www.holismokes.org

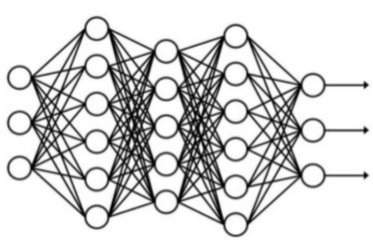
Supervised deep learning methods for Rubin LSST Image classification, mass modeling, and photometric redshift prediction

Raoul Cañameras (MPA/TUM ➜ LAM after 1.12.2023)

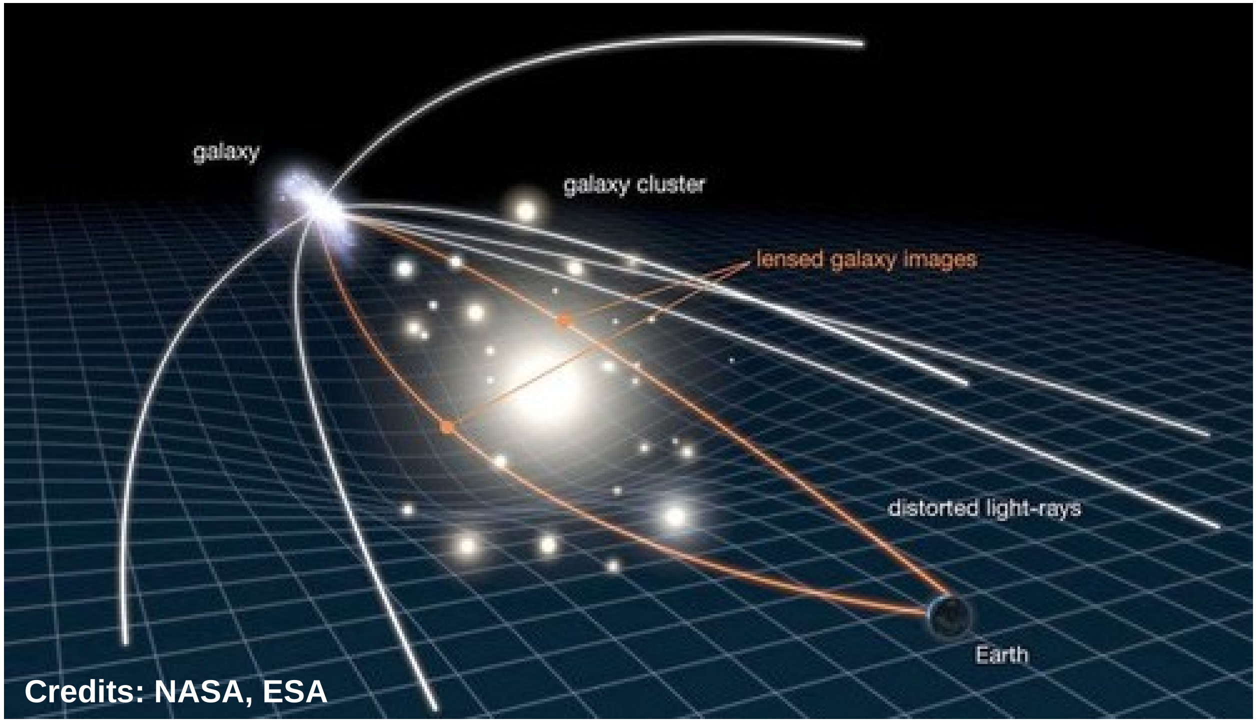
rcanameras@mpa-garching.mpg.de

Seminaire pole ML/DL du CeSAM

November 20, 2023



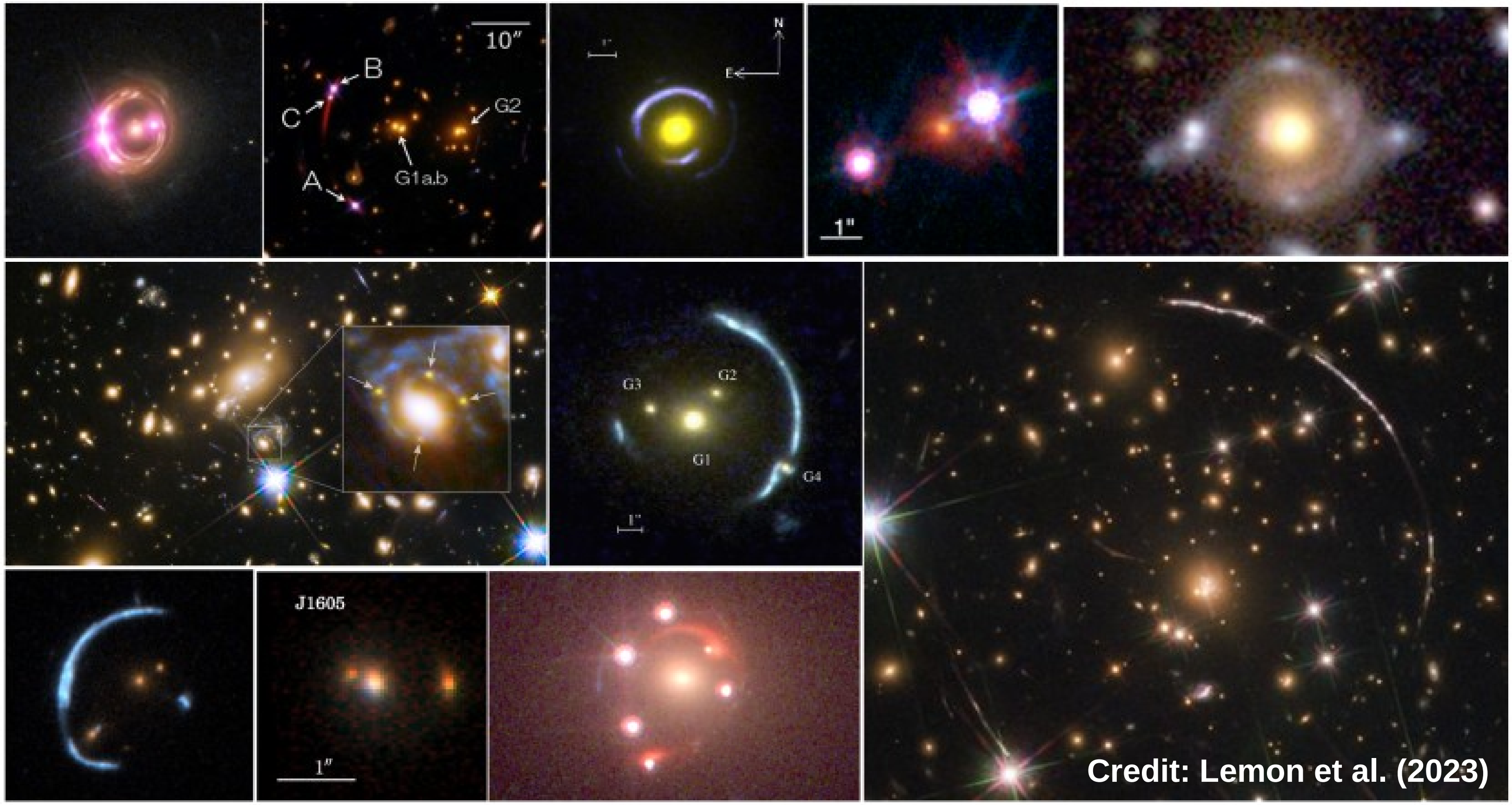
Strong gravitational lensing



Massive galaxy or galaxy cluster bends space-time

