Seminaire pole ML/DL du CeSAM
4 Mai 2021

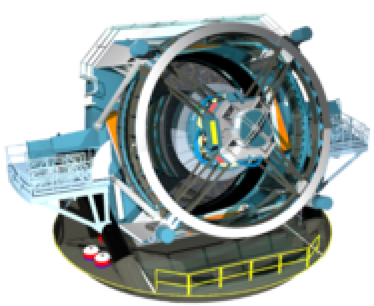
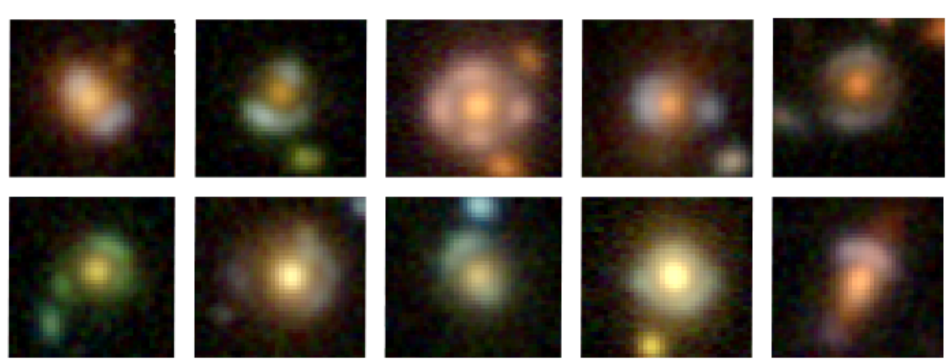
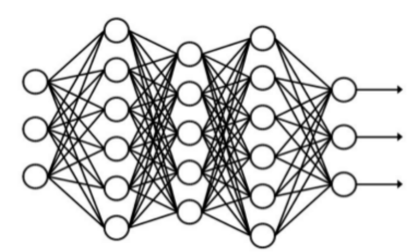
Max-Planck-Institut
für Astrophysik

# Finding and modeling strong gravitational lenses with deep neural networks

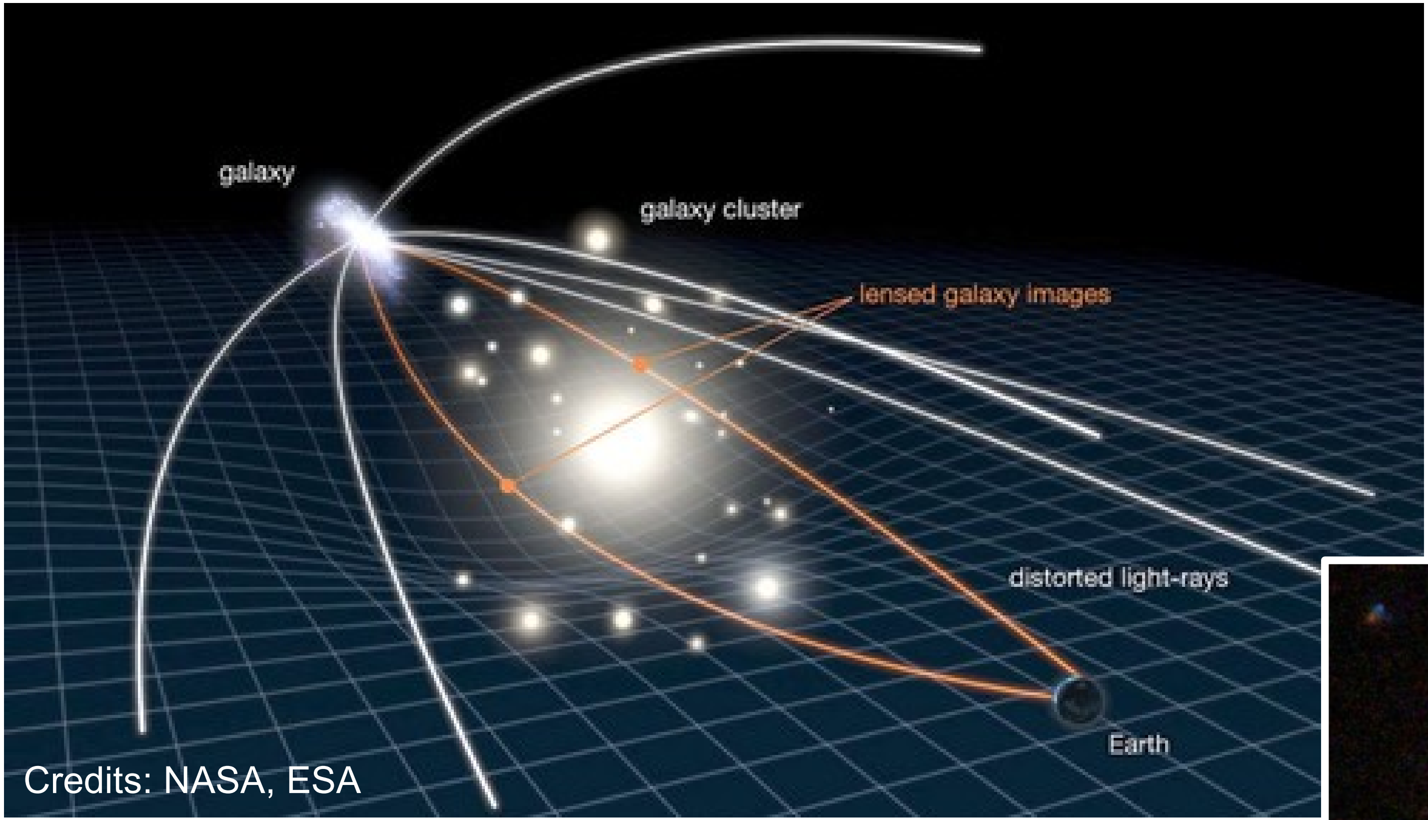
Raoul Cañameras (MPA Garching)

rcanameras@mpa-garching.mpg.de

S. Schuldt, S. Suyu, Y. Shu, S. Taubenberger, T. Meinhardt, L. Leal-Taixe, et al.

# Strong gravitational lensing

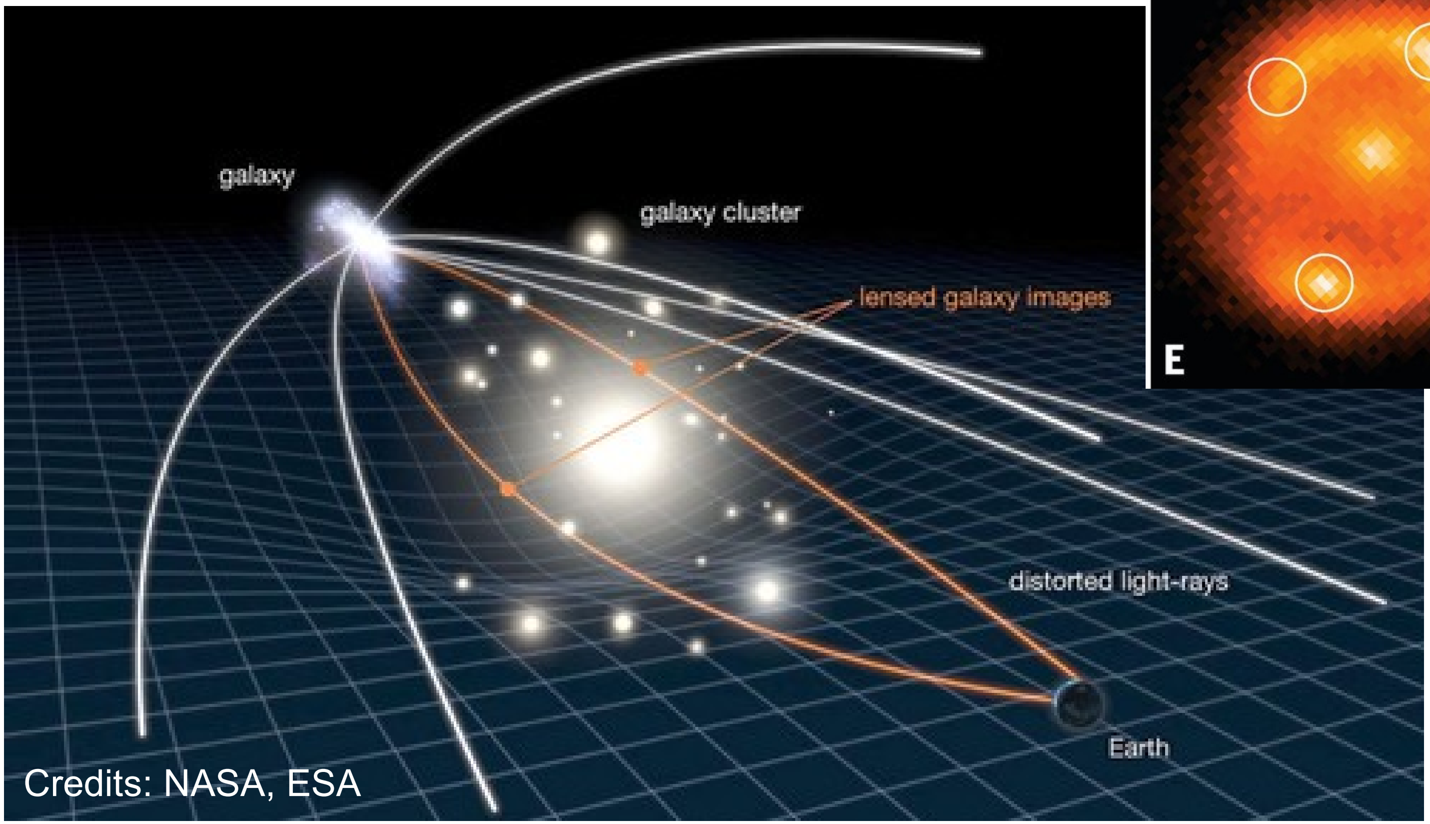




Cosmic Horseshoe (ESA, NASA)

- Strong lensing regime: Elongated arcs and multiple images
- Galaxy evolution and dark-matter:
  - Ideal probe of the total mass in the foreground lens galaxy
  - Detailed studies of strongly magnified background galaxies

# Strong gravitational lensing



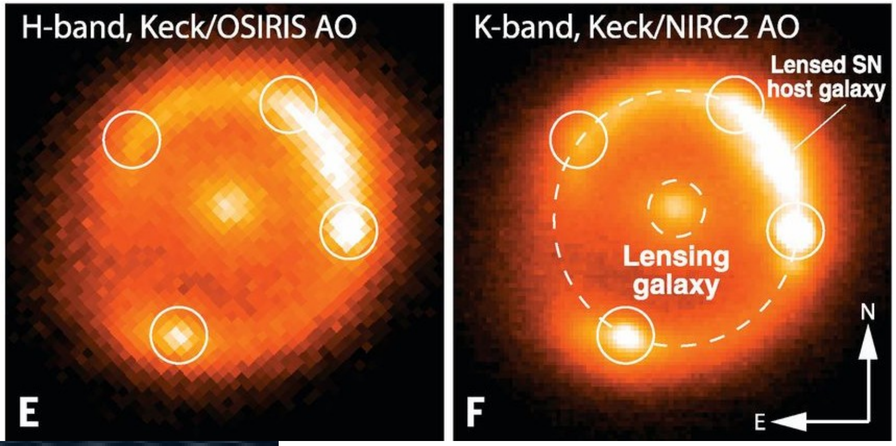




**Fig.** The lensed SN iPTF16geu (Goobar+2017).

*Time delay*

$$t = 1/c \times D_{\Delta t} \times \Phi_{lens}$$

***Time-delay distance:* 1/H$_0$**

- Strongly lensed time-variable sources

  → Time-delays and lens modeling

  → Measure of the Cosmic Expansion rate (Refsdal+1964)

# Morphology of galaxy-scale strong lenses

- Most lens galaxies are massive luminous red galaxies
  → **Good news!**
- Finding galaxy-scale strong lenses
  - Simple binary classification problem?
  - Need to exclude a wide range of contaminants:
    – Spirals, ring galaxies, mergers, etc...
  - Get rid of image artefacts automatically
  - Ensure position/rotation invariance

Cosmic Horseshoe (ESA, NASA)

