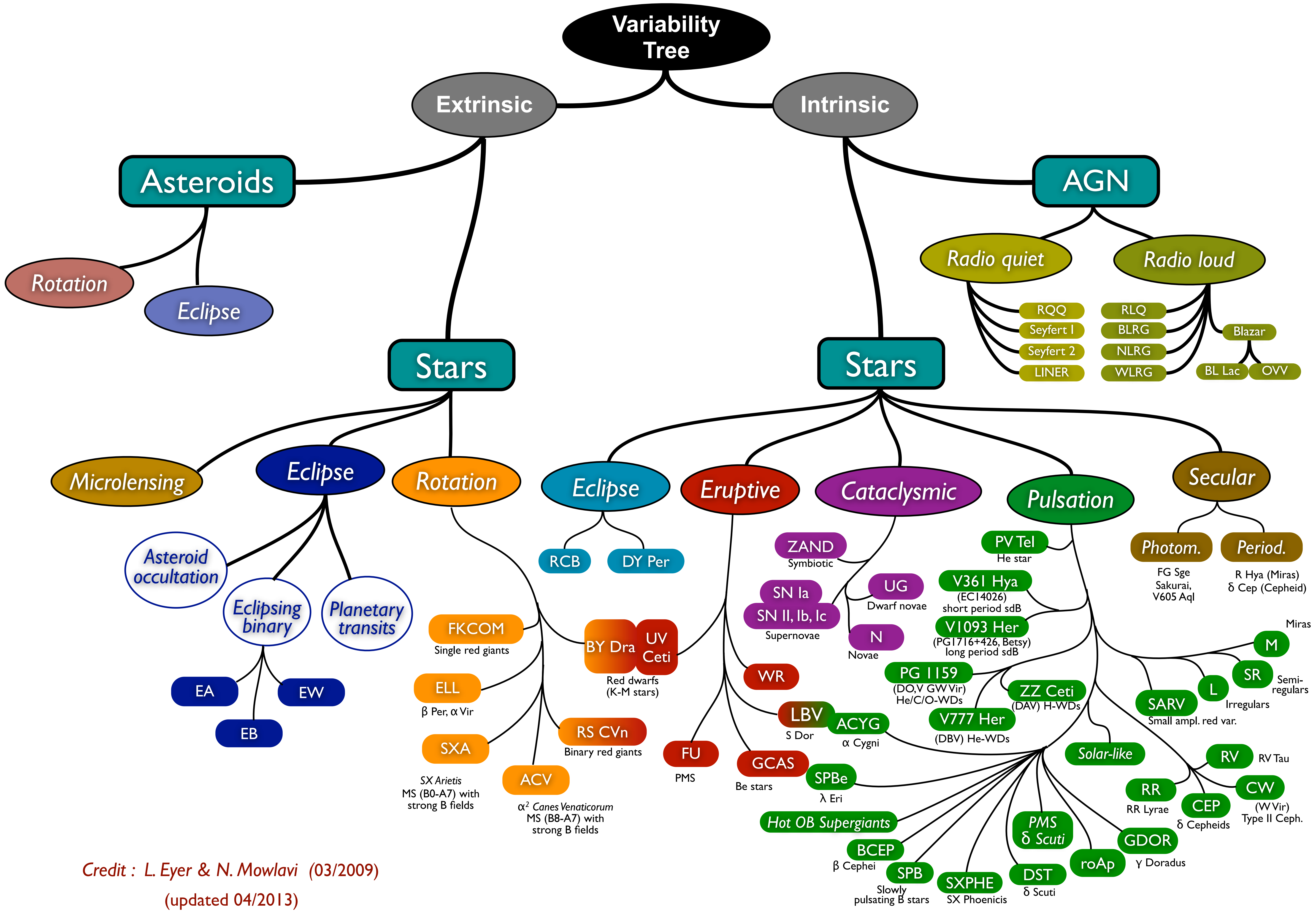
Upload New Spectroscopy

Fainter transients - no spectra

Faint LSST/Rubin objects will have to rely on ML

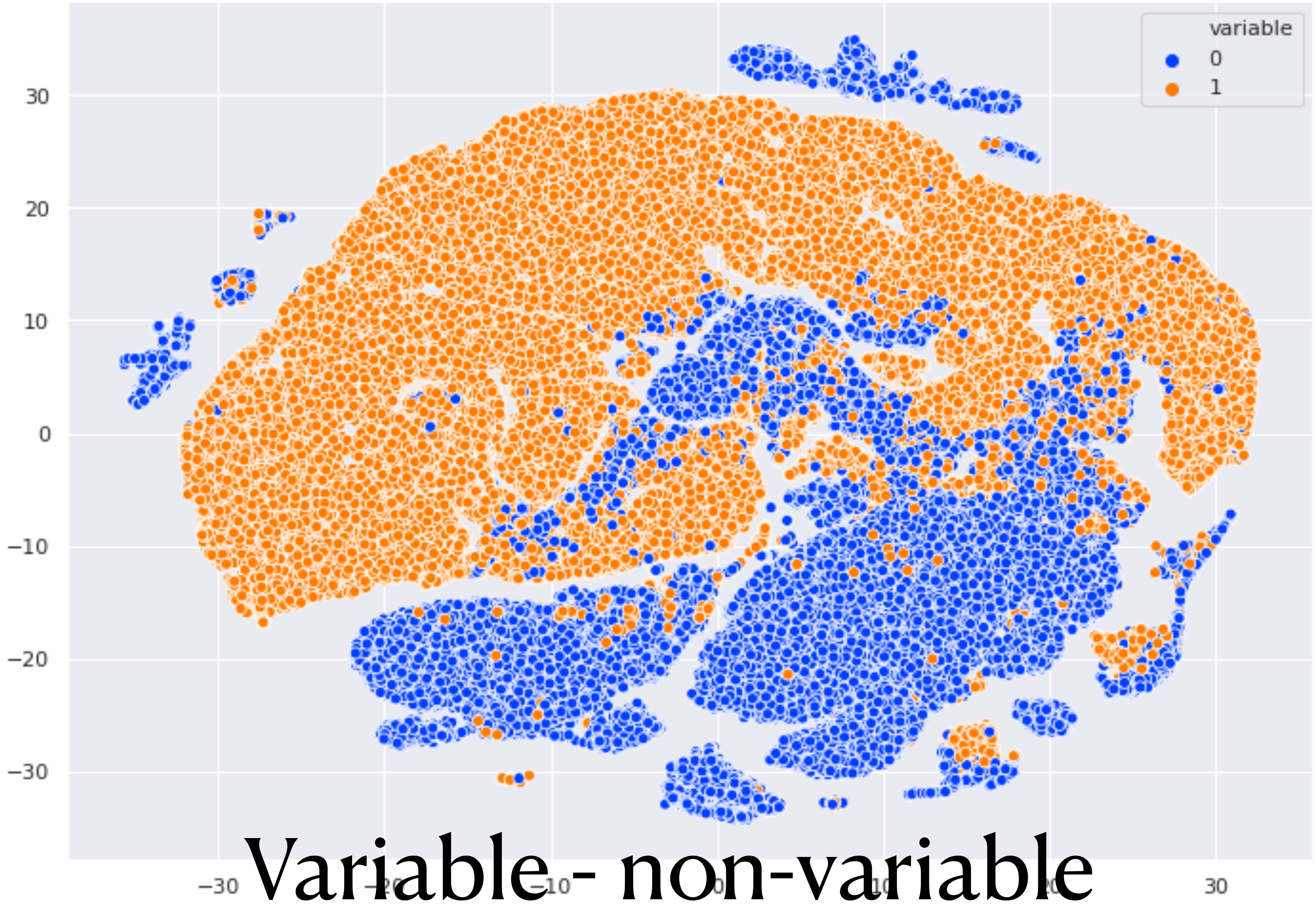
Classification of variables and transients