Strong lensing regime: Elongated arcs + multiple images of the background galaxy

Strong gravitational lensing
1 - Static background sources, e.g. distant galaxies



Studies of galaxy evolution and dark-matter distribution:

- Total mass of the foreground lens galaxy/cluster
- High-resolution studies of magnified background galaxies

Strong gravitational lensing
2 – Time-variable sources, e.g. quasars, supernovae

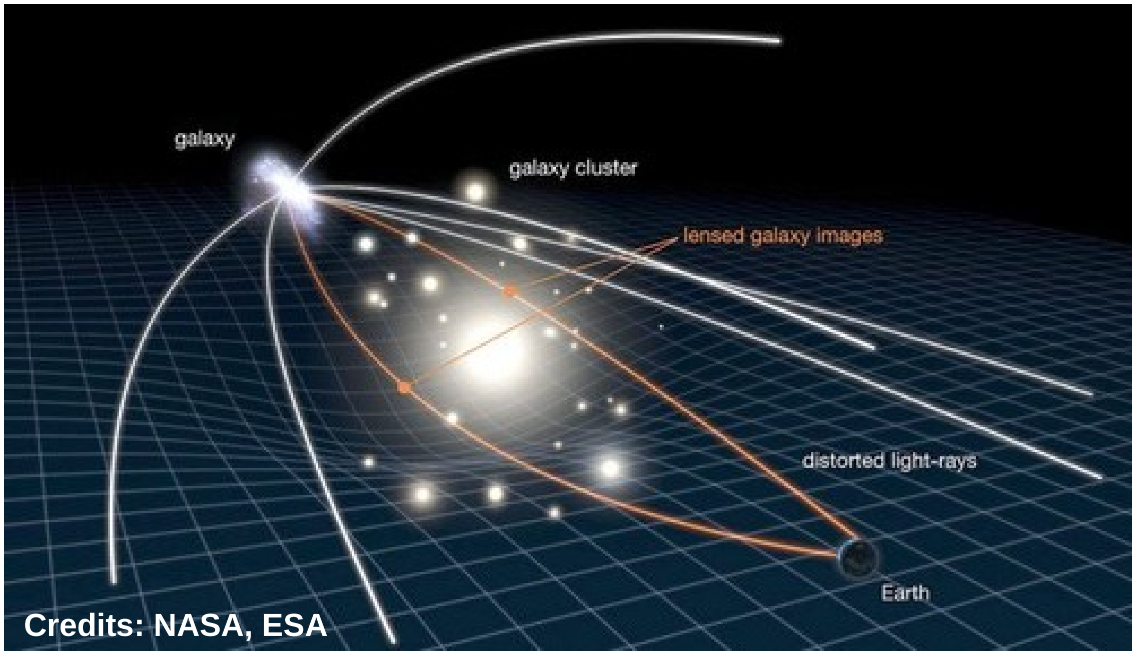


Illustration of lensed supernova event (credit: S. More)

