Identification of problematic epochs in Astronomical Time Series through Transfer Learning

Stefano **Cavuoti**
on behalf of:
De Cicco D., Doorenbos L., Brescia M., Torbaniuk O., Longo G. and Paolillo M.

Variable Objects

An object is considered to be variable if exhibiting a σ^{lc} in excess of the 95th percentile of the σ^{lc} distribution, over a running 0.5 mag wide bin centered on the magnitude of each source, highly conservative threshold, ≈ 2%) turned out to be variable.

blue = non-variable objects

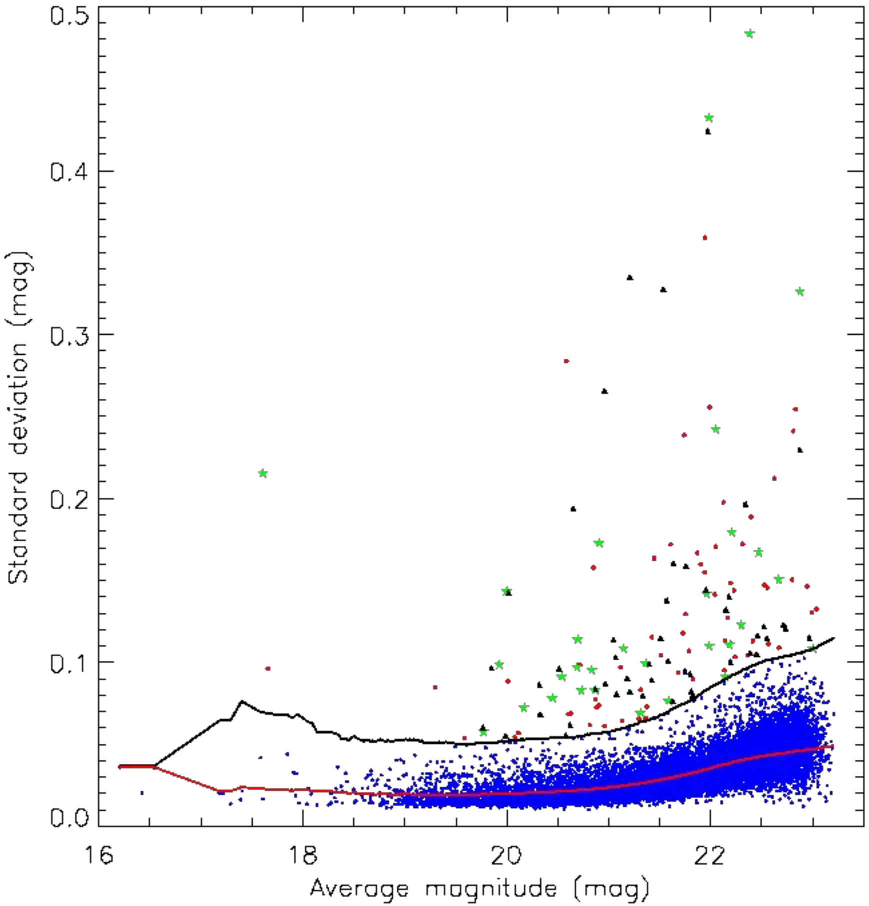
green = stars

red = galaxies

black = ambiguous

— average standard deviation

— variability threshold

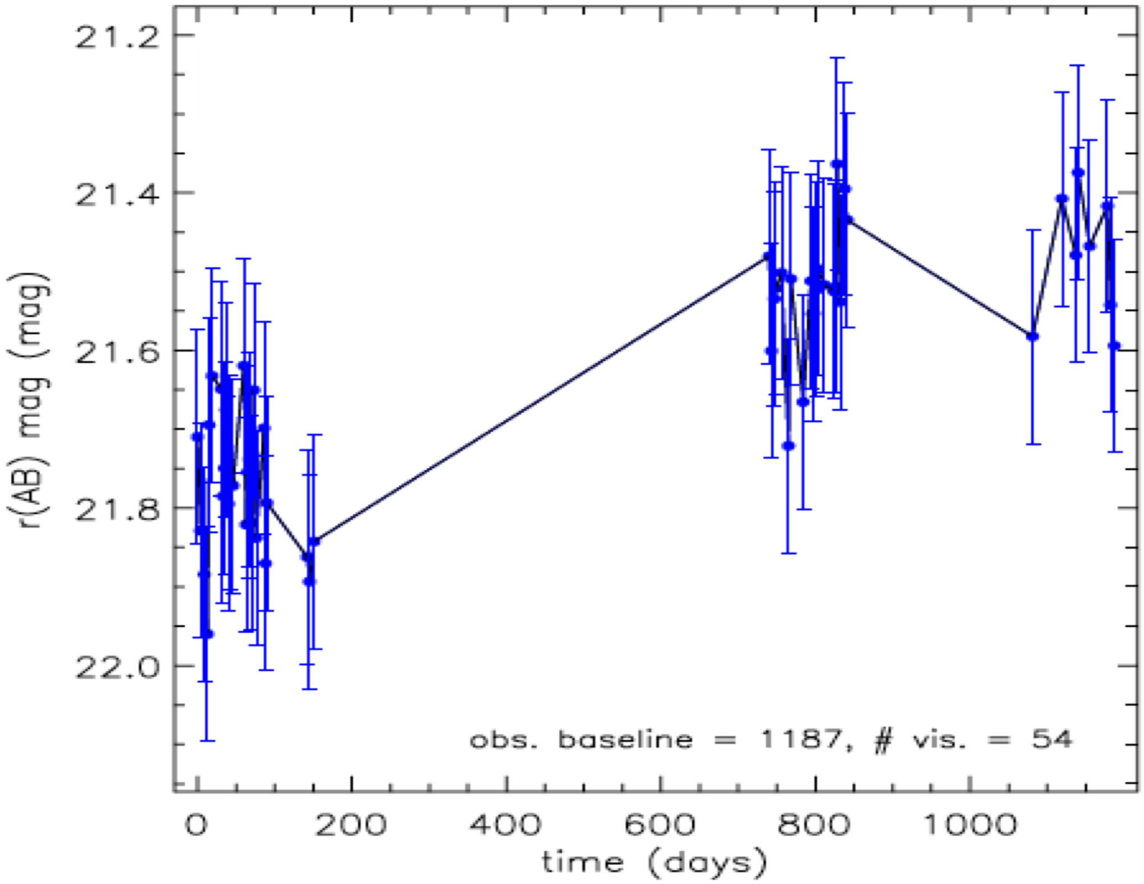


WARNING

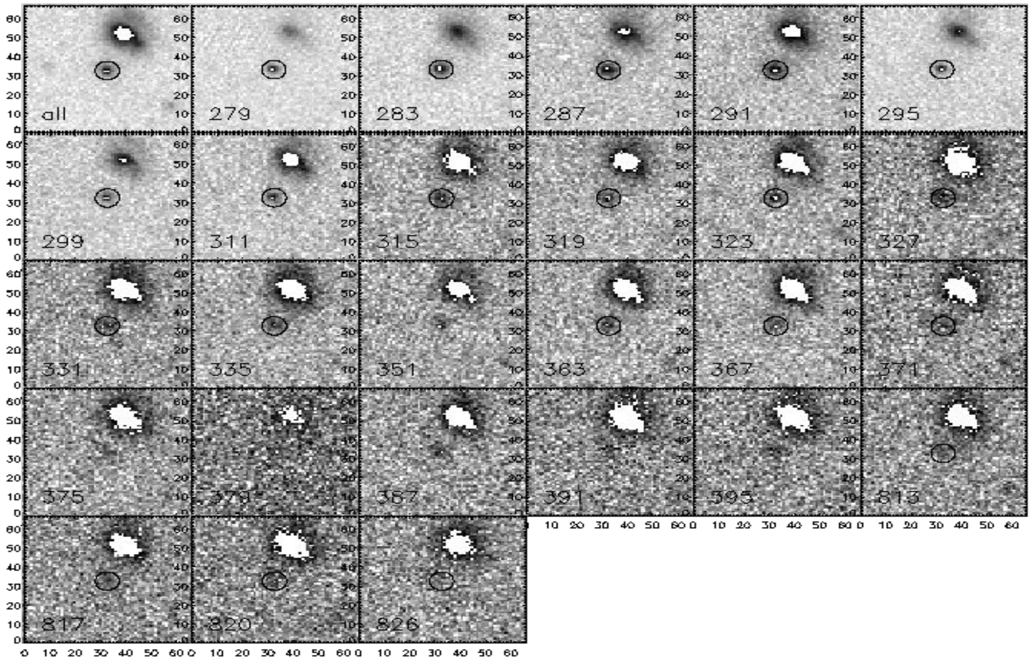
IN THE NEXT SLIDE AN EXAMPLE OF ASTRONOMICAL TIME SERIES IS SHOWN THIS CAN TRIGGER STATISTICIANS IN SEVERAL WAYS