100K alerts a night - published in real-time



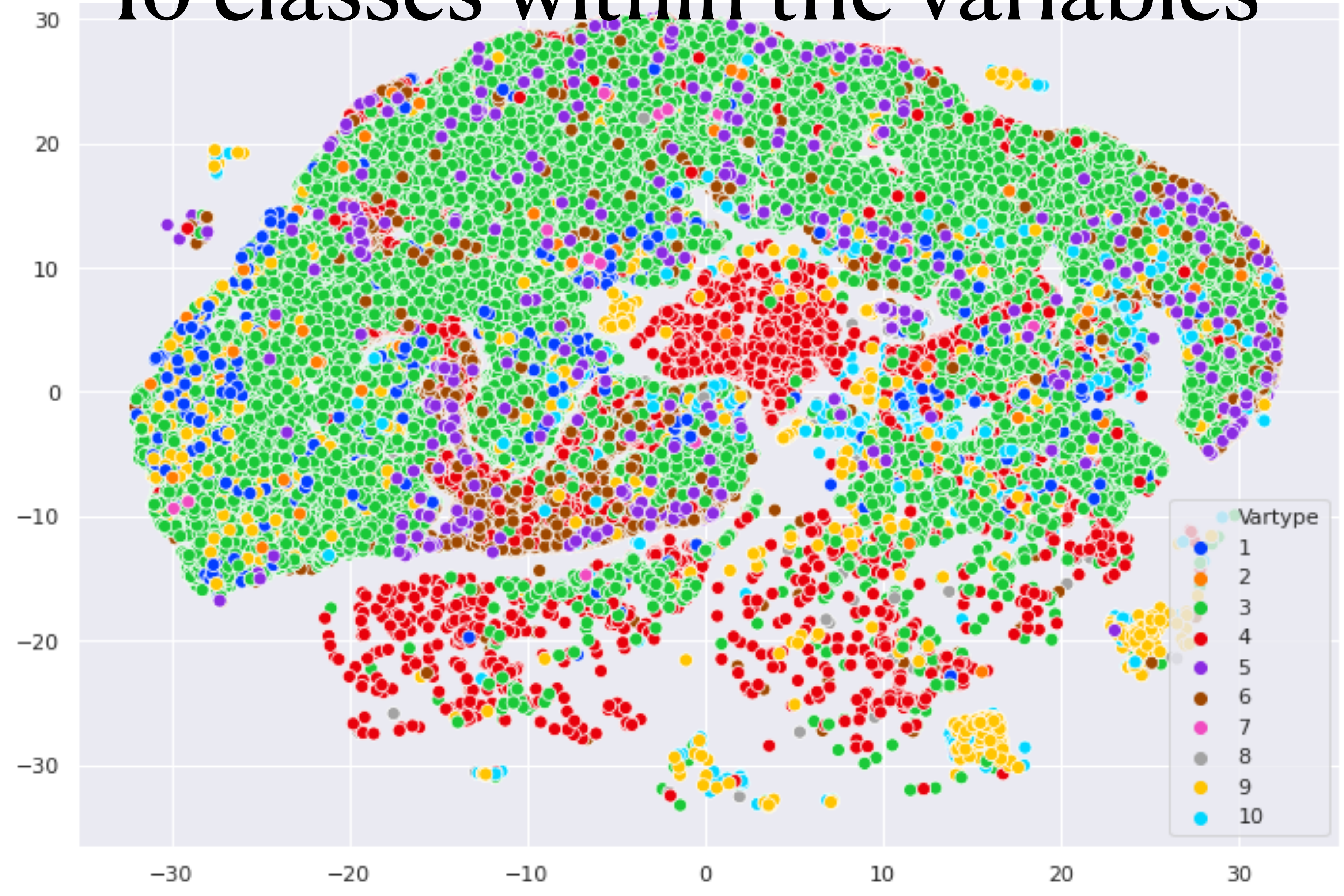
Next campaign: Classifying anomalies

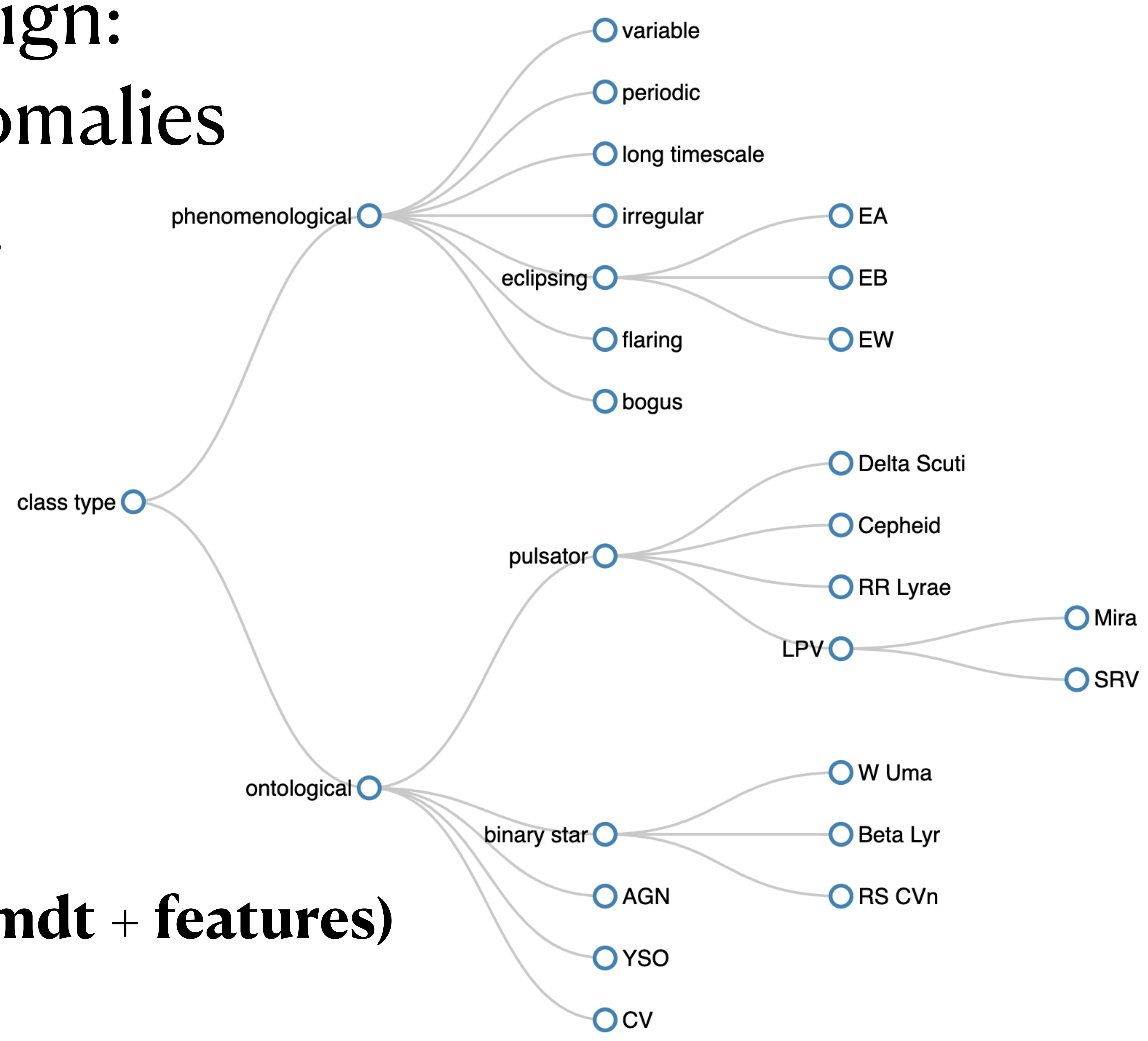
Van Roestel 2020



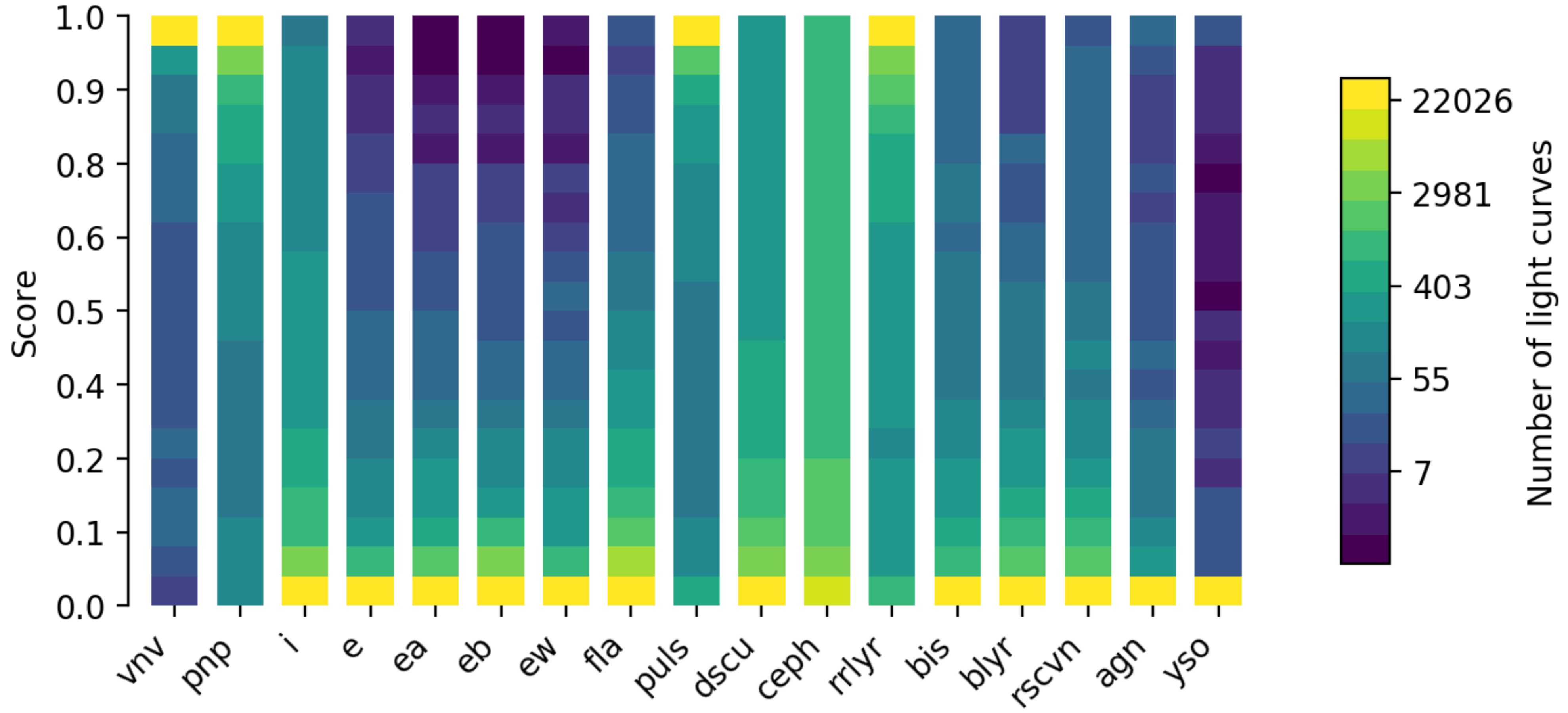
Variable - non-variable

10 classes within the variables





XGBoost, DNN (dmdt + features)



Kushal Tirumala

Internal project

Streaking asteroids

Type	Count
Masked bright star	1526
Skip (Includes 'Not Sure' and seemingly 'Blank Images’)	1051
Yin–Yang (multiple badly subtracted stars)	437
Dementors and ghosts	207
Satellite (long streak – could be partially masked)	89
Satellite flashes	32
Cosmic rays	28
Asteroid (short streak)	3
Plausible Asteroid (short streak)	3
Naked stars	1

ZTF DeepStreaks

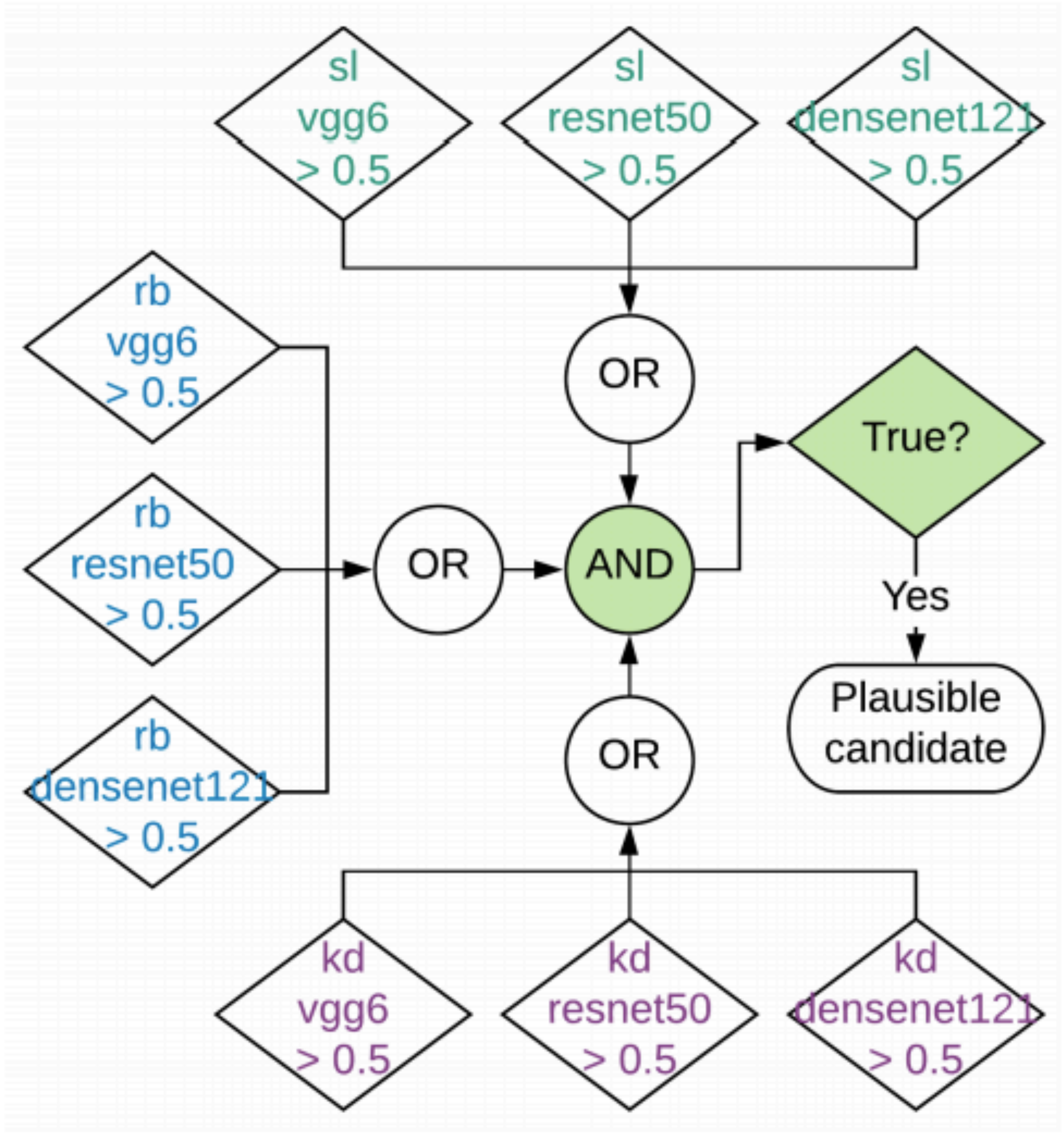
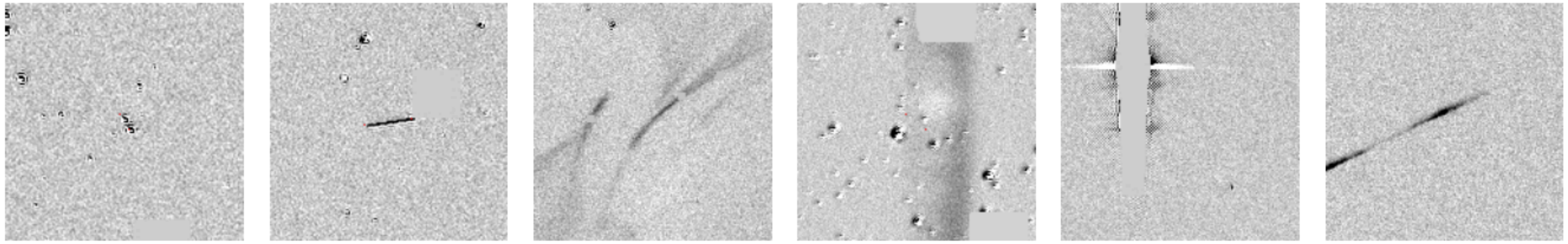
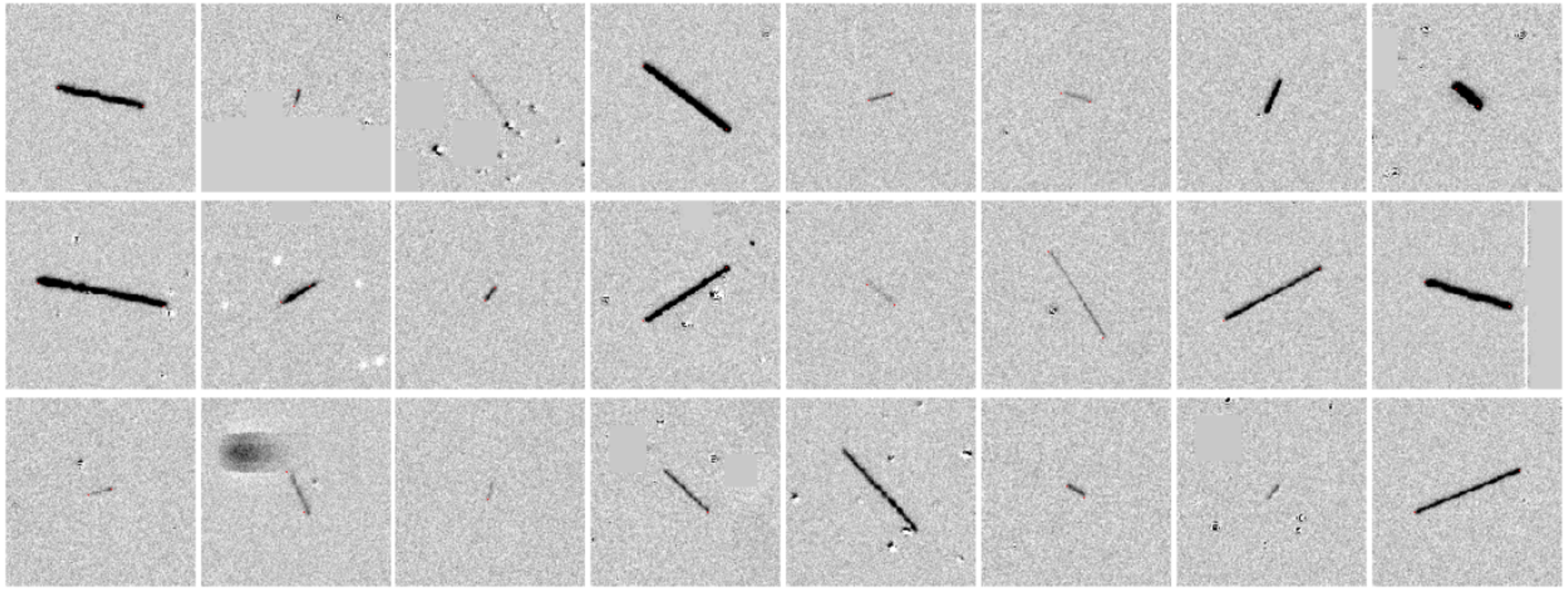


Figure 2. Decision logic used by **DeepStreaks** to identify plausible streaks. The problem is split into three simpler sub-problems, each solved by a dedicated group of classifiers assigning real vs. bogus ("rb"), short vs. long ("sl"), and keep vs. ditch ("kd") scores. At least one member of each group must output a score that passes a pre-defined threshold. See Section 2.1 for details.



(a) Bad subtraction (b) Cosmic ray (c) "Dementor" (d) "Ghost" (e) Masked star (f) Satellite trail