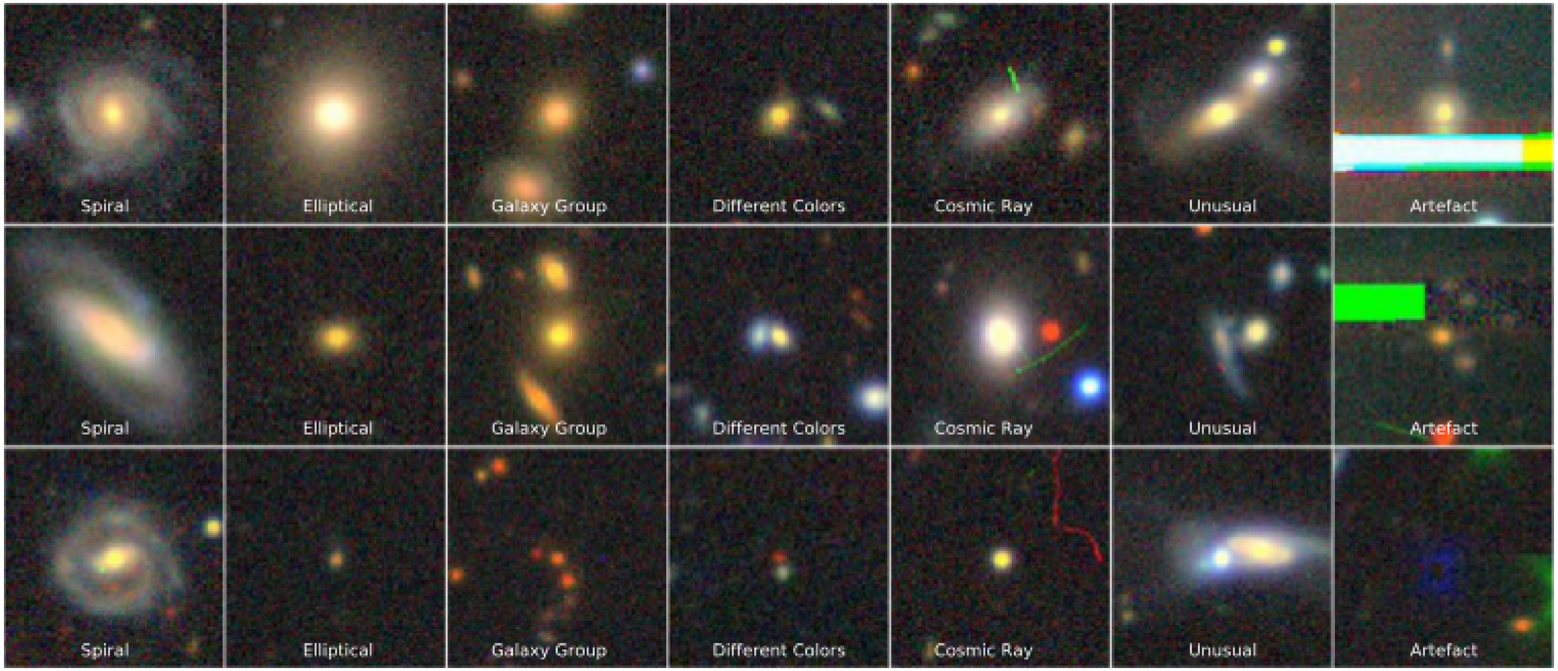



**Fig.** Different galaxy types to be excluded (Huang+2021).

# How to find galaxy-scale strong lenses?

Very rare events → From 1/1000 down to 1/$10^5$

- Using spectroscopy

  e.g. SLACS, BELLS, BELLS-GALLERY, S4TM, SILO samples (Bolton+2006; Treu+2006; Koopmans+2006; Gavazzi+2007; Bolton et al. 2008; Treu et al. 2009; Auger et al. 2009; Shu+2016; Shu+17; Talbot+2021)

- Using single or multi-band imaging
  - Arc-finder algorithms (Gavazzi+2014, Avestruz+2019)
  - Principal component analysis (Joseph+2014; Paraficz+2016)
  - Lens modeling and masking (Sonnenfeld+2018)
  - Citizen-science projects (Marshall+2016, Sonnenfeld+2020)
  - Visual inspection (Diehl+2017, Khullar+2021)
  - Or … **Deep learning**

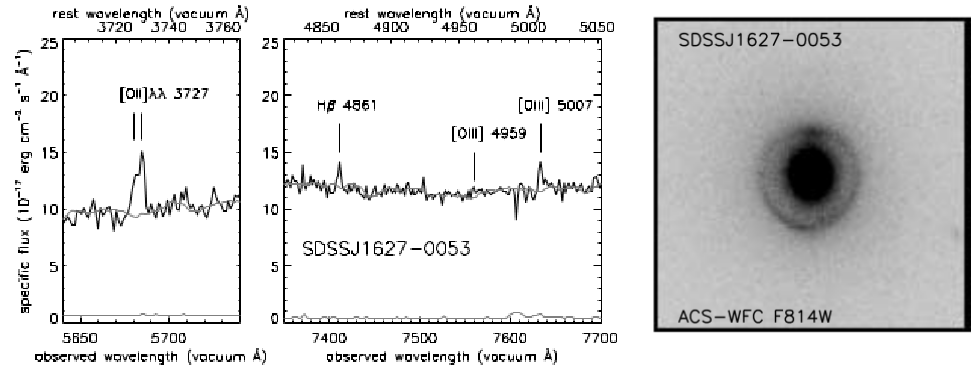



**Fig.** SLACS lens SDSSJ1627-0053 (Bolton+2008).

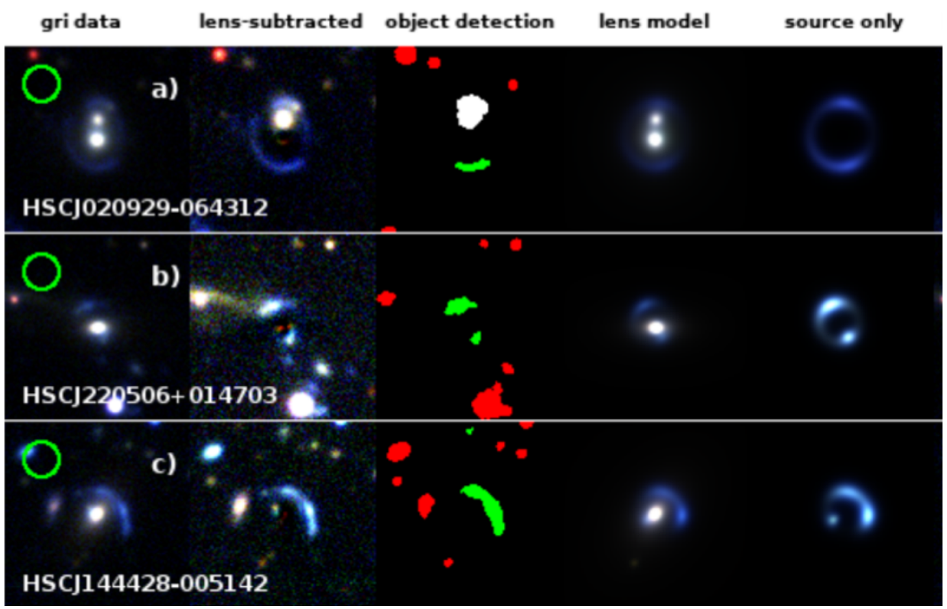
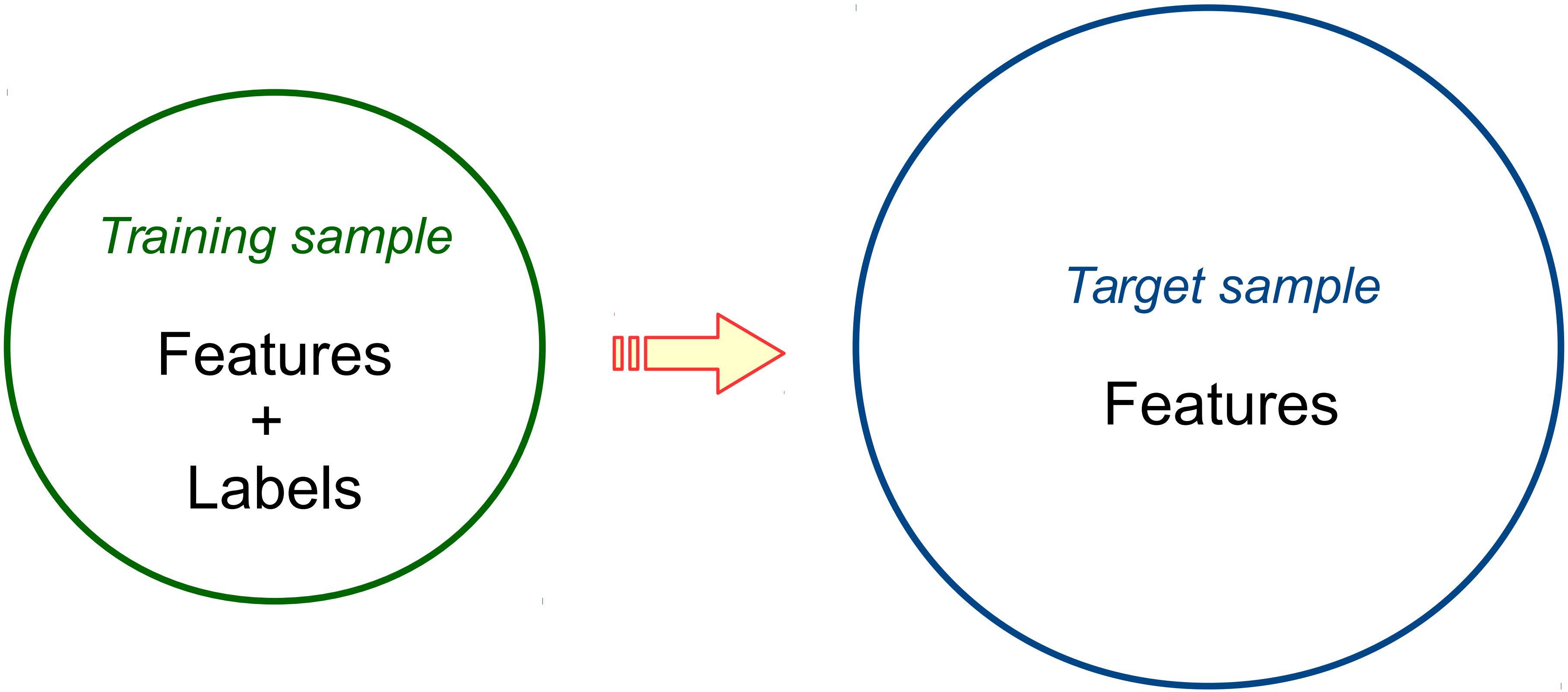



**Fig.** YATTALENS arcfinder applied to HSC (Sonnenfeld+2018).

# Supervised machine learning classification

- CNNs are supervised machine learning techniques optimized for image analysis (LeCun+1998)



# Neural Networks

- Training phase

Loss function (e.g. binary cross-entropy)