Time delays t

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

Measured with light monitoring

Time-delay distance: **$\propto 1/H_0$**

From lens modeling

Multiple images appear at different times

➜ One-step physical measurement of cosmological distance

➜ **Measure of the Cosmic Expansion rate (Refsdal 1964)**

Searching for new strong gravitational lenses
Case of static galaxy-galaxy lensing systems

Simple binary classification problem (lens vs nonlens) BUT

Strong lensing events are very rare ➜ 1 galaxy out of 10^4 or 10^5 in a given data set

Need to exclude a wide range of contaminants: Spirals, ring galaxies, mergers, etc…

Need to get rid of image **artefacts** automatically + Ensure position/rotation **invariance**

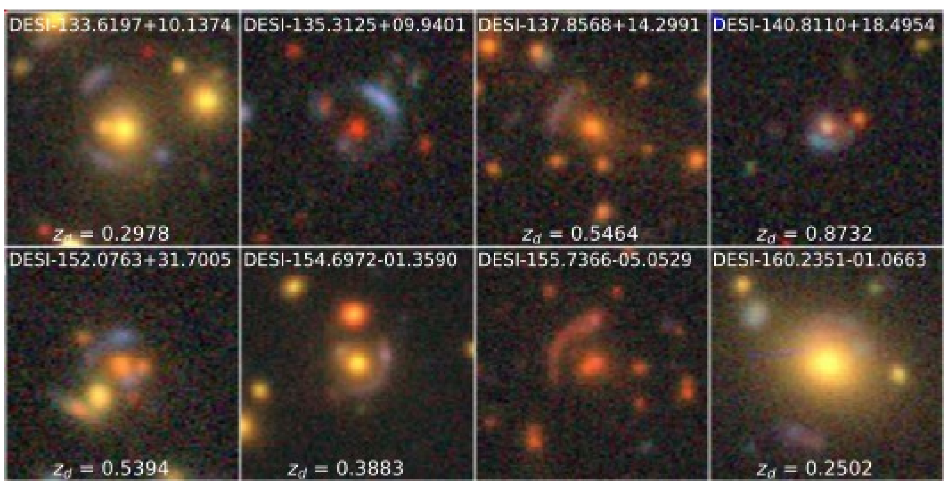
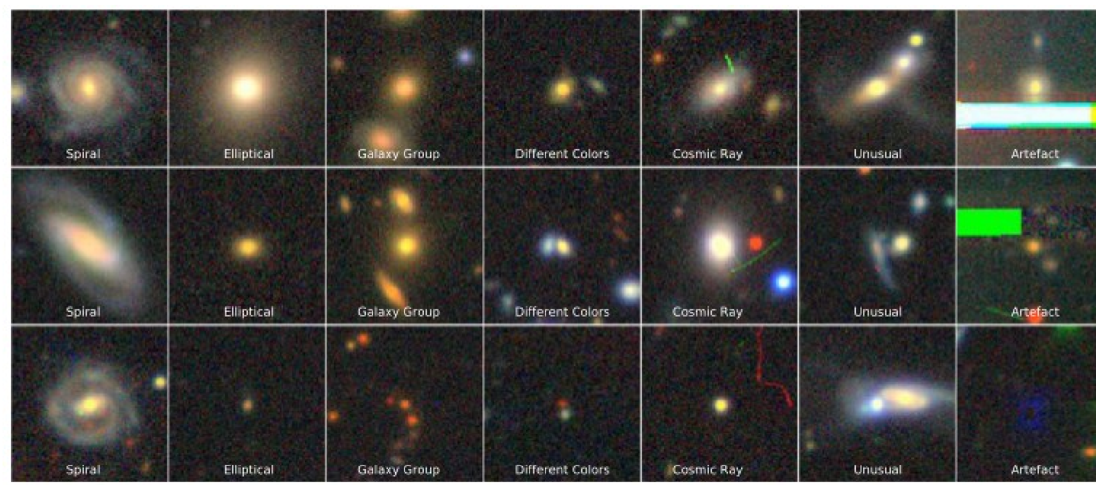


