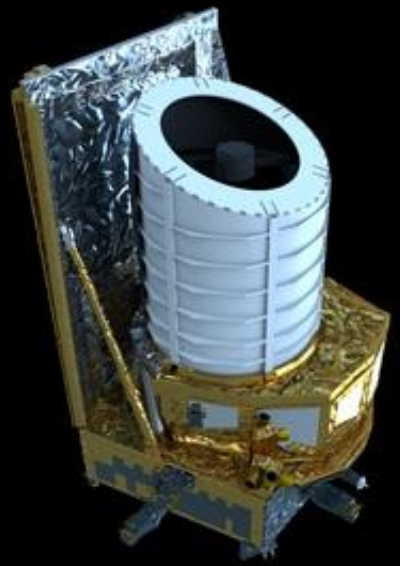
Gravitational lens finding challenge

Elodie Savary

EPFL, Switzerland

Euclid

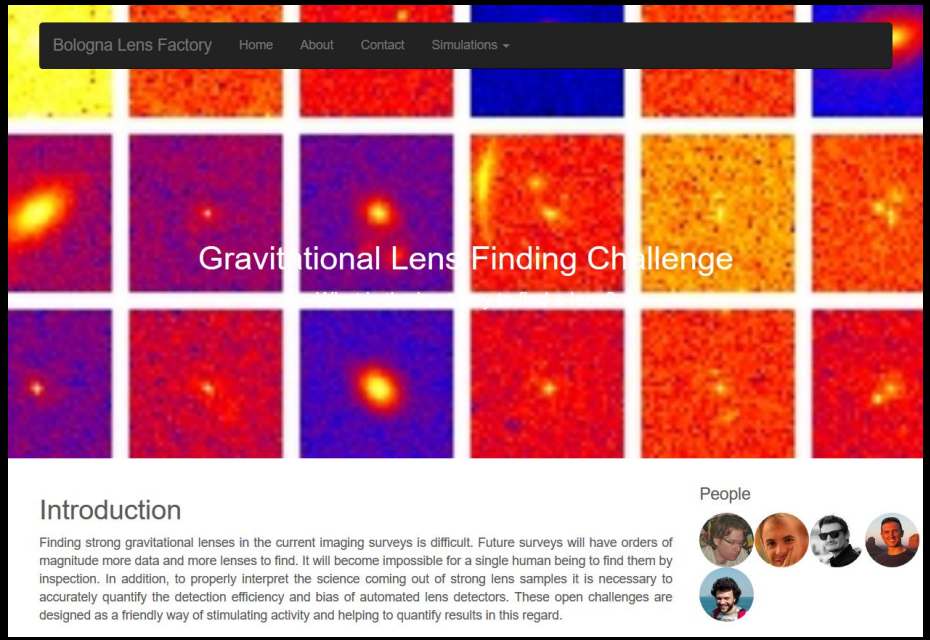
- 15’000 deg^2
- 12 billions sources
- 1.5 billion galaxies



ESA/ATG medialab

Expected : ~ 200’000 galaxy-galaxy lenses

Gravitational Lens Finding Challenge



Bologna Lens Factory

http://metcalf1.difa.unibo.it/blf-portal/gg_challenge.html

Challenge 1.0

Data:

- C1: space-based survey (one band): 100’000 images.
- C2: ground-based survey , 4 bands (I,G,R,U): 100’000 images.

Challenge 1.0

Main results: (Metcalf, et al 2018)

- Machine learning based methods perform better than human inspection except for very complex lenses (Jackpot lenses).
- Color information is very important.
- The most difficult false positives were underrepresented in the simulations.

Challenge 1.0

Main results: (Metcalf, et al 2018)

- Machine learning based methods perform better than human inspection except for very complex lenses (Jackpot lenses) <= ? Same simulations in the training and the test set.
- Color information is very important.
- The most difficult false positives were underrepresented in the simulations.

Challenge 2.0

Data:

Euclid-like observations in VIS and NISP (J, Y and H bands).

- **Training set** :100’000 images + catalog of lens properties
- **Testing set** :100’000 images

Lens Properties catalog

General: index number, source ID

Source properties (before lensing): redshift, EUC_VIS magnitude

Lens properties: redshift, number of pixels above 1x sigma, total flux in those pixels, halo mass, stellar mass, magnitude.