CNNs

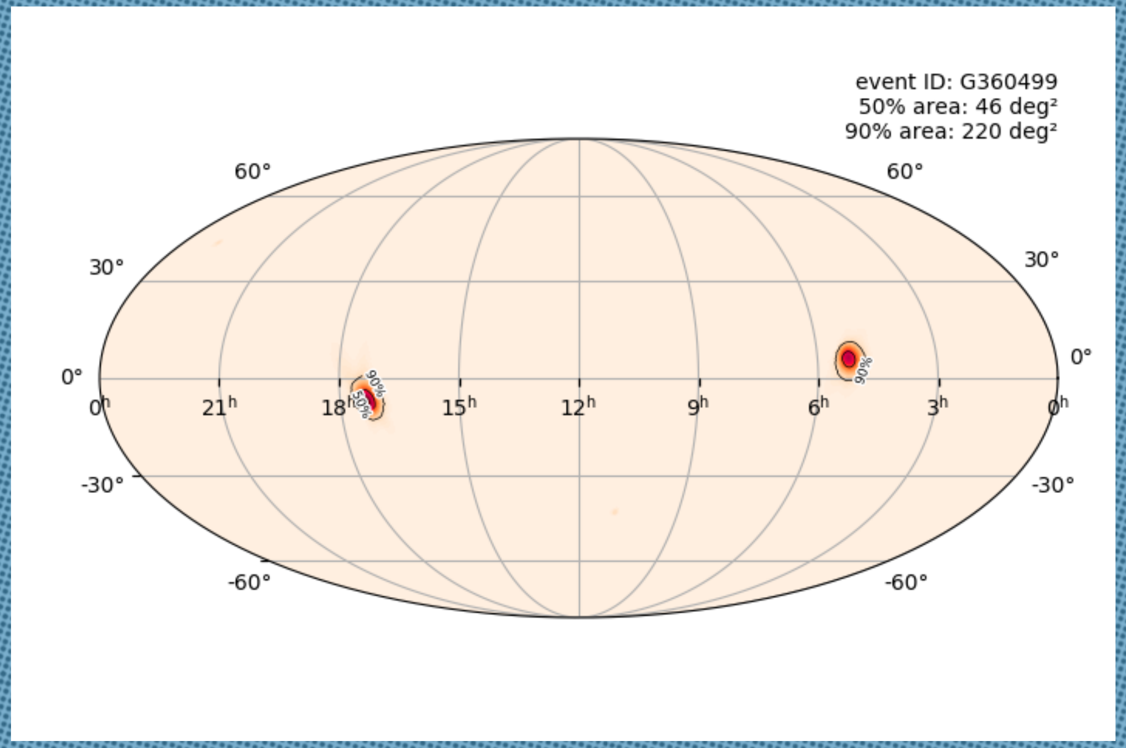
Duev, Mahabal, … arXiv:1904.05920

Multi Messenger Astronomy and GW follow-up

Several GraceDB alerts were later retracted

S200116ah: >99% NSBH

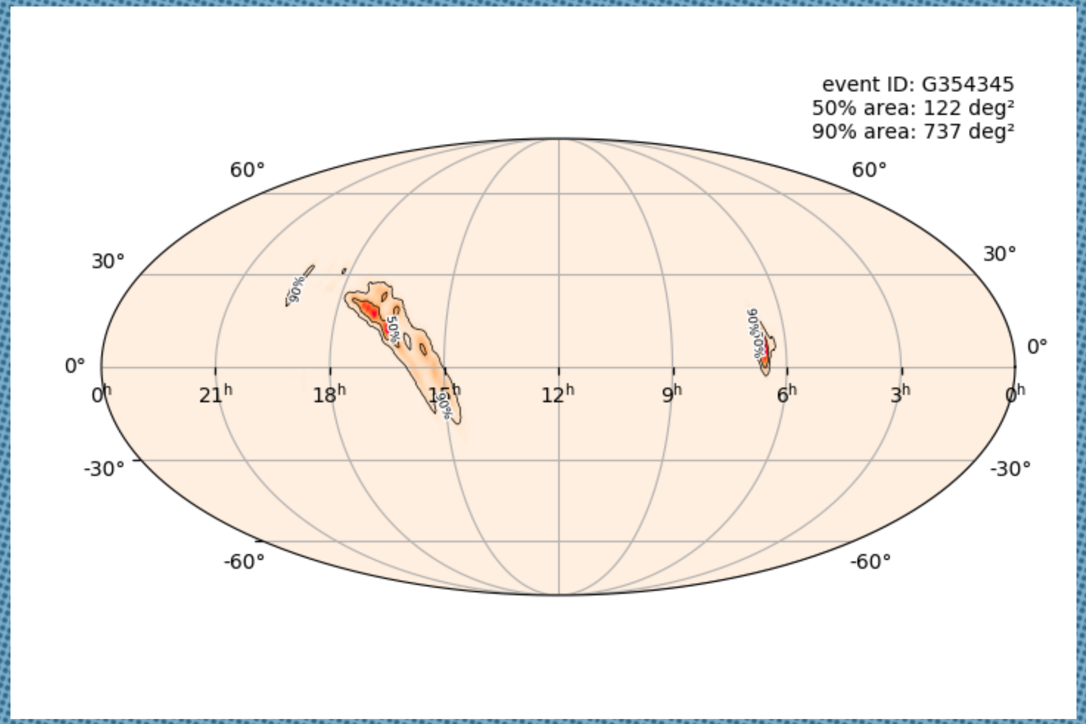
FAR: 1 per 15618 yrs



- 1st followup: 85 s
- Retraction: 19 min

S191110x: >99% mass gap

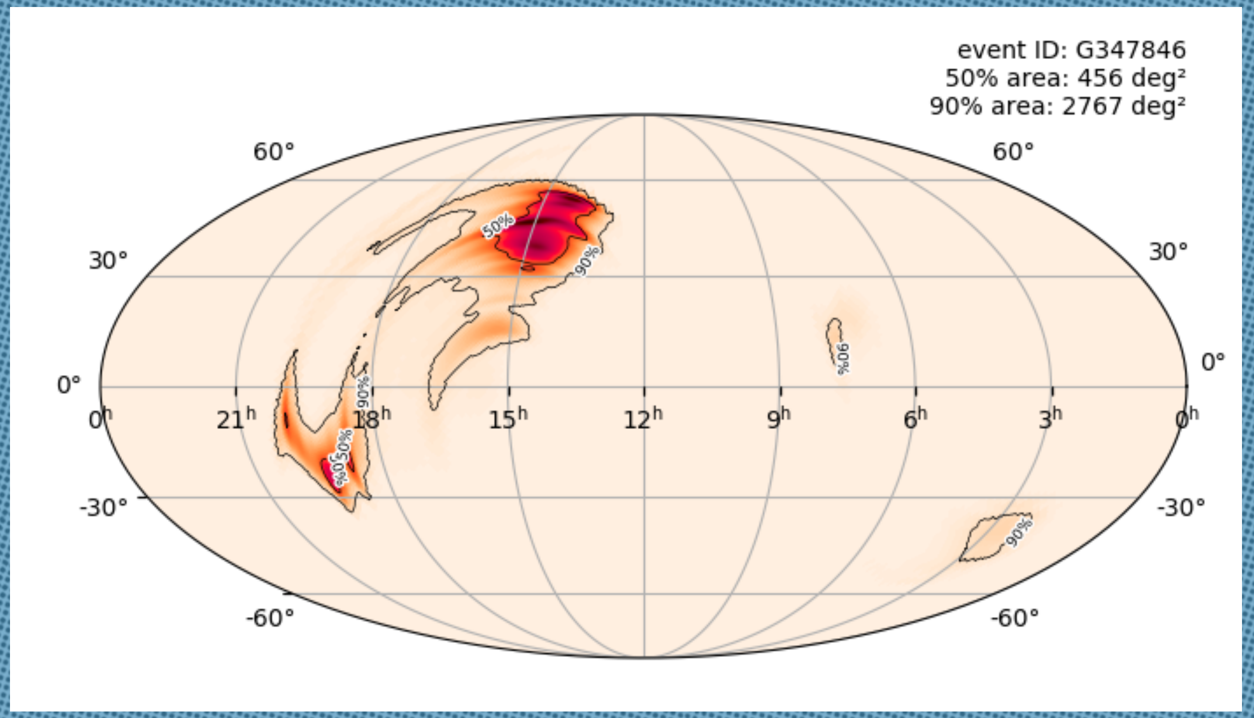
FAR: 1 per 1082 yrs



- 1st followup: 10 mins
- Retraction: 24 mins

S190822c: >99% BNS

FAR: 1 per 5e9 yrs

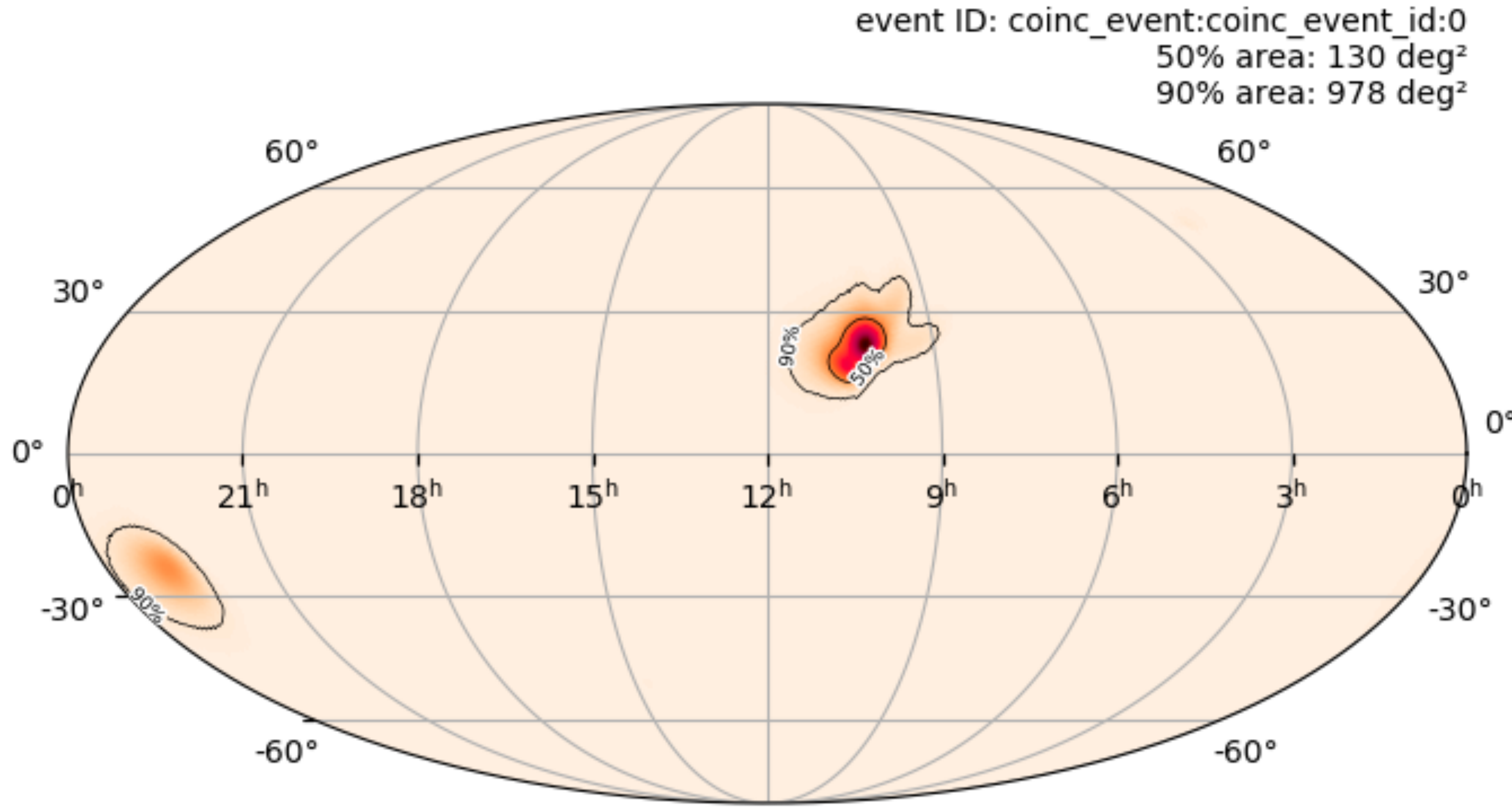


- 1st followup: 23 mins
- Retraction: 1.5 hours

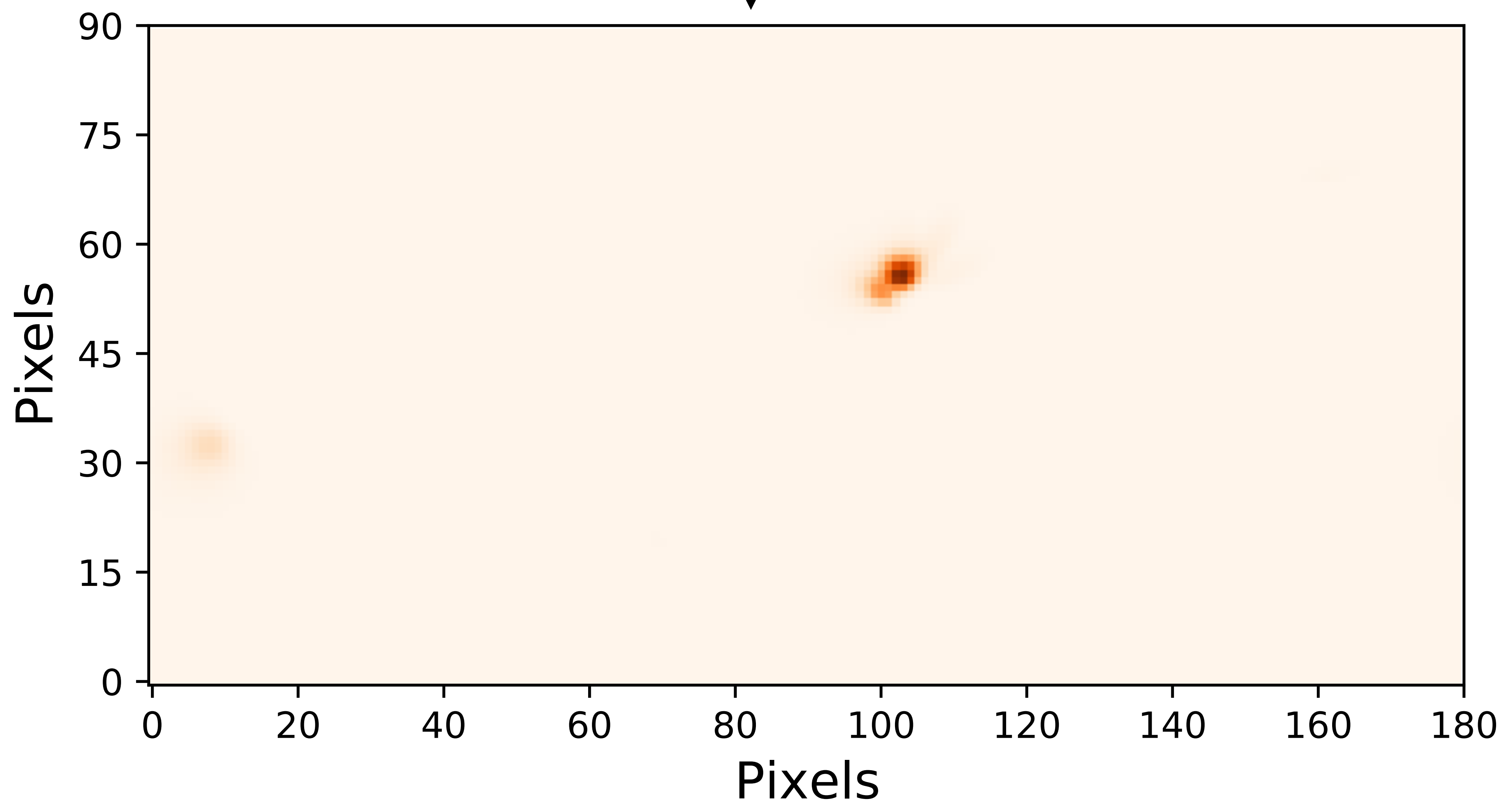
End of O3 GraceDB stats

Candidates: 56
Retracted: 24

Can we classify these based on skymaps?



