Big Data Era Time Domain Astronomy

Pablo A. Estevez

Joint work with

Guillermo Cabrera-Vives[*†‡], Ignacio Reyes[§*], Francisco Förster[†*], Juan-Carlos Maureira[†] and Pablo Huentelemu[§]

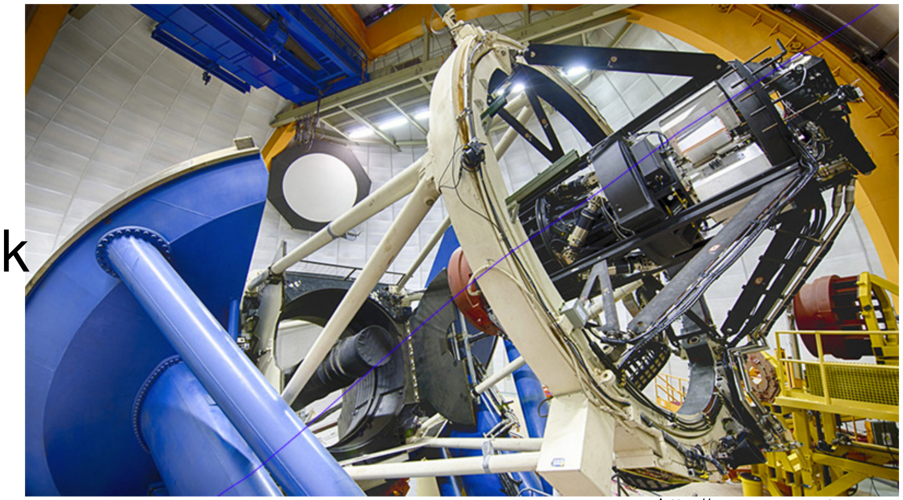
[*]Millennium Institute of Astrophysics, Chile
[†]Center for Mathematical Modeling, Universidad de Chile, Chile
[‡]AURA Observatory in Chile
[§]Department of Electrical Engineering, Universidad de Chile, Chile

Contents

- High cadence Transient Survey (HiTS)
- Deep Learning: Convolutional Neural Networks
- Correntropy Filter for Detecting Transients in Image Sequences
- Conclusions

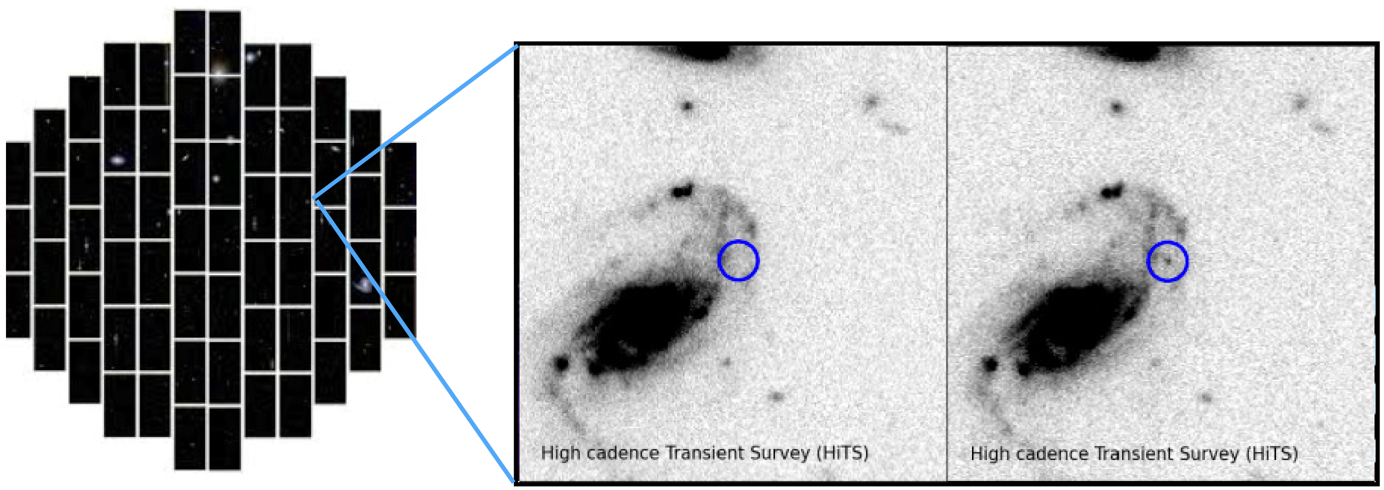
HiTS: High cadence Transient Survey (F.Förster et al.)