$$L(y,p) = -\frac{1}{N}\sum_{i=0}^{N} y_i \log(p_i) + (1-y_i)\log(1-p_i)$$

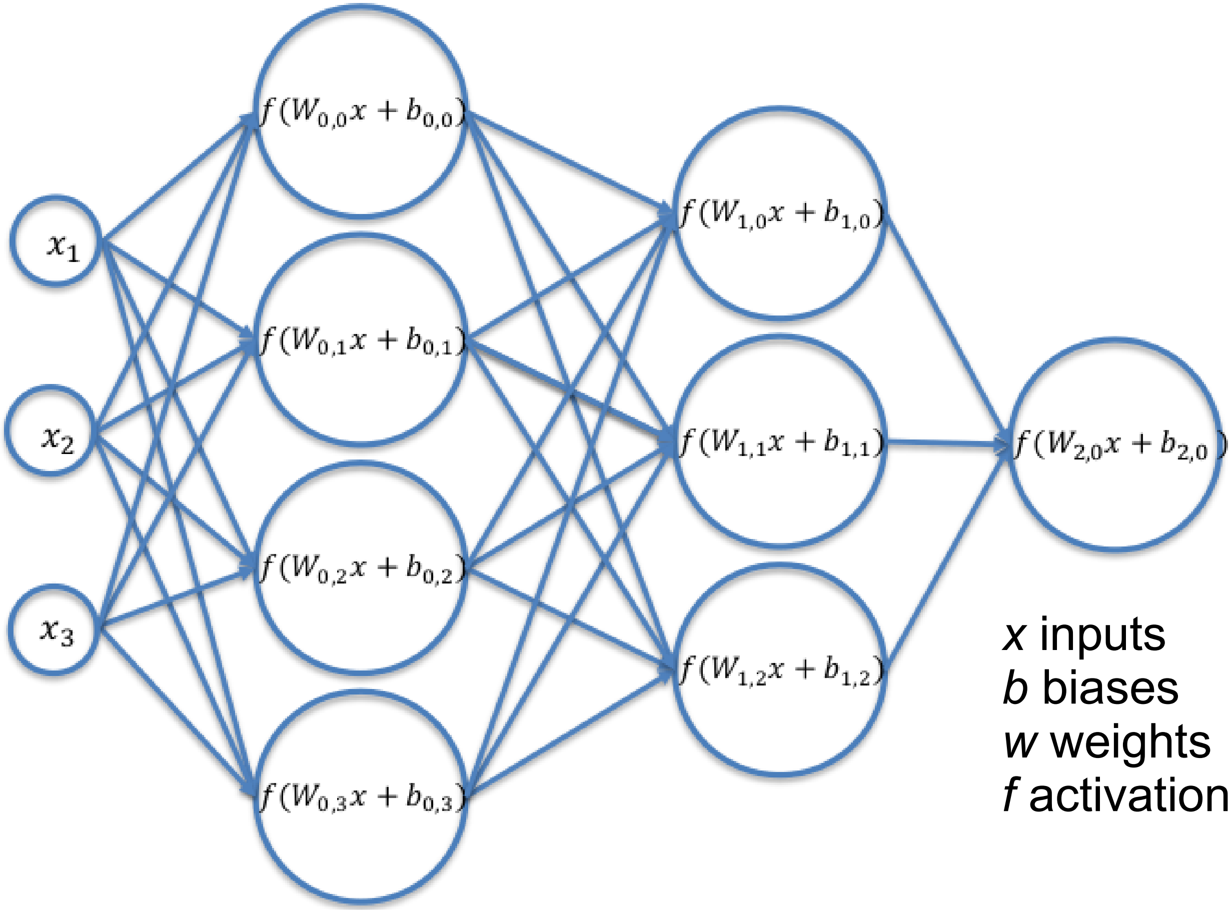



**Fig.** Credit Leal-Taixe, Niessner

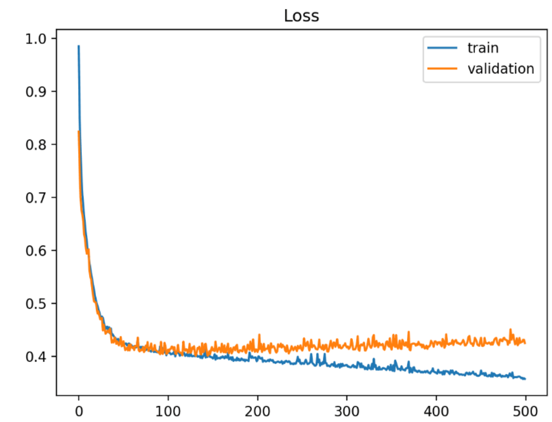
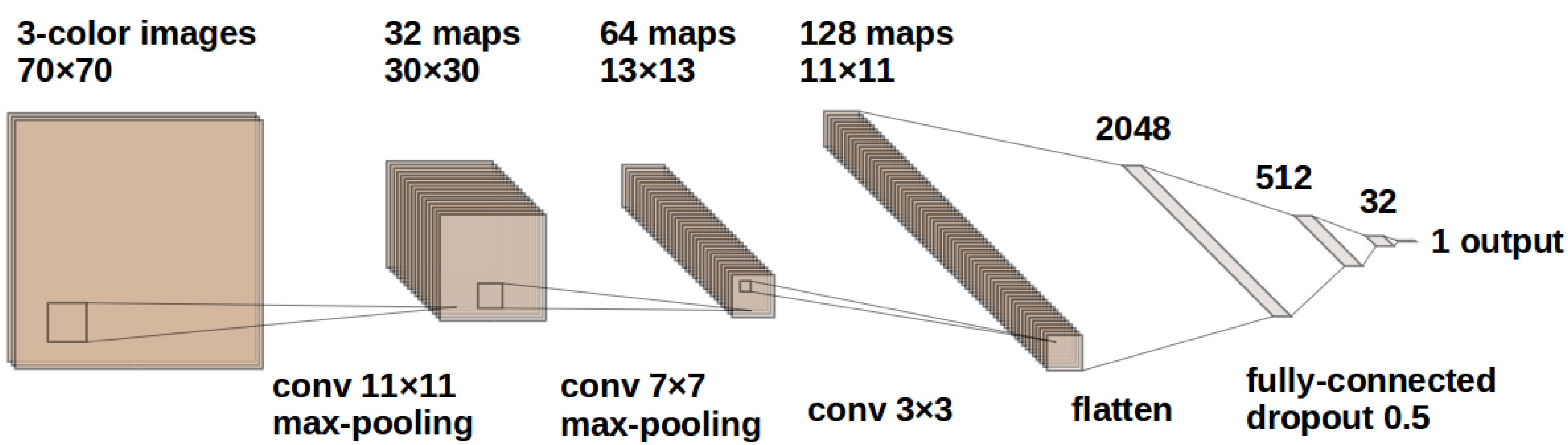



**Fig.** Credit J. Brownlee

*x* inputs
*b* biases
*w* weights
*f* activation functions

# Convolutional Neural Networks

- CNNs are supervised machine learning techniques optimized for image analysis (LeCun+1998)
- Capture image characteristics by learning the coefficients of convolutional kernels
- Need at least $10^4$ labelled images for training BUT only ~$10^3$ lenses known



# Lens finding with CNNs

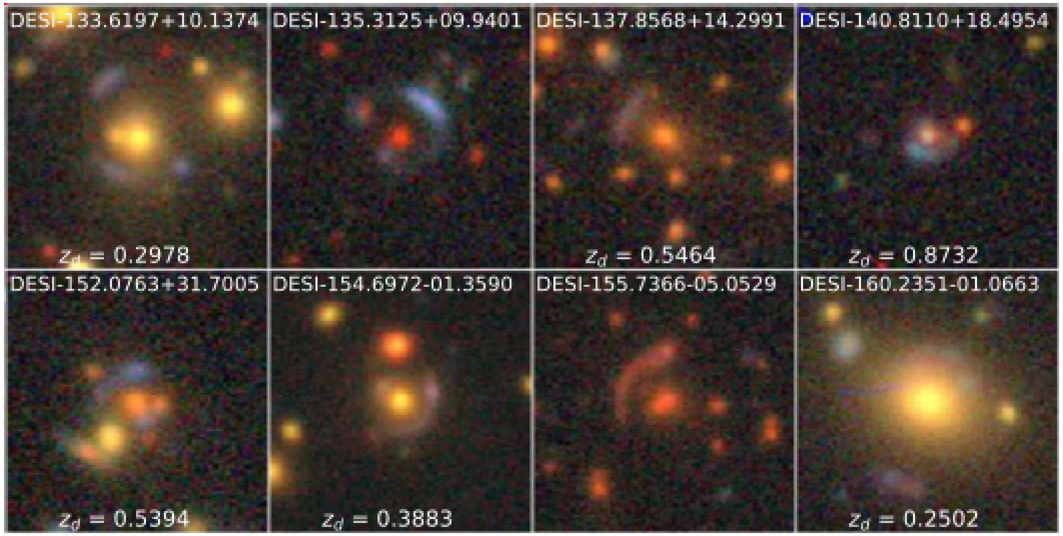
- Successfully applied over the last five years to
  - CFHTLS (Jacobs+2017)
  - COSMOS HST (Pourrahmani+2018)
  - KiDS (Petrillo+2017;+2019; Li+2020)
  - DES (Jacobs+2019a,b)
  - DECaLS (Huang+2020;+2021)

(Need a visual inspection stage)

→ **Several hundred high-quality strong lens candidates**

Lens confirmation currently on-going

→ Systematically outperform non-ML techniques (Metcalf+2019)




**DECaLS, Huang+2021**

**KiDS, Petrillo+2019**

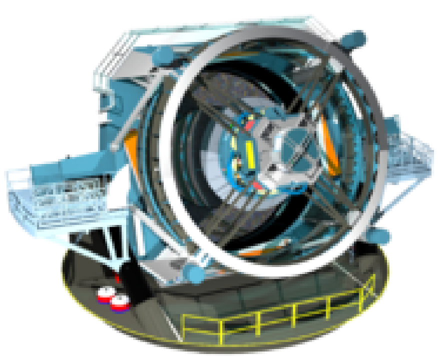
| Name | Type | AUROC | $TPR_0$ | $TPR_{10}$ | Short description |
|---|---|---|---|---|---|
| CMU-DeepLens-Resnet-ground3 | Ground-based | 0.98 | 0.09 | 0.45 | CNN |
| CMU-DeepLens-Resnet-Voting | Ground-based | 0.98 | 0.02 | 0.10 | CNN |
| LASTRO EPFL | Ground-based | 0.97 | 0.07 | 0.11 | CNN |
| CAS Swinburne Melb | Ground-based | 0.96 | 0.02 | 0.08 | CNN |
| AstrOmatic | Ground-based | 0.96 | 0.00 | 0.01 | CNN |
| Manchester SVM | Ground-based | 0.93 | 0.22 | 0.35 | SVM/Gabor |
| Manchester2 | Ground-based | 0.89 | 0.00 | 0.01 | Human Inspection |
| ALL-star | Ground-based | 0.84 | 0.01 | 0.02 | Edges/gradiants and Logistic Reg. |
| CAST | Ground-based | 0.83 | 0.00 | 0.00 | CNN/SVM |
| YattaLensLite | Ground-based | 0.82 | 0.00 | 0.00 | SExtractor |

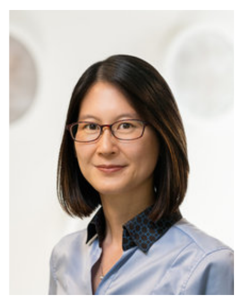
**Lens finding challenge, Metcalf+2019**

# Automated pipelines for wide-field surveys

Our main goals are

- Build lens finding pipelines for systematic searches
- Test extensively and prepare for LSST and Euclid
- Extend to strong lens modeling and photometric redshift estimation





**Sherry Suyu**

**Stefan Schuldt**

**Stefan Taubenberger**

**Yiping Shu**

# Lens finding in PanSTARRS

Cañameras et al. 2020, A&A 644, 163

- Systematic search over the 3 billion sources detected by the Pan-STARRS 3π survey (30 000 deg$^2$) → 3 filters *gri*
- Simple cuts to exclude the Milky Way plane, stars, very faint galaxies
- Two-step approach optimized for wide-separation galaxy-scale lenses

  1) a catalog-based neural network classification of source photometry,

  2) a CNN trained on multi-band images

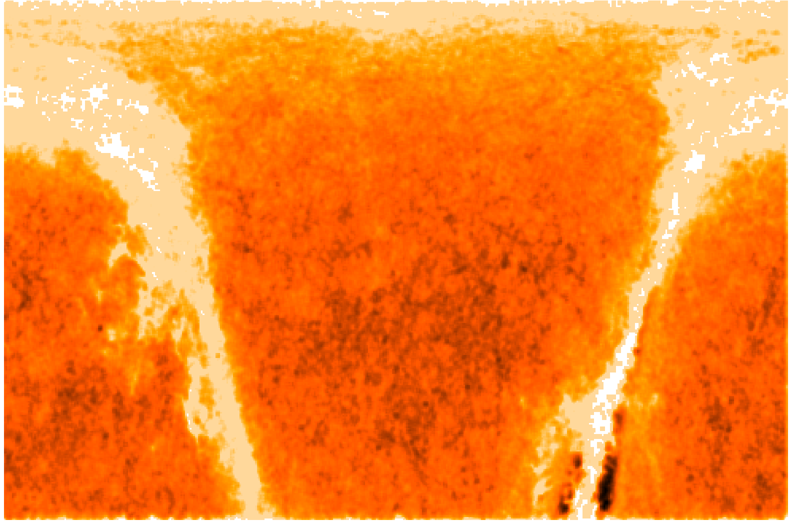


**Fig.** PS1 sources after removing stars.

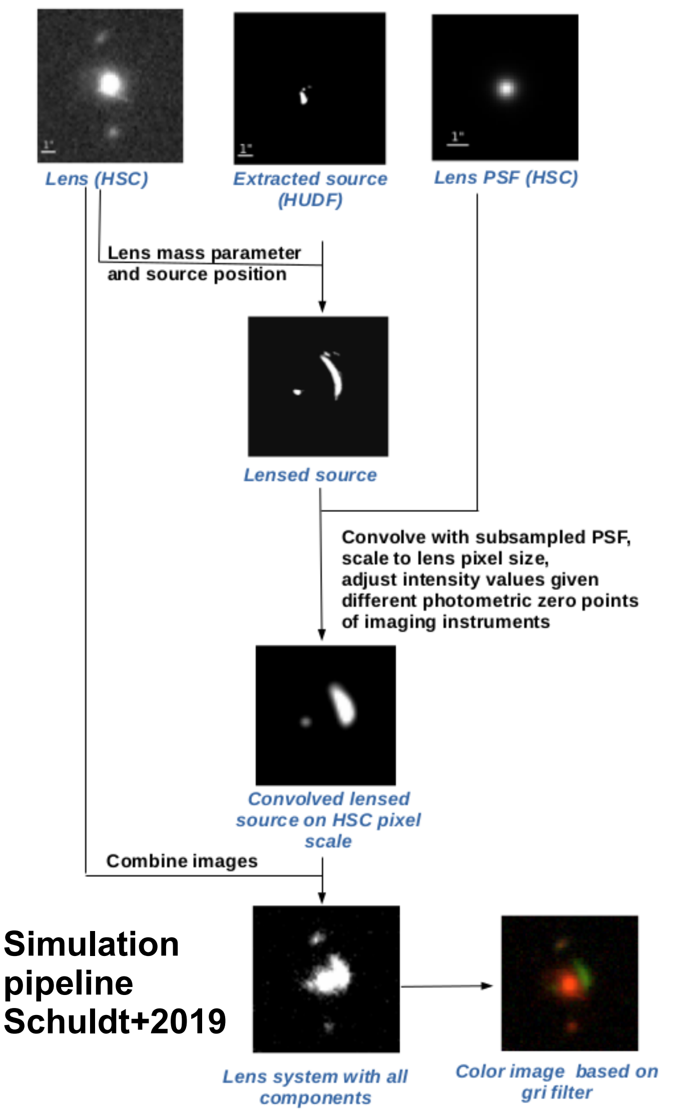
# Design of PanSTARRS lens simulations

Cañameras et al. 2020, A&A 644, 163

Realistic lens simulations → the main ingredient for higher accuracies

- Major aspects
  - realistic lens galaxies
  - good proxies of lens mass
  - Einstein radius distributions
  - number of multiple images
  - source colors and morphologies
  - inclusion of neighbours/artifacts
  - good PSF models
- Match properties of PanSTARRS coadds
- ~~Fully simulated data~~