Reproject and Bin



The data

NS $m \in [1.2, 2.2] M_\odot\,;\ \chi \leq 0.05$

BH $m \in [3, 100] M_\odot\,;\ \chi \leq 0.99$

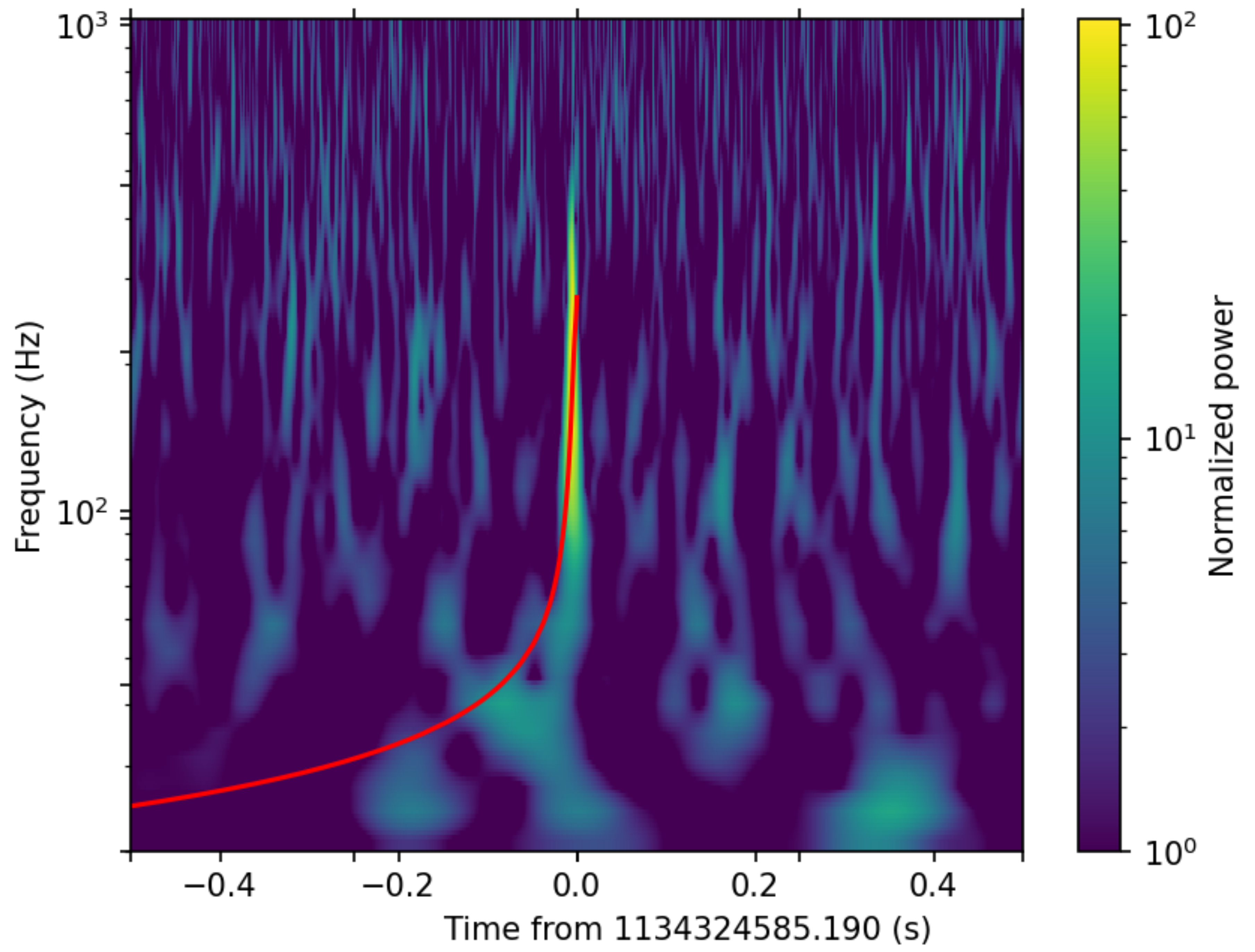
	O1	O2a	O2b
Network	H1, L1	H1, L1	H1, L1, V1 (4 options)
PSD	GW151012	GW170104	GW170729

Single detector SNR ≥ 4.5
Network SNR ≥ 7

2857 examples

Bogus set: glitches

Combine glitches with 2-OGC triggers



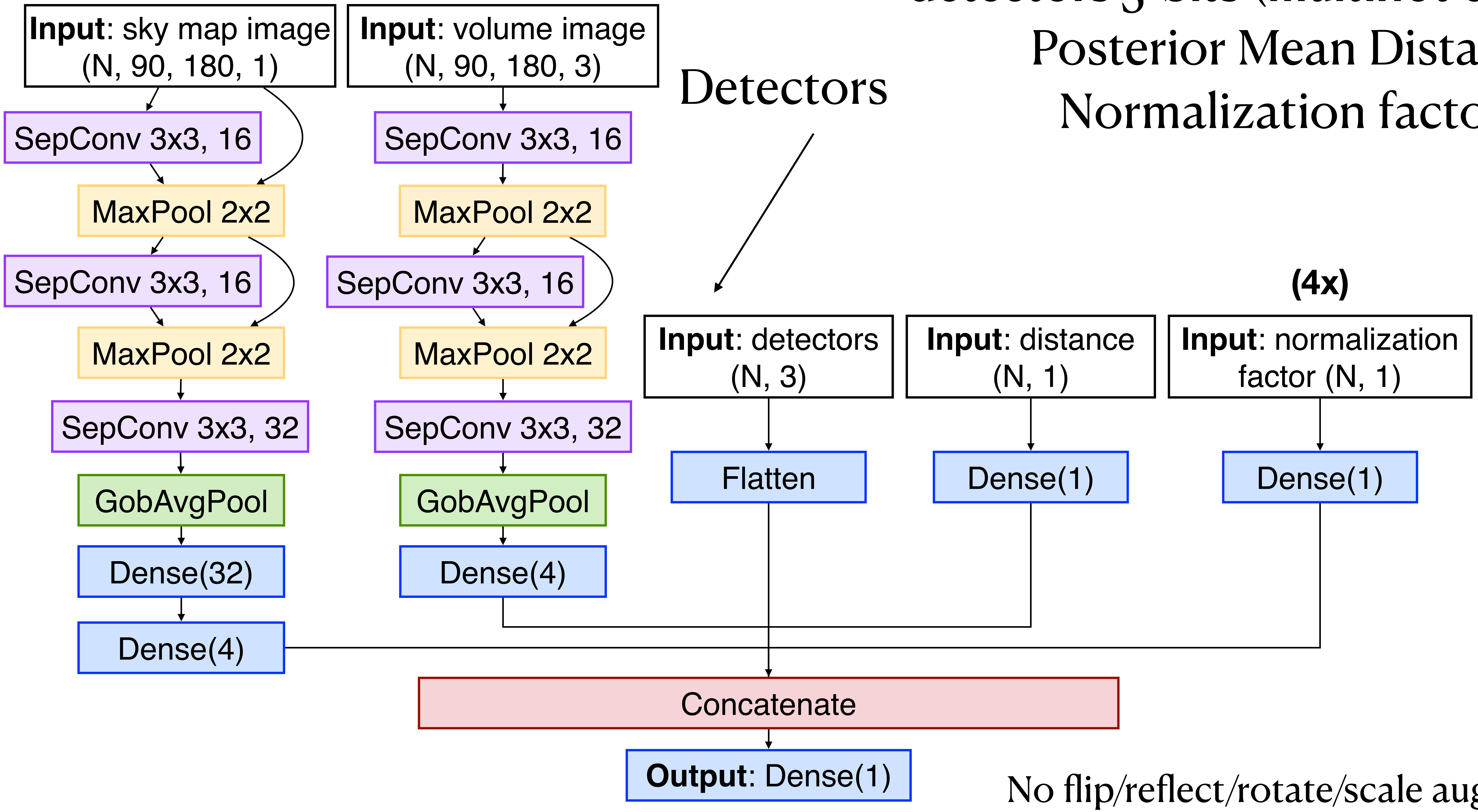
1267 examples

The model

Sky map: 90 x 180
Stacked volume: 90 x 180 x 3
detectors 3-bits (multihot encoding)
Posterior Mean Distance
Normalization factors

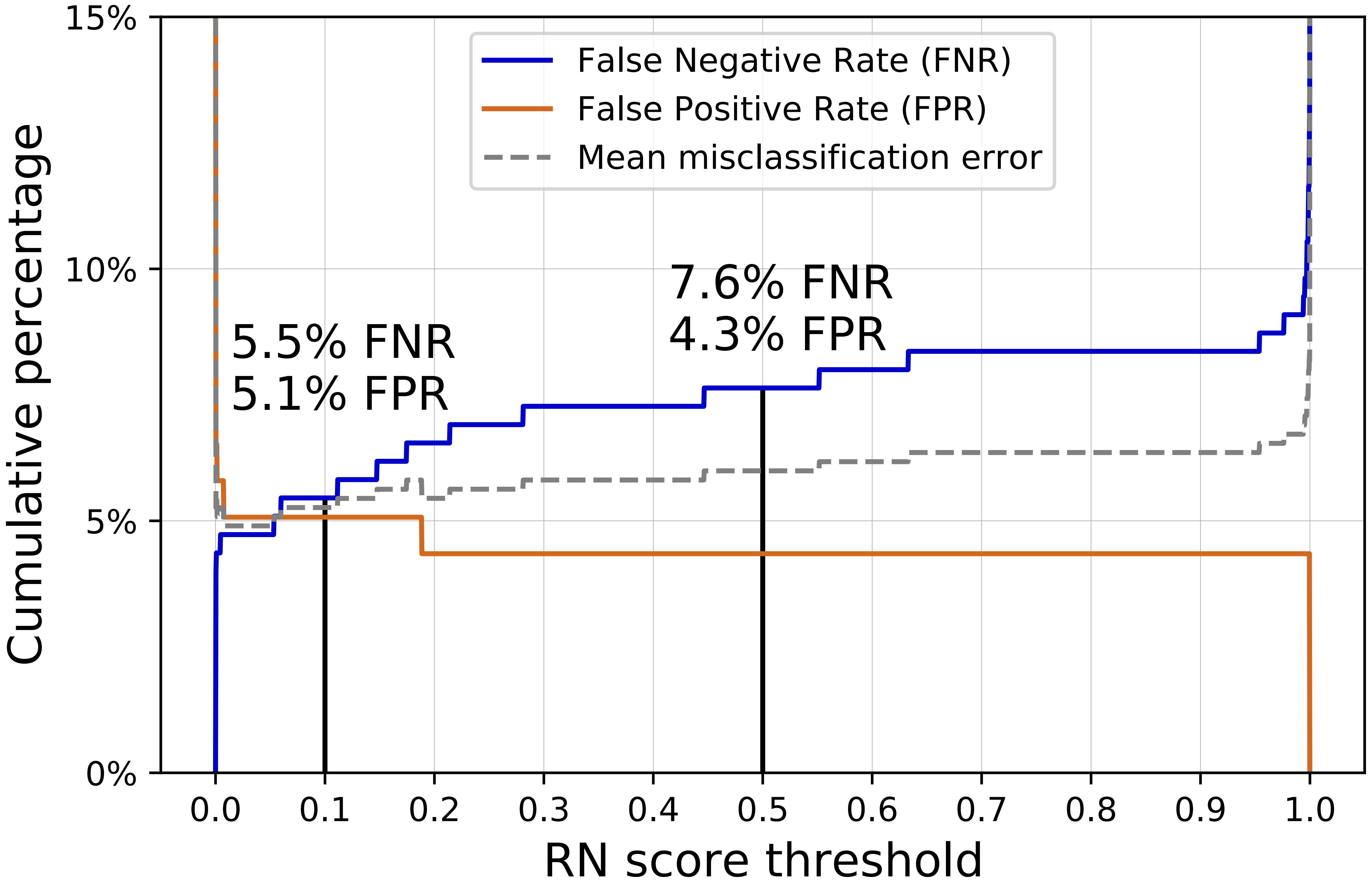
SkyMap

Detectors



No flip/reflect/rotate/scale augmentations

GW Sky Map Network Performance



$$A = \frac{TP + TN}{TP + FP + TN + FN}$$

Total predictions

How often is the model correct?

$$P = \frac{TP}{TP + FP}$$

Predicted positives

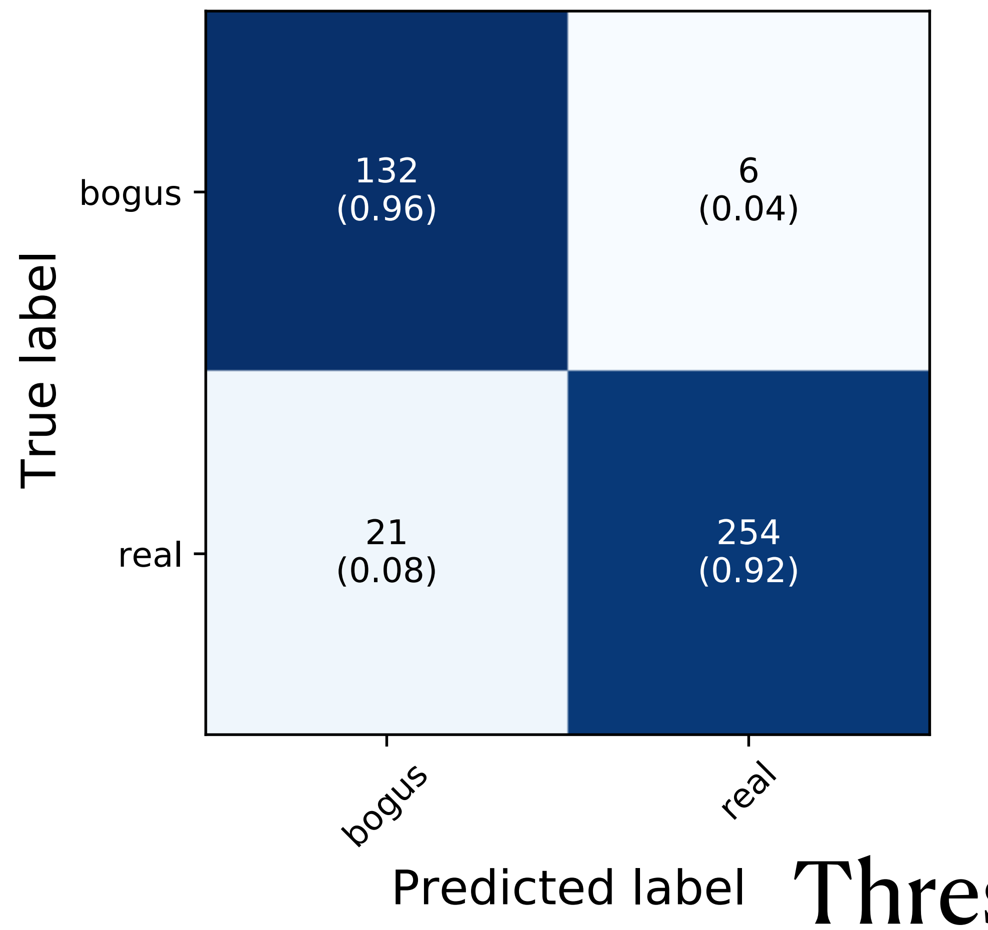
How often is the model correct when it predicts that the candidate is real?

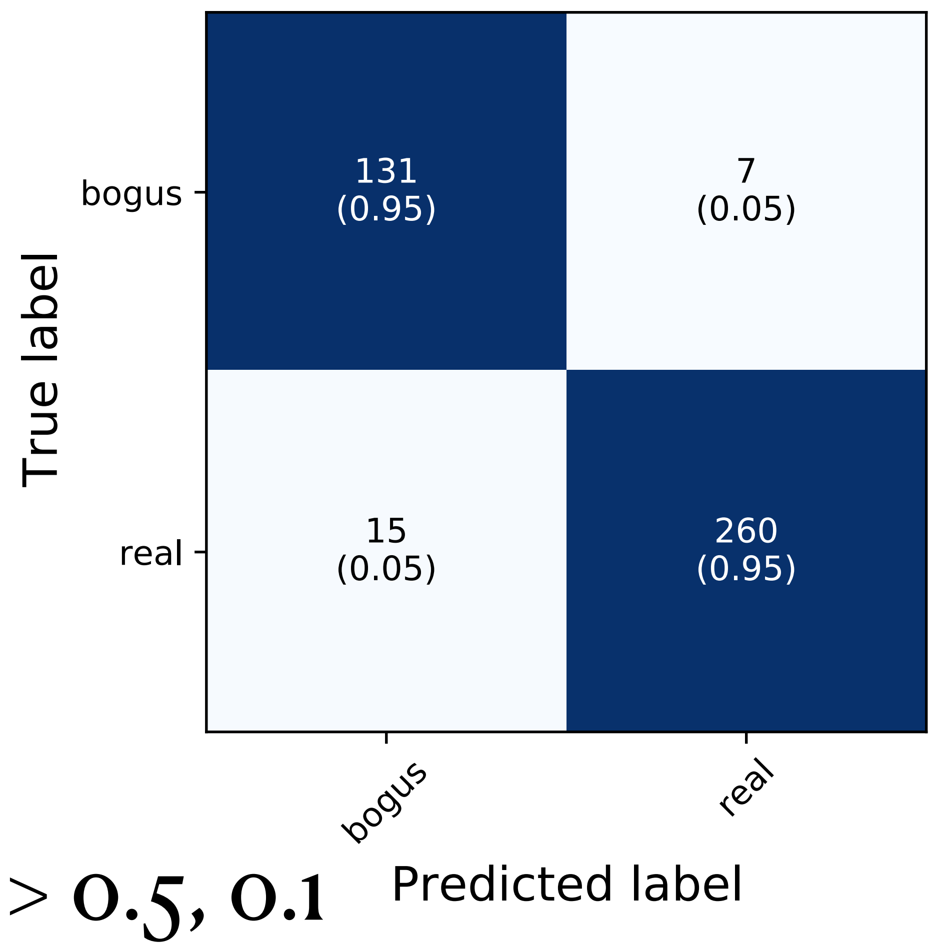
$$R = \frac{TP}{TP + FN}$$

Actual positives

How many real candidates are predicted correctly?