Scientific Objective:
Find evidence of Shock Breakout



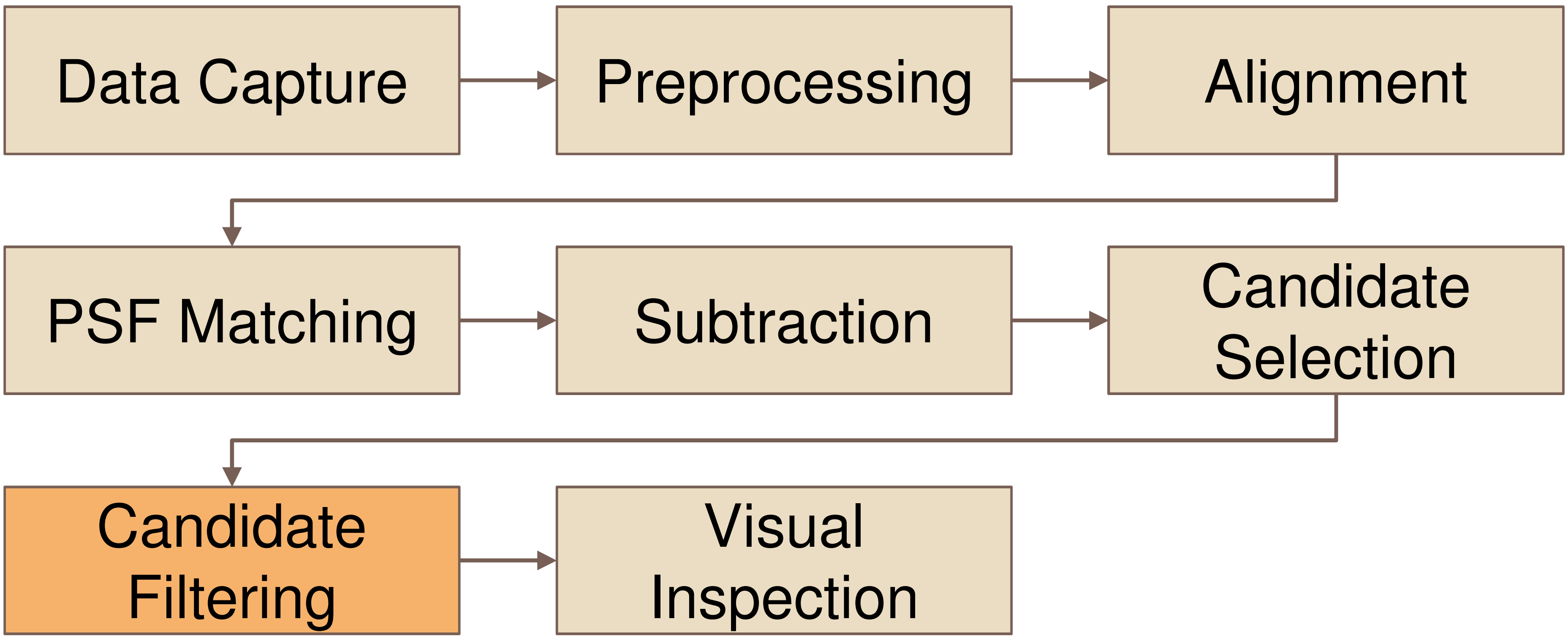
http://www.symmetrymagazine.org/

Dark Energy Camera (DECAM)
at Cerro Tololo, Chile
512 Mpixels
64 CCDs



1 HiTS field = 64 CCD arrays ~512 Megapixels per field and epoch

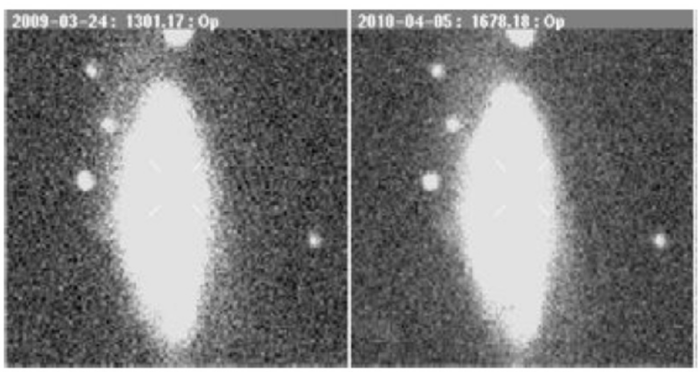
HiTS Image Reduction Pipeline

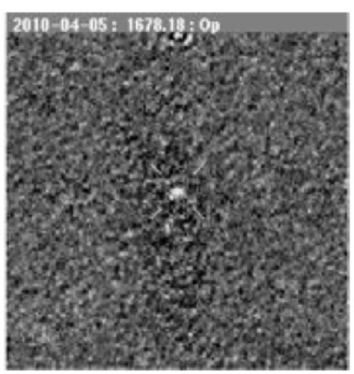


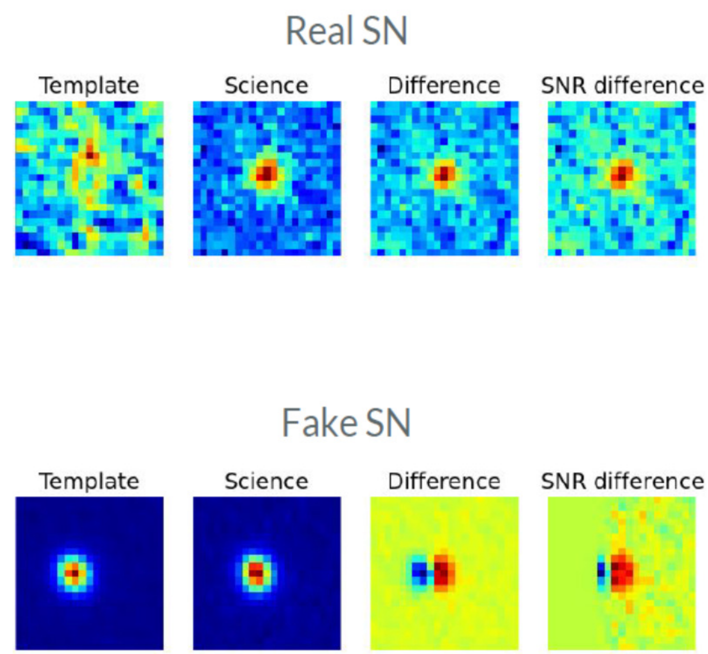
- At this point, candidates are dominated by artifacts 1:10K
- ML to find the needles in the haystack

Image Differencing

Problem description



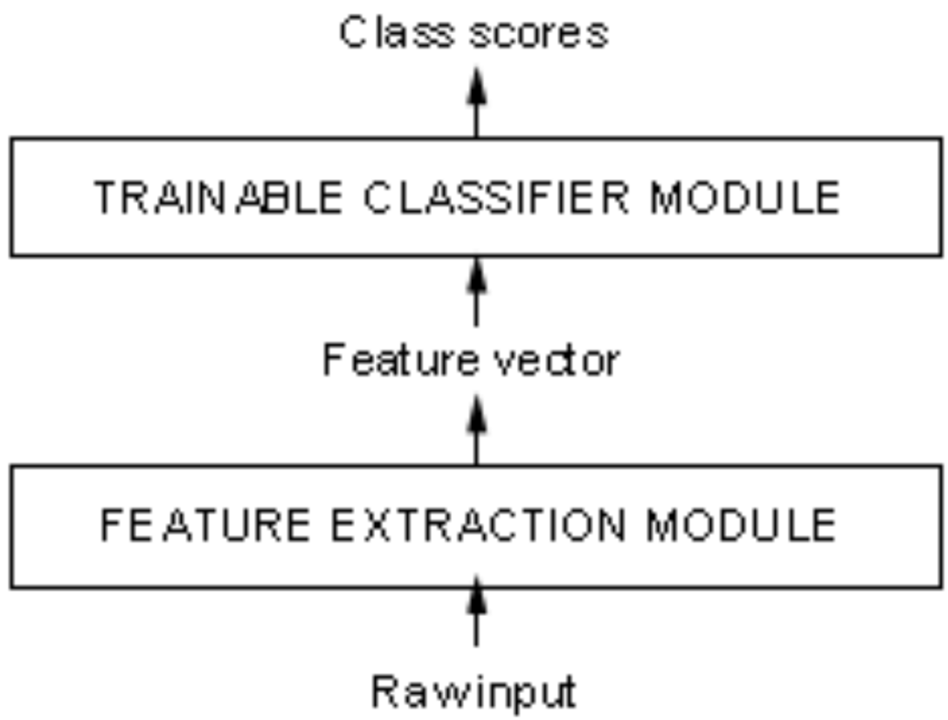




- Example of template (reference), science (current image), difference image, SNR difference image
- Image stamps of 21x21 pixels

Traditional Pattern Recognition Model

Random Forest classifier



Feature engineering:
56 handmade features

Inspired in Movie “Inception”