→ **Observed images + ray-tracing = Paint lensed arcs on survey stacks**



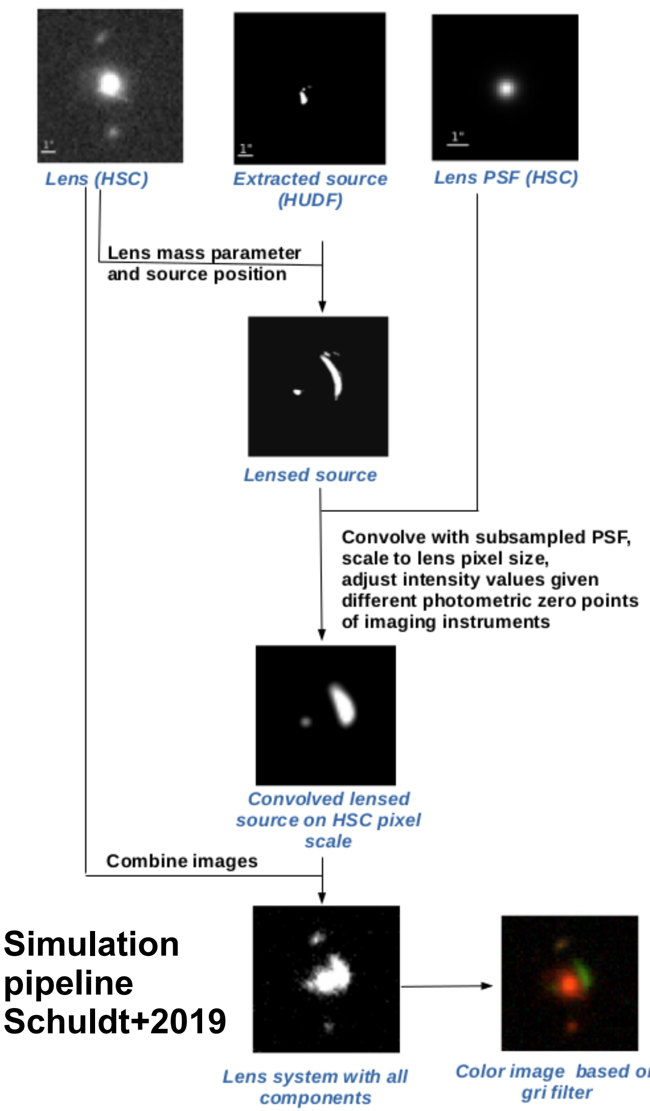

**Simulation pipeline Schuldt+2019**

# Design of PanSTARRS lens simulations

Cañameras et al. 2020, A&A 644, 163

Realistic lens simulations → the main ingredient for higher accuracies

- Major aspects
  - realistic lens galaxies
  - good proxies of lens mass
  - Einstein radius distributions
  - number of multiple images
  - source colors and morphologies
  - **inclusion of neighbours/artifacts**
  - good PSF models




**Simulation pipeline Schuldt+2019**

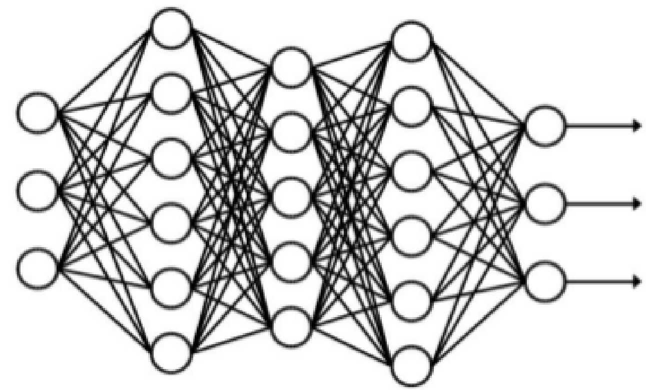
# Step 1- Catalog-level neural network

Cañameras et al. 2020, A&A 644, 163

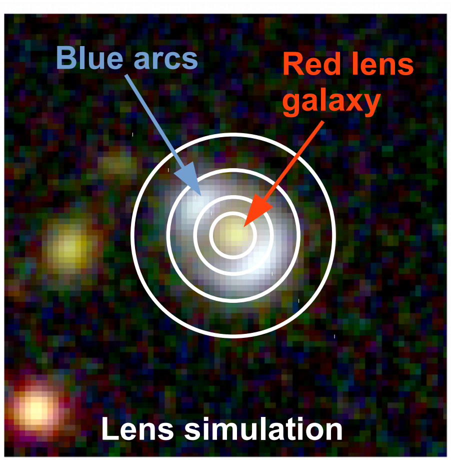
1) Aperture photometry of mocks in *gri* bands
→ 1.04”, 1.76”, 3.00”,  and 4.64” radii

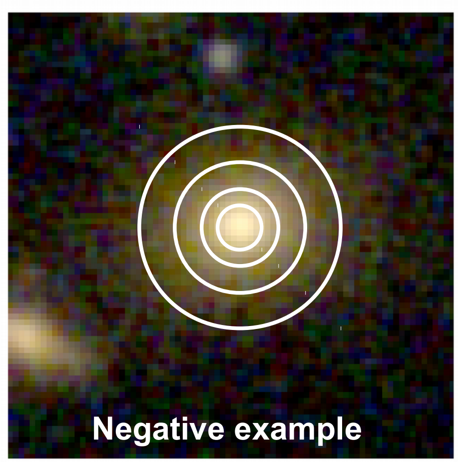
→ color variations and radial gradients

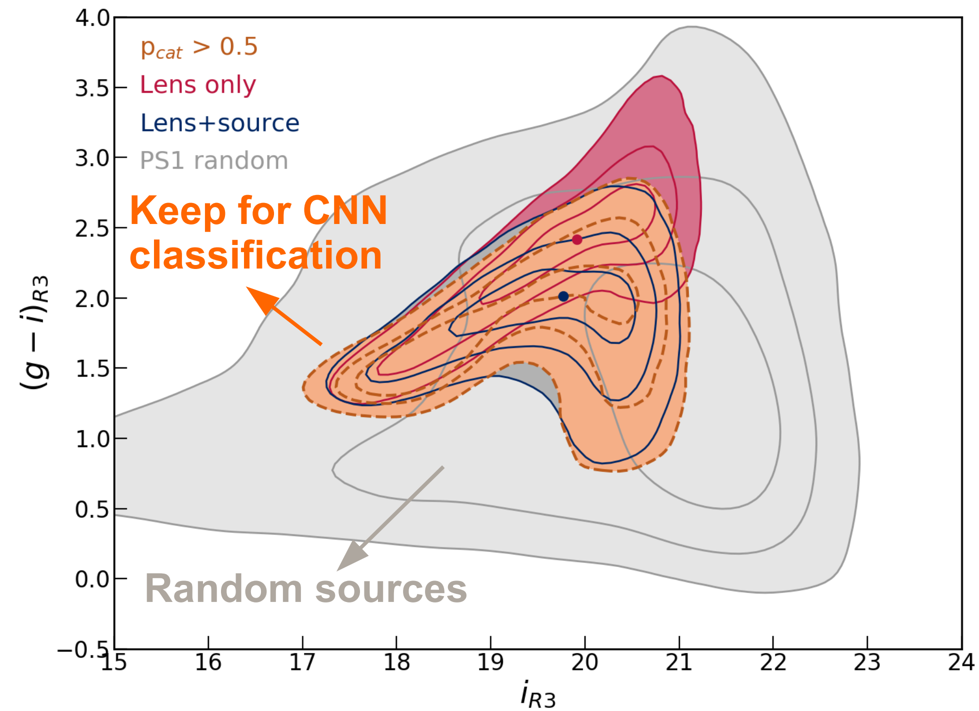
2) Photometry of $10^5$ negative examples



- $10^5 + 10^5$ Labelled examples
- Classify with a fully-connected network
- Safe: 0 known lenses excluded

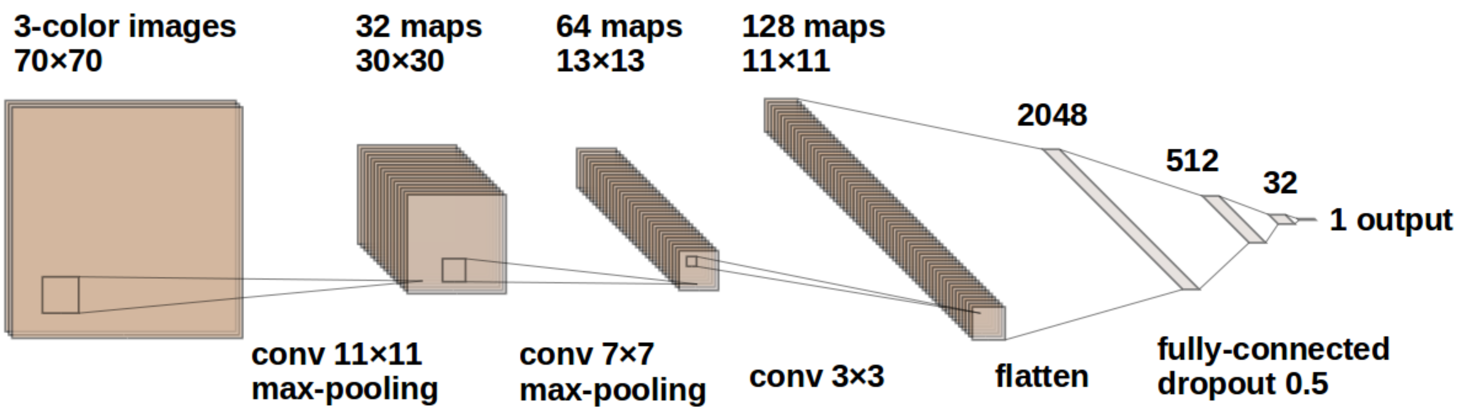


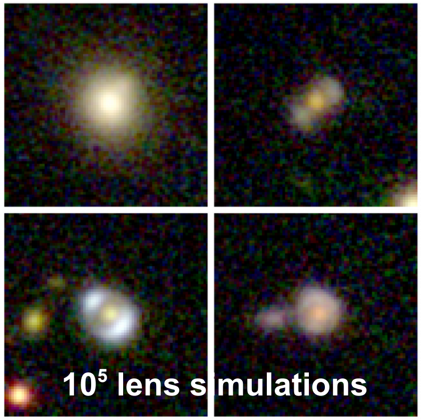



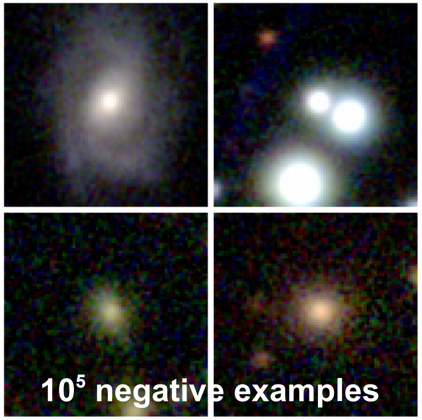



# Step 2- Convolutional neural network

Cañameras et al. 2020, A&A 644, 163

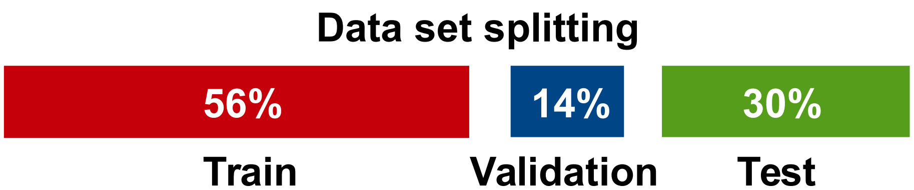


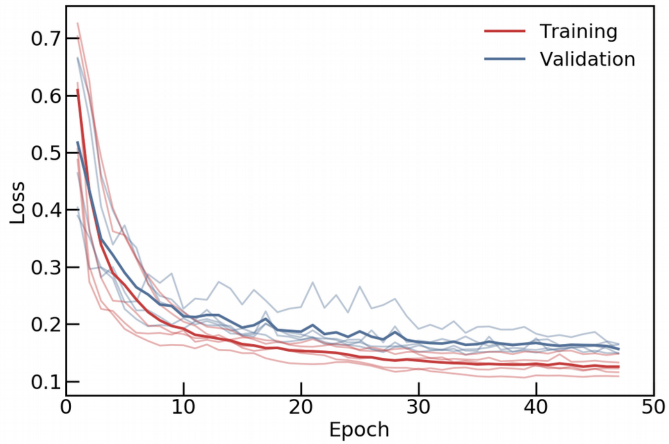




Classify image cutouts in gri bands

- Negative examples: LRGs, face-on spirals, rings, groups from GalaxyZoo + different fractions
- Tests on the CNN architecture
- Hyperparameter optimization
- Cross-validation and best epoch
- 12000 network candidates



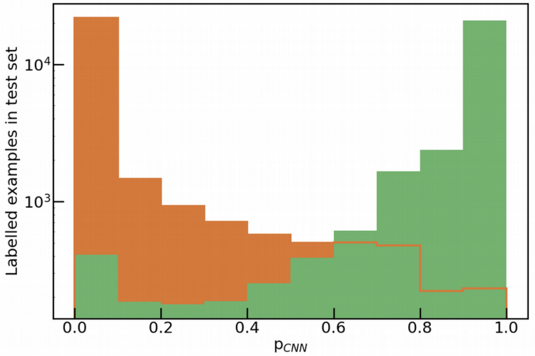



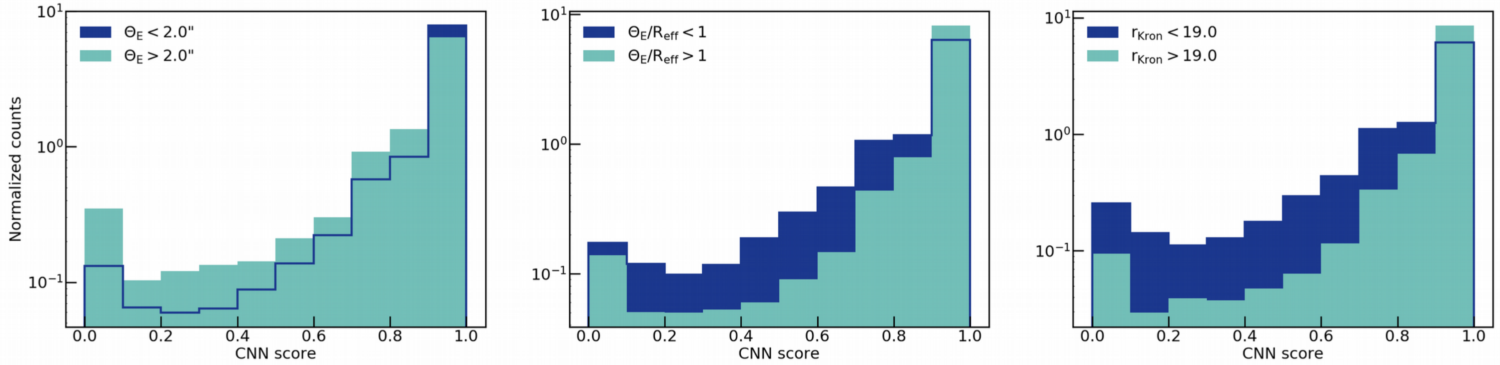
# Step 2- Convolutional neural network

Cañameras et al. 2020, A&A 644, 163

## Testing the network predictions

- Using our test set
  - – distribution of scores as a function of Einstein radius, lens magnitude, lens effective radius
  
    → depends on the data set construction...
- Using an independent set
  - – are known lenses recovered by the CNN?
  