Accuracy: 93.5%
Precision: 97.7%
Recall: 92.4%





Threshold > 0.5, 0.1

arXiv:2010.11829 Cabero, Mahabal, McIver

Predictions (Threshold > 0.5)

O3 GraceDB events from CBC pipelines

Prediction	Real	Bogus
4 published GW	4	-
22 retracted events	12	10
51 non-published GW candidates	41	10

Misclassifications: Many were identified by the GstLal and MBTAOnline pipelines, training noise examples are from PyCBC

70% retracted had Virgo data

GWTC-2
the O3a catalog

4 Published GW
13 new
22 confirmed
7 retracted

GWSkyNet verdict:
23 real
6 non-astrophysical

28/29 correct
no reals called bogus

Name	Inst.	cWB		GstLAL			PyCBC			PyCBC BBH		
		FAR (yr^{-1})	SNR*	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR*	p_{astro}	FAR (yr^{-1})	SNR*	p_{astro}
GW190408_181802	HLV	$< 9.5 \times 10^{-4}$	14.8	$< 1.0 \times 10^{-5}$	14.7	1.00	$< 2.5 \times 10^{-5}$	13.5	1.00	$< 7.9 \times 10^{-5}$	13.6	1.00
GW190412	HLV	$< 9.5 \times 10^{-4}$	19.7	$< 1.0 \times 10^{-5}$	18.9	1.00	$< 3.1 \times 10^{-5}$	17.9	1.00	$< 7.9 \times 10^{-5}$	17.8	1.00
GW190413_052954	HLV	–	–	–	–	–	–	–	–	7.2×10^{-2}	8.6	0.98
GW190413_134308	HLV	–	–	3.8×10^{-1}	10.0	0.95	–	–	–	4.4×10^{-2}	9.0	0.98
GW190421_213856	HL	3.0×10^{-1}	9.3	7.7×10^{-4}	10.6	1.00	1.9×10^{0}	10.2	0.89	6.6×10^{-3}	10.2	1.00
GW190424_180648	L			$7.8 \times 10^{-1\dagger}$	10.0	0.91						
GW190425	LV			$7.5 \times 10^{-4\dagger}$	13.0	–						
GW190426_152155	HLV	–	–	1.4×10^{0}	10.1	–	–	–	–	–	–	–
GW190503_185404	HLV	1.8×10^{-3}	11.5	$< 1.0 \times 10^{-5}$	12.1	1.00	3.7×10^{-2}	12.2	1.00	$< 7.9 \times 10^{-5}$	12.2	1.00
GW190512_180714	HLV	8.8×10^{-1}	10.7	$< 1.0 \times 10^{-5}$	12.3	1.00	3.8×10^{-5}	12.2	1.00	$< 5.7 \times 10^{-5}$	12.2	1.00
GW190513_205428	HLV	–	–	$< 1.0 \times 10^{-5}$	12.3	1.00	3.7×10^{-4}	11.8	1.00	$< 5.7 \times 10^{-5}$	11.9	1.00
GW190514_065416	HL	–	–	–	–	–	–	–	–	5.3×10^{-1}	8.3	0.96
GW190517_055101	HLV	6.5×10^{-3}	10.7	9.6×10^{-4}	10.6	1.00	1.8×10^{-2}	10.4	1.00	$< 5.7 \times 10^{-5}$	10.2	1.00
GW190519_153544	HLV	3.1×10^{-4}	14.0	$< 1.0 \times 10^{-5}$	12.0	1.00	$< 1.8 \times 10^{-5}$	13.0	1.00	$< 5.7 \times 10^{-5}$	13.0	1.00
GW190521	HLV	2.0×10^{-4}	14.4	1.2×10^{-3}	14.7	1.00	1.1×10^{0}	12.6	0.93	–	–	–
GW190521_074359	HL	$< 1.0 \times 10^{-4}$	24.7	$< 1.0 \times 10^{-5}$	24.4	1.00	$< 1.8 \times 10^{-5}$	24.0	1.00	$< 5.7 \times 10^{-5}$	24.0	1.00
GW190527_092055	HL	–	–	6.2×10^{-2}	8.9	0.99	–	–	–	–	–	–
GW190602_175927	HLV	1.5×10^{-2}	11.1	1.1×10^{-5}	12.1	1.00	–	–	–	1.5×10^{-2}	11.4	1.00
GW190620_030421	LV			$2.9 \times 10^{-3\dagger}$	10.9	1.00						
GW190630_185205	LV			$< 1.0 \times 10^{-5}$	15.6	1.00						
GW190701_203306	HLV	5.5×10^{-1}	10.2	1.1×10^{-2}	11.6	1.00	–	–	–	–	–	–
GW190706_222641	HLV	$< 1.0 \times 10^{-3}$	12.7	$< 1.0 \times 10^{-5}$	12.3	1.00	6.7×10^{-5}	11.7	1.00	$< 4.6 \times 10^{-5}$	12.3	1.00
GW190707_093326	HL	–	–	$< 1.0 \times 10^{-5}$	13.0	1.00	$< 1.0 \times 10^{-5}$	12.8	1.00	$< 4.6 \times 10^{-5}$	12.8	1.00
GW190708_232457	LV			$2.8 \times 10^{-5\dagger}$	13.1	1.00						
GW190719_215514	HL	–	–	–	–	–	–	–	–	1.6×10^{0}	8.0	0.82
GW190720_000836	HLV			$< 1.0 \times 10^{-5}$	11.7	1.00	$< 2.0 \times 10^{-5}$	10.6	1.00	$< 3.7 \times 10^{-5}$	10.5	1.00
GW190727_060333	HLV	8.8×10^{-2}	11.4	$< 1.0 \times 10^{-5}$	12.3	1.00	3.5×10^{-3}	11.5	1.00	$< 3.7 \times 10^{-5}$	11.8	1.00
GW190728_064510	HLV	–	–	$< 1.0 \times 10^{-5}$	13.6	1.00	$< 1.6 \times 10^{-5}$	13.4	1.00	$< 3.7 \times 10^{-5}$	13.4	1.00
GW190731_140936	HL	–	–	2.1×10^{-1}	8.5	0.97	–	–	–	2.8×10^{-1}	8.2	0.96
GW190803_022701	HLV	–	–	3.2×10^{-2}	9.0	0.99	–	–	–	2.7×10^{-2}	8.6	0.99
GW190814	LV			$< 1.0 \times 10^{-5}$	22.2	1.00						
GW190828_063405	HLV	$< 9.6 \times 10^{-4}$	16.6	$< 1.0 \times 10^{-5}$	16.0	1.00	$< 1.5 \times 10^{-5}$	15.3	1.00	$< 3.3 \times 10^{-5}$	15.3	1.00
GW190828_065509	HLV	–	–	$< 1.0 \times 10^{-5}$	11.1	1.00	5.8×10^{-5}	10.8	1.00	$< 3.3 \times 10^{-5}$	10.8	1.00
GW190909_114149	HL	–	–	1.1×10^{0}	8.5	0.89	–	–	–	–	–	–
GW190910_112807	LV			$1.9 \times 10^{-5\dagger}$	13.4	1.00						
GW190915_235702	HLV	$< 1.0 \times 10^{-3}$	12.3	$< 1.0 \times 10^{-5}$	13.1	1.00	8.6×10^{-4}	13.0	1.00	$< 3.3 \times 10^{-5}$	12.7	1.00
GW190924_021846	HLV	–	–	$< 1.0 \times 10^{-5}$	13.2	1.00	$< 6.3 \times 10^{-5}$	12.5	1.00	$< 3.3 \times 10^{-5}$	12.4	1.00
GW190929_012149	HLV	–	–	2.0×10^{-2}	9.9	1.00	–	–	–	–	–	–
GW190930_133541	HL	–	–	5.8×10^{-1}	10.0	0.92	3.4×10^{-2}	9.7	1.00	3.3×10^{-2}	9.8	0.99

TABLE IV. Gravitational wave candidate event list. We find 39 candidate events passing the FAR threshold of 2.0 yr^{-1} in at

Table 4

Bogus:
S190510g
S190718y
S190901ap
S190910d
S190910h
S190930t

Missed:
S190923y

Several brokers

LCO MARS
Make Alerts Really Simple

Fink

Lasair

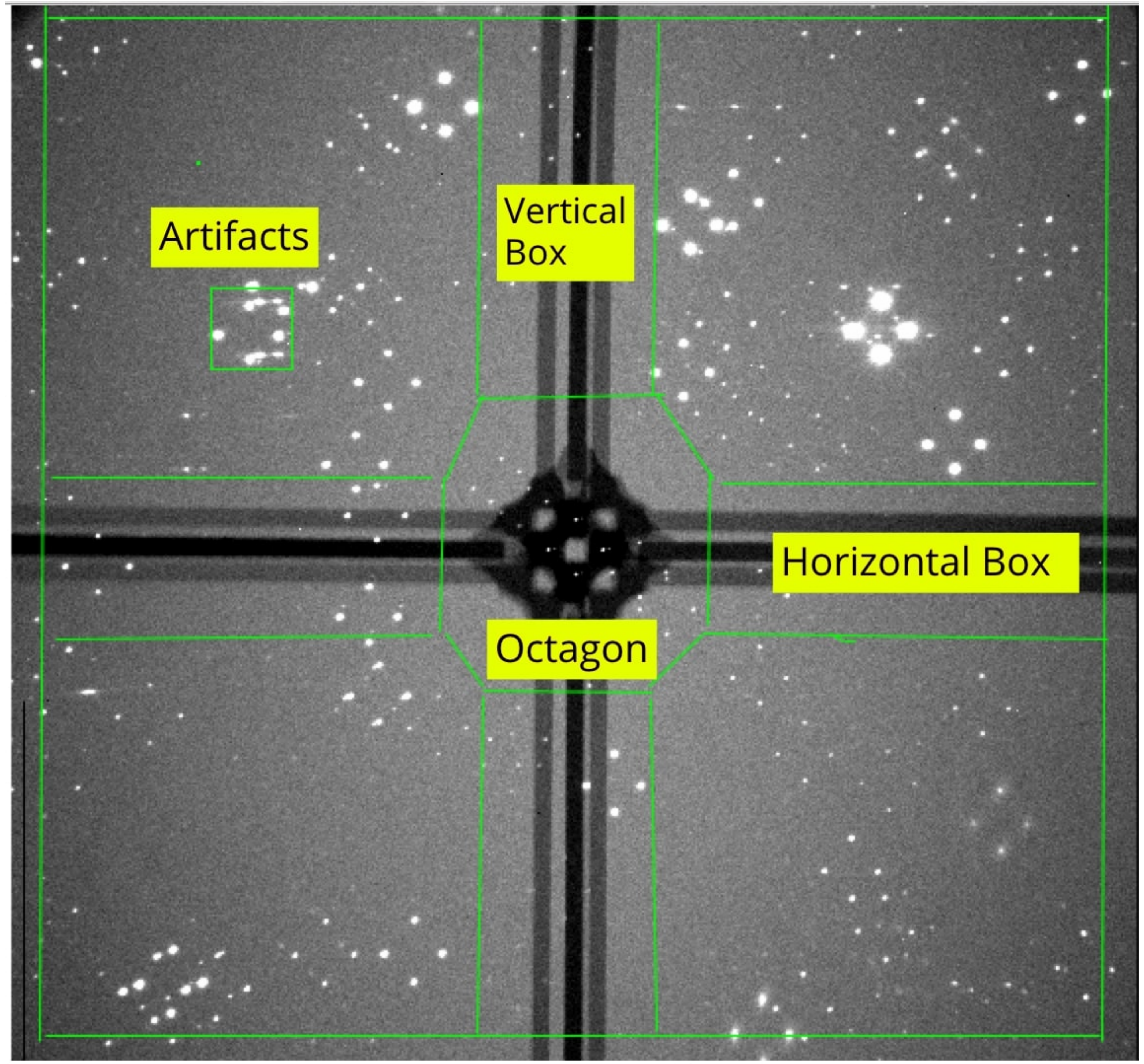
Getting ready for LSST/VRO with ZTF data