Lensing parameters: critical curve center coordinates, Einstein area of large critical curves, number of critical curves, distance of source from center of caustic, area of caustic, average surface brightness contrast between lens and source in pixels above 1 sigma, number of sources added

Lensed source properties: number of pixels above 1x sigma, total flux in those pixels, number of images, effective magnification of source

Other informations: number of companions m < 25, within 3, 5, 10 arcsec

Lens definition

Lenses: n_source_im > 0, mag_eff > 1.6 and n_pix_source > 20:

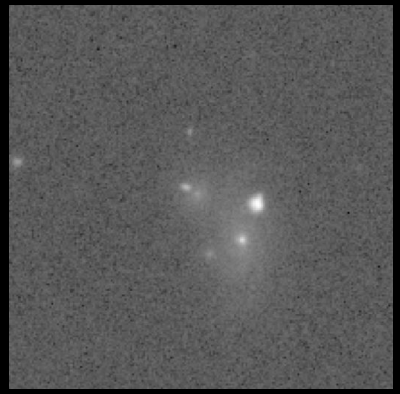
49218 images

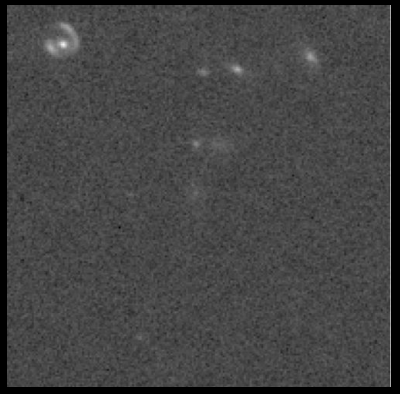
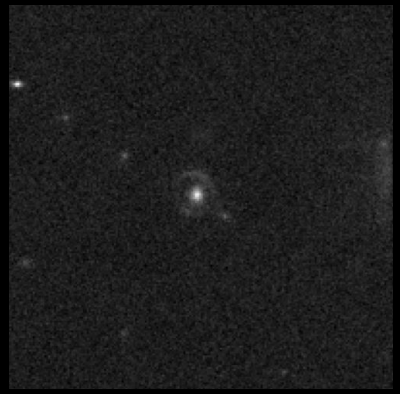
Non lenses: n_source_im = 0:

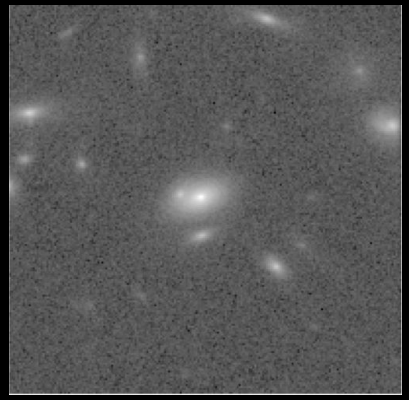
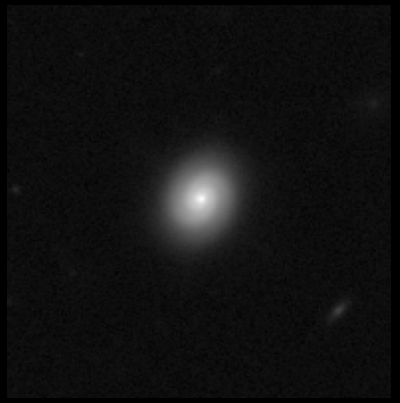
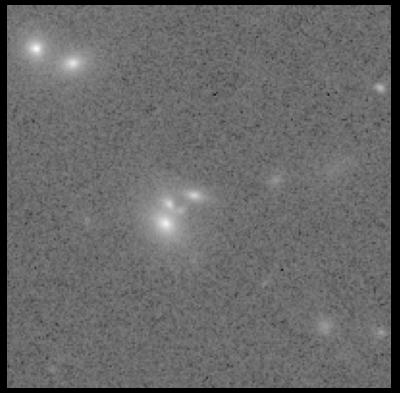
17033 images