    → 14/16 + higher scores when similar to mocks

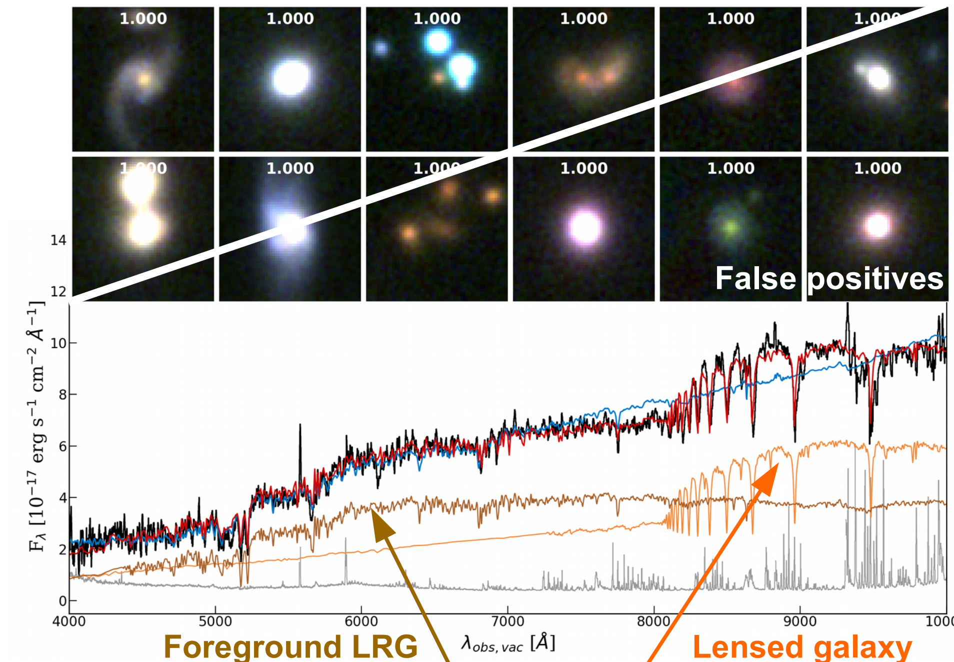



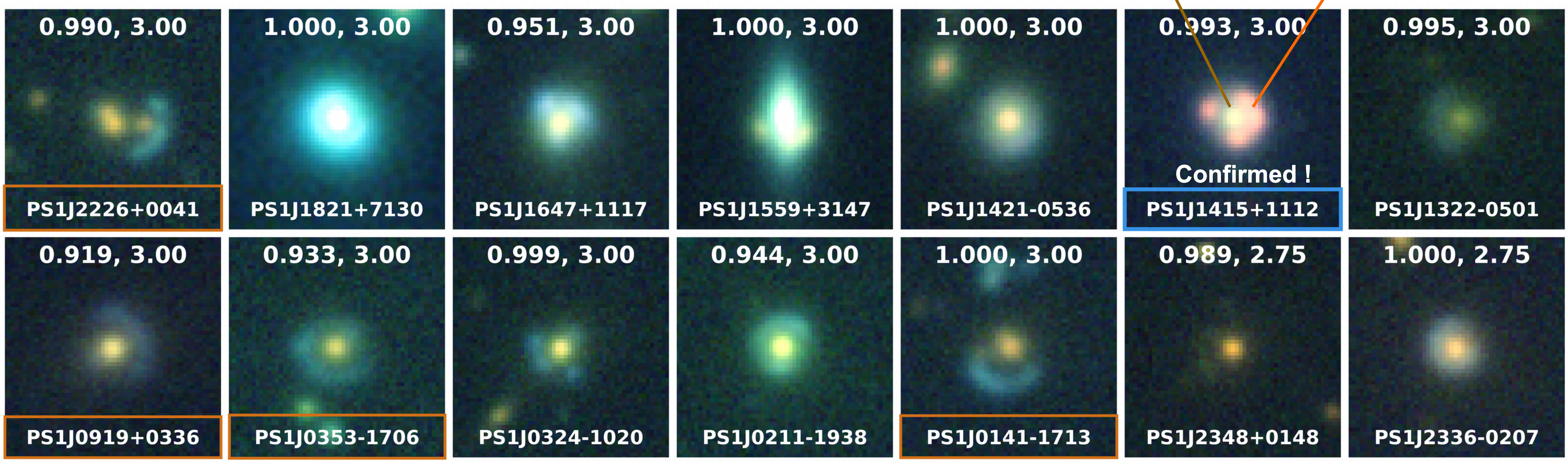
# PanSTARRS lens search results

Cañameras et al. 2020, A&A 644, 163

**330 new high-quality lens candidates**

- Recover known lenses
- One system confirmed in spectro
- Sample spectroscopic follow-up on-going
- PanSTARRS seeing and depth are major limitations
- False positives are problematic



# Testing lens finding pipelines

- Construction of the ground truth data is arbitrary → Need to carefully test the influence of the training set design on output classifications
- Performances measured on simulated data sets (Metcalf+2019)
- Imperfect generalization (see Lanusse+2018; Schaefer+2018; Davies+2019)
- Solutions are 1) more realistic lens finding challenge data or 2) observed data sets

→ **An independent test set from real survey data**

→ Use existing Subaru Hyper Suprime-Cam imaging similar to forthcoming LSST

| **Wide** | $g$ | $r$ | $i$ | $z$ | $y$ |
|---|---|---|---|---|---|
| exposure (min) | 10 | 10 | 16 | 20 | 16 |
| seeing (arcsec) | 0.77 | 0.76 | 0.58 | 0.68 | 0.68 |
| depth (mag) | 26.6 | 26.2 | 26.2 | 25.3 | 24.5 |
| saturation (mag) | 17.6 | 17.4 | 18.0 | 17.5 | 17.3 |
| area (deg$^2$) | 942 | 1022 | 796 | 905 | 924 |
| **Deep+UltraDeep** | $g$ | $r$ | $i$ | $z$ | $y$ |
| exposure (min) | 49 | 45 | 65 | 130 | 88 |
| seeing (arcsec) | 0.81 | 0.74 | 0.62 | 0.63 | 0.71 |
| depth (mag) | 27.3 | 26.9 | 26.7 | 26.3 | 25.3 |
| saturation (mag) | 18.1 | 18.2 | 18.7 | 17.7 | 17.3 |
| area (deg$^2$) | 35 | 35 | 35 | 36 | 36 |

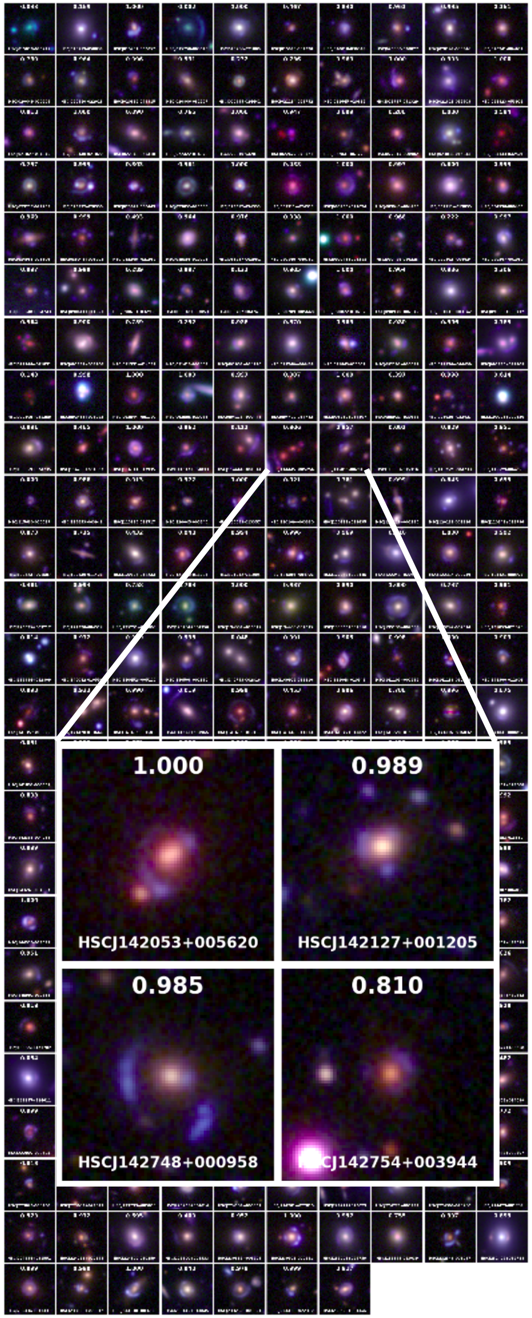
**HSC survey status, PDR2, Aihara+2019**

| Quantity | Baseline Design Specification |
|---|---|
| Optical config. | Three-mirror modified Paul-Baker |
| Mount config. | Alt-azimuth |
| Final $f$-ratio, aperture | $f/1.234$, 8.4 m |
| Field of view, étendue | 9.6 deg$^2$, 319 m$^2$deg$^2$ |
| Plate scale | 50.9 $\mu$m/arcsec (0.″2 pix) |
| Pixel count | 3.2 gigapixels |
| Wavelength coverage | 320–1050 nm, *ugrizy* |
| Single-visit depths, design[a] | 23.9, 25.0, 24.7, 24.0, 23.3, 22.1 |
| Single-visit depths, min.[b] | 23.4, 24.6, 24.3, 23.6, 22.9, 21.7 |
| Mean number of visits[c] | 56, 80, 184, 184, 160, 160 |
| Final (co-added) depths[d] | 26.1, 27.4, 27.5, 26.8, 26.1, 24.9 |

**LSST baseline design, Ivezic+2019**

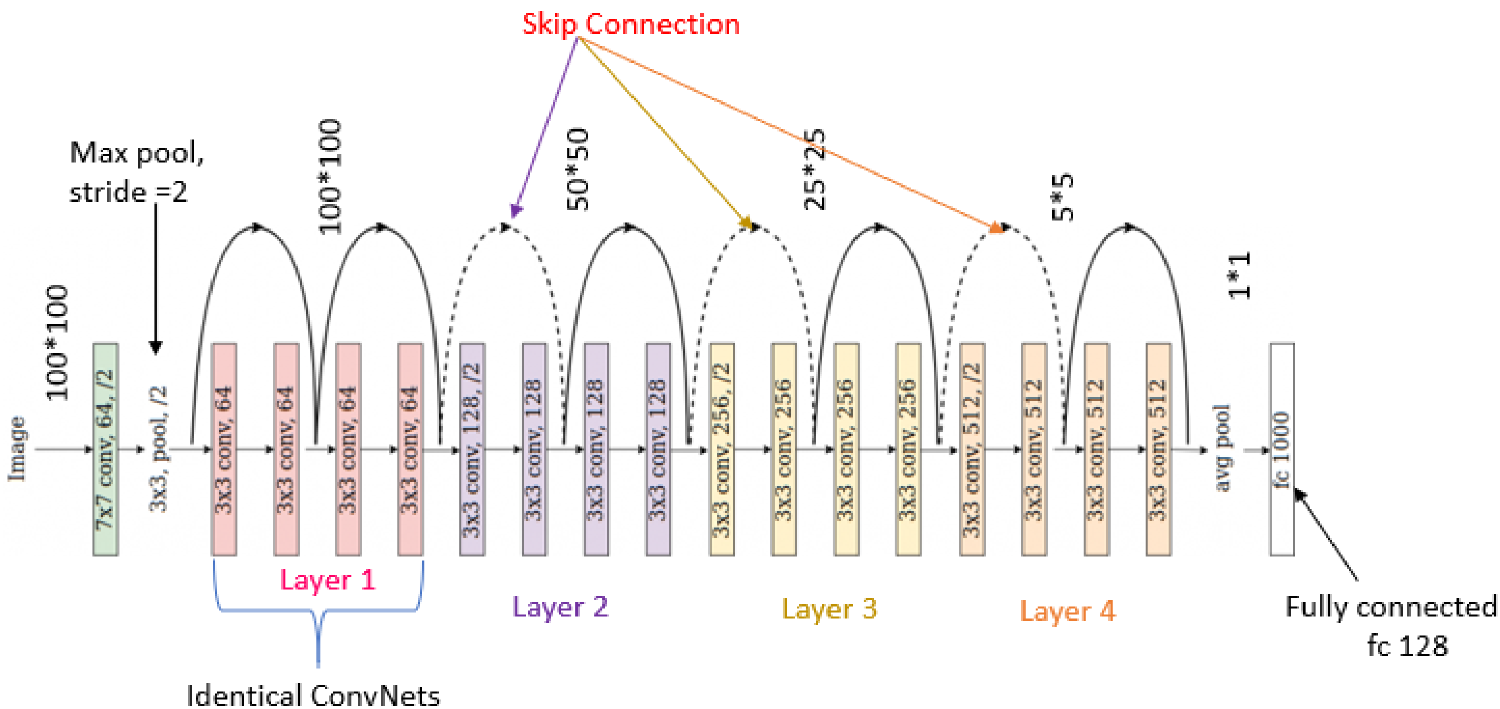
# Testing lens finding pipelines

- Train/Validate + test on galaxy sets from HSC, same depth, same background properties, etc ...
- 220 known galaxy-scale lenses from HSC (SuGOHI)
  - Test completeness for different configurations (arc morphology, source colors, etc ...)
- 50,000 non-lenses in COSMOS
  - Quantify the number of false positives → representative of final classification on real data
- 1000 ambiguous cases in SpaceWarps (Sonnenfeld+2020)