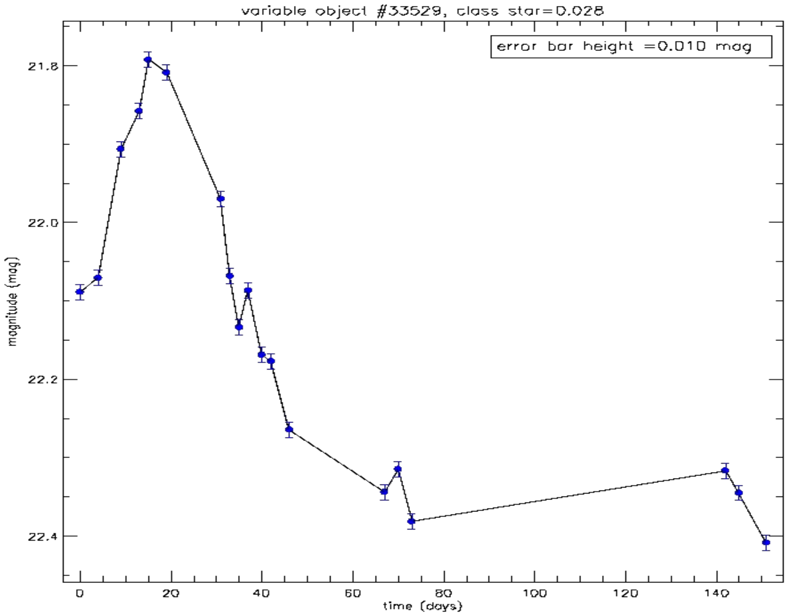
VIEWER DISCRETION IS ADVISED

Example of Light Curve



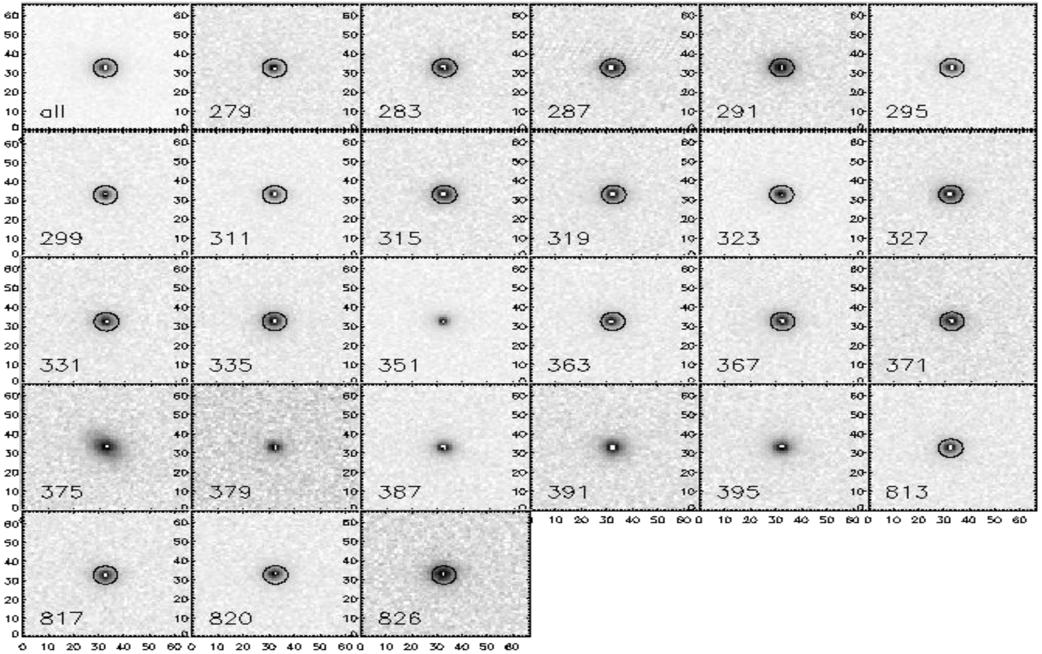
Example of Supernovae

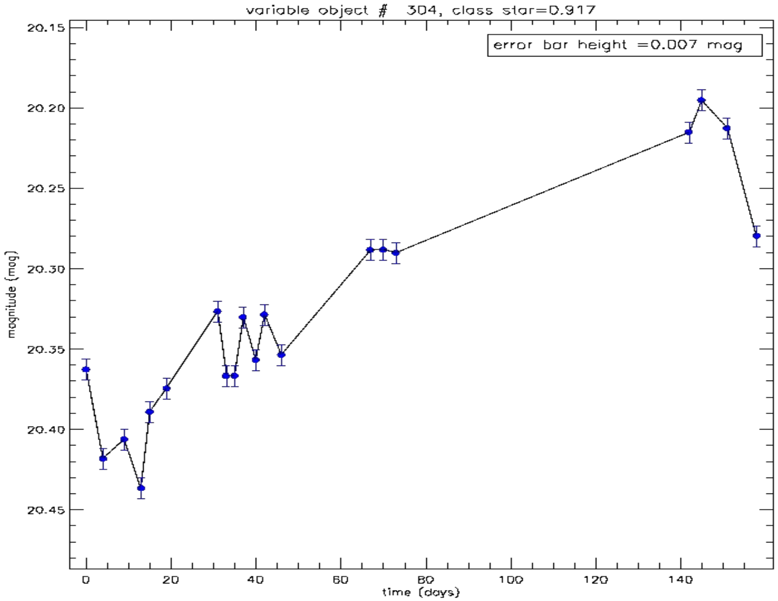




multi-epoch snapshots and light curve of a supernova candidate

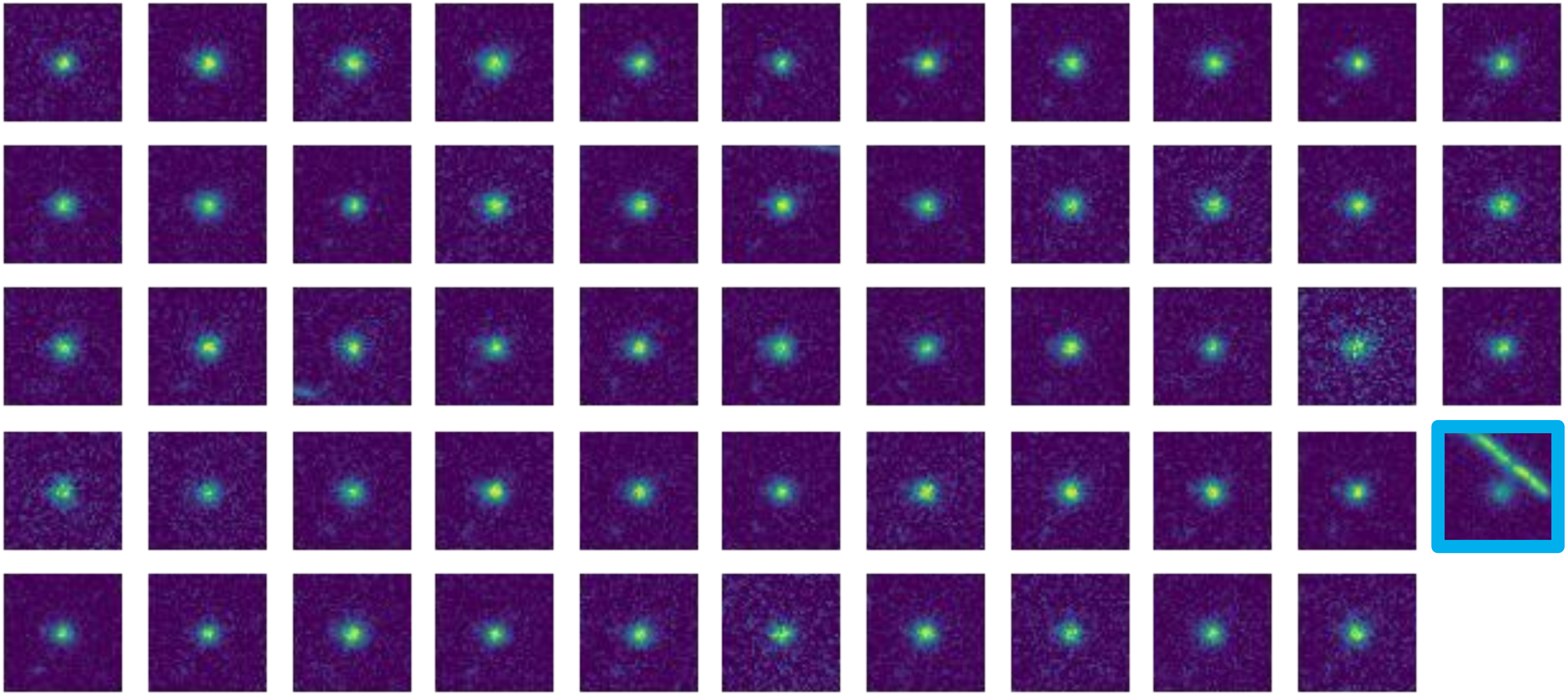
Example of AGN Candidate





multi-epoch snapshots and light curve of an AGN candidate

Example of a Problematic Image in Time Series

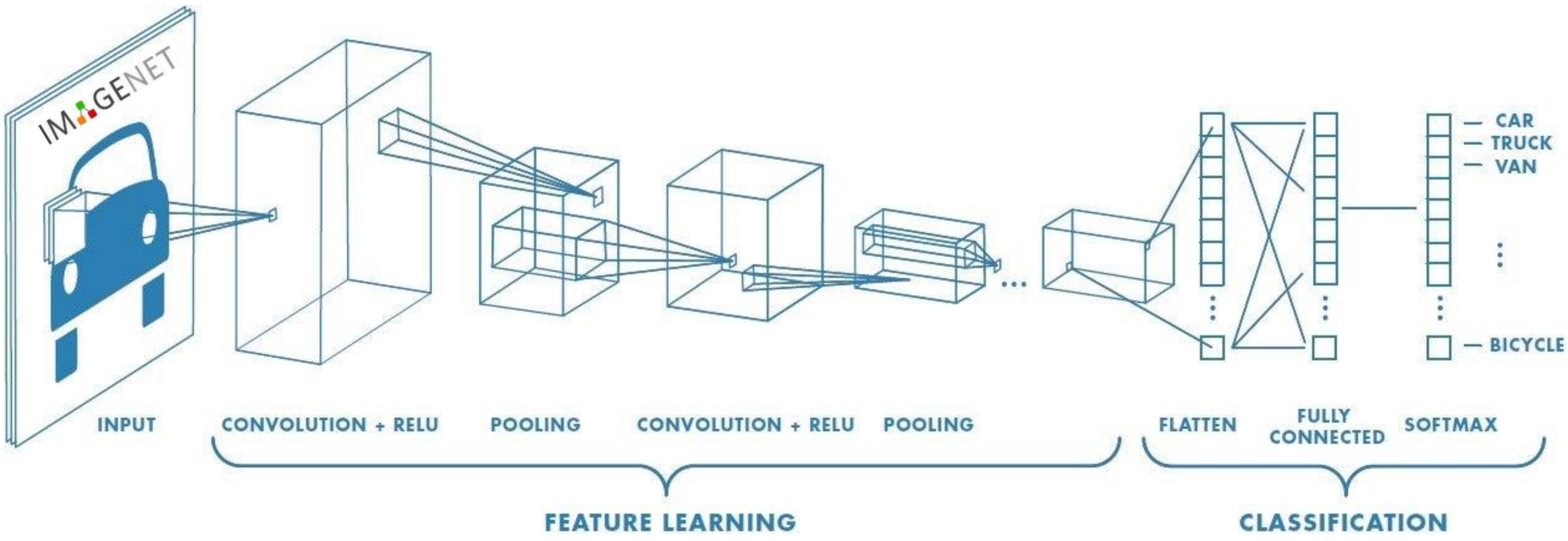


How to determine features without a train?