Lenses

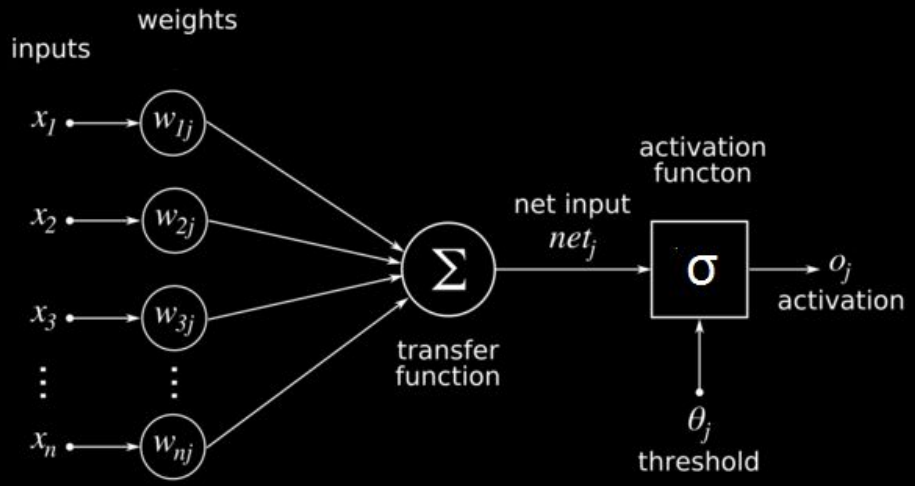
Non lenses

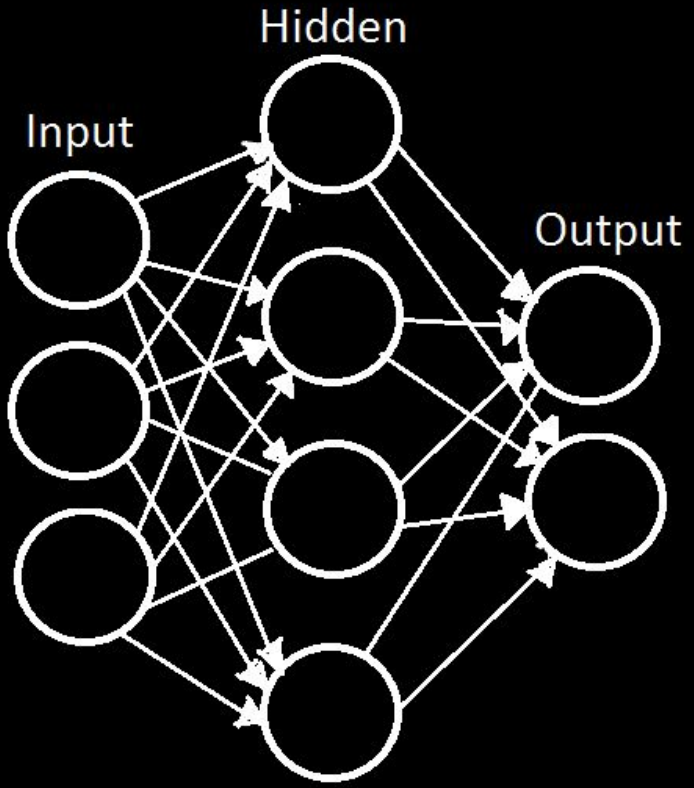


Artificial neural networks



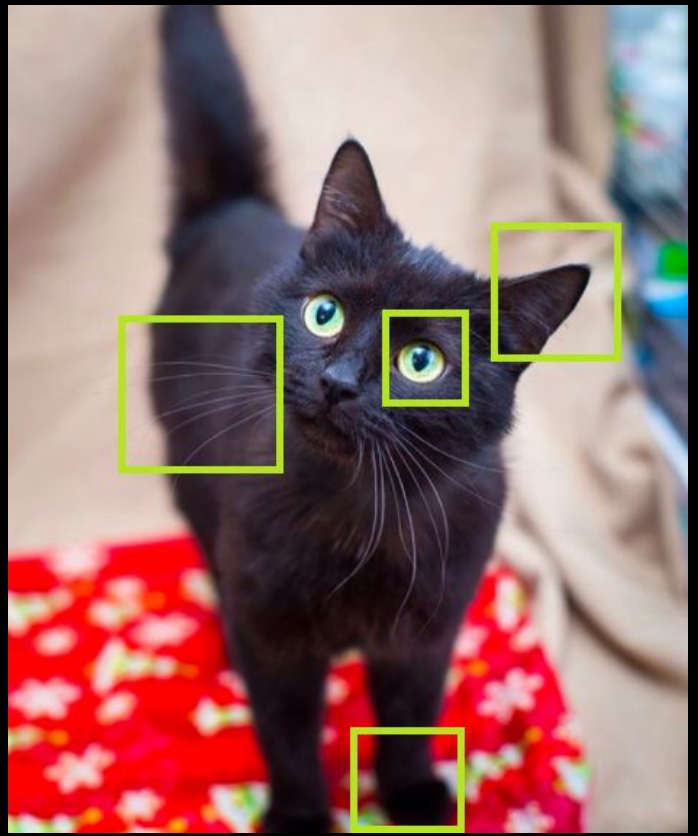


Convolutional neural networks

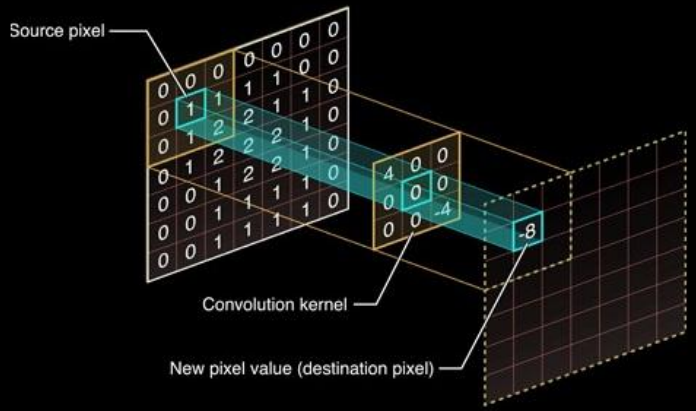
Does the image contain a cat?