Fig. Examples of strong lens candidates from DECaLS

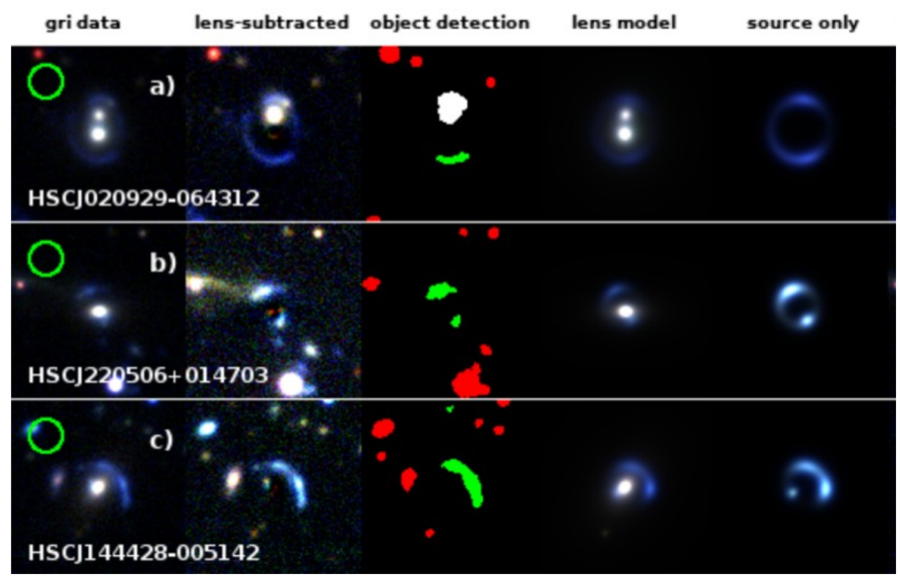


Various types of nonlens contaminants (from Huang et al. 2021)

Searching for new strong gravitational lenses

Case of single or multi-band imaging data sets

Rubin LSST
10^5 new lenses
(Collett 2015)



YATTALENS arcfinder applied to HSC (Sonnenfeld+2018)

Supervised Deep Learning classification

e.g. in CFHTLS (Jacobs+2017); COSMOS HST (Pourrahmani+2018); KiDS (e.g., Petrillo+2017; +2019; Li+2020, Li+2021); DES (e.g., Jacobs+2019a,b, Rojas+2022); DECaLS (e.g., Huang+2020; +2021, Storfer+2022); CFIS (e.g., Savary+2022); DELVE (e.g., Zaborowski+2023)