ImageNet

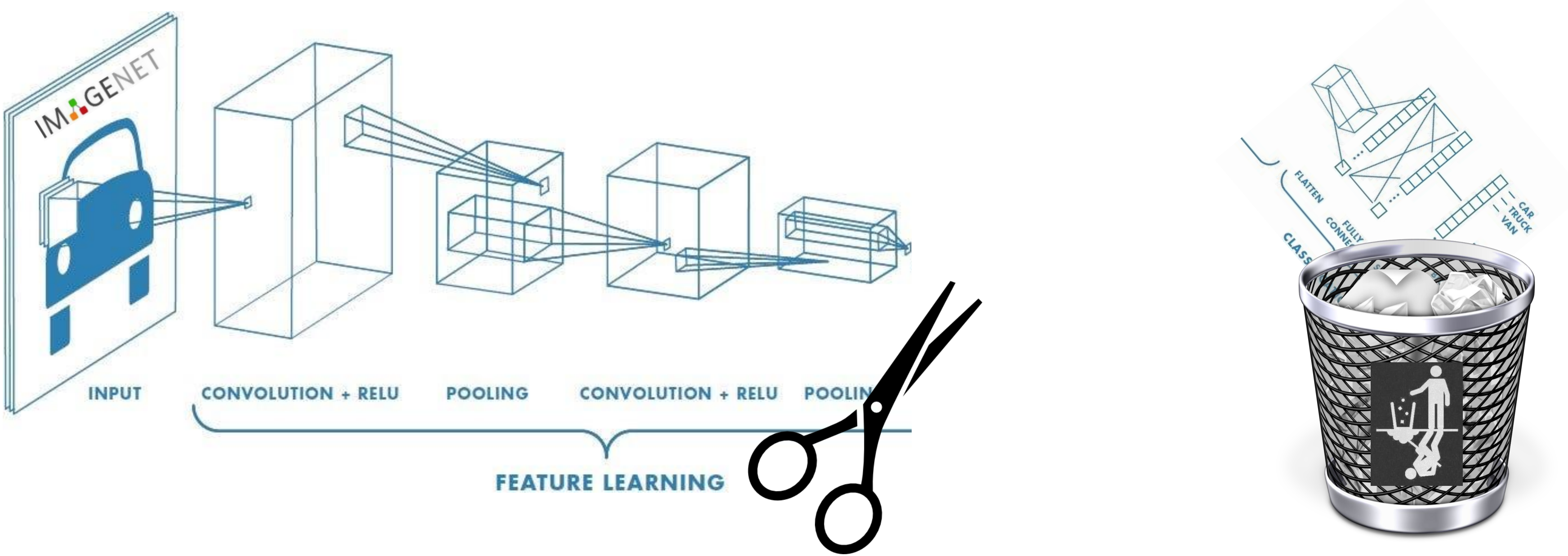
1.2 M objects
1000 categories

Convolutional Neural Networks

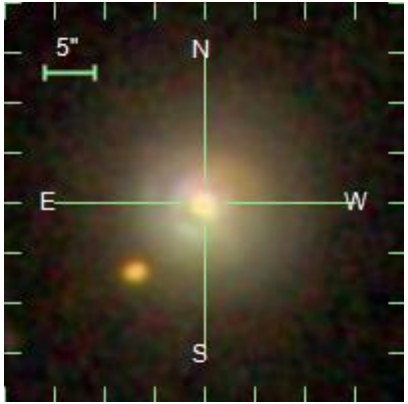


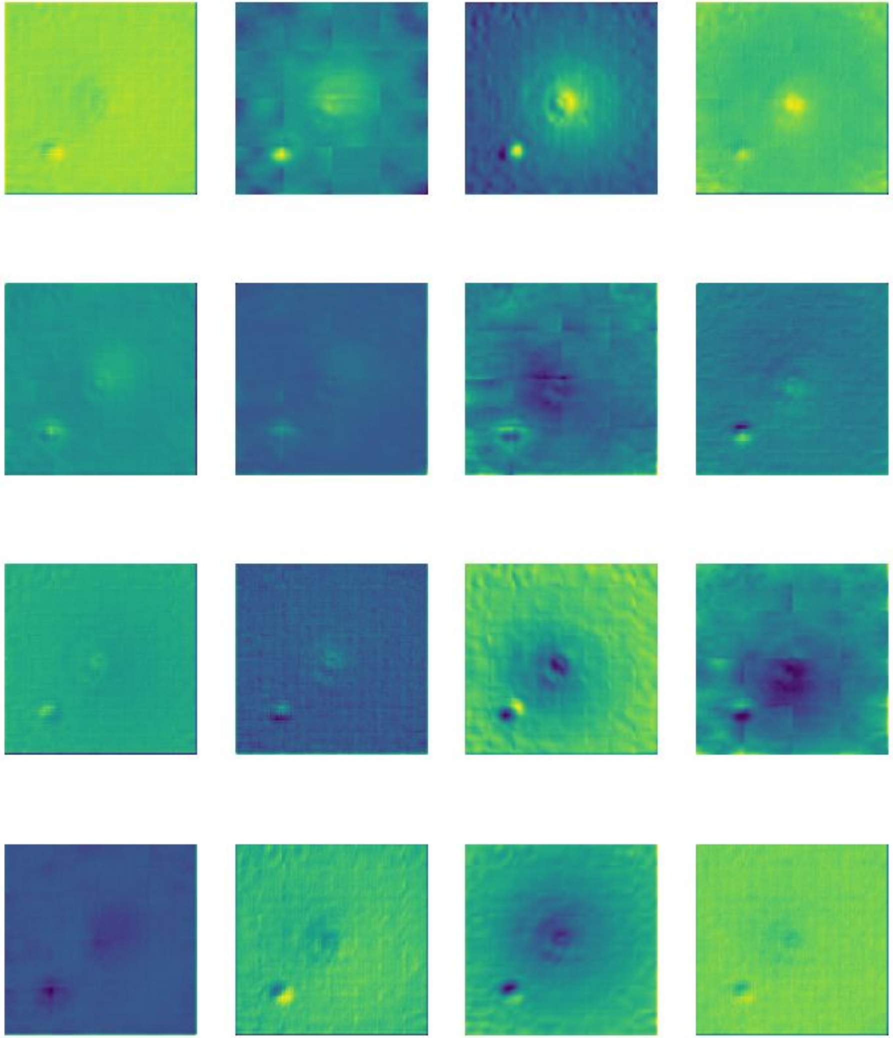
Convolutional Neural Networks

One man’s trash is another man’s treasure.
- English Proverb

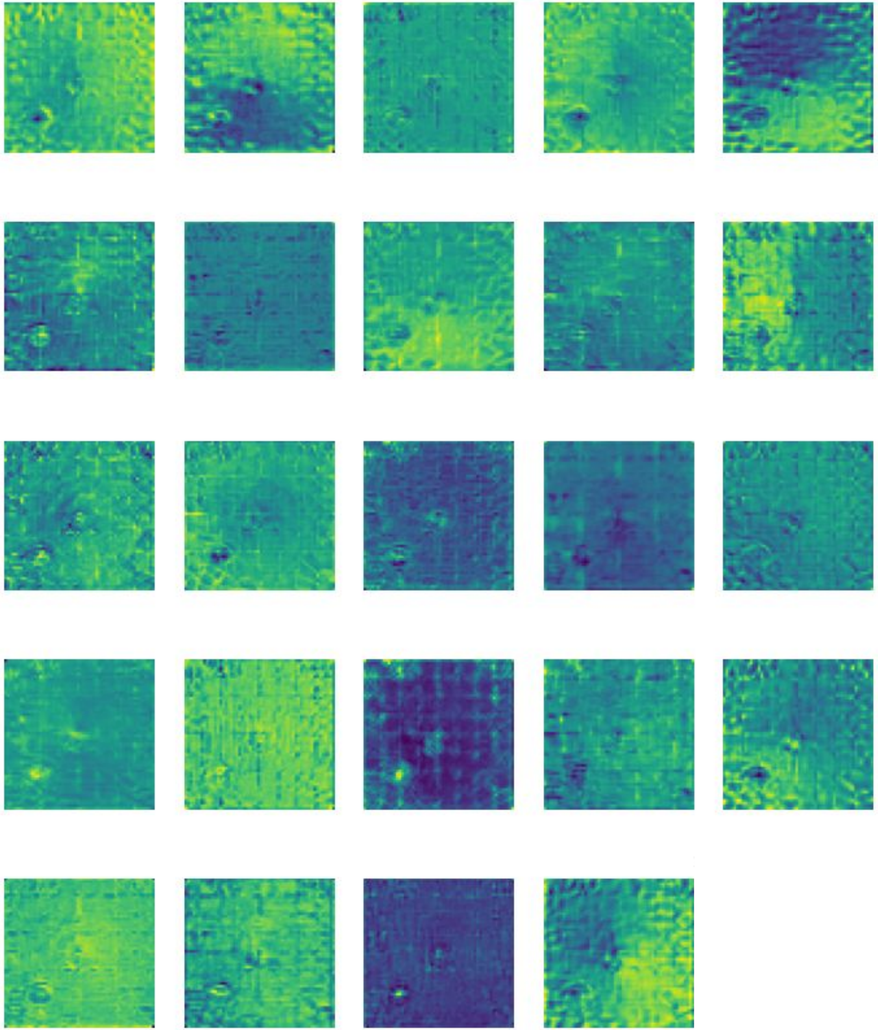


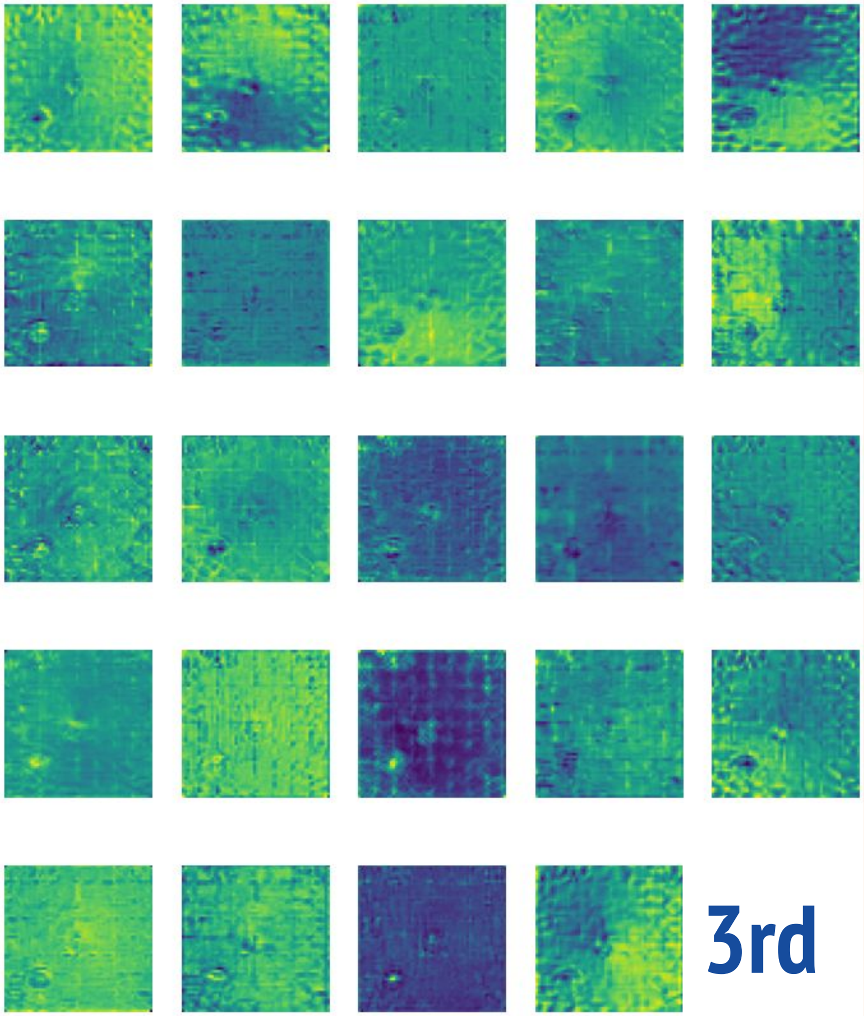
What happens to an astronomical image?