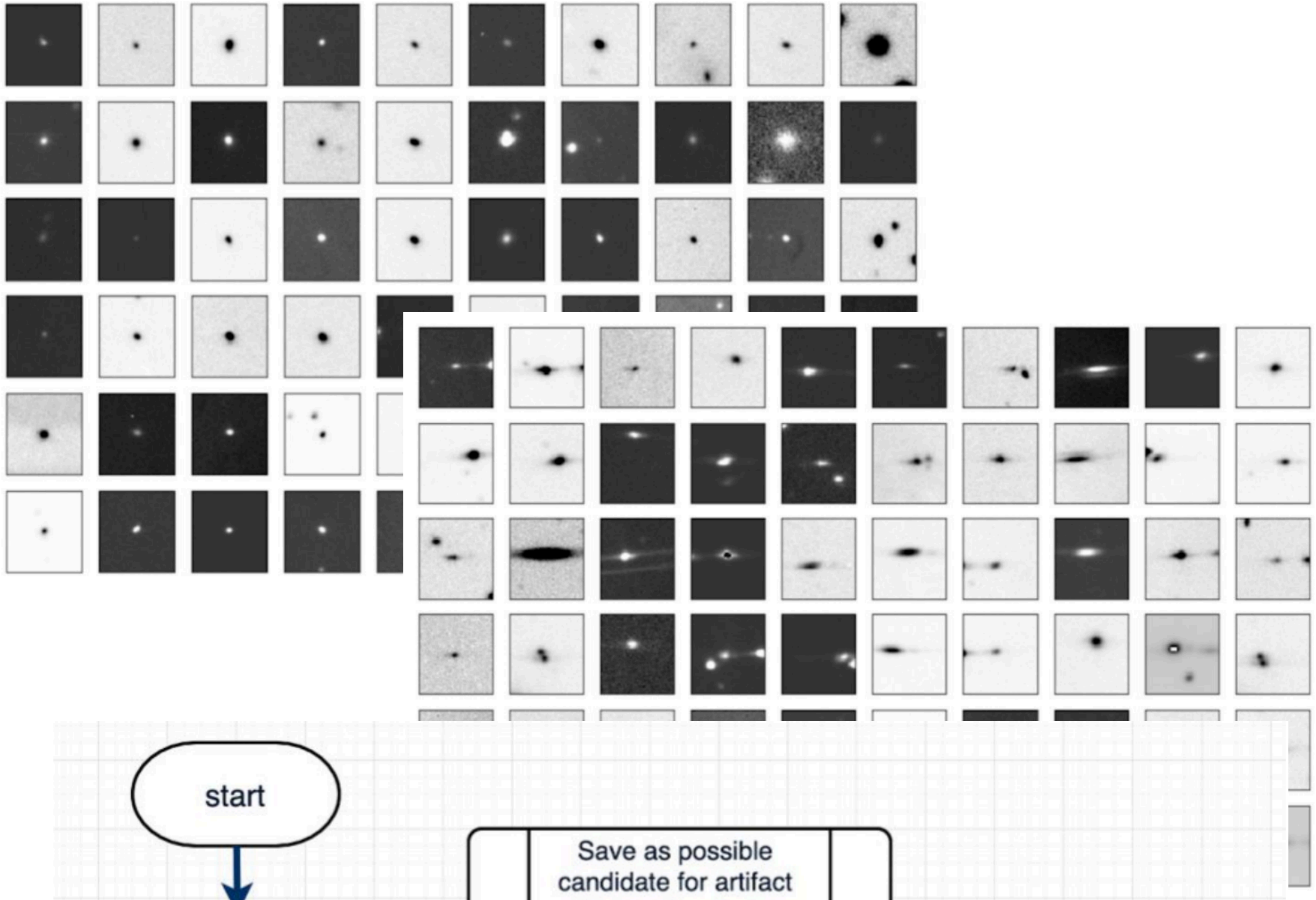
ANTARES beta

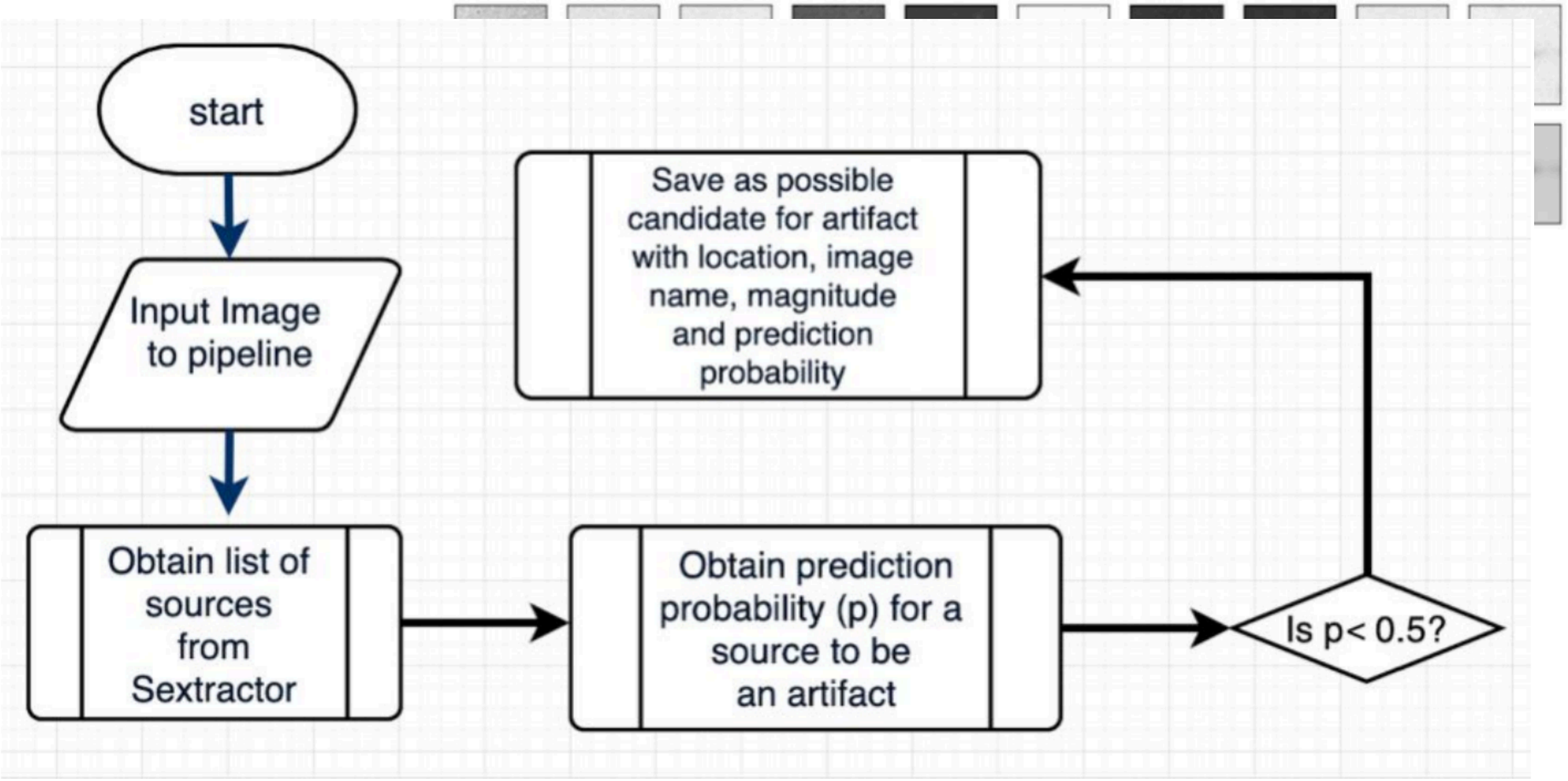
ALeRCE
Automatic Learning for the Rapid Classification of Events