Francisco Forster

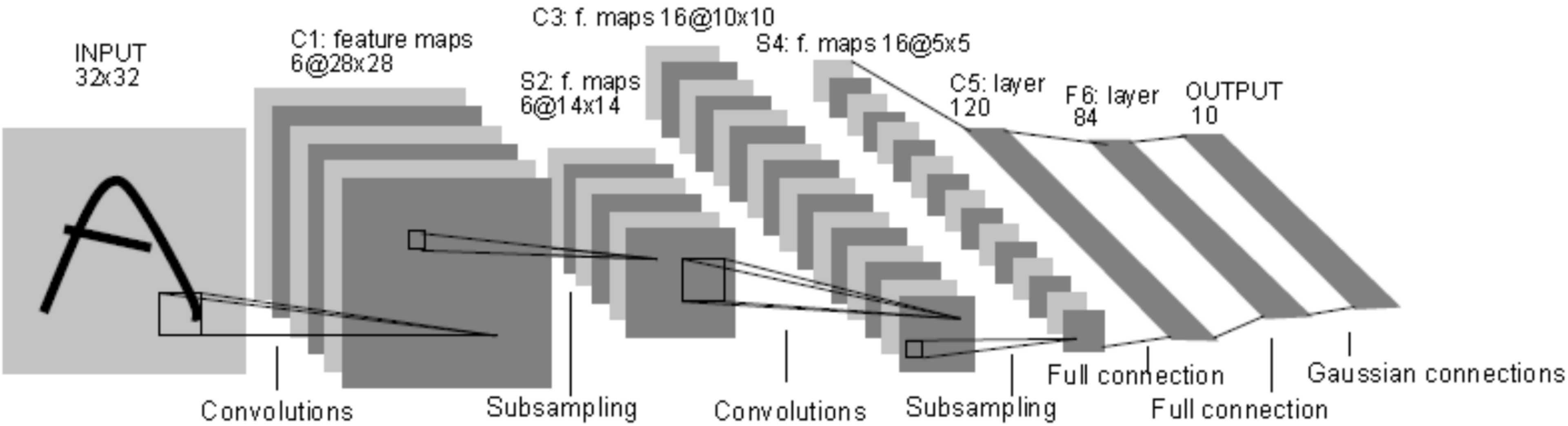
Guillermo Cabrera

Deep Learning

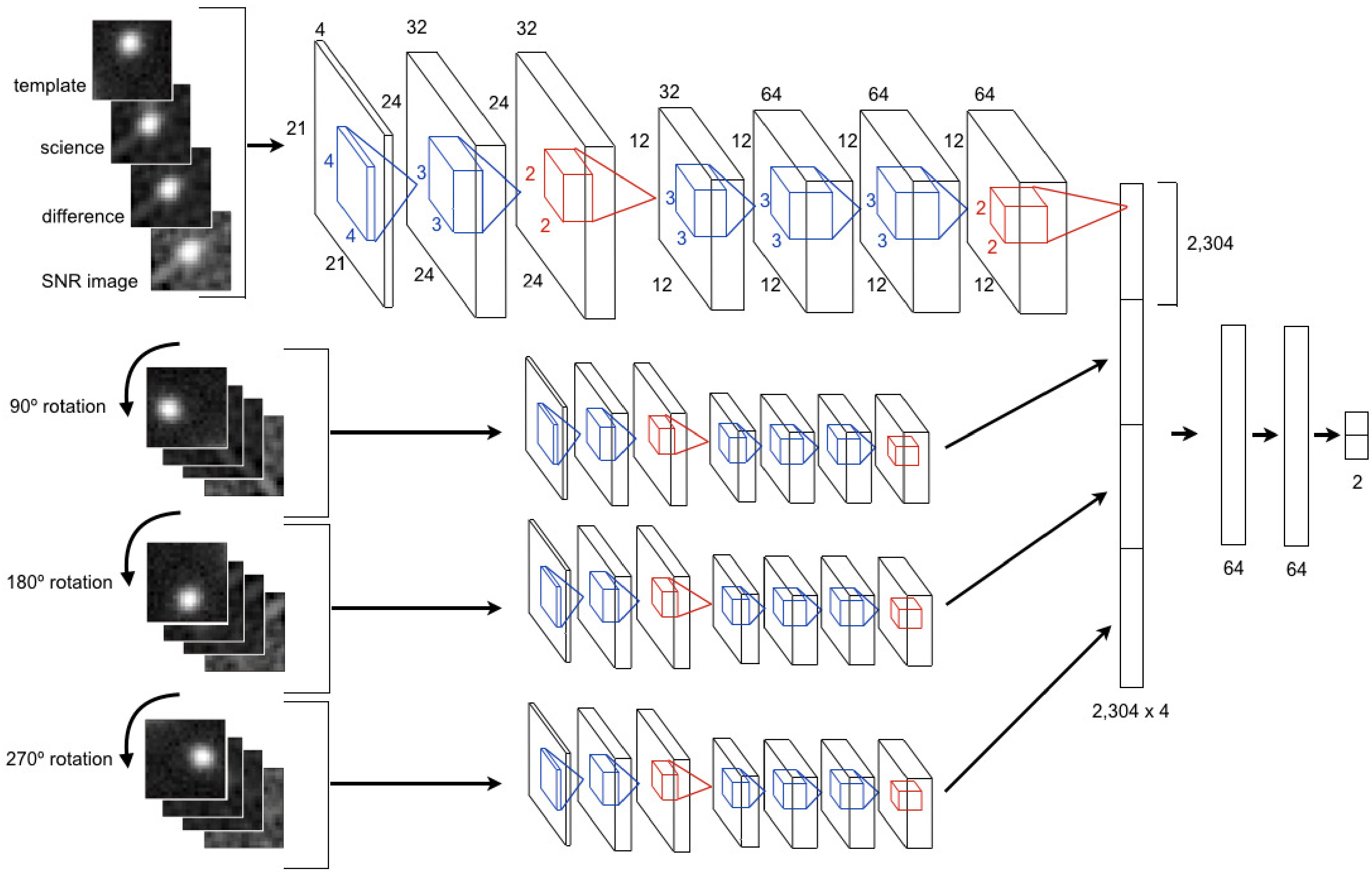
- Deep Learning is achieving impressive results in data science.
- They are based on neural networks, and recent breakthroughs are due to:
  - Large datasets (e.g. millions of images)
  - Faster algorithms and machines
  - New way of dealing with overfitting
- Most common type: Convolutional Neural Nets (ConvNets)

LeNet5 (1998)
Y. LeCun et al.

- Architecture designed to process images (handwritten numbers) including invariances to traslation, scaling and distortion
- It combines convolutional and subsampling (pooling) layers



Convolutional Neural Nets Applied to HiTS



Transients vs Artifacts

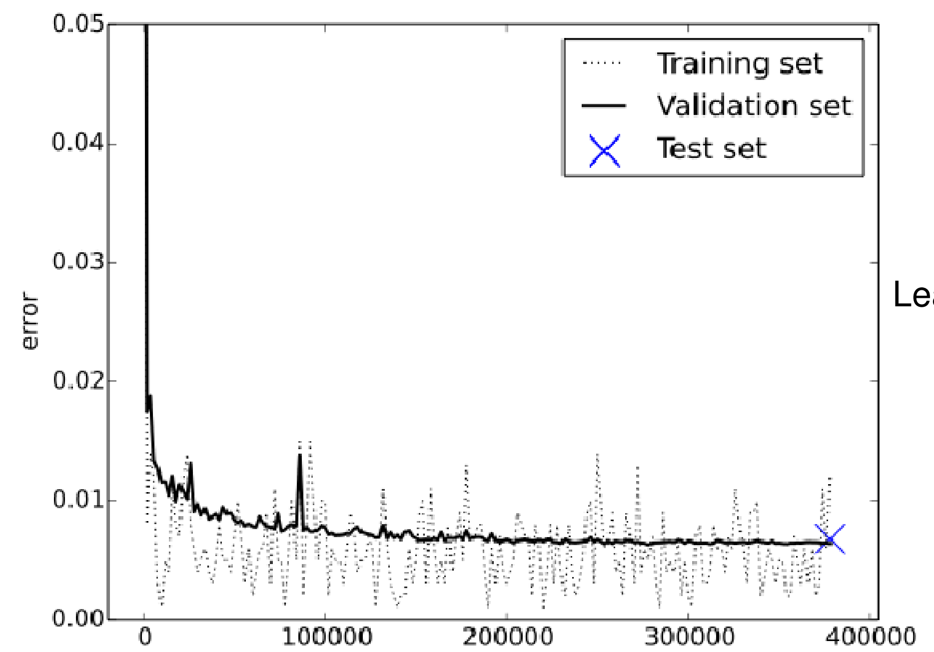
Training: Simulated and Real Data from HiTS 2013