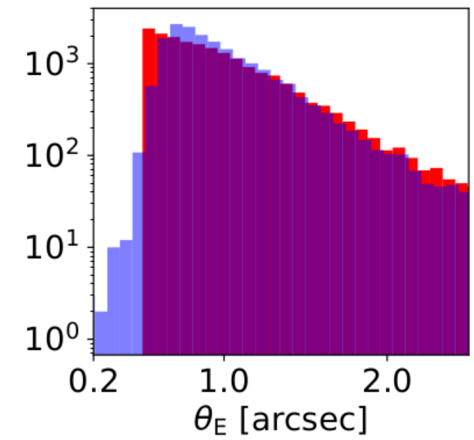
# Influence of network architecture and data processing

- Architectures previously improved using LSST/Euclid simulations (Lanusse+2018)
- We have tested different CNN architectures by varying number of layers, of filters per layer, convolutional kernel sizes, etc...
- Deeper ResNet *generally* help → depends on the data set
- sqrt stretching *always* helps → other data pre-processing and augmentation have little influence
- Remaining problems → difficult to recover >80% SuGOHI lenses while maintaining FPR < 0.01% + network predictions not perfectly rotation invariant



# Influence of lens simulations

We have tested multiple combinations of positive/negative examples

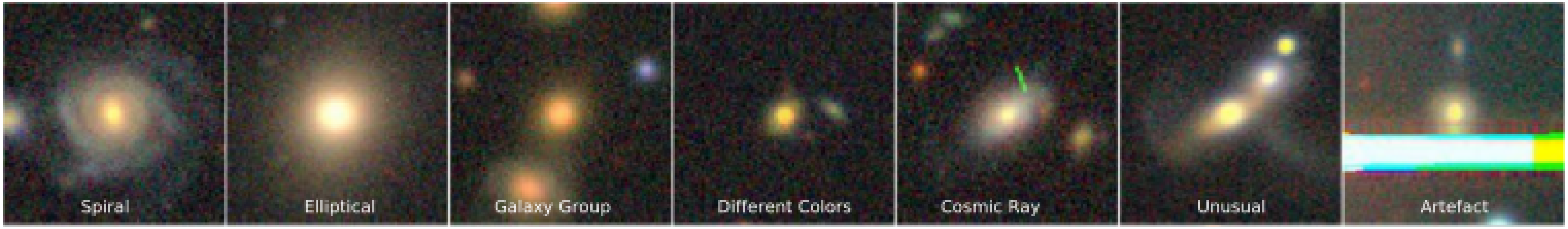



- Highly-realistic lens simulation with
  - Various distributions on physical parameters (e.g. natural/flat distributions in Einstein radius?)
  - Various selections of lens and source galaxies (colors, redshifts, ...)
  - Various configurations (ratio of doubles/quads), min S/N, min μ

→ Parameter distributions play a major role (do not need to follow nature)

- Negative examples including
  - Random non-lens galaxies, or boosted fractions of usual interlopers (spirals, rings, isolated LRGs, groups, etc…)
  - Draw interlopers from GalaxyZoo + *Unsupervised classifications*

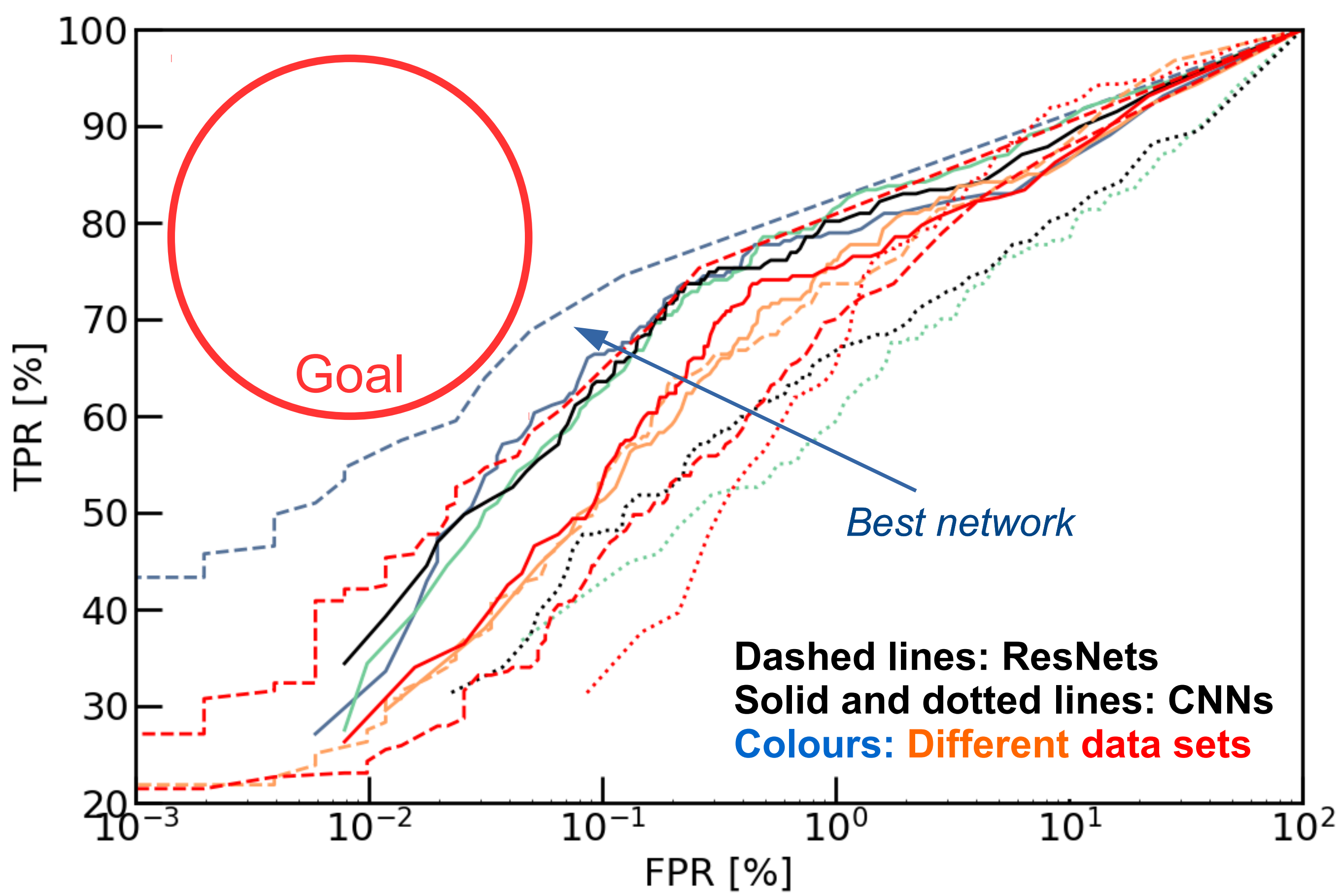
→ Need to include sufficient examples in each class for training



# Testing lens finding pipelines

Receiver Operating Characteristic (ROC) curves using SuGOHI known lenses and COSMOS non-lenses:

- Major improvements for specific networks and data sets
- Performance directly indicate behaviour on real survey data



# New galaxy-scale lens candidates from HSC Wide

Cañameras et al. 2021, in prep.

Best ResNet applied to all extended sources (>0.8’’) from HSC Wide DR2

- ~6000 network recommendations → recover SuGOHI + several new candidates
- Brute force approach without strict catalog-level pre-selection works!
- Different sets of candidates from different methods with little overlap

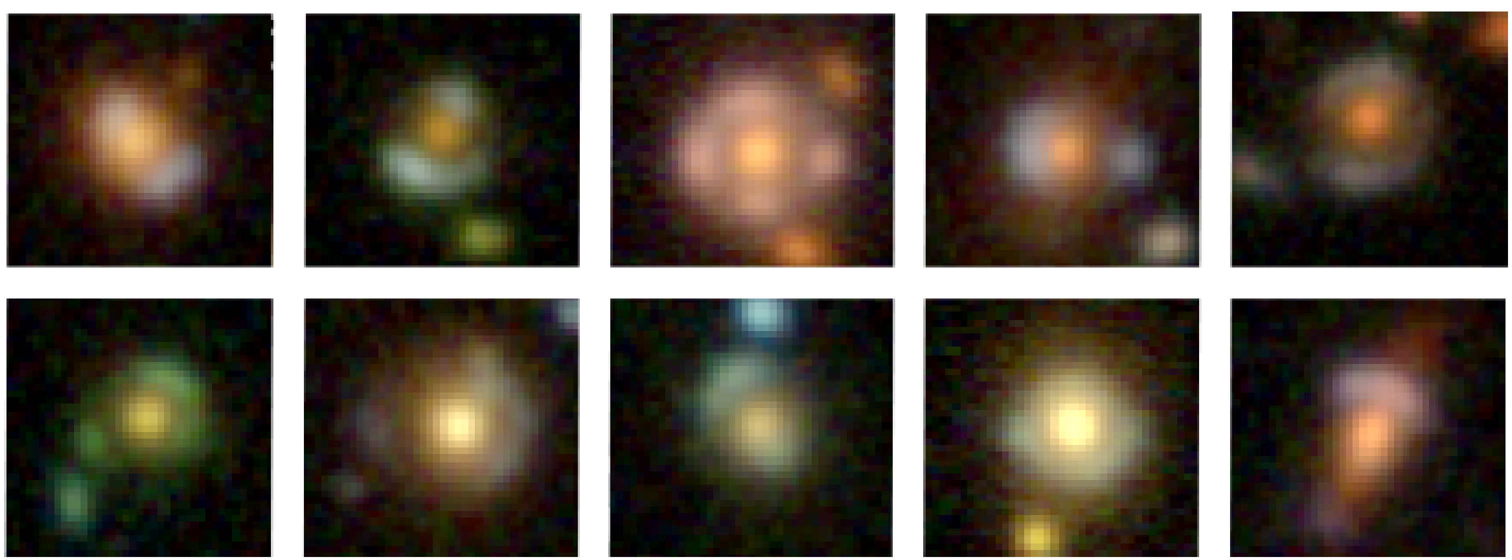


**Fig.** New ResNet high-quality lens candidates from HSC DR2.

# Towards a systematic pipeline for LSST

## Preparation for LSST

- Current ResNet sufficient for our lensed SN search
- Human inspection for HSC: 6000 network candidates = 1.5–3 hours
  → x 50 for LSST (only for *simple* wide-separation lenses)
- General lens search vs targeted lens search
  - Still too many false positives for clean lens selection + completeness not ideal
  - Impossible to bypass visual inspection? → 1-10% of neural network recommendations are good candidates