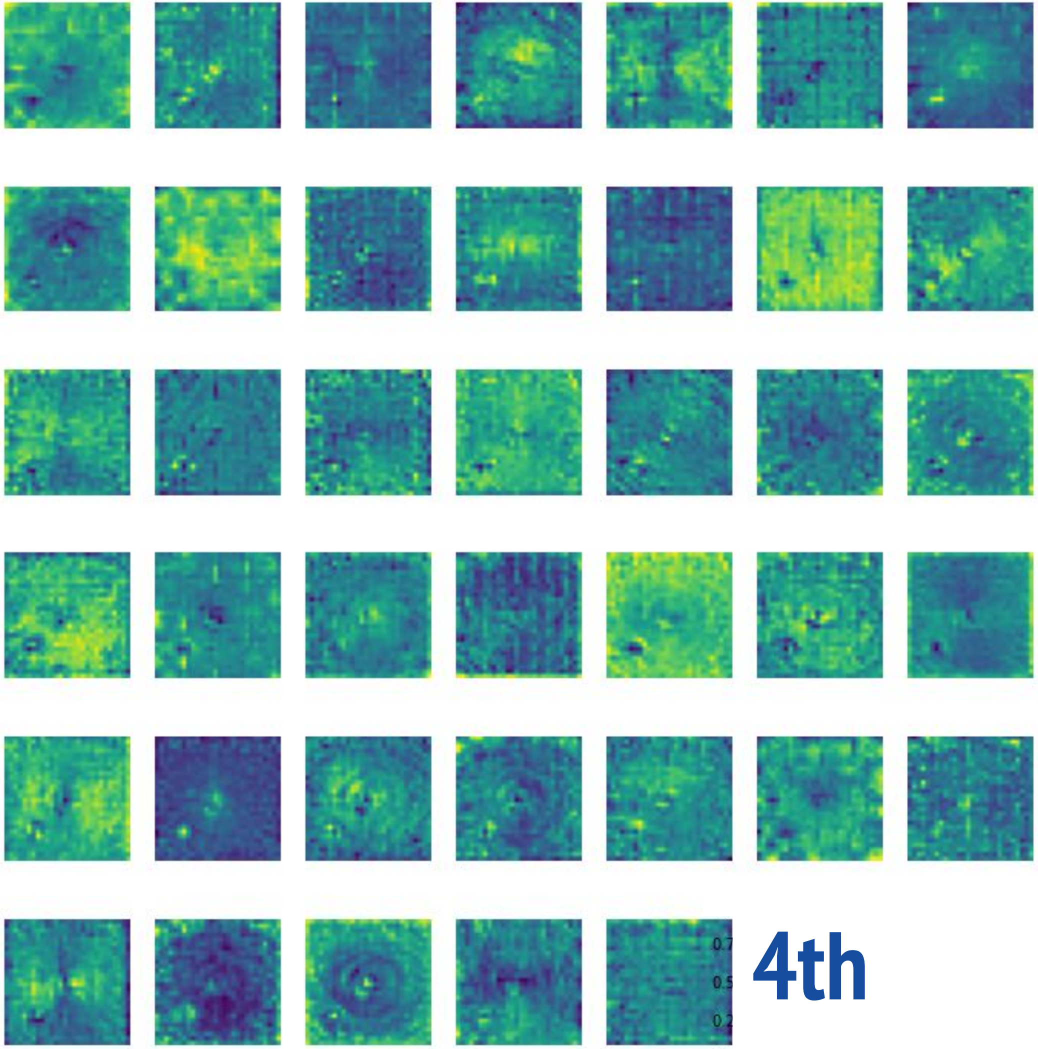




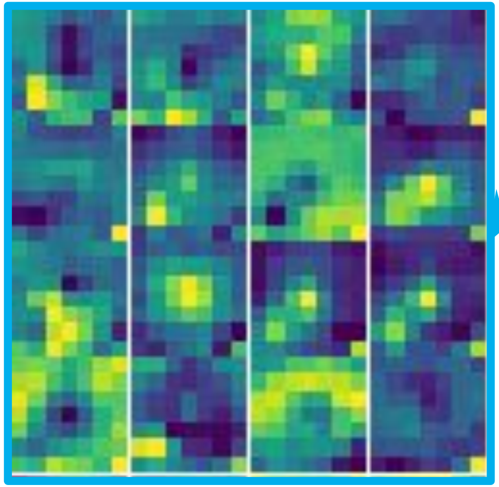
2nd layer

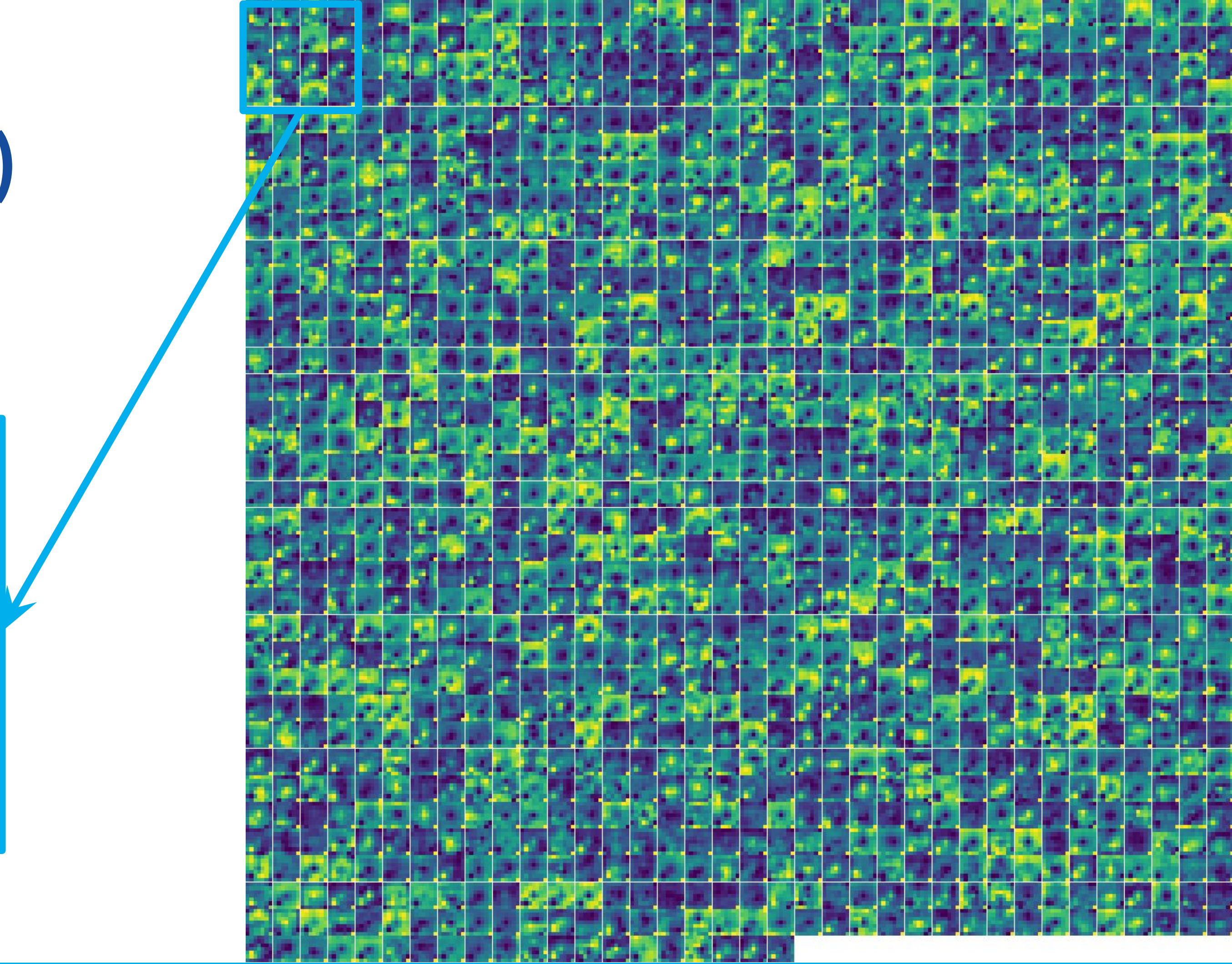



3rd

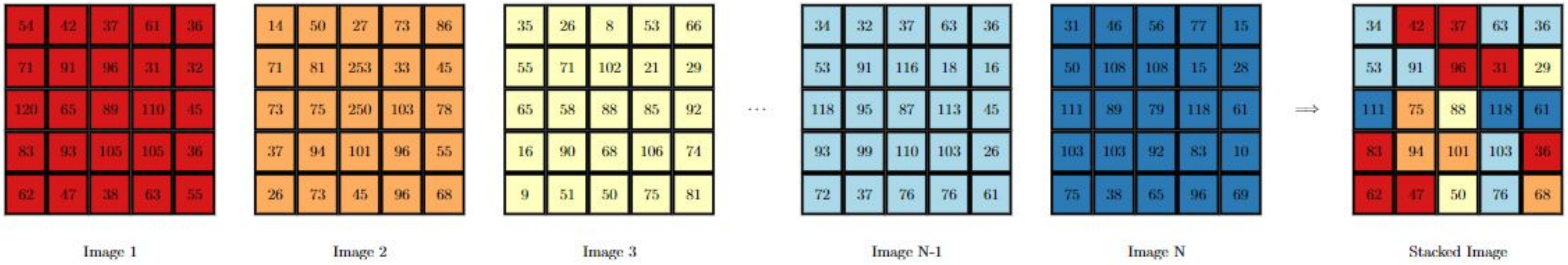

4th

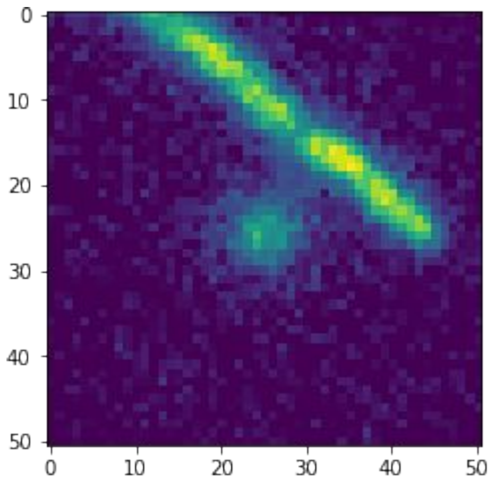
Last layer
(~1300 7x7 maps)