Data:

- 802,087 non-transients (negatives) + 802,087 simulated transients (positives).
- 1,250,000 for training, 100,000 for validating, 100,000 for testing.

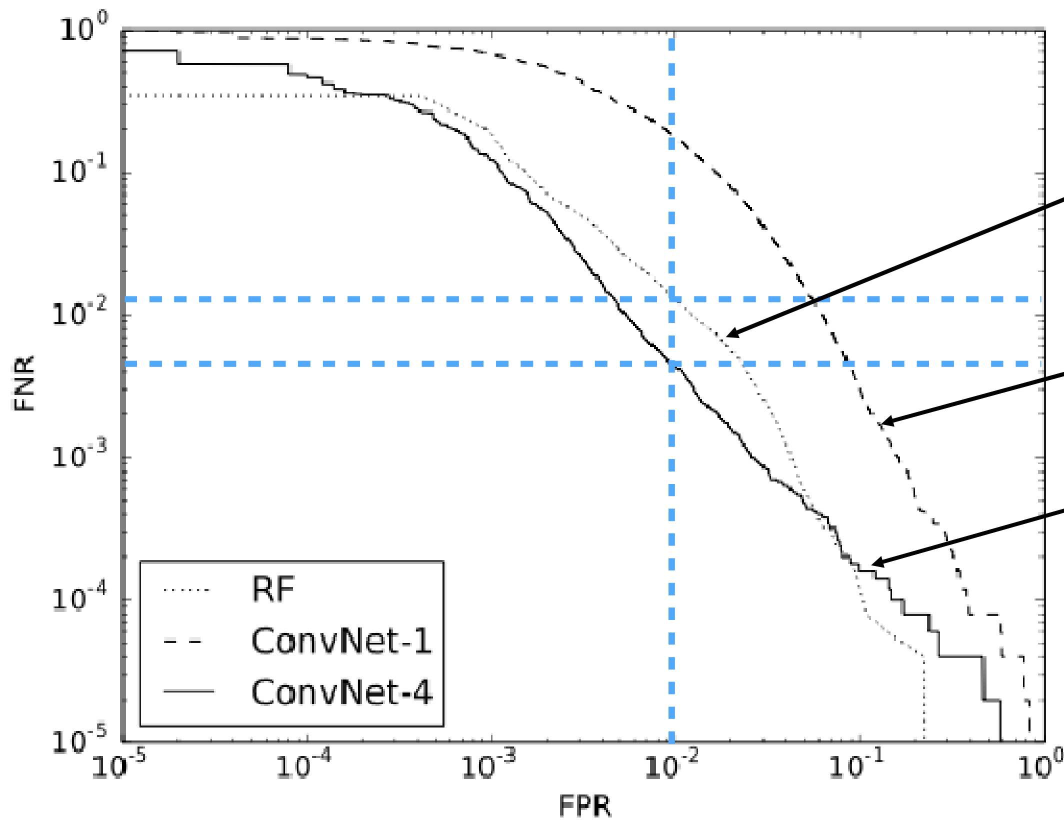
Training:

- Stochastic gradient descent (SGD) with batches of 50 examples.
- Learning rate reduced to half every 50,000 iterations
- Implemented using Theano and took approximately 37 hours to train on a NVIDIA TESLA K20 GPU.



Learning Curve

Detection Error Tradeoff (DET)



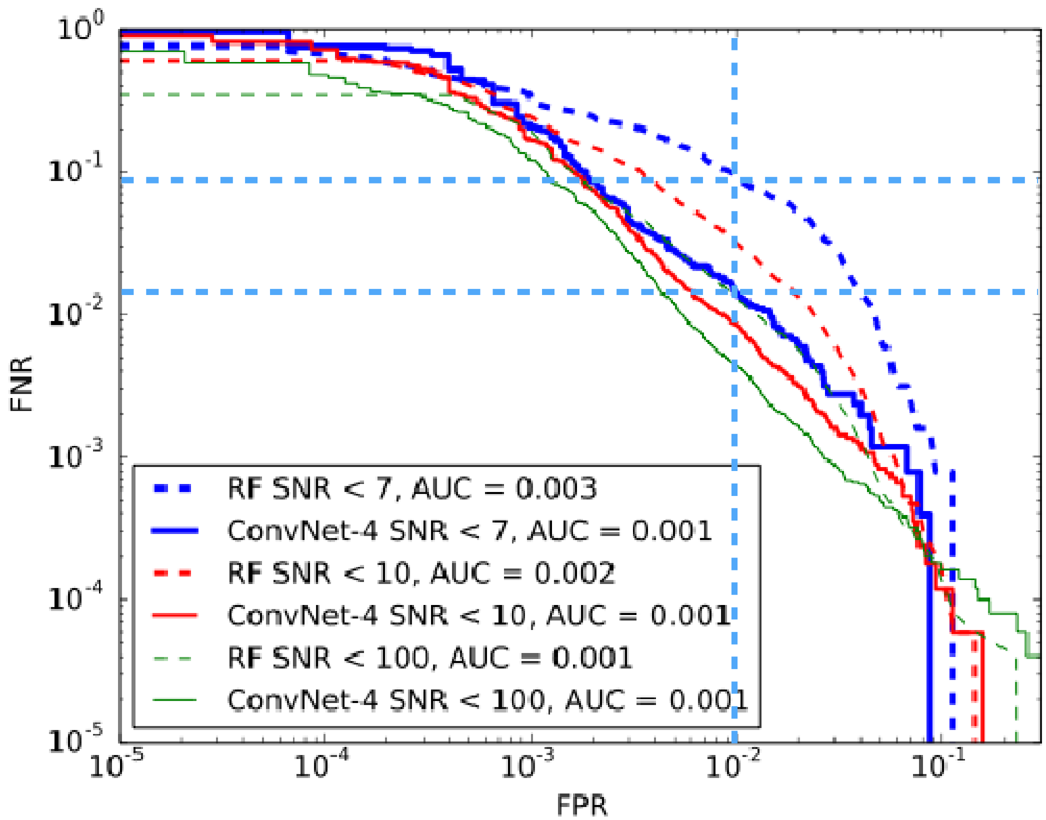
RF: random forests +feature engineering approach (56 handmade features)

ConvNet-1: only the difference image.