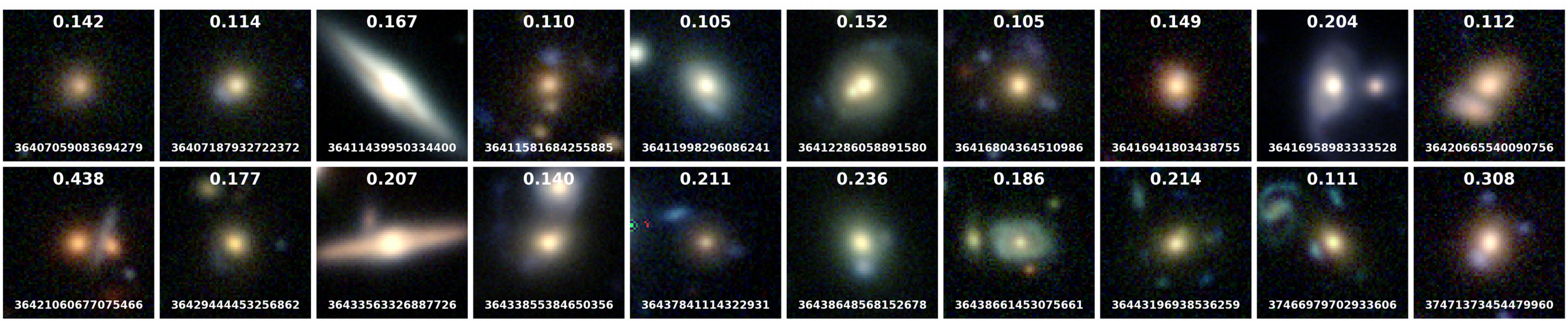



**Fig.** False positives from HSC DR2.

# Towards a systematic pipeline for LSST

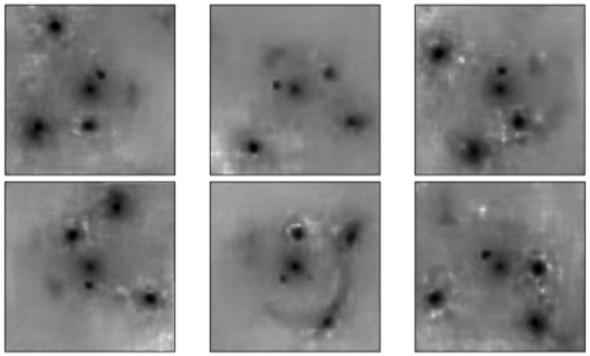
## Ideas to be tested

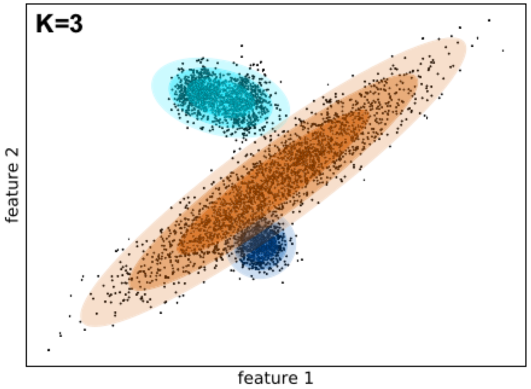
– Calibrate neural network scores as probabilities (Guo+2017)

– Combine with citizen science projects (Marshall+2016)

– Architecture level
  - Multiclass classification (Teimoorinia+2020) ❌
  - ResNet pre-trained in ImageNet database
  - Invariant architecture (Schaefer+2018)
  - Committees of networks: train multiple CNNs and combine to increase prediction stability (Schaefer+2018) ✓
  - Outlier detection (Margalef-Bentabol et al. 2020)
  - Unsupervised learning (Cheng+2020) → for exotic lenses or in combination with supervised algorithms

– Data set level
  - Masking neighbours ❌
  - Lens light subtraction
  - Denoising images ❌
  - Adding more bands
  - Classification & modeling





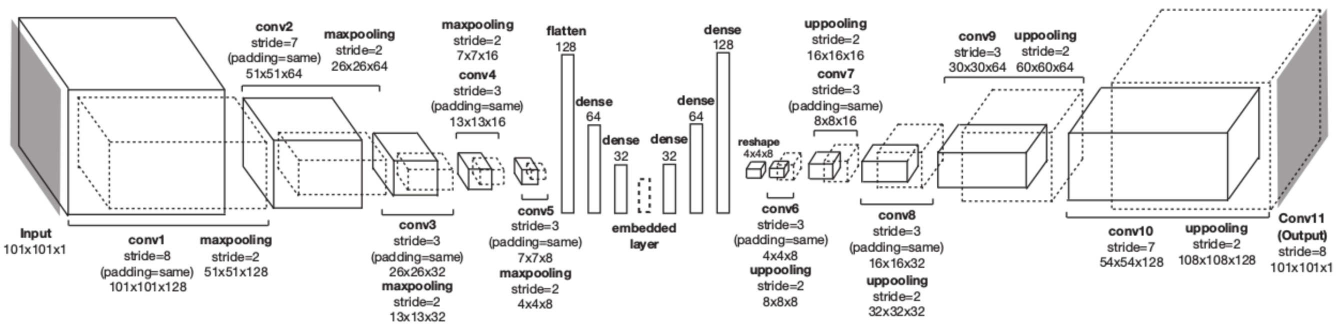




**Fig.** Unsupervised lens finding with (1) a convolutional autoencoder, and (2) a Bayesian Gaussian mixture model (Cheng+2020).

# Efficient strong lens modeling

## Predict lens mass profile parameters

- Traditional parameter fitting techniques
- Regression convolutional neural network
  - Start simple = Singular Isothermal Ellipsoid (position, ellipticity, axis ratio, Einstein radius)
  - – Hezaveh+2017, Perreault-Levasseur+2017, Bom+2019, Madireddy+2019, Park+2020, Pearson+2019, Pearson+2021

→ Trained and tested mostly on fully-simulated data, or idealistic S/N or configurations

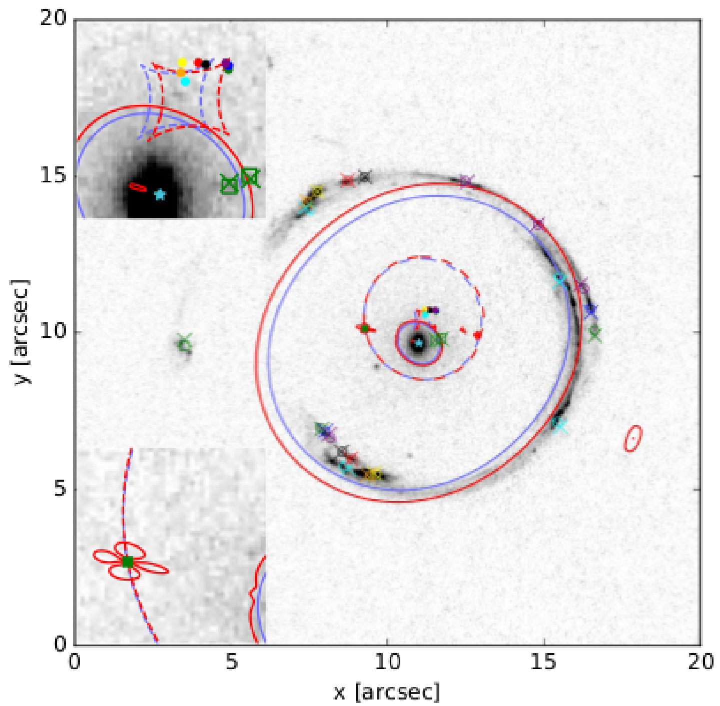
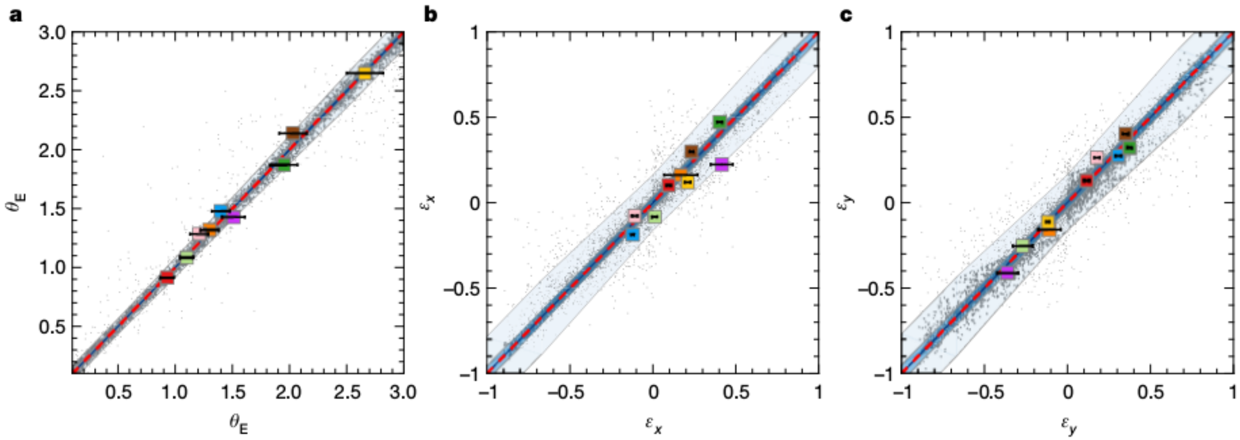



**Fig.** MCMC lens modeling for the Cosmic Horseshoe (Schuldt+2019).



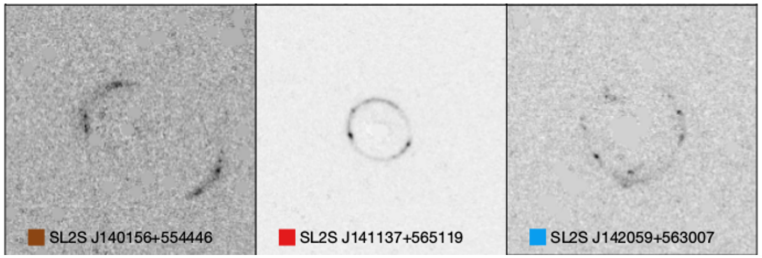




**Fig.** Lens modeling with deep learning (Hezaveh+2017).

# Efficient strong lens modeling

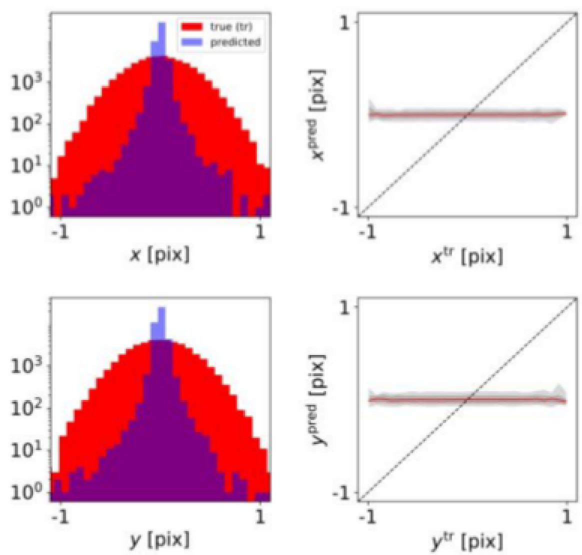
**Stefan Schuldt** – schuldt@mpa-garching.mpg.de

Schuldt et al. 2021, A&A 646, 126

## Predict lens mass profile parameters

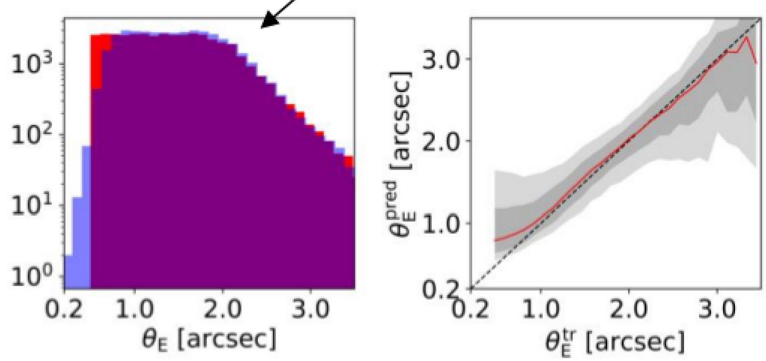
- Regression convolutional neural network
  - Singular Isothermal Ellipsoid (position, ellipticity, axis ratio, Einstein radius)
  - **Realistic lens simulations** – train and test on HSC Wide *griz* to prepare for LSST

Lens center

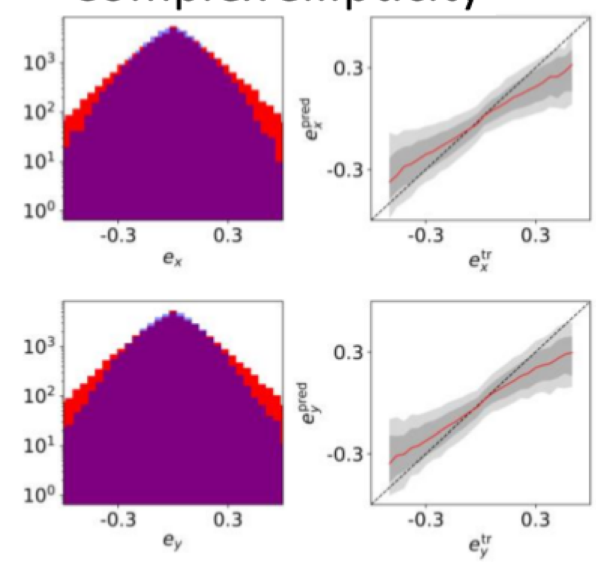


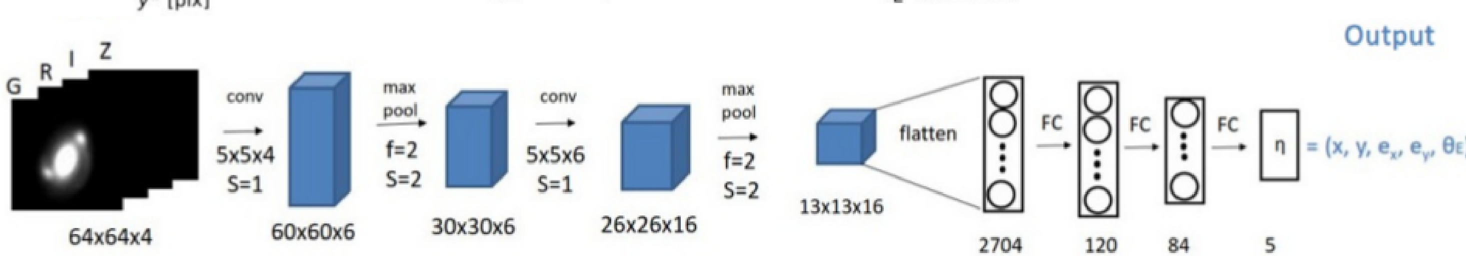

Einstein radius

Flat up to ~2"