- Cross-matches
- Value additions
- Classifications

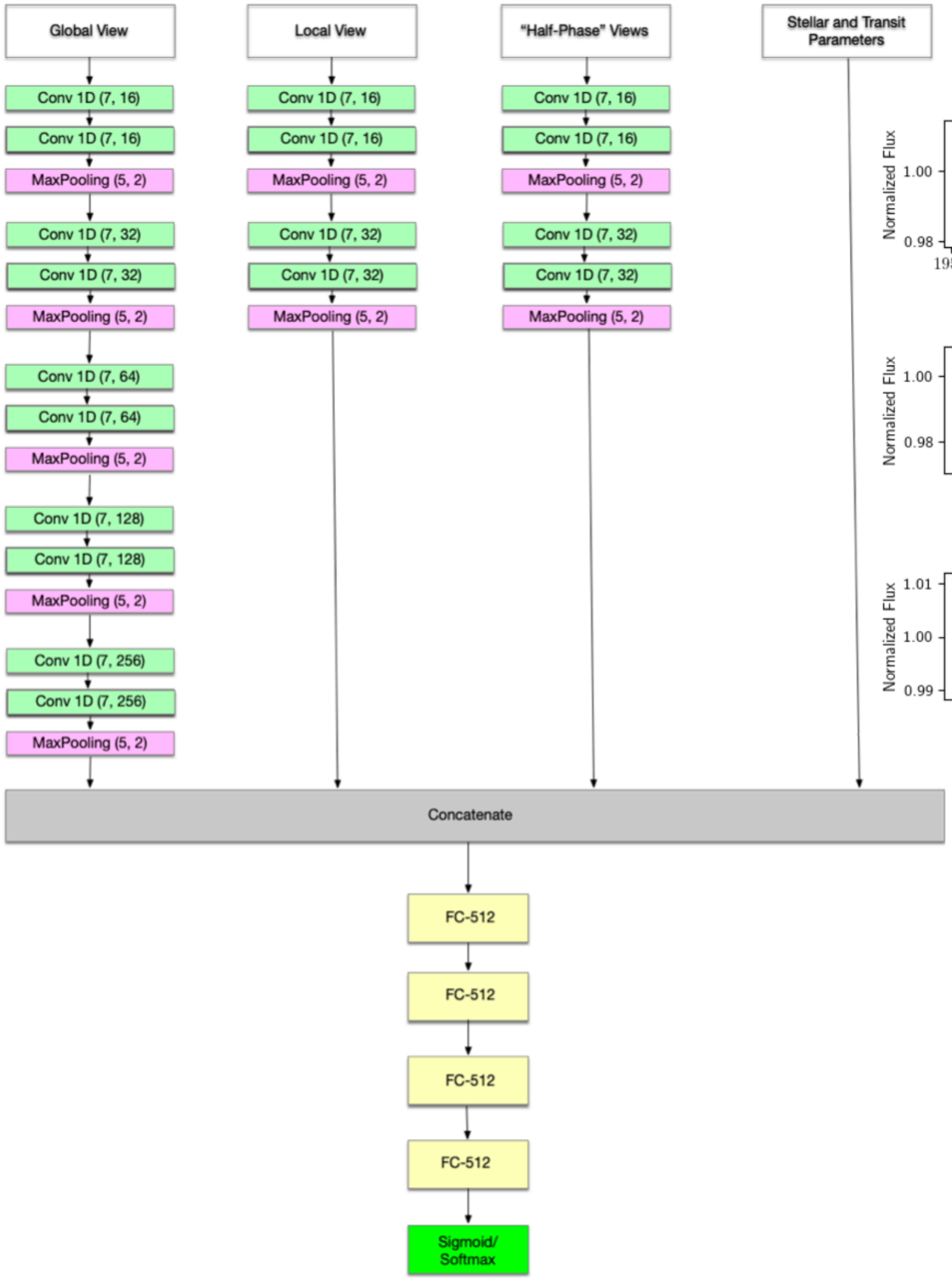
Robopol



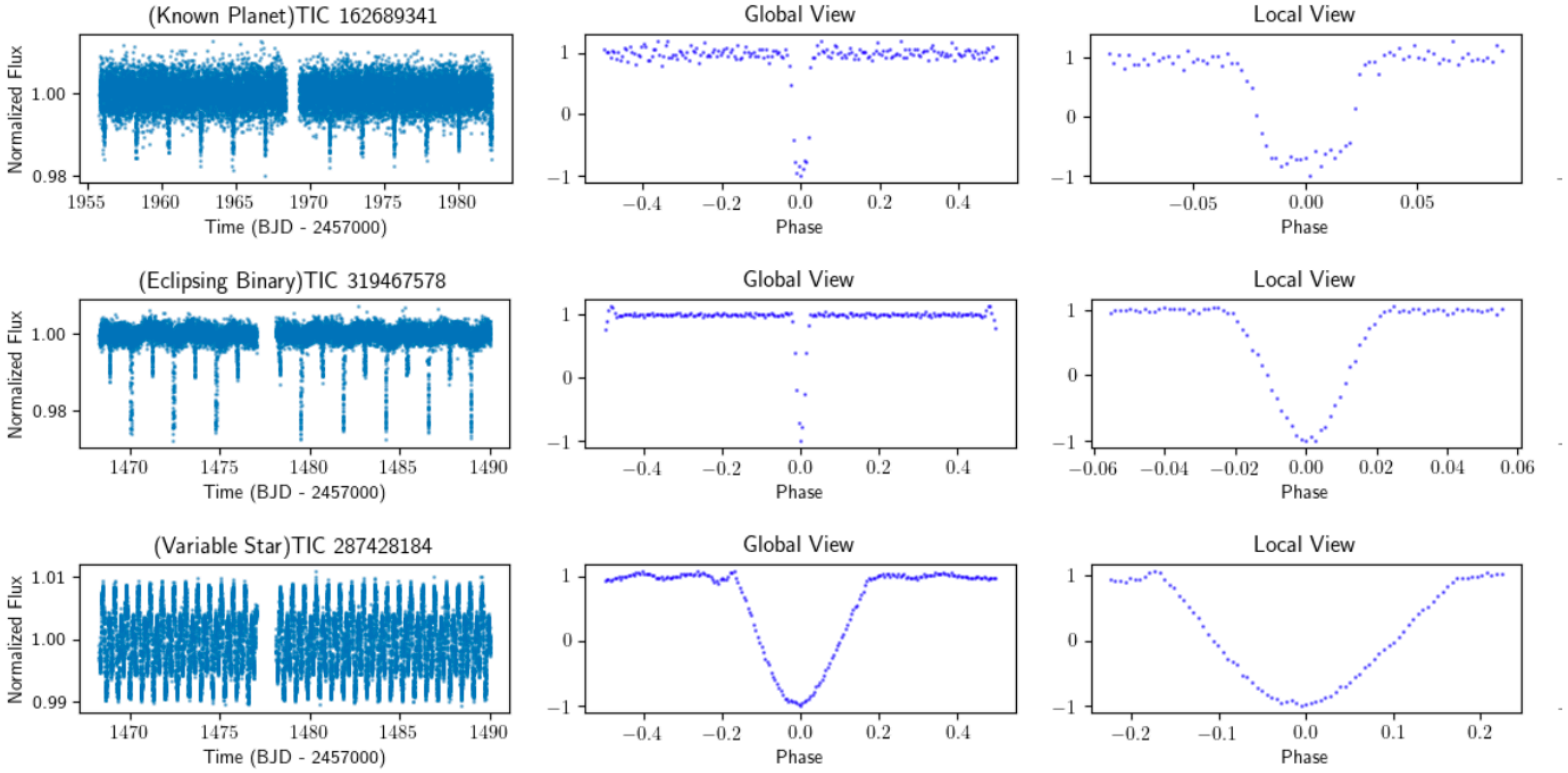




Paranjpye, Mahabal et al. arXiv:1911.08327

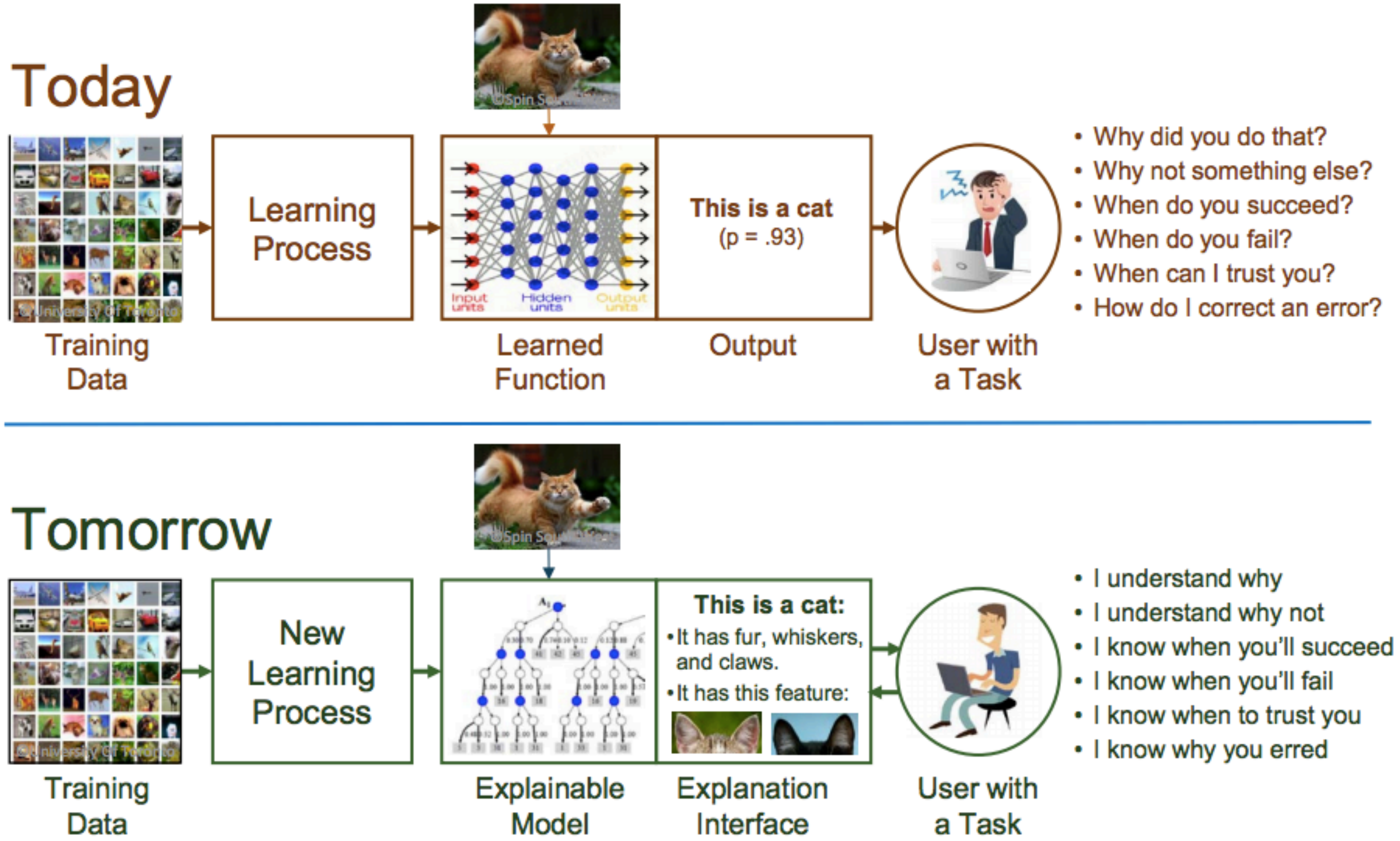


Metadata



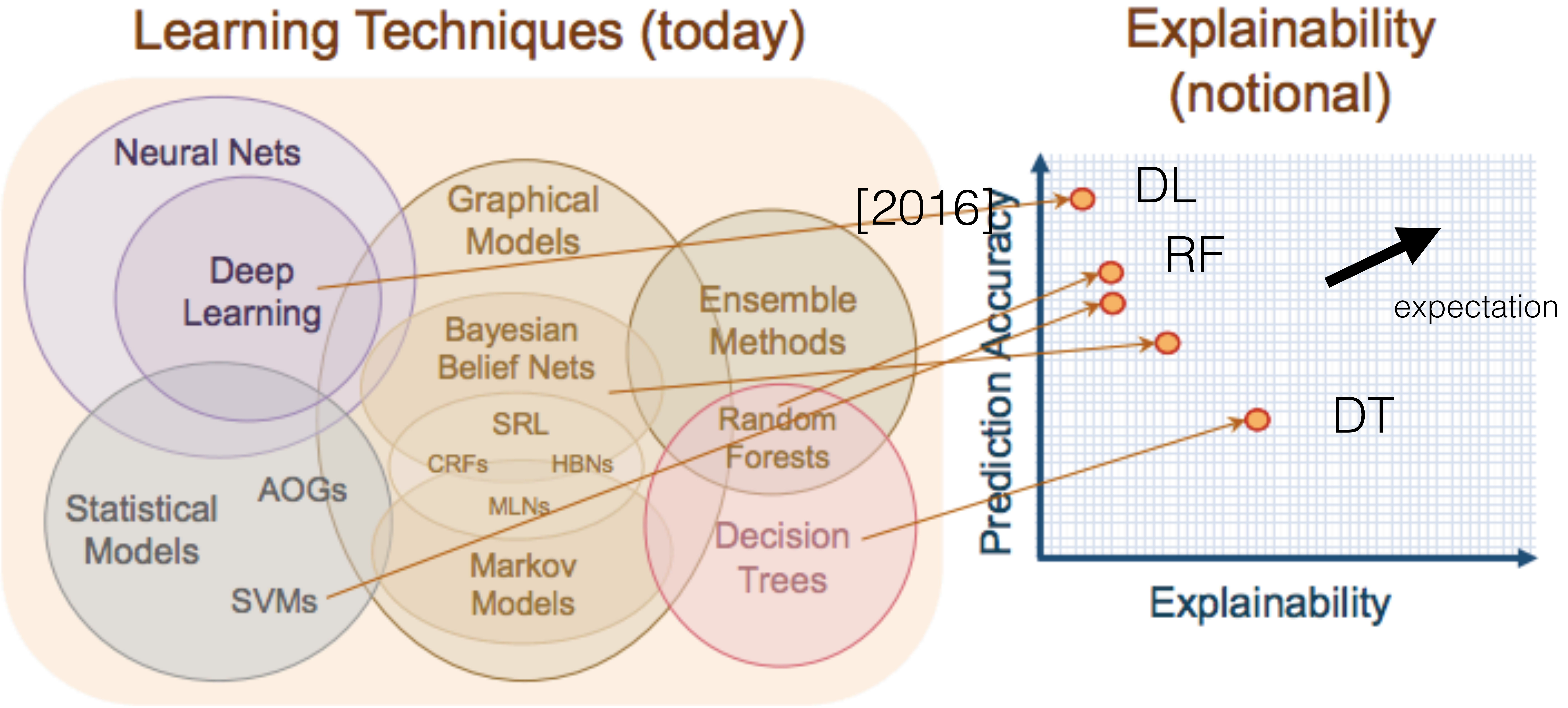
Hunting for Exoplanets
S Rao, Mahabal, ...

Interpretability



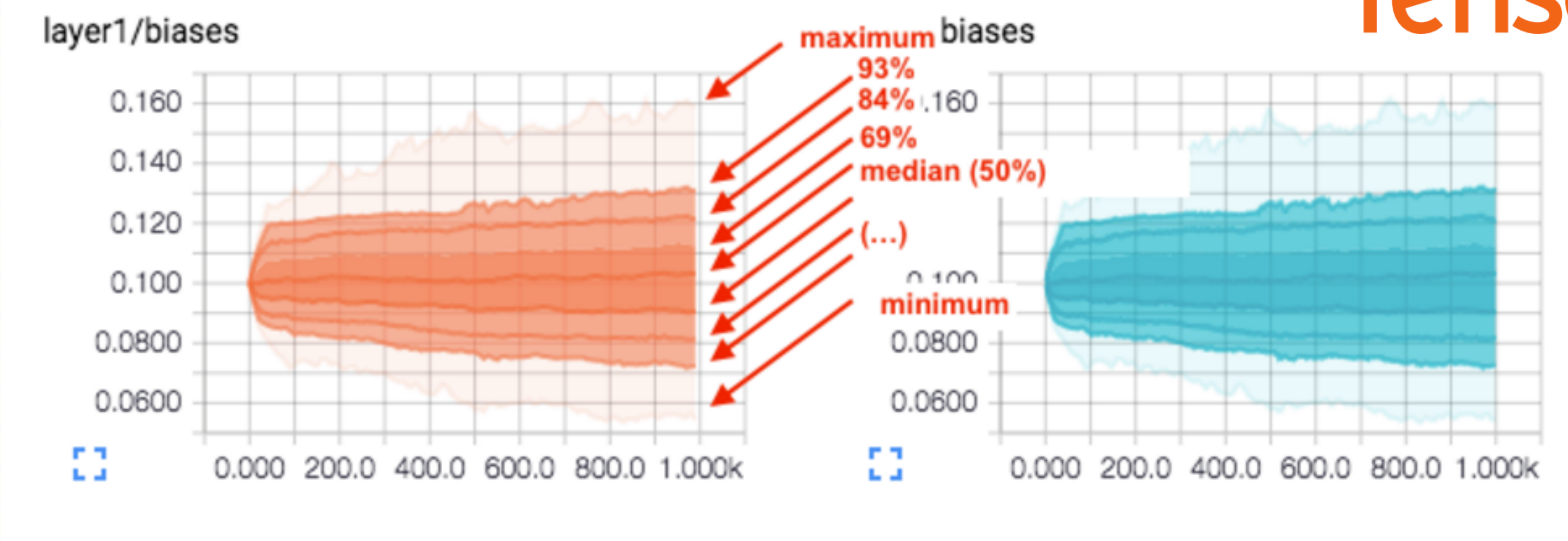
David Gunning (DARPA/I2O)

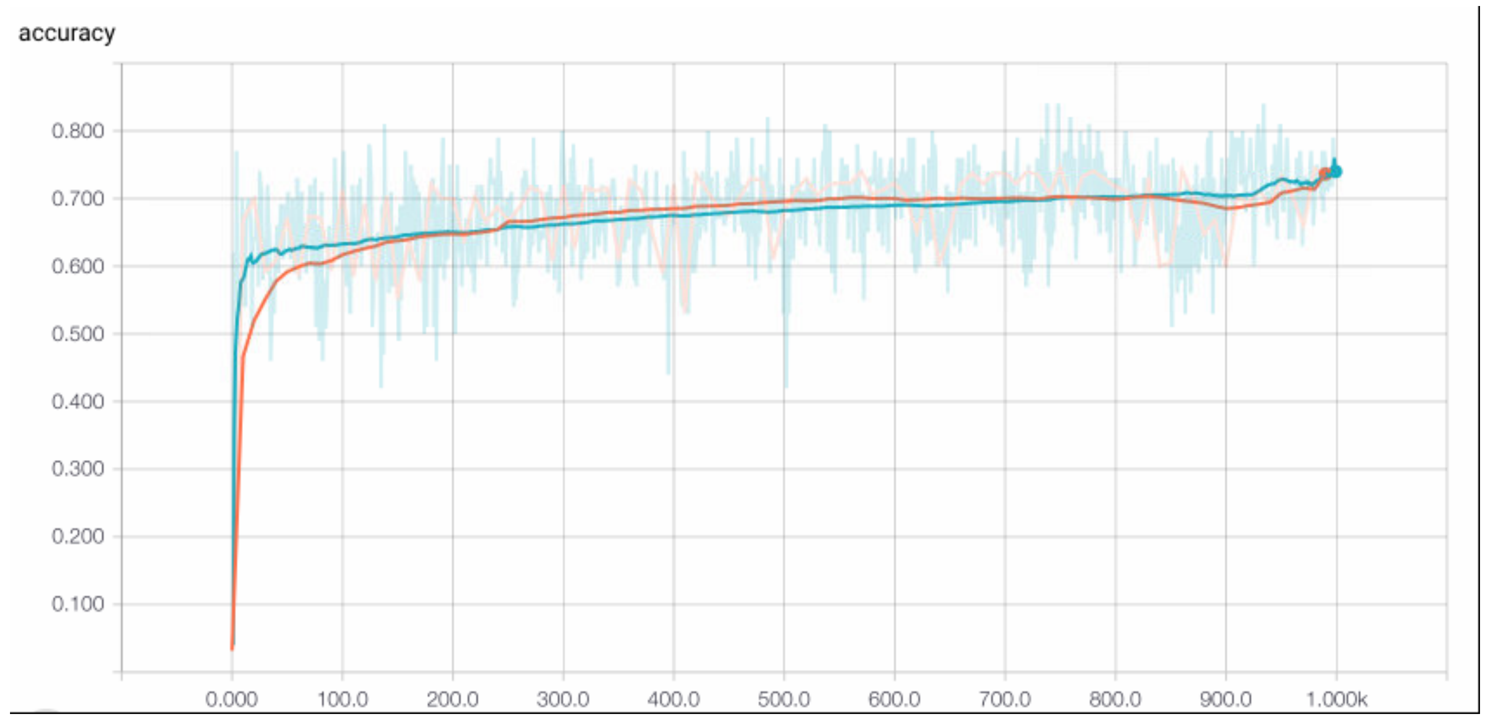
https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf



David Gunning (DARPA/I2O)

Distribution Summaries





Percentile distributions over the data: max, 93, 84, 69, 50, 31, 16, 7, min

Before

After

fromthegrapevine.com

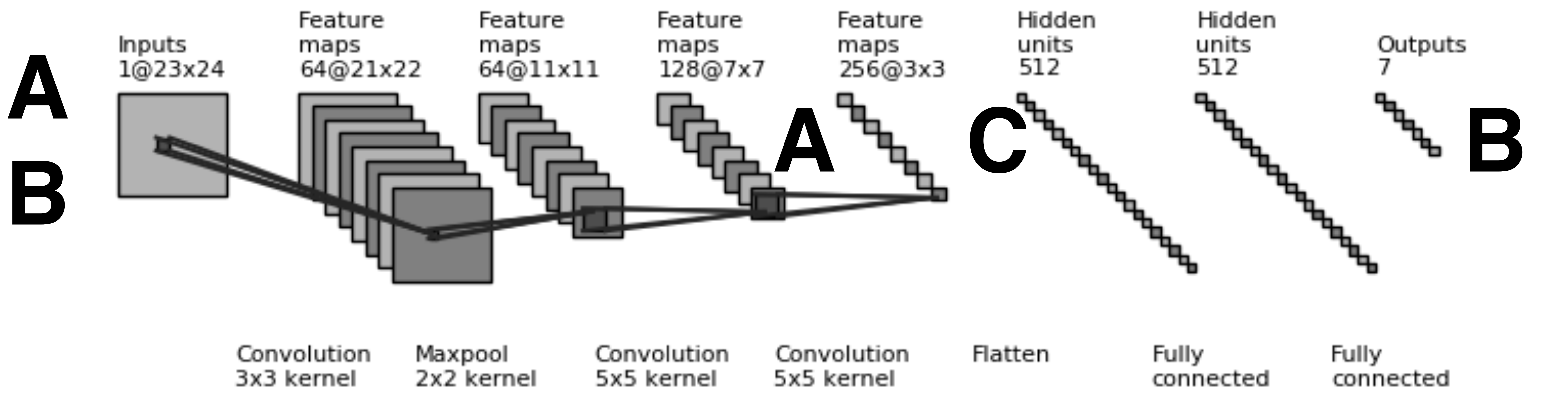
telegraph.co.uk

Visualization for interpretability

A. Activation Maximization
   - Initial layer filters easy to visualize
   - Generate input image that activates later filters