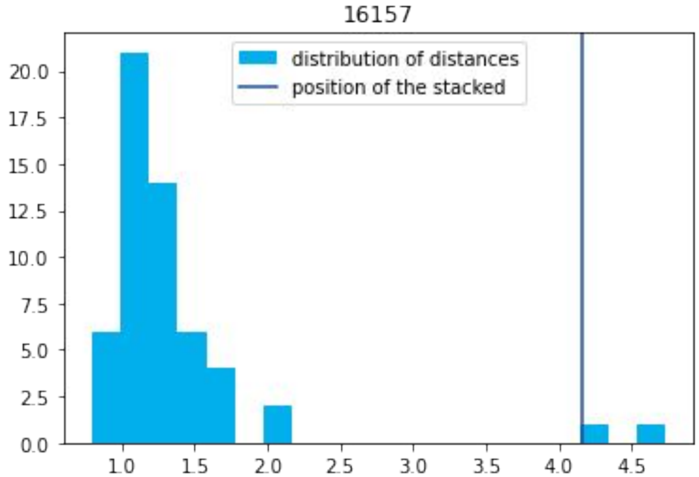


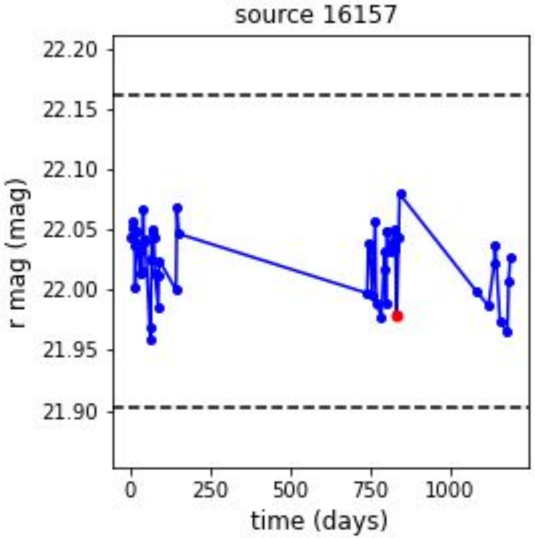


Stacked Image as Reference

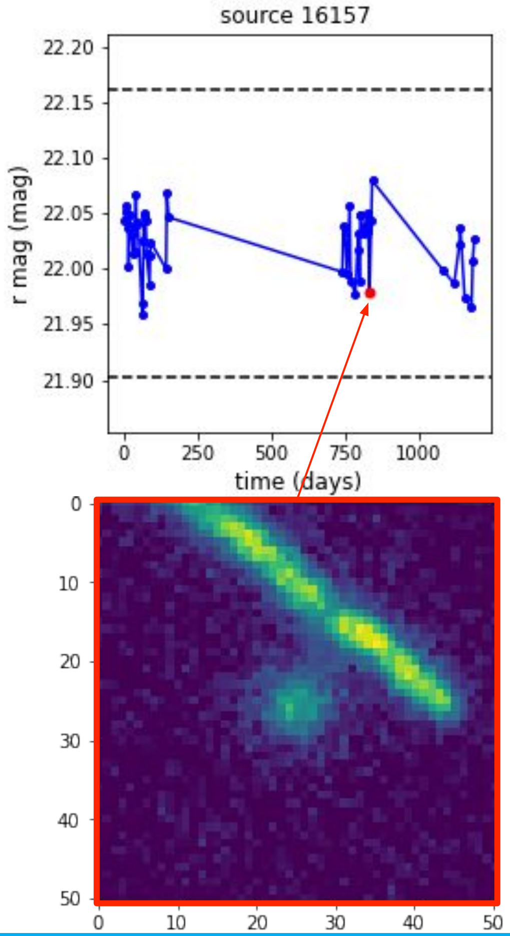








Traditional Approach: Sigma Clip



Pro:

- Simple,
- Fast,
- Efficient in removing the worst images

Cons:

- There are still points that could not be removed in that way

VST Overview

The VLT Survey Telescope (VST), located at Cerro Paranal, Chile, is a joint venture between the European Southern Observatory (ESO) and the Osservatorio Astronomico di Capodimonte (OAC) in Napoli, Italy.

primary mirror: 2.65 m

secondary mirror: 0.94 m

f#: 5.5

detector: OmegaCAM

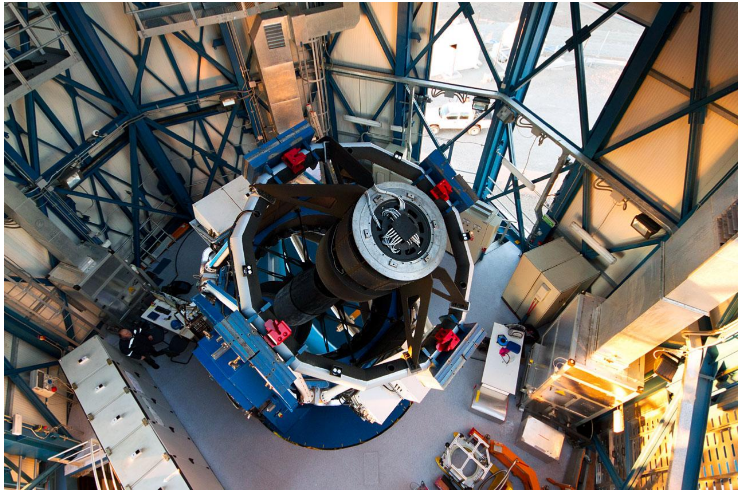
(≈268 Mpx, 26 x 26 cm^2,

≈ 50 GB of raw data per night)

Field of View (FoV): 1° x 1°

Focal Plane Scale (FPS): 0.21''/pixel

wavelength coverage: 0.3 – 1.0 μm



The VST-COSMOS Dataset (I)

visit	obs. date	seeing (FWHM) (arcsec)
1	2011-Dec-18	0.64
2	2011-Dec-22	0.94
3	2011-Dec-27	1.04
4	2011-Dec-31	1.15
5	2012-Jan-02	0.67
6	2012-Jan-06	0.58
7	2012-Jan-18	0.62
8	2012-Jan-20	0.88
9	2012-Jan-22	0.81
10	2012-Jan-24	0.67
11	2012-Jan-27	0.98
12	2012-Jan-29	0.86
13	2012-Feb-02	0.86
14	2012-Feb-16	0.50
15	2012-Feb-19	0.99
16	2012-Feb-21	0.79
17	2012-Feb-23	0.73
18	2012-Feb-26	0.83
19	2012-Feb-29	0.90
20	2012-Mar-03	0.97
21	2012-Mar-13	0.70
22	2012-Mar-15	1.08
23	2012-Mar-17	0.91
24	2012-May-08	0.74

1 yr
7 month-gap

8 month-gap

visit	obs. date	seeing (FWHM) (arcsec)
25	2012-May-11	0.85
26	2012-May-17	0.77
27	2013-Dec-27	0.72
28	2013-Dec-30	1.00
29	2014-Jan-03	0.86
30	2014-Jan-05	0.81
31	2014-Jan-12	0.73
32	2014-Jan-21	1.18
33	2014-Jan-24	0.80
34	2014-Feb-09	1.28
35	2014-Feb-19	0.89
36	2014-Feb-21	0.93
37	2014-Feb-23	0.81
38	2014-Feb-26	0.81
39	2014-Feb-28	0.77
40	2014-Mar-08	0.91
41	2014-Mar-21	0.96
42	2014-Mar-23	0.92
43	2014-Mar-25	0.66
44	2014-Mar-29	0.89
45	2014-Apr-04	0.58
46	2014-Apr-07	0.61
47	2014-Dec-03	1.00

6 yr
9 month-gap

8 month-gap

visit	obs. date	seeing (FWHM) (arcsec)
48	2015-Jan-10	0.71
49	2015-Jan-28	0.90
50	2015-Jan-31	0.73
51	2015-Feb-15	0.70
52	2015-Mar-10	0.80
53	2015-Mar-14	0.84
54	2015-Mar-19	1.00
55	2021-Dec-28	0.57
56	2022-Jan-05	0.58
57	2022-Jan-13	0.69
58	2022-Jan-25	0.71
59	2022-Feb-03	0.96
60	2022-Feb-10	0.62
61	2022-Feb-24	0.56
62	2022-Mar-11	0.75
63	2022-Mar-24	0.68
64	2022-Mar-31	0.70
65	2022-Nov-23	0.91
66	2022-Dec-10	0.56
67	2022-Dec-30	0.63
68	2022-Dec-31	1.07
69	2023-Jan-02	0.78
70	2023-Jan-03	0.84

The VST-COSMOS Dataset (II)

De Cicco+21; De Cicco+19; De Cicco+15; De Cicco+22, De Cicco+ in prep.

Test-ground survey for AGN variability studies

- ✦ **wide coverage**: ~1 sq. deg. with VST, vs. ~9.6 sq. deg. with LSST
- ✦ **long baseline**: >10 yr with VST, vs. ~10 yr with LSST
- ✦ **multi-band observations**: *gri* bands available with VST vs. *ugrizy* with LSST
- ✦ **dense sampling**: 3 day cadence, (*r*-band, 54 visits) with VST vs. expected 2-day 2-filter cadence with LSST
- ✦ **single-visit depth**: r ~ 24.6 mag (point sources, 5σ) with VST vs. ~24.7 mag with LSST

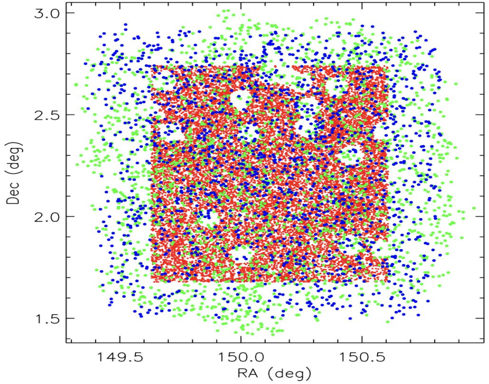
The VST-COSMOS dataset is currently one of the few datasets taking advantage of considerable depth, long baseline, and high observing cadence

Some Variability Studies with VST

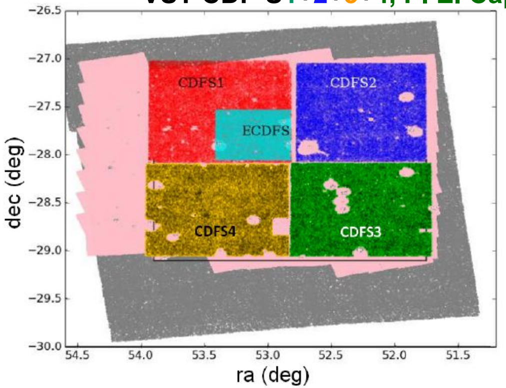
1. Analysis of optical variability, COSMOS, 1 sq. deg., 5 month baseline - *De Cicco+15*
2. Analysis of optical variability, CDF-S, 1+1 sq. deg., 5+3 month baselines - *Falocco+15*
3. Extension of the analysis for COSMOS to 3.3 yr - *De Cicco+19*
4. Complementary analysis for 2 more sq. deg. for the CDF-S - *Poulain+20*
5. Machine learning-based analysis of optical variability in COSMOS - *De Cicco+21*
6. Analysis of the ensemble variability properties in COSMOS - *De Cicco+22*

VST-COSMOS, PI G. Pignata

Chandra
XMM



VST-CDF-S1+2+3+4, PI E. Cappellaro



Results

17,995 sources

989,725 individual observations.