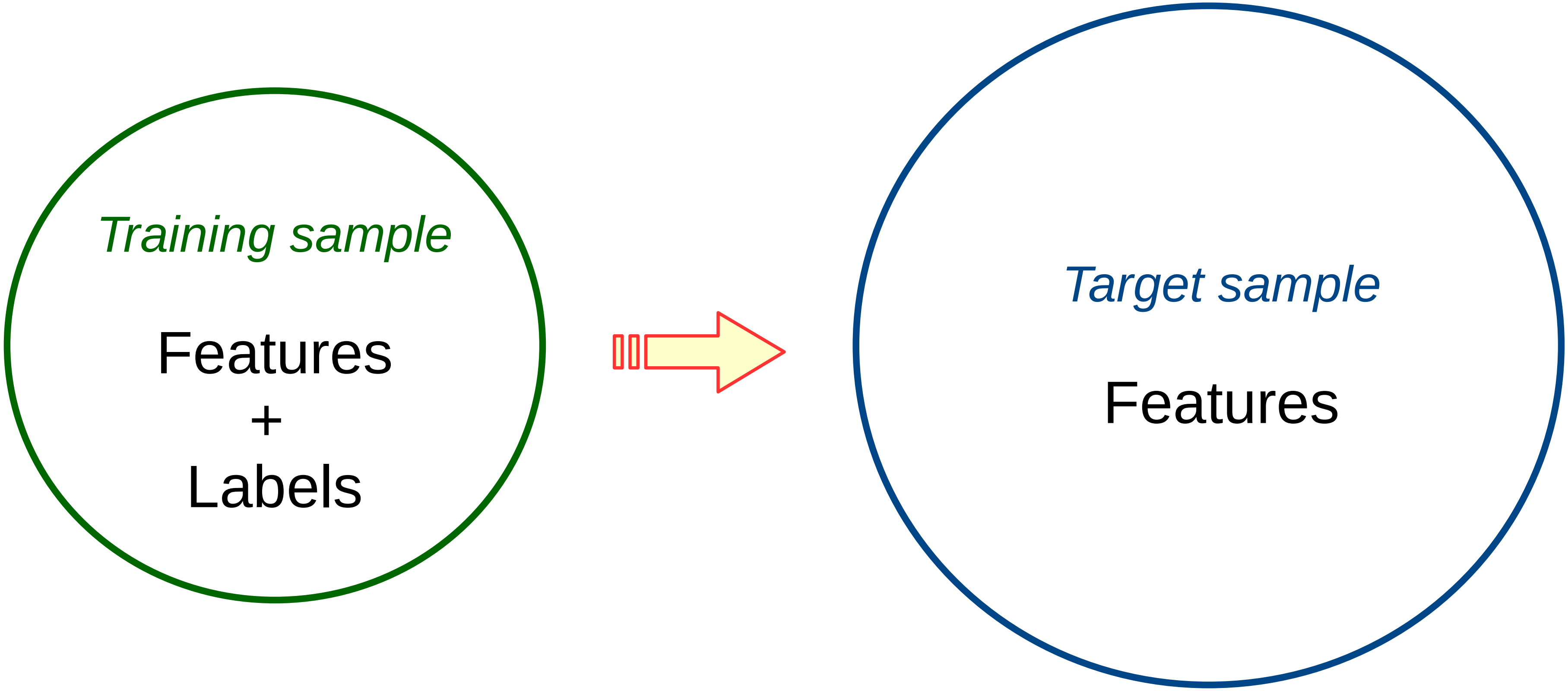
systematically outperforms non-ML algos

e.g. Arc-finder algorithms (e.g., Gavazzi+2014, Avestruz+2019); Principal component analyses (e.g., Joseph+2014; Paraficz+2016); Lens modeling and masking (e.g., Sonnenfeld+2018); Citizen-science projects (e.g., Marshall+2016, Sonnenfeld+2020); Visual inspection (e.g., Diehl+2017, Khullar+2021)

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

Table. Results of lens-finding challenge (Metcalf+2019)