Convolutional neural networks

Does the image contain a cat?

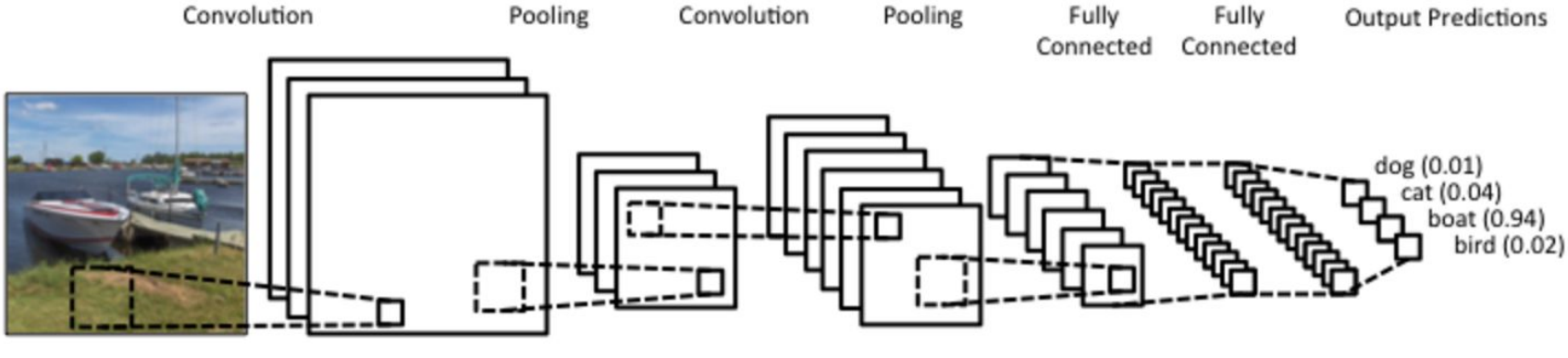


Convolutional neural networks



Credit : Deep learning tutorial CPVR2012

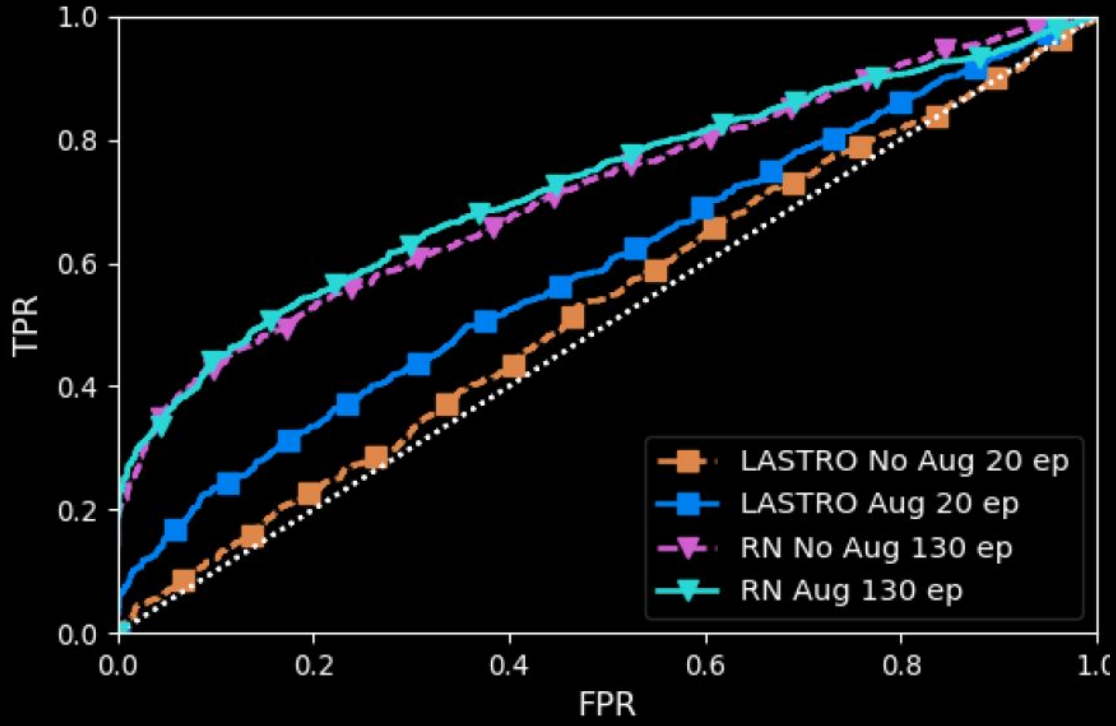
Convolutional neural networks



CNN architectures

- Deep convolutional neural networks as strong lens detectors, Schaeffer et al. 2017
- Deep residual learning for image recognition, Kaiming He et al. 2016
- EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks Tan et al. 2019