914,156 after Sigma Clip

336 extra anomalous epochs from our algorithm

Results

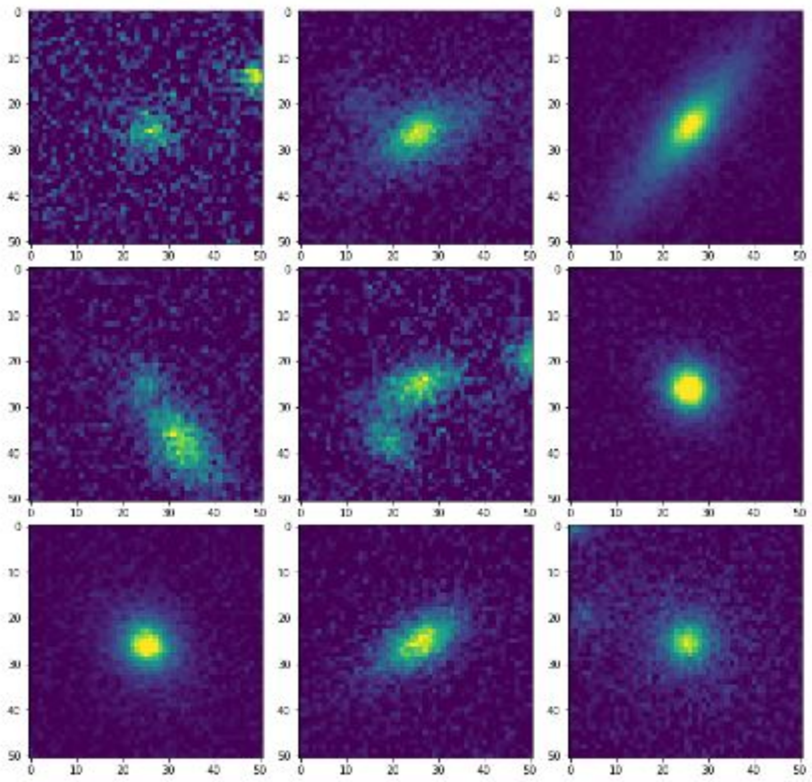
17,995 sources

98

91

33

48 Low Quality images



Results

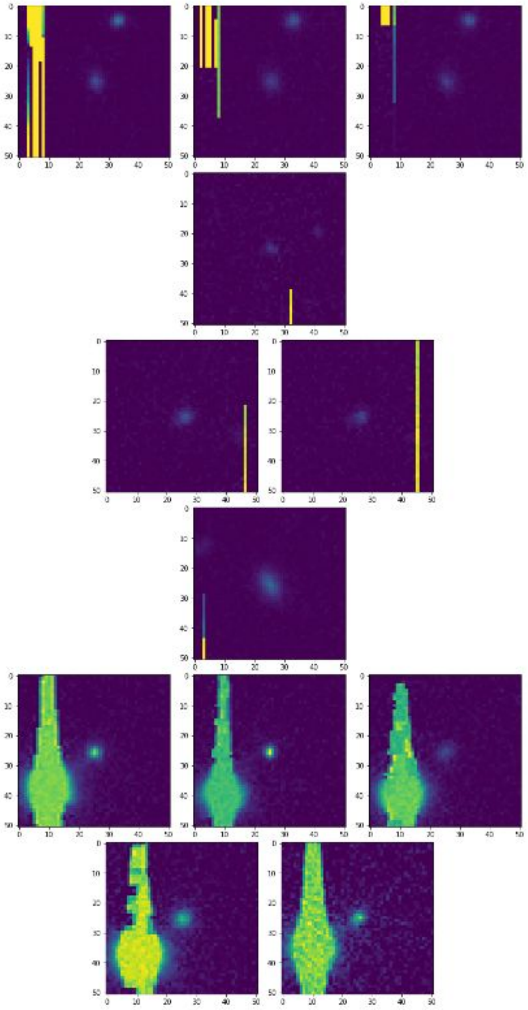
17,995 sources

98

91

33

21 Objects Contaminated by a nearby bright object



Results

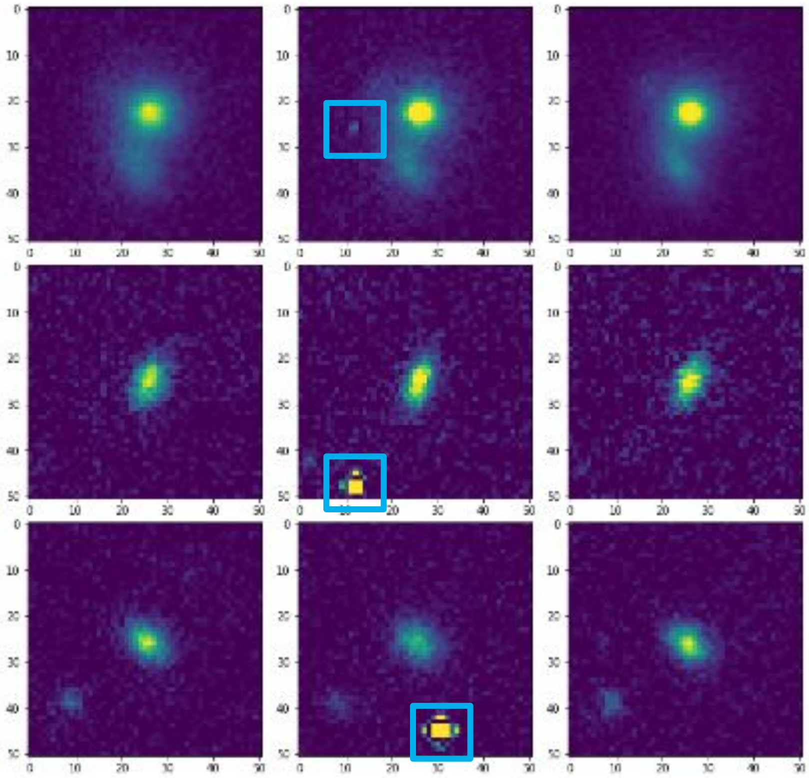
17,995 sources

98

91

33

3 Transients



Results

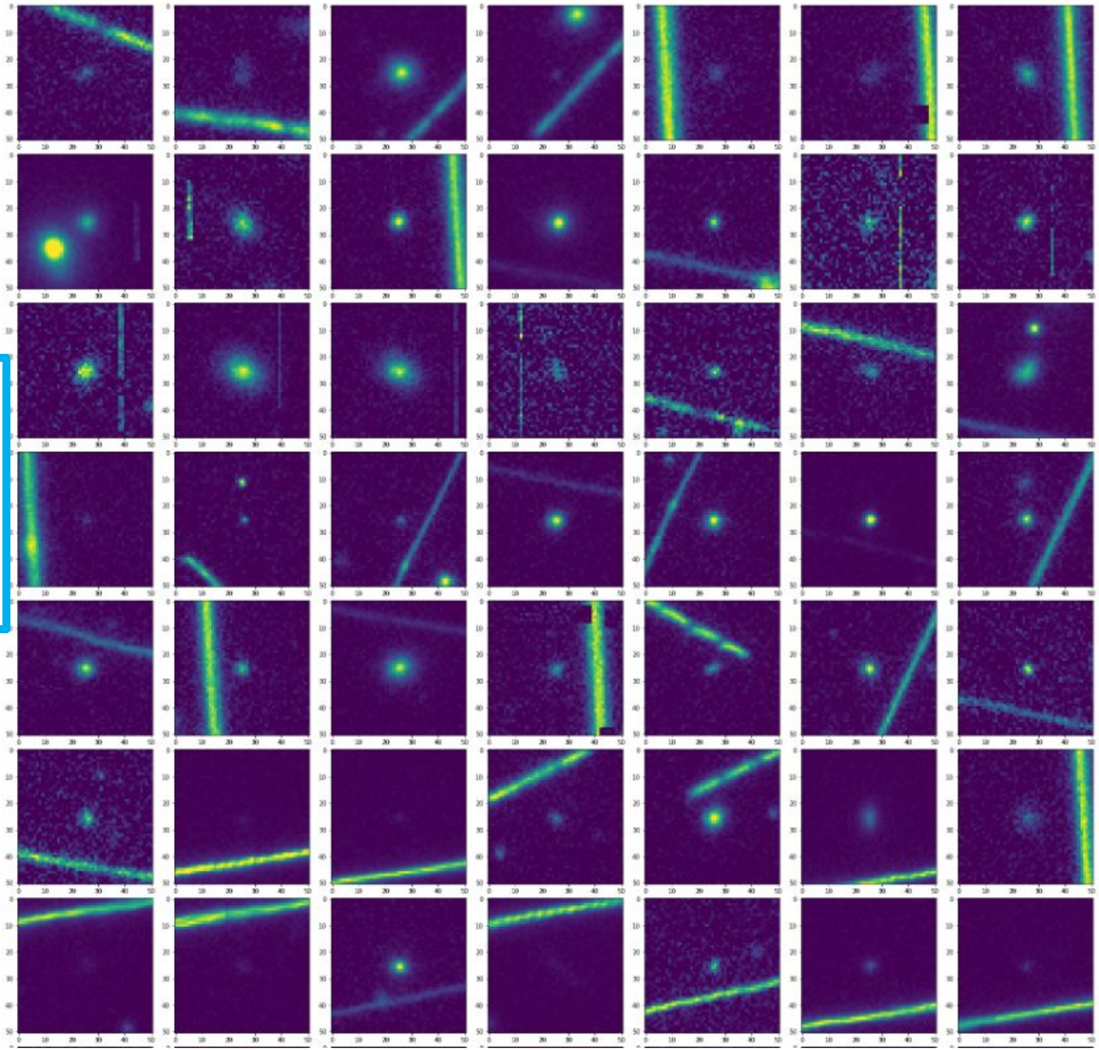
17,995 sources

98

91

33

185 Transits



Results

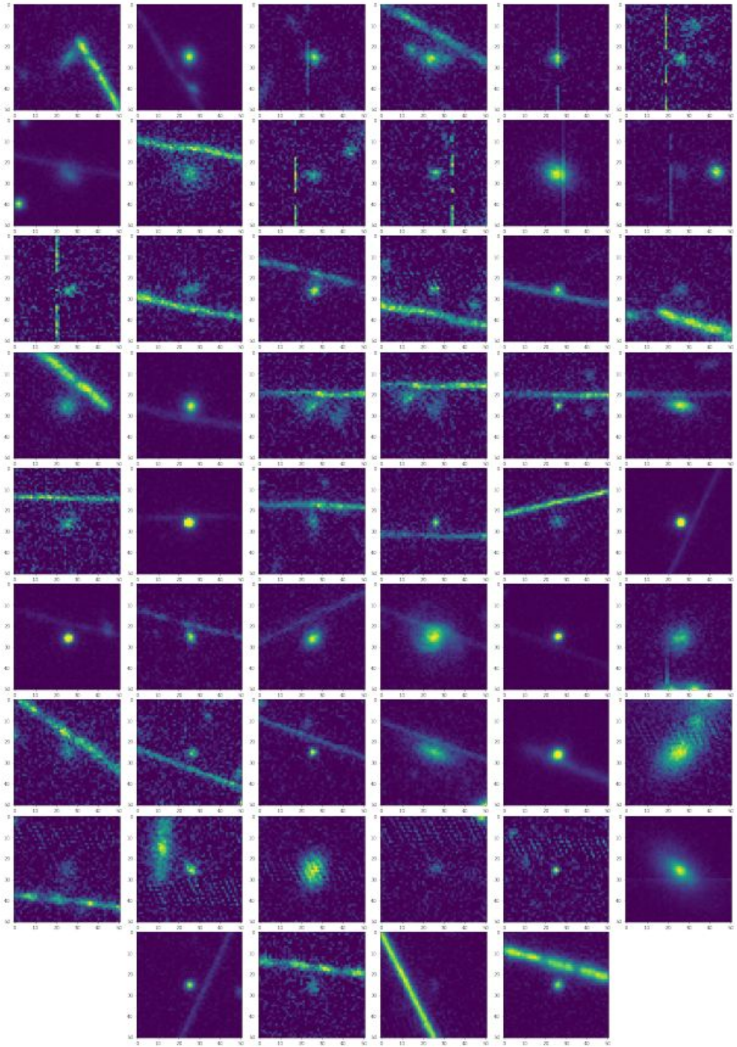
17,995 sources

98

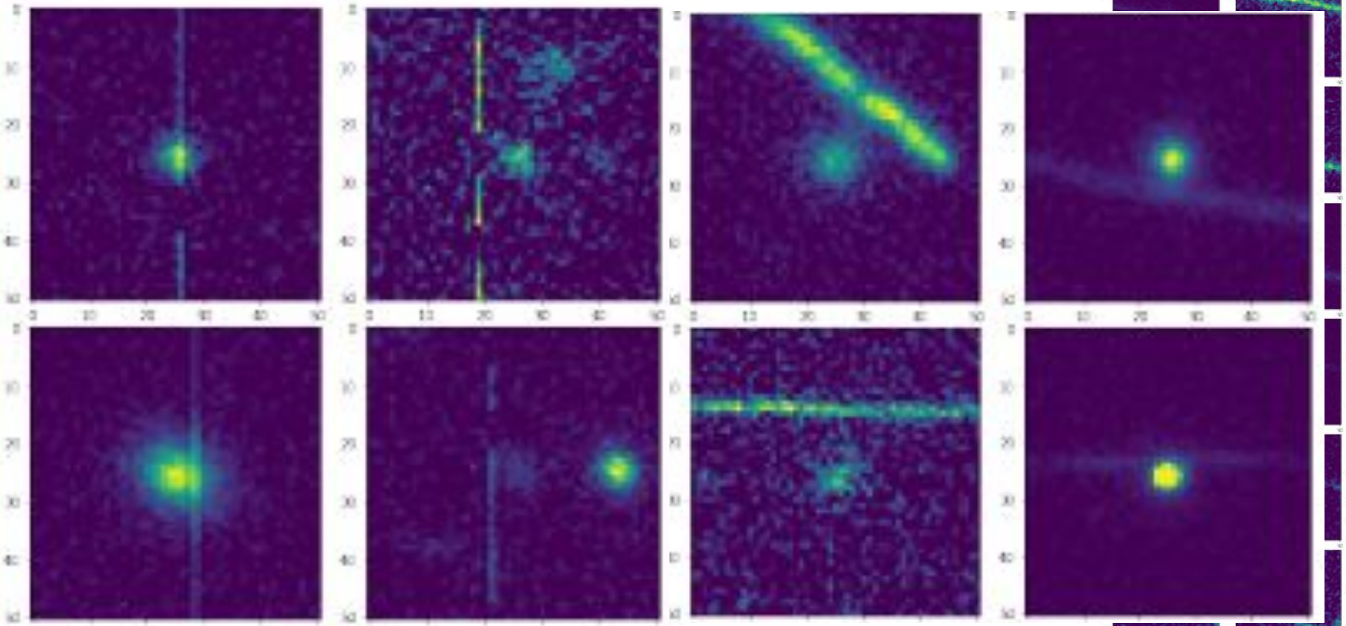
91

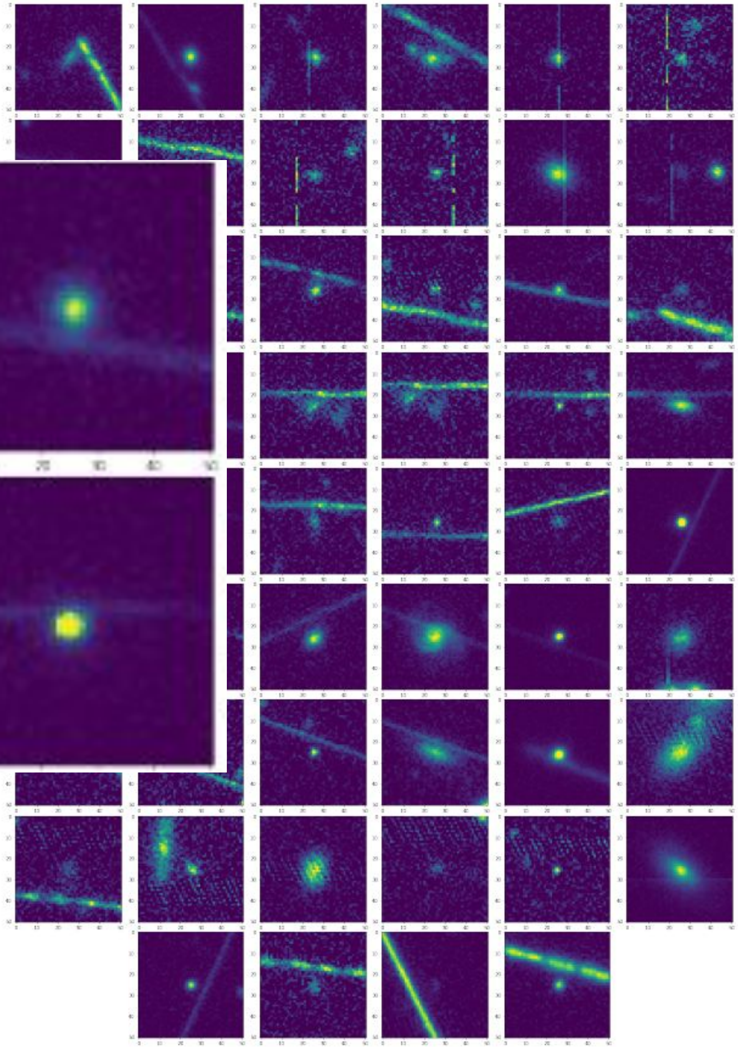
33

52 Really Problematic Images

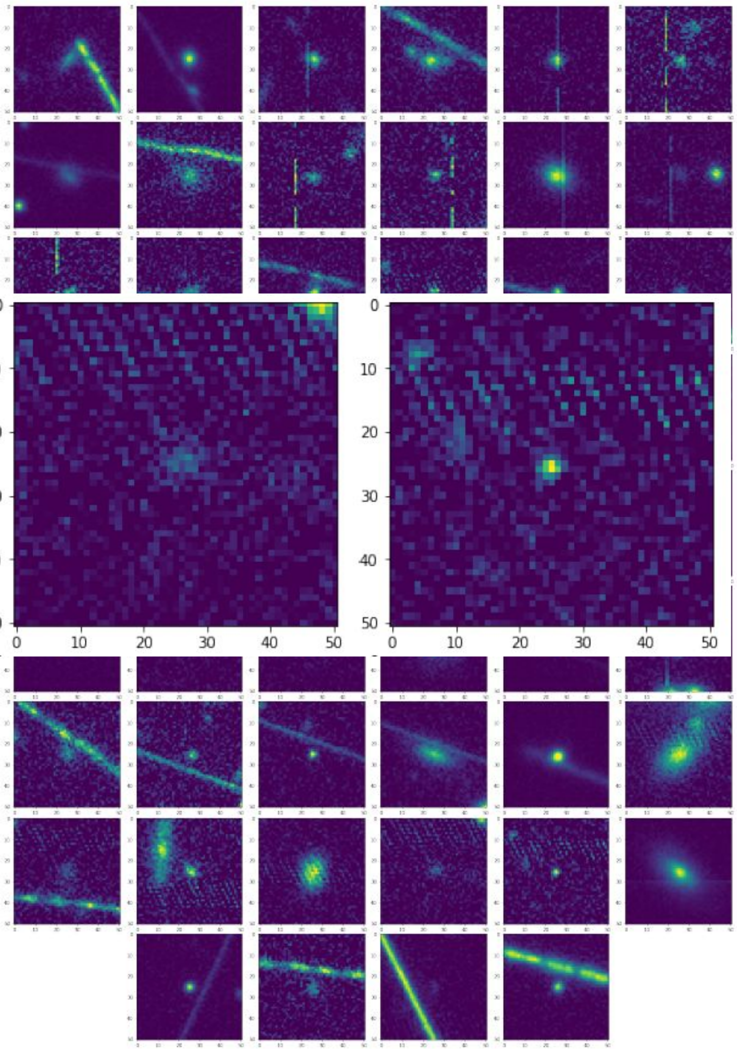


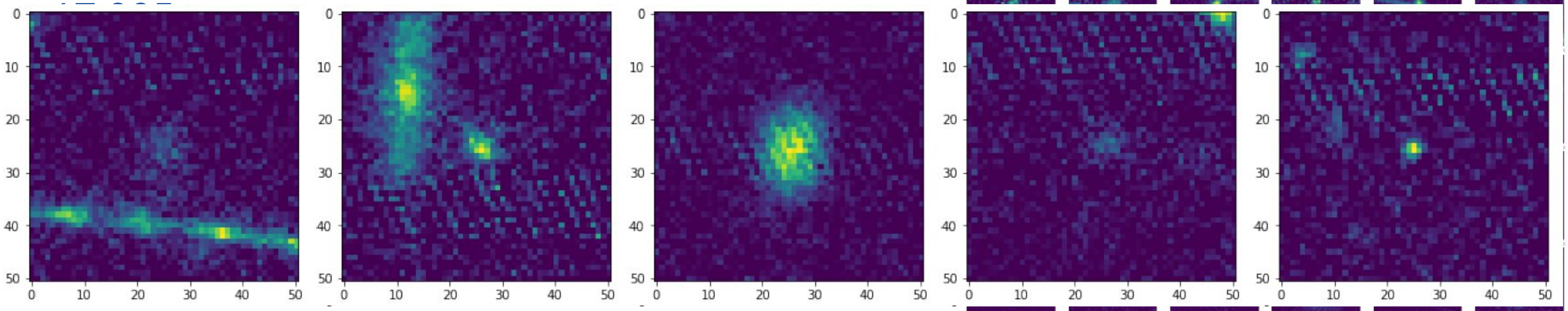
Results

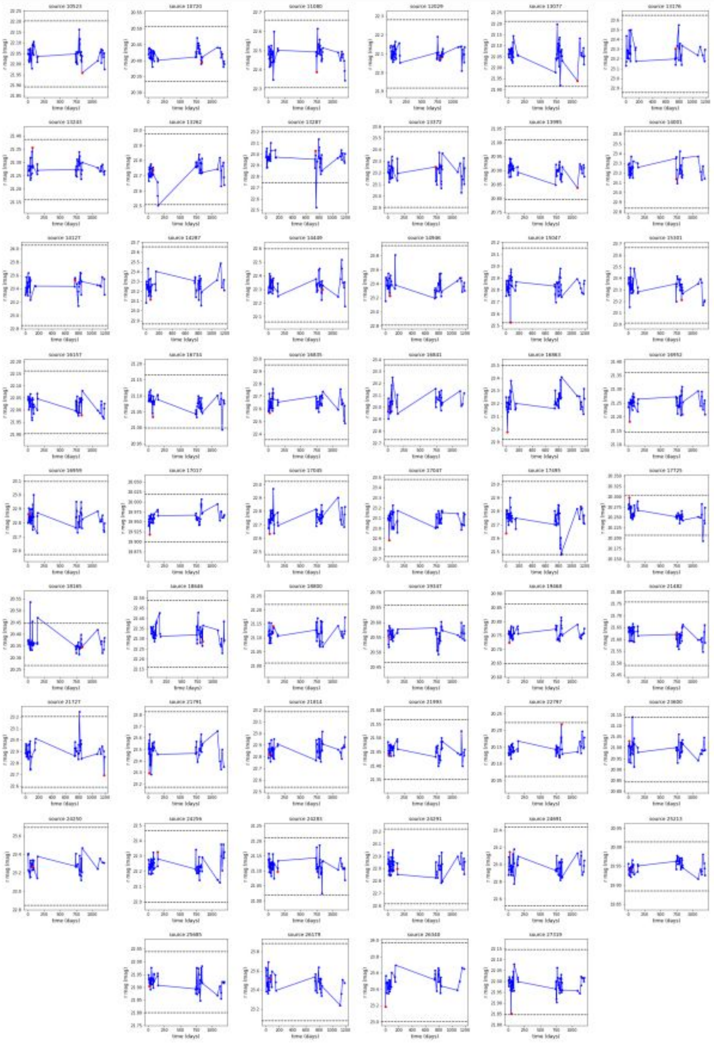




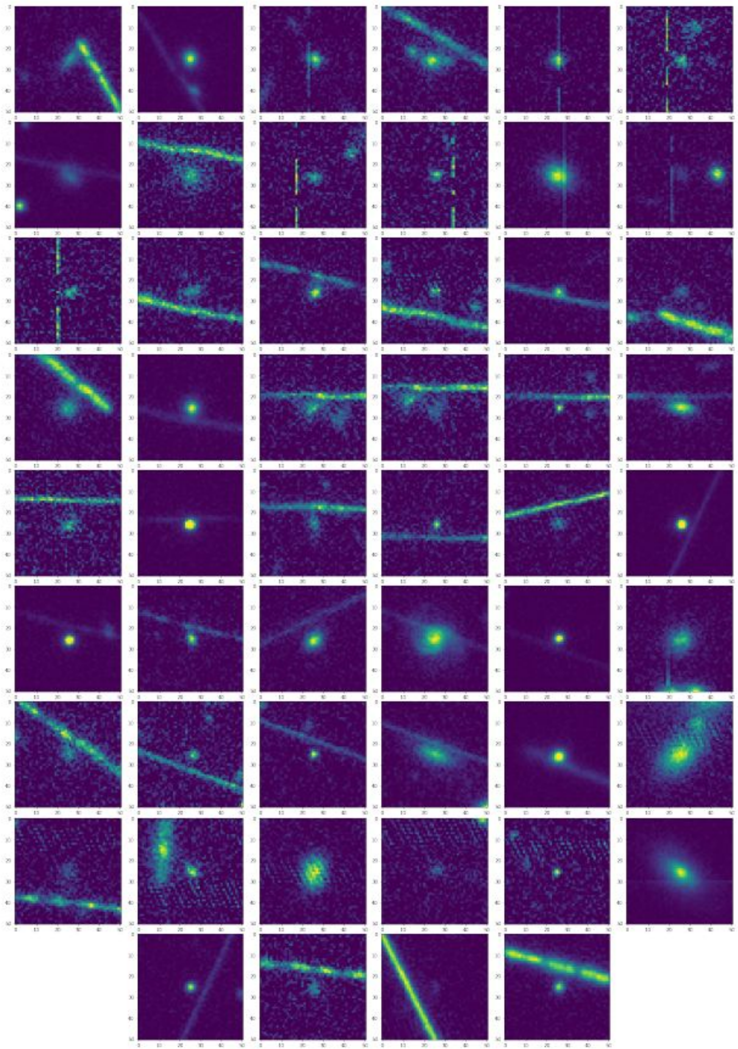
Results