Supervised machine learning classification



Neural Networks

Training phase

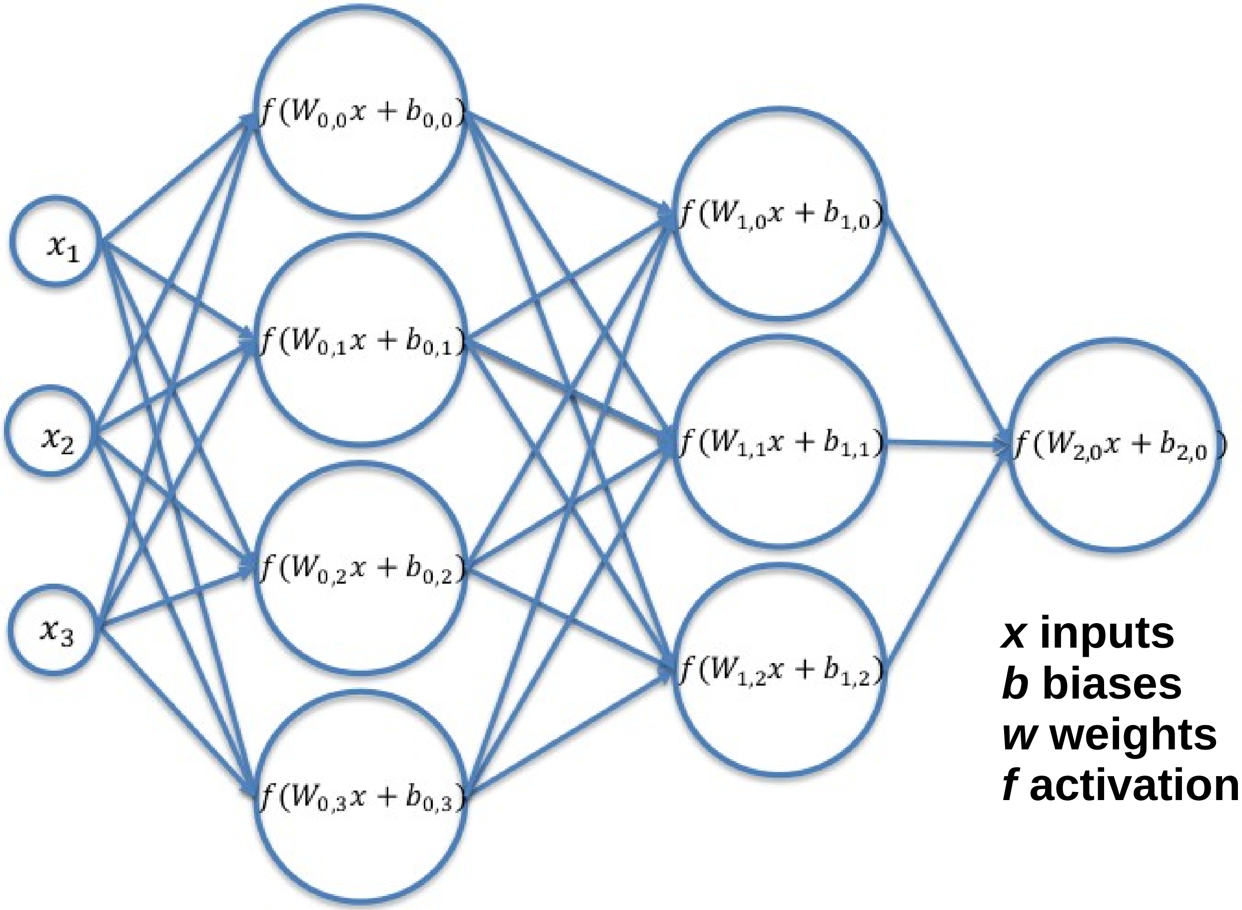


Fig. Credit Leal-Taixe, Niessner

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

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

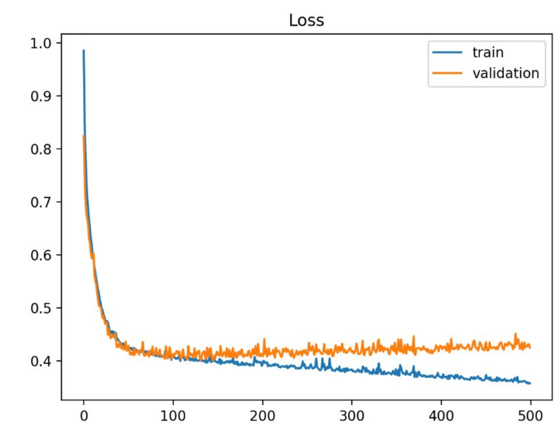


Fig. Credit J. Brownlee

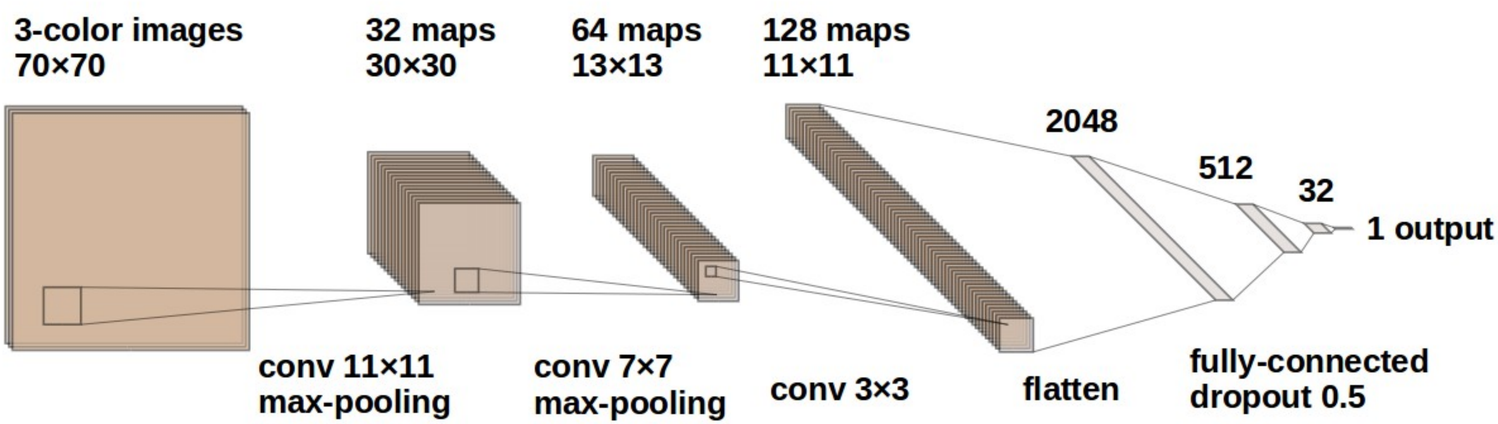
x **inputs**
b **biases**
w **weights**
f **activation functions**

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



Automated pipelines for wide-field imaging surveys

Several hundred candidates from CNNs in the literature

Elevated confirmation rate >80% (Tran et al. 2022)

BUT drastic pre-selection to cope with data volumes

We need fully automated, all-sky searches
for current surveys and for Rubin LSST

+ Extend deep learning methods to mass modeling
and photometric redshift estimation

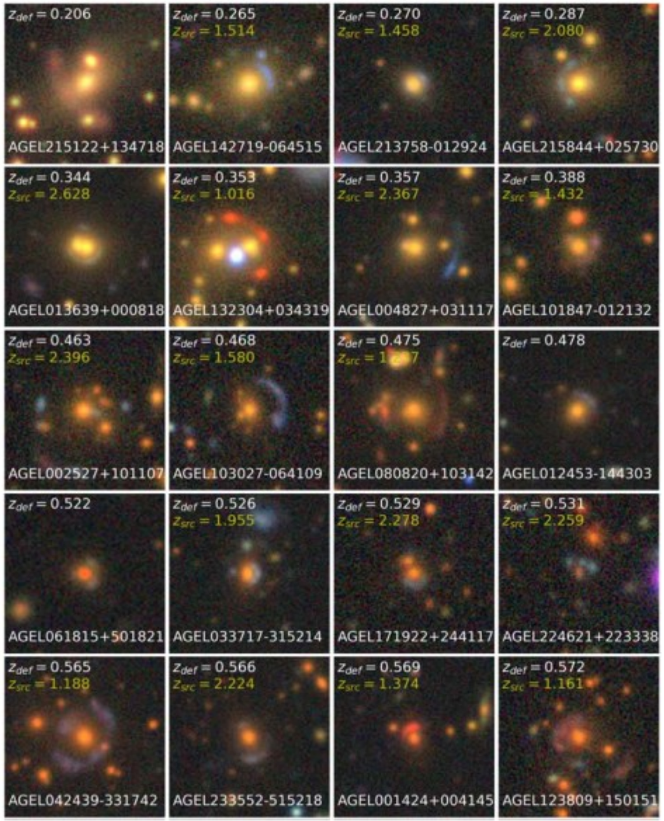


Fig. Follow-up of AGEL lenses (Tran+2022)