B. Saliency Maps
   - Gradient of o/p category wrt input image
   - Understanding attention of the classifier

C. Class Activation Maps
   - Gradients based on first dense layer
   - Spatial information still intact



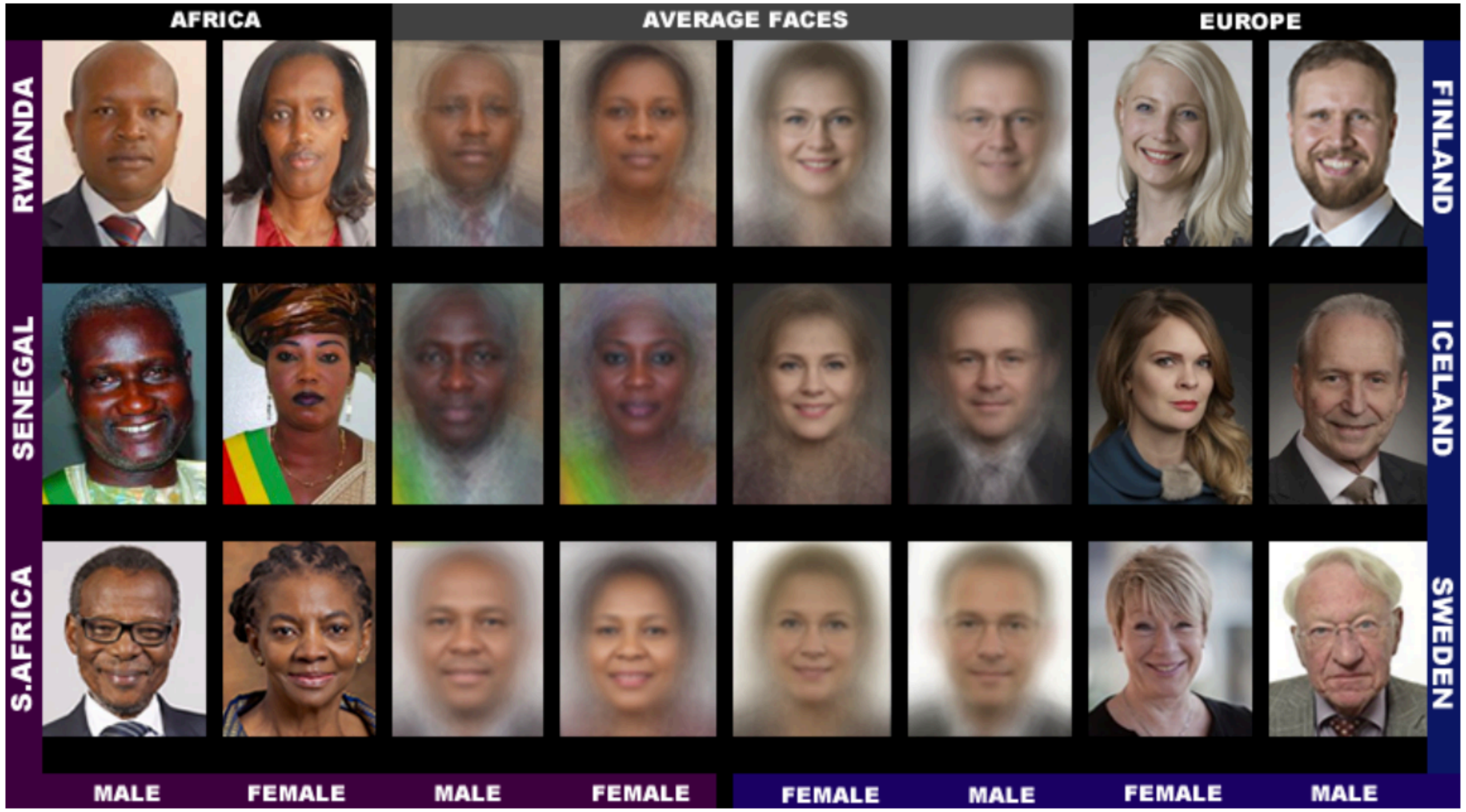
https://raghakot.github.io/keras-vis/

Bias in Machine Learning

Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

Automated systems are not inherently neutral. They reflect the priorities, preferences, and prejudices—the coded gaze—of those who have the power to mold artificial intelligence.
(MIT Media Lab)

Buolamwini and Gebru

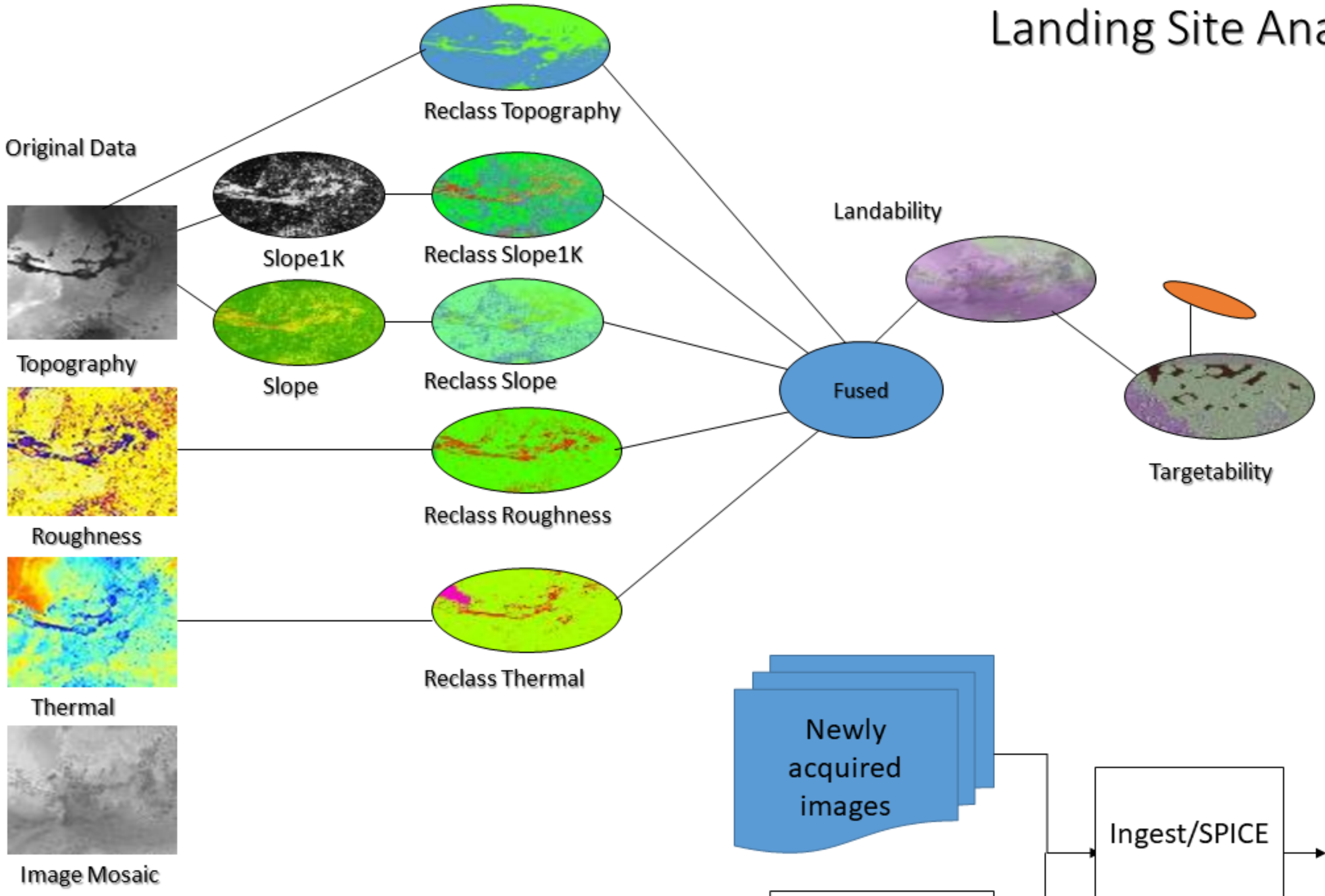


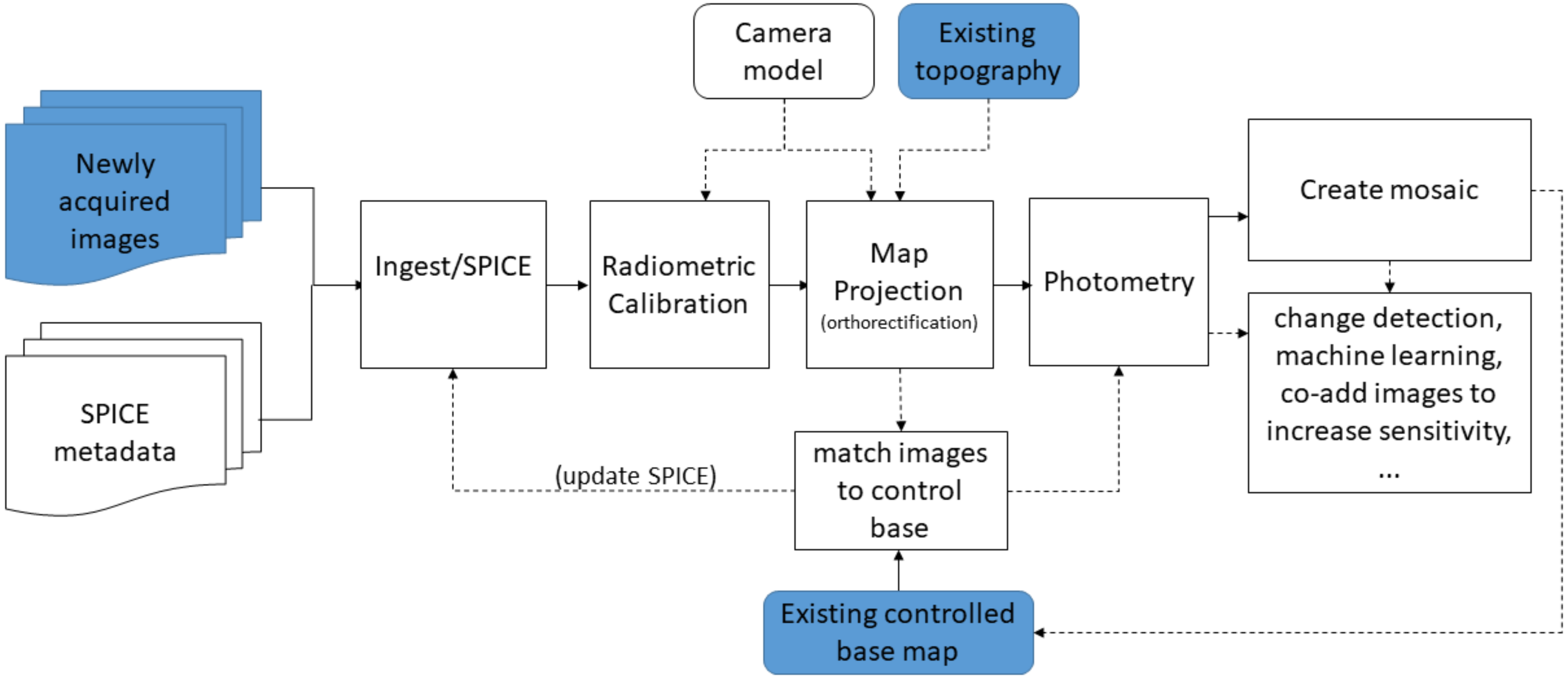
Pilot Parliaments Benchmark (PPB)

Gebru - Yann Lecun debate

We should all be wary of the bias in our data and the methodology

Landing Site Analysis





From Data Fusion for In-space Nebulae
Mahabal, Hare, Fox, Hallinan 2020
Based on a Keck Institute for Space Studies workshop

Summary

- Deep Learning has enetered almost every aspect of astronomy
- Good training sets are critical
- Models should be as simple as possible
- Building in interpretability is critical (though not easy yet)
- Many examples, codes available

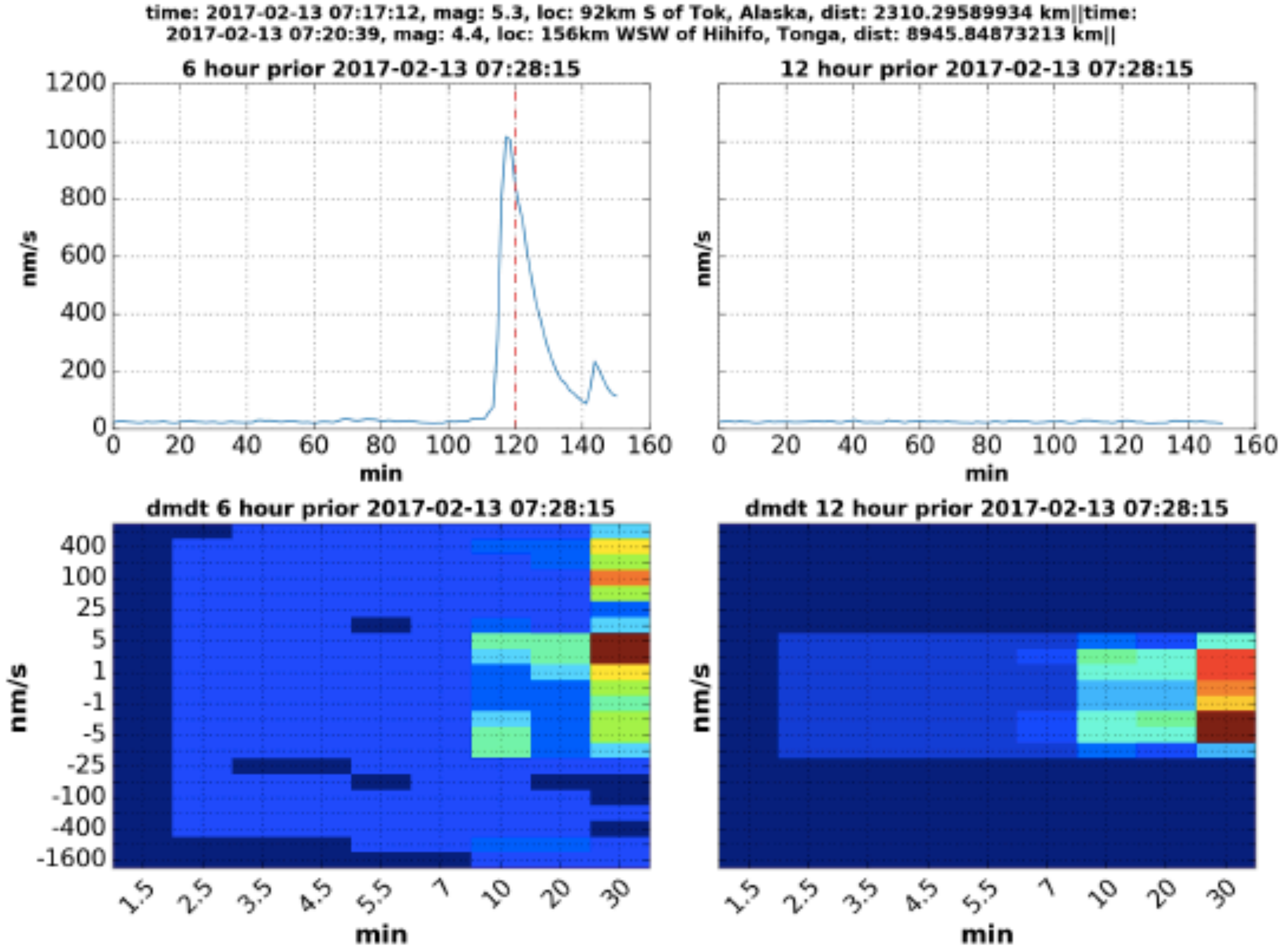
Don't do this!

Outline

- Intro to ZTF (4)
- Images -> diff images -> catalogs -> variability (5)
- diff images -> real-bogus (5)
- diff images -> transinet for transients (4)
- transients -> spectra (3)
- multiple catalogs -> time series -> dmdt (5)
- catalogs -> asteroids -> streaking (near earth) asteroids (3)
- Multi Messenger Astronomy -> GW skynet (5)

- LSTM/autoencoders (3)
- LSST/Rubin brokers, transfer learning (4)
- Good training data, ensure no overfitting, balanced sets (3)
- Tuned hyperparameters, retraining, interpretability/explainability (4)
- Many tricks/technques, active learning/adaptive training
- Exoplanets (5)
- Data fusion/nebulae (5)

Effect of earthquakes



How to choose bins?
Histogram equalization in both axes?