Most of them could not be cut by any sigma clip



Results

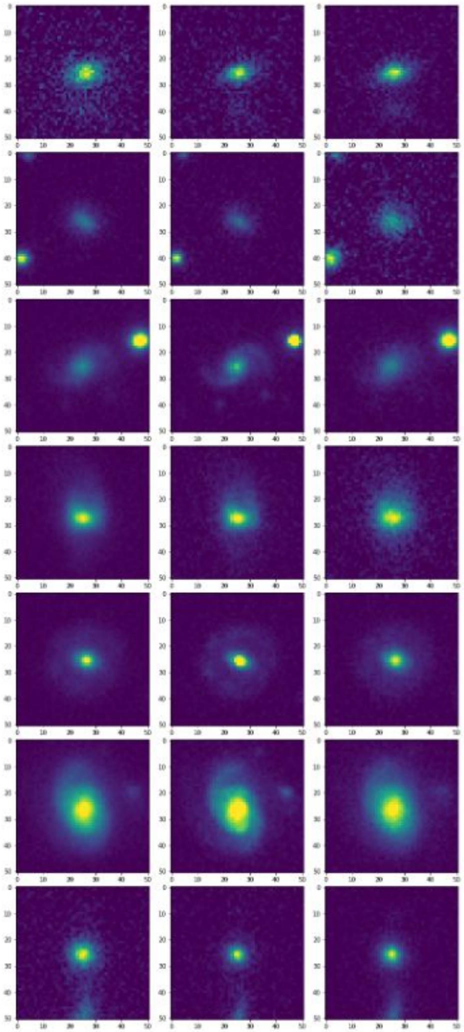
17,995 sources

98

91

33

27 without a proper explanation for some of them probably there is still a problem with the quality but is not really explicit



Concluding without having concluded anything (yet)

This kind of method can allow us to identify problematic epochs in large surveys.

Pro:

- It find problematic epochs neglected by the sigma clip allowing to improve the reliability of the results
- Fast (less than 1 second per lightcurve)
- Could be applied to several tasks and several surveys since it does not requires specific tuning

Cons:

- Sigma clip is still necessary since the method is not capable to recover everything
- There are contaminants (good epochs flagged) although they are really few

Thank you for your Attention


Take-Home Message