ConvNet-4: using all 4 images.

- Consider 100,000 candidates per night, from which around 10 are real transients, and we want to visually inspect 1,000.   FPR~10^{-2}
- By using our ConvNet model we reduce the FNR from ~10^{-2} to ~$3x10^{-3}$

Detection Error Tradeoff (DET)



- Results for SNR < 7 are shown in blue
- FNR is reduced from ~10-1 to ~3x10-2 for very faint sources.

Test on 2015 HiTS Campaign (real data)

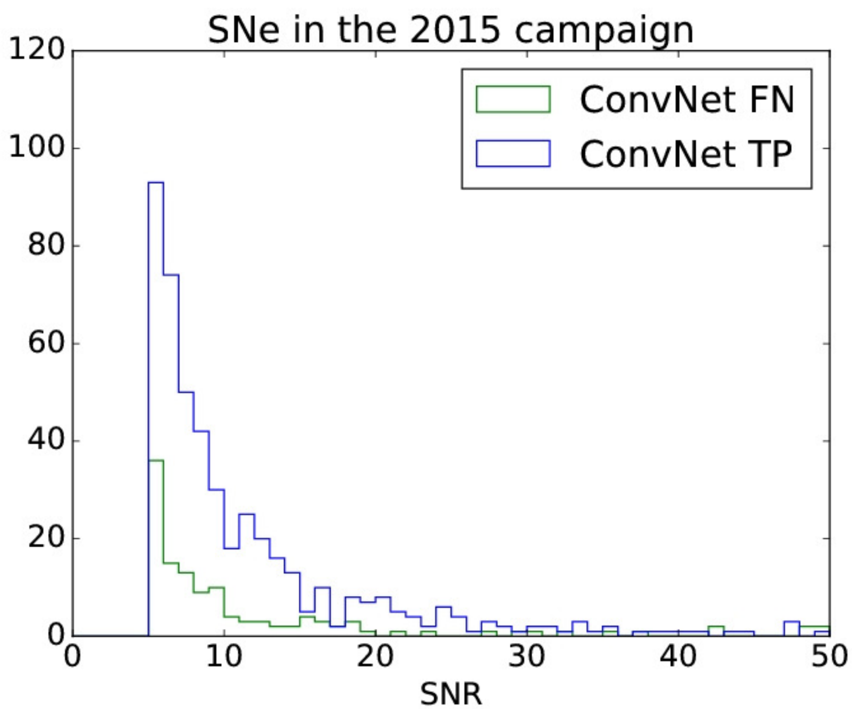
- All the difference images with SNe candidates were analyzed (Total: 628)