Projects part of **HOLISMOKES** (Suyu+2020)

(Highly Optimized Lensing Investigations
of Supernovae, Microlensing Objects, and
Kinematics of Ellipticals and Spirals)

Technische Universität München TUM

HOLI SMOKES

All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163

Systematic strong-lens search over the Pan-STARRS 3π survey (30 000 deg^2)

→ 3×10^9 sources to be classified

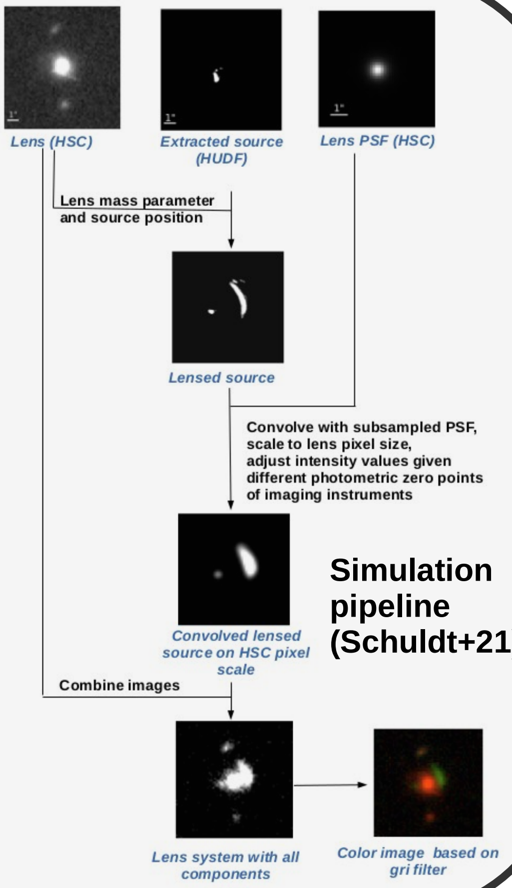
Realistic lens simulations for high classification accuracies

Need realistic lens galaxies, good proxies of lens mass, Einstein radius distributions, number of multiple images, source colors and morphologies

+ match properties of PanSTARRS coadds (sky background, inclusion of neighbours and artifacts, good PSF models, etc)

→ **Paint lensed arcs on survey stacks**

→ **10^5 labelled examples**



Simulation pipeline (Schuldt+21)

All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163

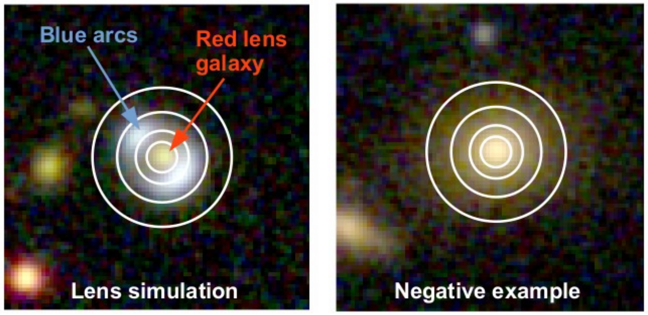
Systematic strong-lens search over the Pan-STARRS 3π survey (30 000 deg²)

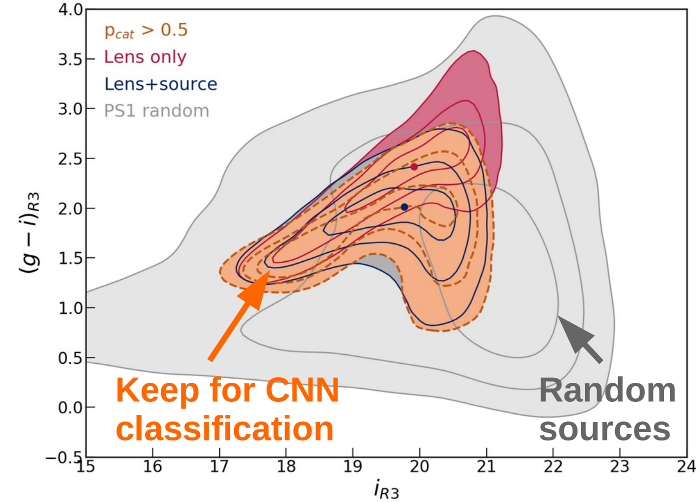
➜ $3\text{x}10^9$ sources to be classified