ROC original training set

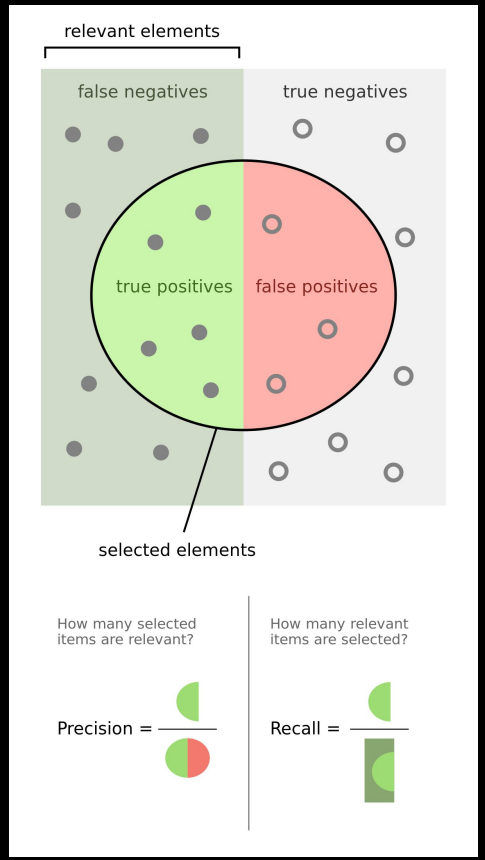


Credit: A. Variu, D. Forero Sanchez, Y Ilichenko

Back to the definition of lenses

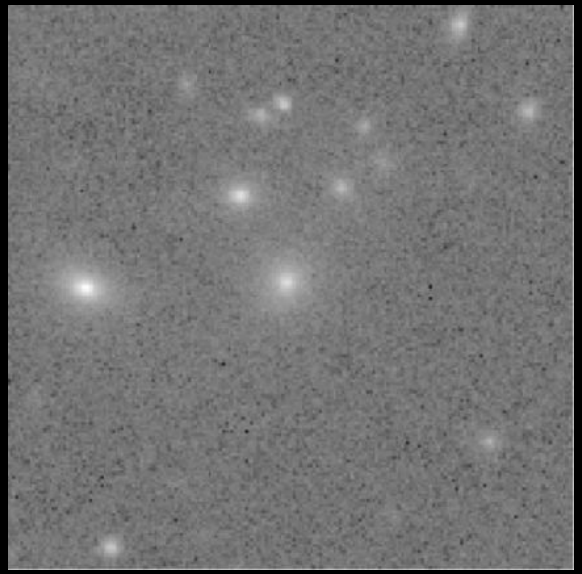
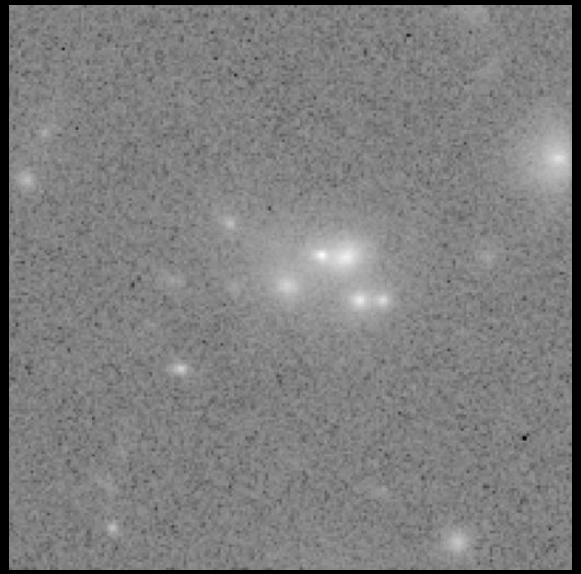
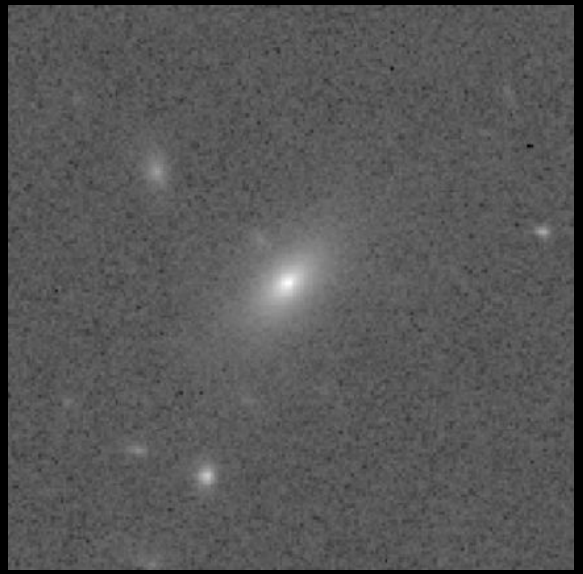
Metric of the Challenge : Max (F_β), $\beta^2 = 0.001$