Complex ellipticity



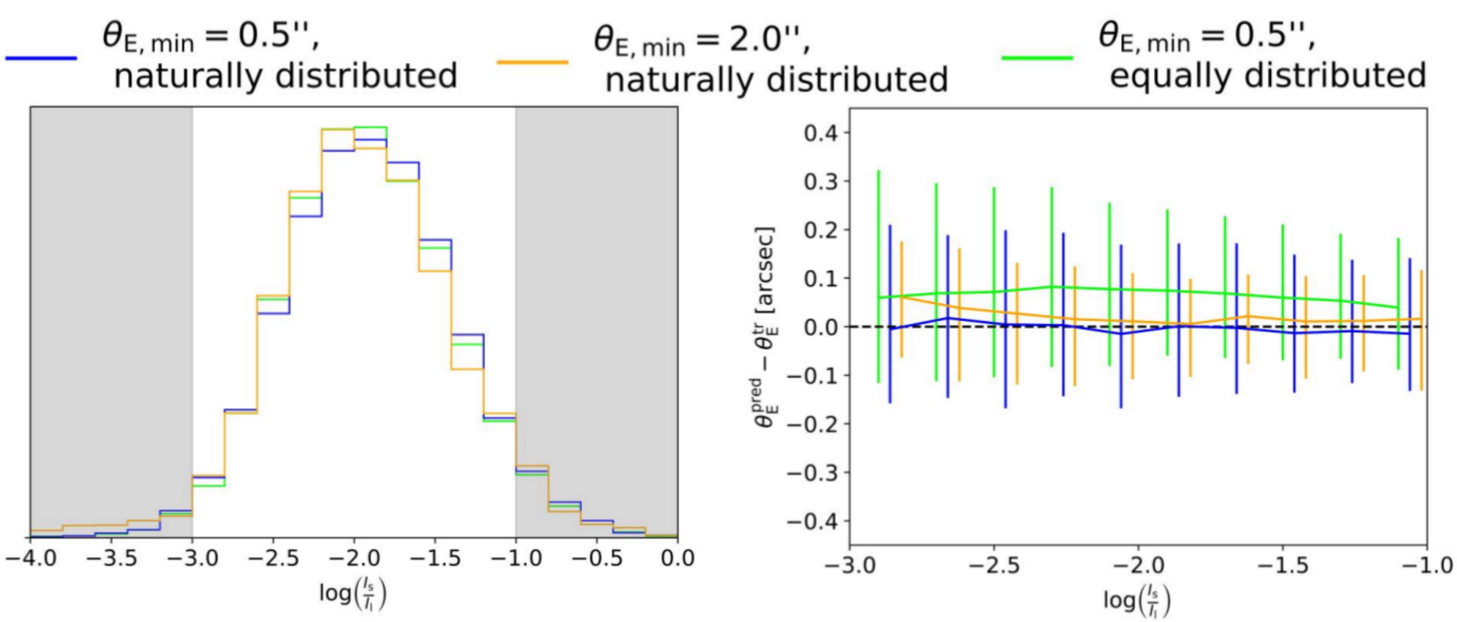
# Efficient strong lens modeling

**Stefan Schuldt** – schuldt@mpa-garching.mpg.de

Schuldt et al. 2021, A&A 646, 126

## Predict lens mass profile parameters

- Results are stable, e.g. for fainter lensed sources
- Future prospects (Schuldt et al., in prep.)
  - Test deeper networks, model SIE + external shear, parameter uncertainties
  - Direct comparison between neural networks and traditional MCMC modeling



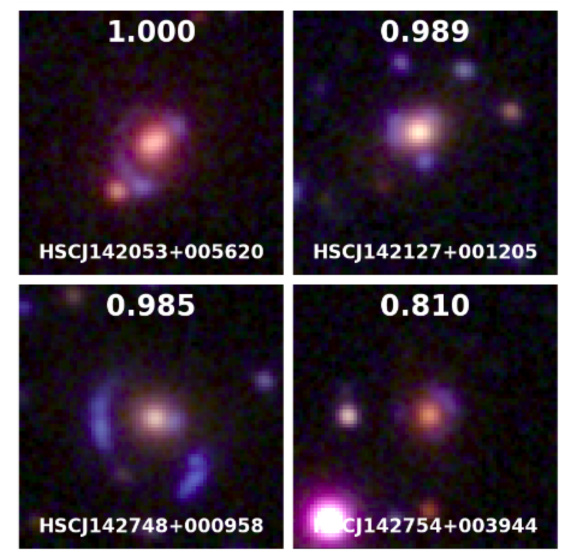




**Fig.** HSC lenses in SuGOHI for validating the DL pipeline.

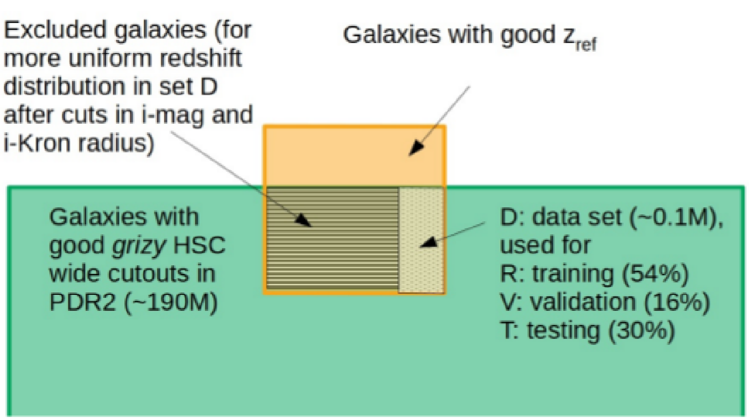
# Photometric redshift estimation

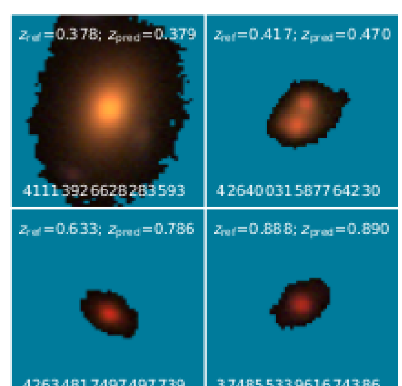
**Stefan Schuldt** – schuldt@mpa-garching.mpg.de

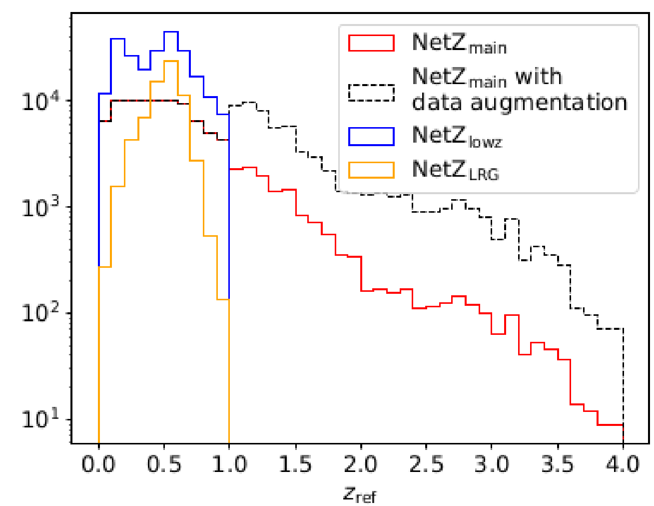
Schuldt et al. 2021b, accepted

## Predict photometric redshifts

- Regression convolutional neural network (d'Isanto+2018, Pasquet+2019)
- **More systematic pipeline** – train and test on HSC Wide *grizy* to prepare for LSST
    - Data set: galaxies without imaging artifacts and with ground truth redshifts from (1) spectro surveys, (2) reliable photo-z in COSMOS (30 bands, Laigle+2016)
    - Limit to mag < 25 and Kron radius >0.8" in i-band + masking + balanced data set → $10^5$ examples for training a simple CNN



# Photometric redshift estimation

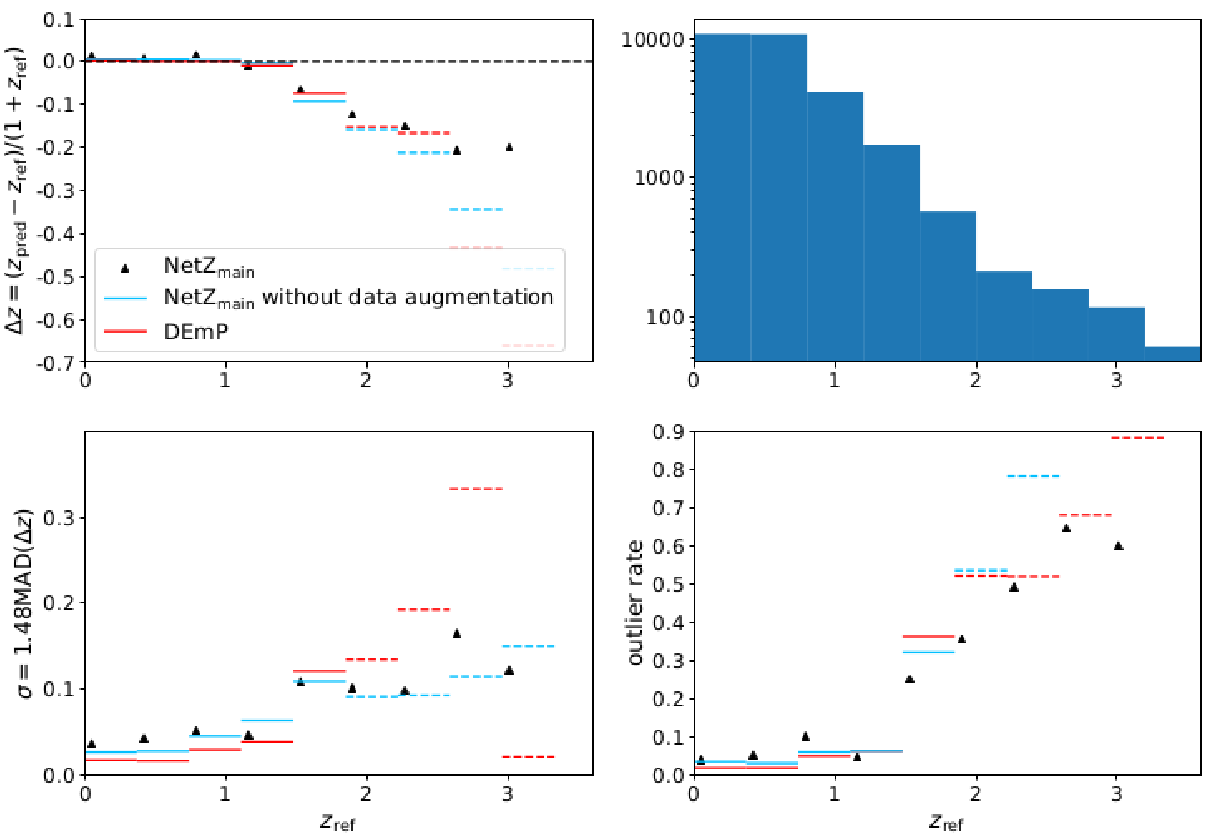
**Stefan Schuldt** – schuldt@mpa-garching.mpg.de

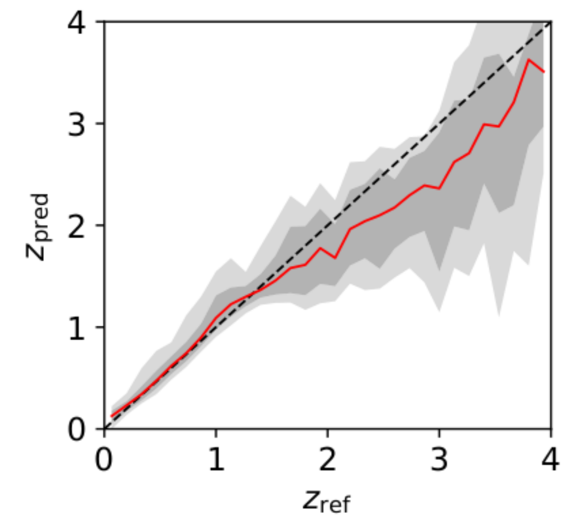
Schuldt et al. 2021b, accepted

## Predict photometric redshifts

- Training over 0 < z < 4, good performance, larger bias at z > 2
- Comparison with DEmP (Hsieh+2014), best method from HSC photo-z team (Nishizawa+2020) → *identical test set*

**CNN estimates based on image cutouts are competitive**




**Bias** $\mathrm{Median}\,(\Delta z_i) = \mathrm{Median}\left(\frac{z_{\mathrm{pred},i} - z_{\mathrm{ref},i}}{1 + z_{\mathrm{ref},i}}\right)$

**Dispersion** $\sigma = 1.48 \times \mathrm{MAD}(\Delta z_i) =$
$= 1.48 \times \mathrm{Median}\,(|\Delta z_i - \mathrm{Median}(\Delta z_i)|)$

**Outlier rate** $f_{\mathrm{outlier}} = \frac{N(|\Delta z_i| > 0.15)}{N_{\mathrm{bin}}}$

# Summary

- Supervised machine learning greatly helps identify strong lenses
- Many false positives, visual inspection needed → impossible with LSST?
- Measure performance: to be tested on independent sets of observed images
- CNNs for automated lens modeling: looks very promising, to be validated on real strong lens systems
- CNNs for photo-z estimates: competitive approach with broad applications, e.g. Rubin Observatory LSST, only requires magnitude and Kron radius cuts → now combine CNNs with catalog-based photometric quantities?