Method	# Correct Detections	FNR
RF	439	0.300
ConvNet	487	0.224

Comparison ConvNet vs RF

- At low SNR, ConvNets has a much lower FNR than RF

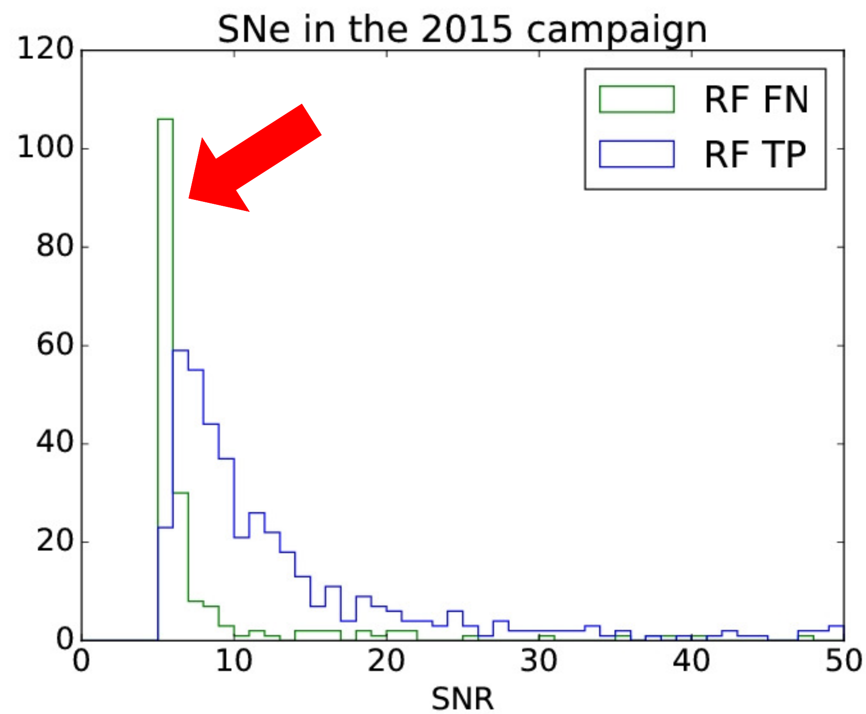
ConvNet



Cases

SNR

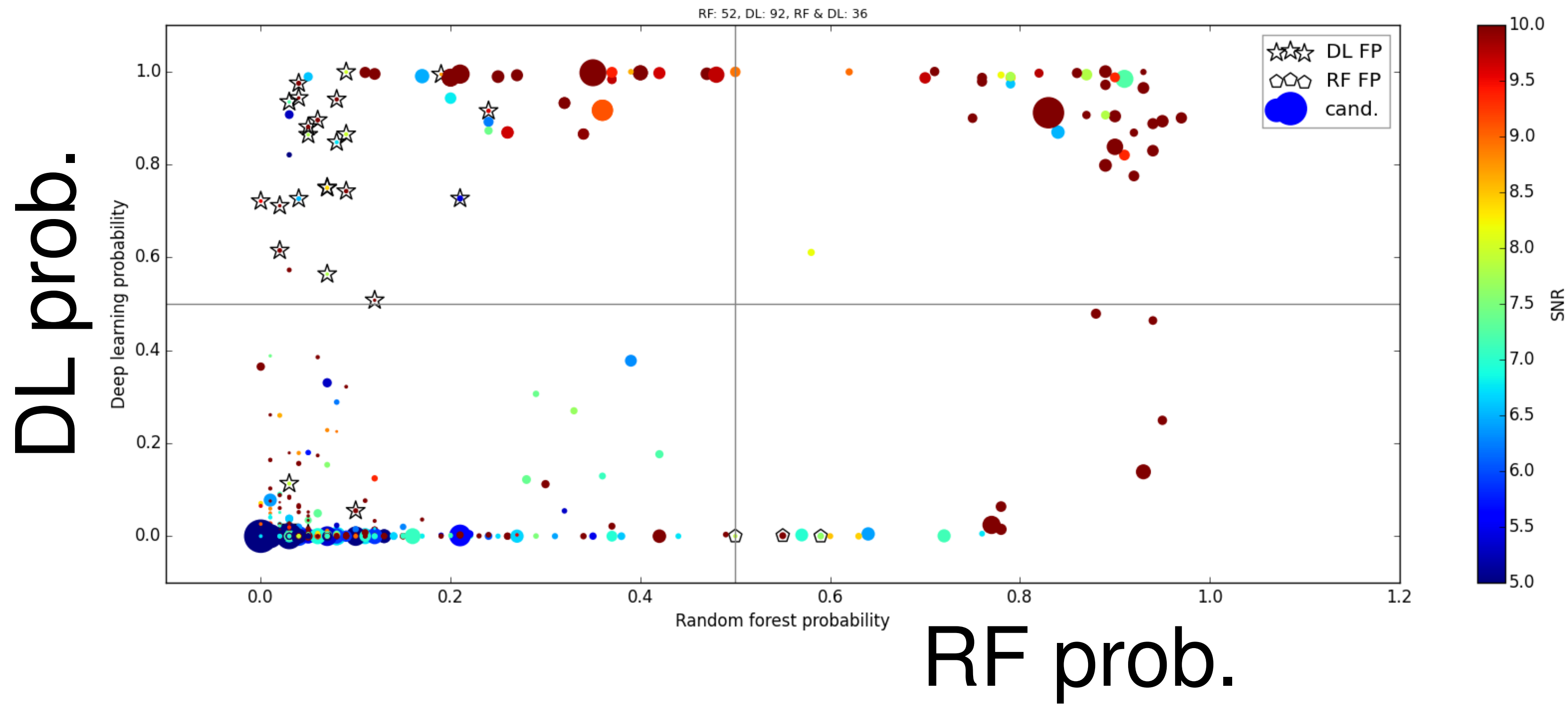
RF



Cases

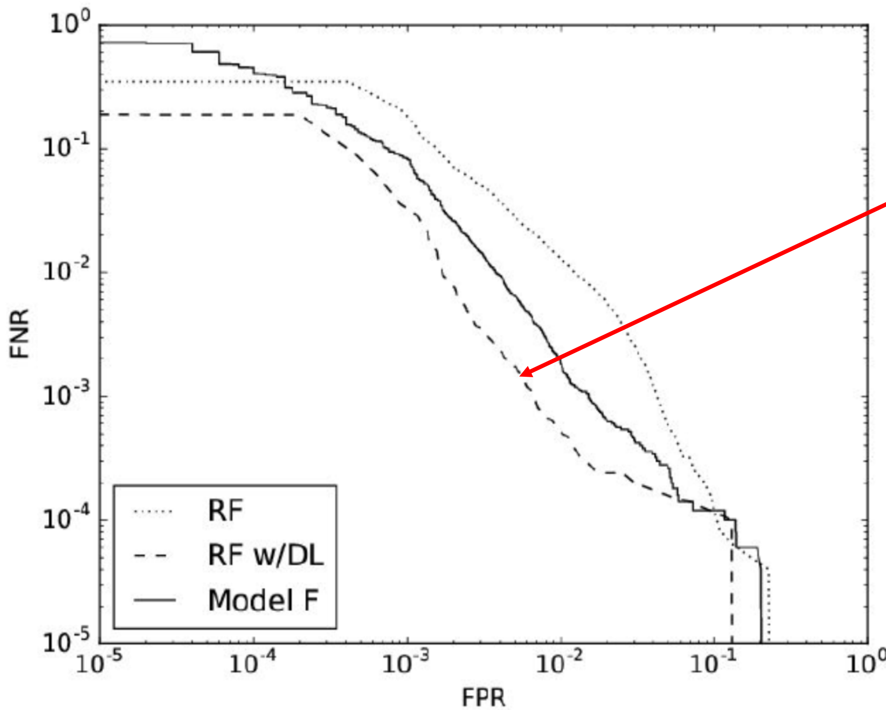
SNR

Synergy between Classifier´s Outputs



- Circles show transient candidates
- Color indicates SNR level.
- This plot is for a given feature

Deep Learning + Random Forests



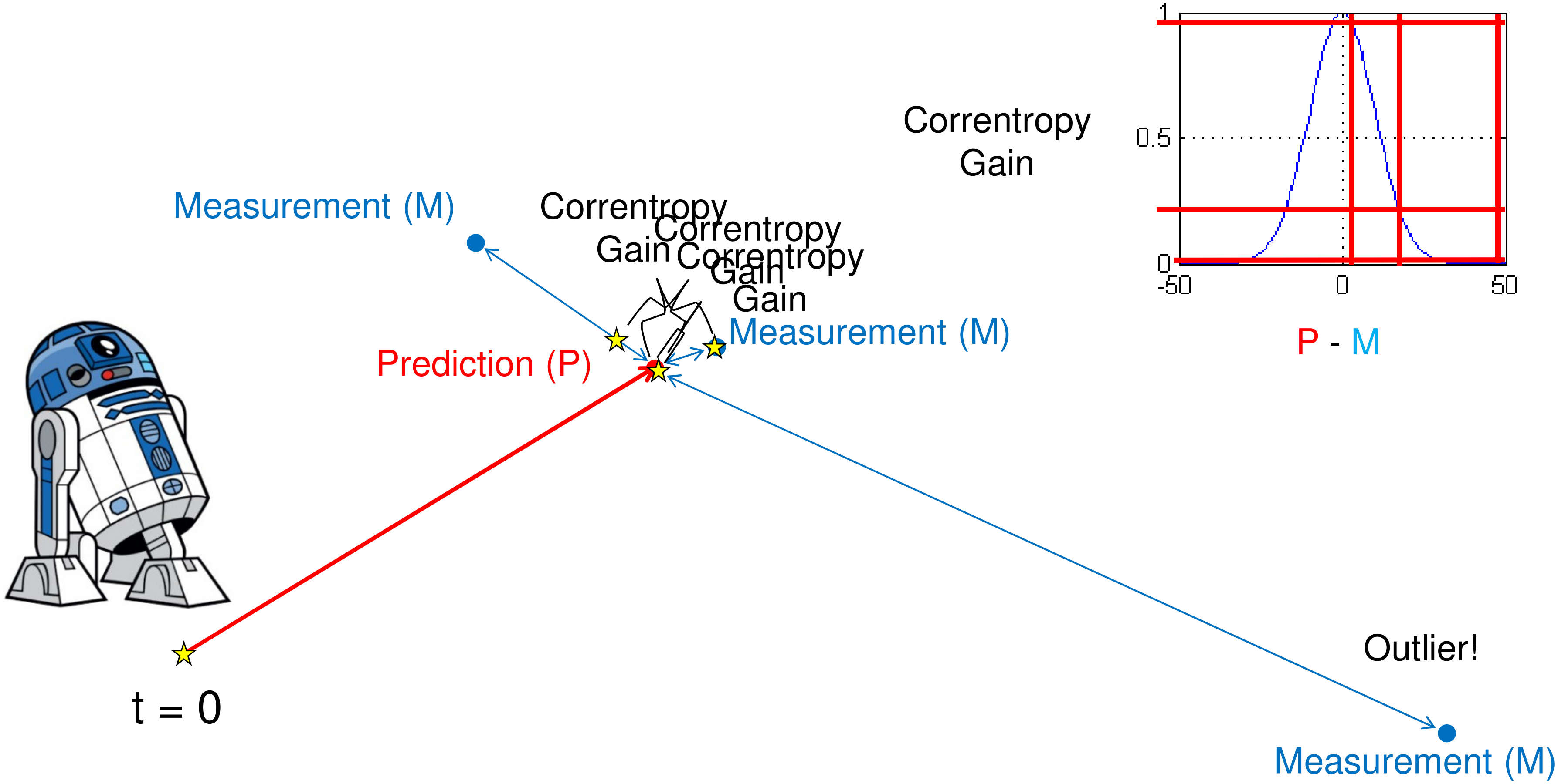
DL+RF:

The DL network output is fed into the RF model.