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}$$



Back to the definition of lenses

n_source_im > 0, mag_eff > 1.6 and n_pix_source > 20

Lens Properties catalog

General: index number, source ID

Source properties (before lensing): redshift, EUC_VIS magnitude

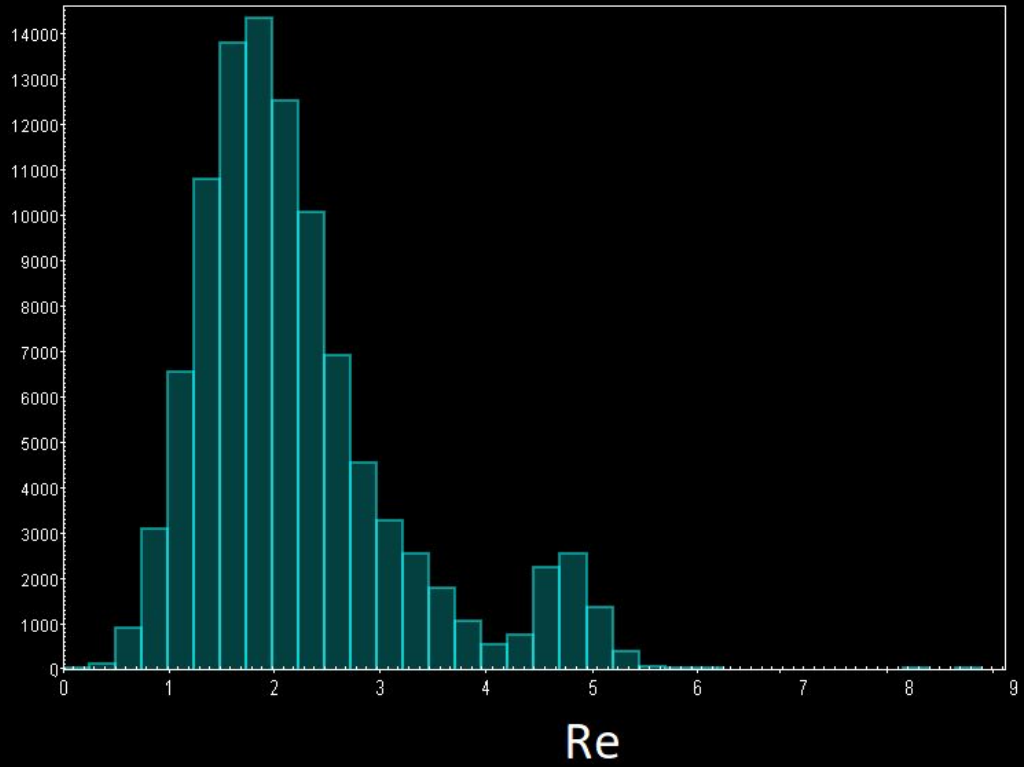
Lens properties: redshift, number of pixels above 1x sigma, total flux in those pixels, halo mass, stellar mass, **magnitude**.

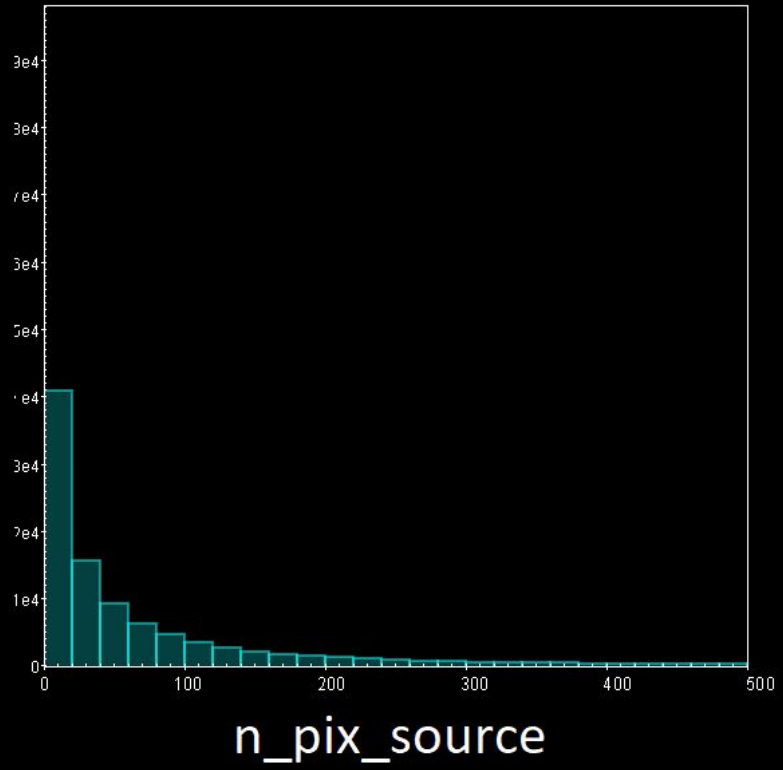
Lensing parameters: critical curve center coordinates, **Einstein area of large critical curves**, number of critical curves, distance of source from center of caustic, area of caustic, average surface brightness contrast between lens and source in pixels above 1 sigma, number of sources added