Despite the complexity of the task this kind of anomaly detection allow us to flag problematic epochs easily

Base Kernels

The **blur** kernel de-emphasizes differences in adjacent pixel values.

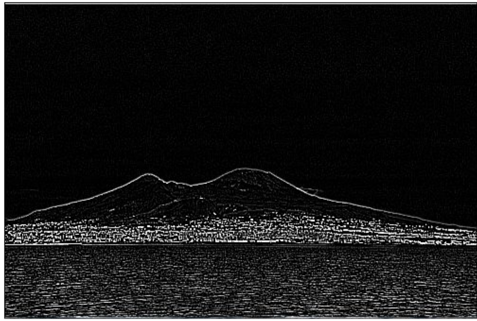
0.0625	0.125	0.0625
0.125	0.25	0.125
0.0625	0.125	0.0625

The **sharpen** kernel emphasizes differences in adjacent pixel values. This makes the image look more vivid.

0	-1	0
-1	5	-1
0	-1	0

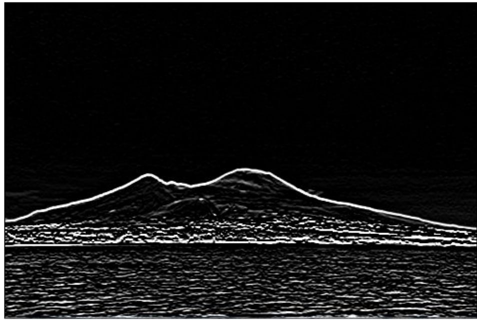
Outline (aka **edge**) highlight large differences in pixel values. A pixel next to a neighbor with close to the same intensity will appear black

-1	-1	-1
-1	8	-1
-1	-1	-1



Sobel kernels are used to show only the differences in adjacent pixel values in a particular direction.

1	2	1
0	0	0
-1	-2	-1



Emboss (similar to sobel) givens the illusion of depth by emphasizing the differences of pixels in a given direction.

-2	-1	1
-1	1	1
0	1	2