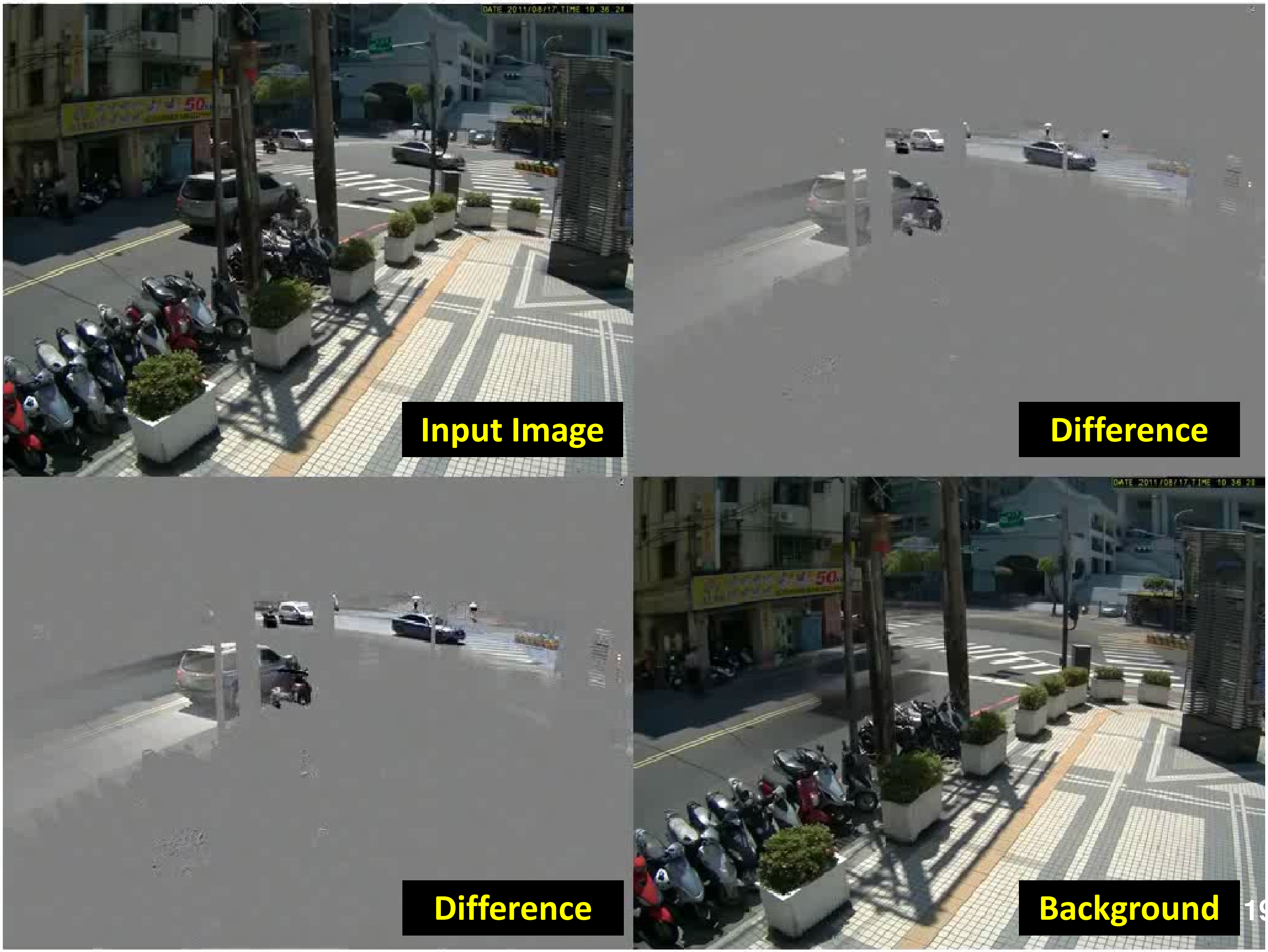
Correntropy Filter

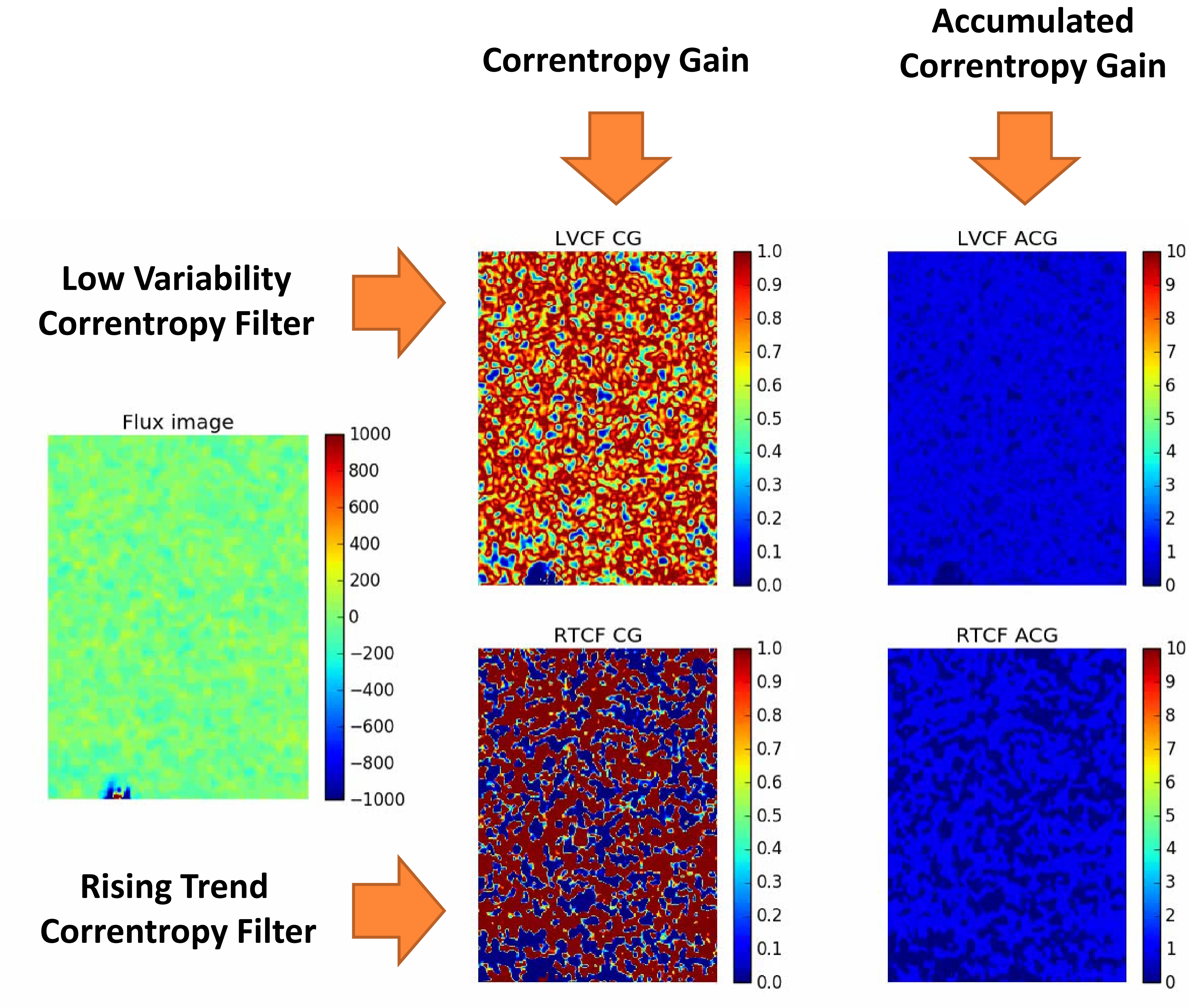
- Kalman Gain is replaced with the Correntropy Gain with Kernel Bandwidth σ:

$$\hat{x}_t = \hat{x}_{t-1} + (z_t - \hat{x}_{t-1})\, G_\sigma(z_t - \hat{x}_{t-1})$$

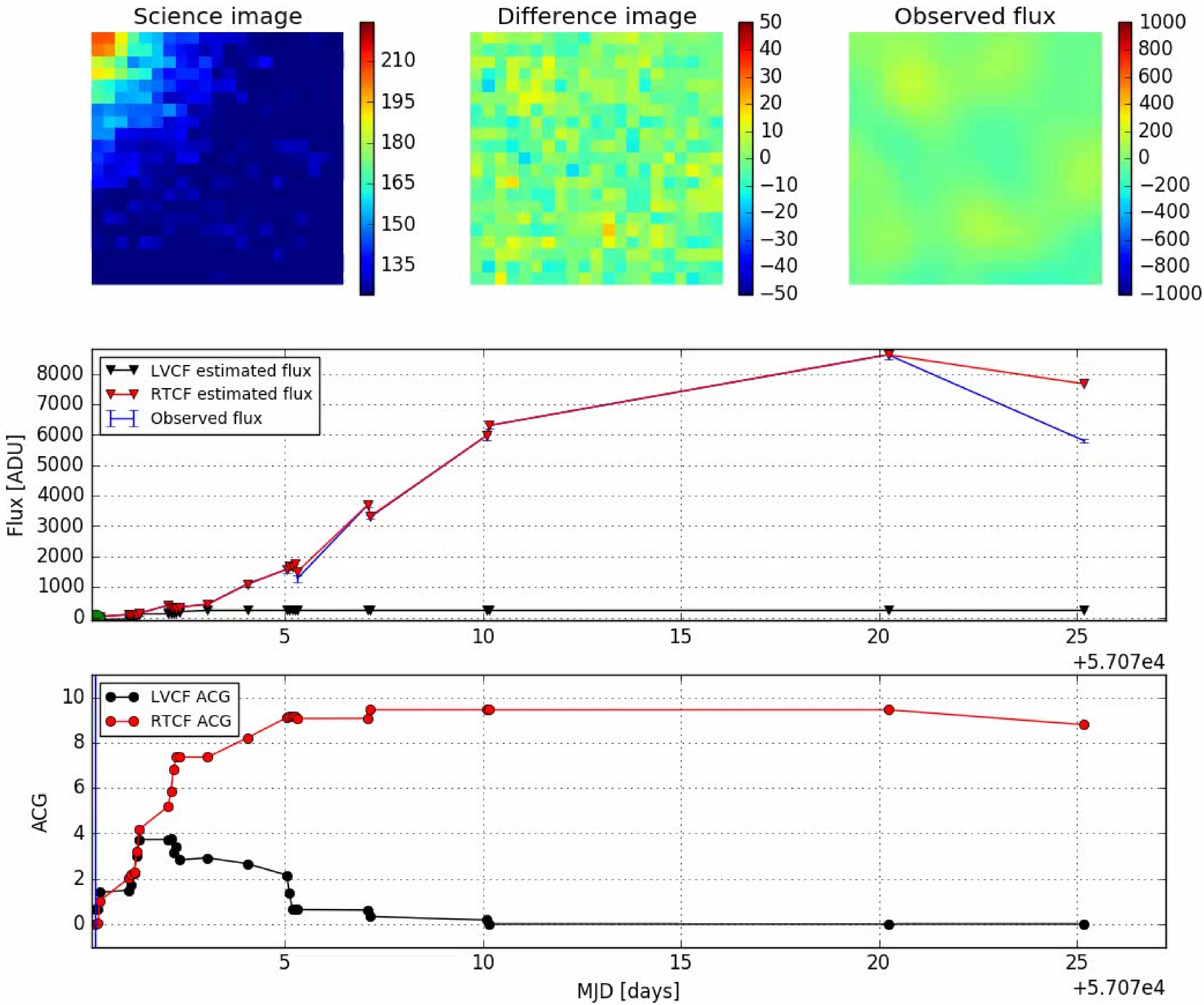


Frame 100/500





Correntropy Gain
Accumulated Correntropy Gain
Low Variability Correntropy Filter
Rising Trend Correntropy Filter
Flux image
1000
800
600
400
200
0
−200
−400
−600
−800
−1000
LVCF CG
RTCF CG
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
LVCF ACG
RTCF ACG
10
9
8
7
6
5
4
3
2
1
0



Science image
210
195
180
165
150
135
Difference image
50
40
30
20
10
0
−10
−20
−30
−40
−50
Observed flux
1000
800
600
400
200
0
−200
−400
−600
−800
−1000
LVCF estimated flux
RTCF estimated flux
Observed flux
Flux [ADU]
8000
7000
6000
5000
4000
3000
2000
1000
0
5
10
15
20
25
+5.707e4
LVCF ACG
RTCF ACG
ACG
10
8
6
4
2
0
MJD [days]
+5.707e4

Results

HiTS Campaign 2014

Total	32
New candidates	**7**
New variable stars	1
False positives	15
Rediscovered HiTS candidates	9

HiTS Campaign 2015

Total	78
New candidates	**10**
New variable stars	16
False positives	52

Conclusions

- The proposed convolutional neural network (ConvNet) approach is useful to detect supernovae.
- Our approach outperforms a previous method based on feature engineering and a random forest (RF) classifier, particularly at low SNR.
- Both models maybe complementary and further research is needed
- Image sequence analysis useful to reduce error rates at the expense of late detections

Thank you

FACULTY OF PHYSICAL AND MATHEMATICAL SCIENCES
UNIVERSIDAD DE CHILE

MILLENNIUM INSTITUTE OF ASTROPHYSICS