Lensed source properties: **number of pixels above 1x sigma**, total flux in those pixels, number of images, effective magnification of source

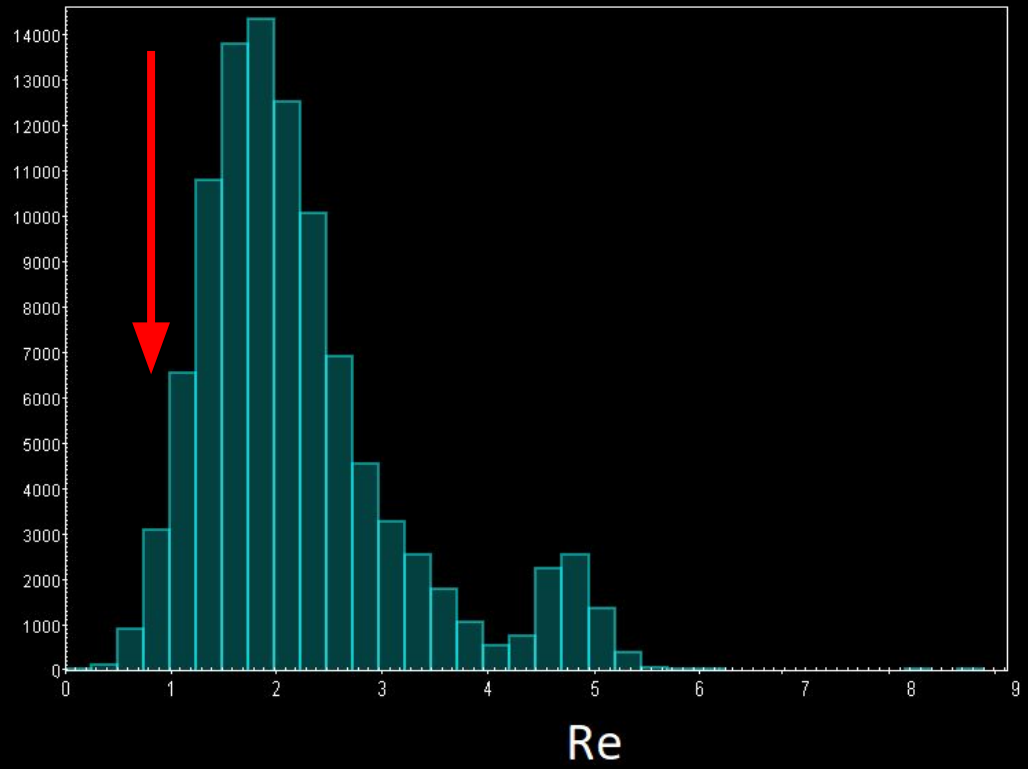
Other informations: number of companions m < 25, within 3, 5, 10 arcsec

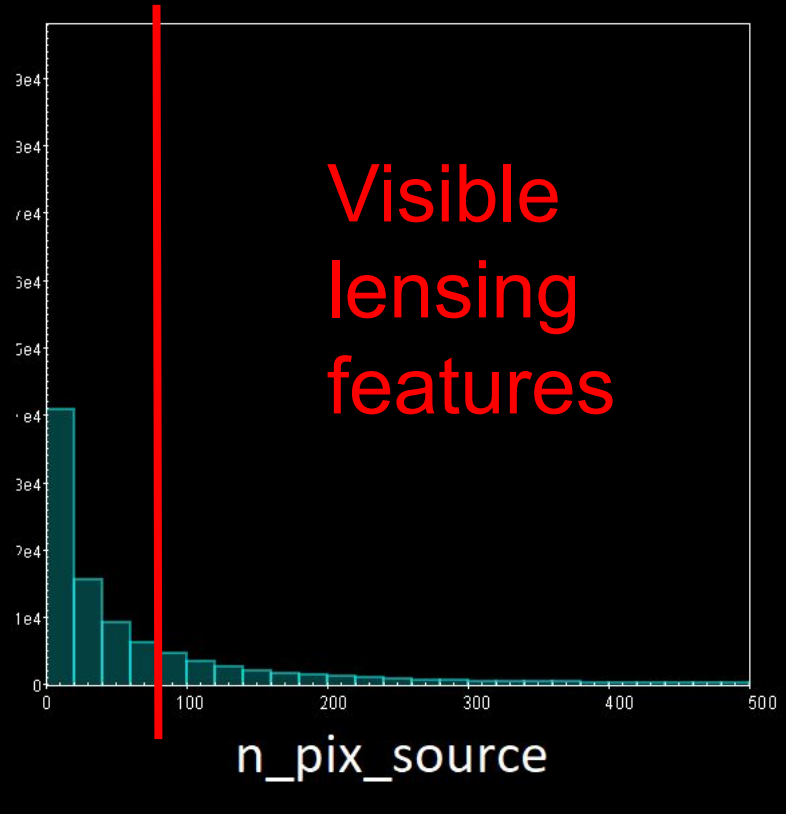
Back to the definition of lenses





Back to the definition of lenses





Visible lensing features

New definition of lenses

- Re < 2 => n_pix_source > 90
- 2 < Re < 2.5 => n_pix_source > 145
- 2.5 < Re < 3 => n_pix_source > 150
- 3 < Re < 3.5 => n_pix_source > 300
- 3.5 < Re < 4 => n_pix_source > 400
- 4 < Re => n_pix_source > 800

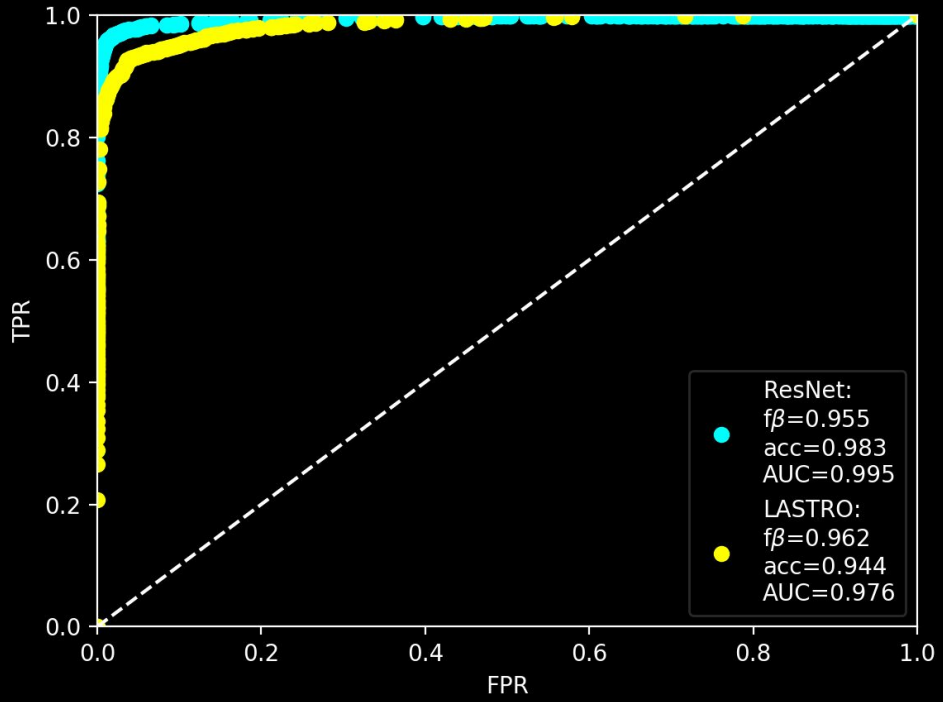
And

- Mag_lens > 20

4337 lenses

ROC modified training set

VIS only



Credit: A. Variu, D. Forero Sanchez, Y Ilichenko

Results with modified definition

Only VIS:

ResNet: Accuracy = 0.983 and FBETA=0.955

LASTRO: Accuracy=0.933 and FBETA =0.962

EfficientNet: Accuracy=0.98 and FBETA =0.99

Results with initial definition

VIS only:

EfficientNet: Accuracy=0.62 and FBETA =0.90

VIS + NIR:

EfficientNet: Accuracy=0.59 and FBETA =0.89

Conclusion & Future Work

The content of the training set is crucial for the quality of the classification.

Future work:

- Is the network is able to find less visible lenses even if trained only on evident lenses?
- Test the classifier on real data.