➜ $2.3\text{x}10^7$ after simple photometric cuts, star removal

Two-step approach to cope with huge data volume

➜ $1.0\text{x}10^6$ after apply neural network on photometry





All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163

Systematic strong-lens search over the Pan-STARRS 3π survey (30 000 deg^2)

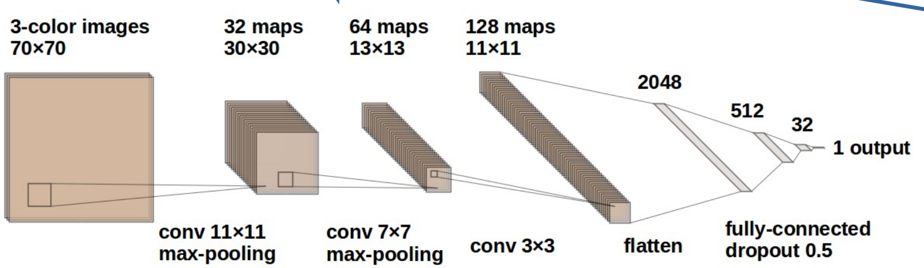
➜ 3x10^9 sources to be classified

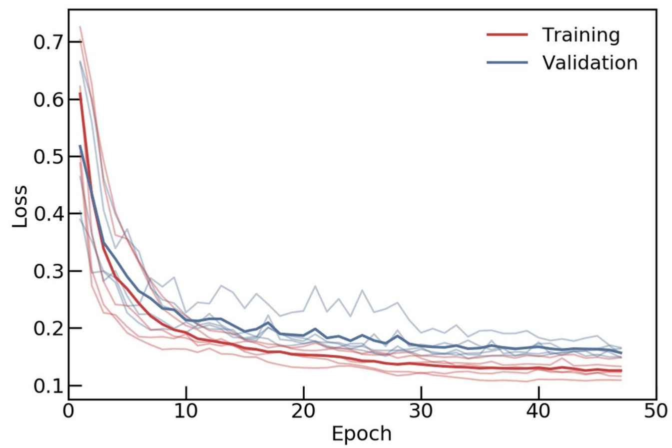
➜ 2.3x10^7 after simple photometric cuts, star removal

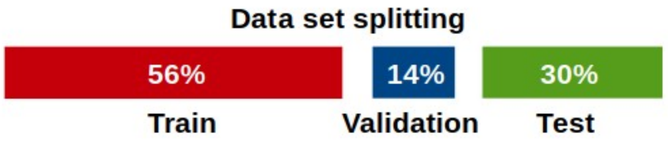
Two-step approach to cope with huge data volume

➜ 1.0x10^6 after apply neural network on photometry

➜ 1.2x10^4 after apply convolutional neural network on g, r, i-band image cutouts

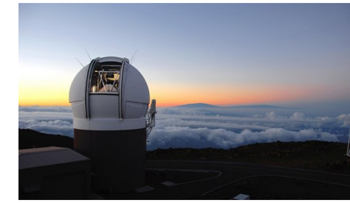


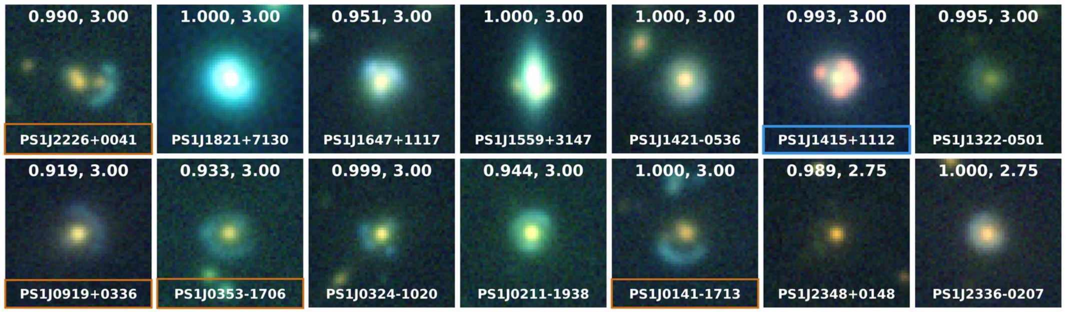




All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163; Cañameras et al. 2021, A&A 6, L6





330 new high-quality lens candidates from Pan-STARRS, follow-up on-going (Taubenberger et al., in prep.)

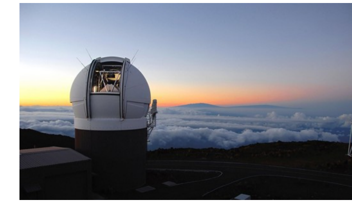
Also successful on ~4 mag deeper imaging from HSC Wide survey (Cañameras et al. 2021, Shu et al. 2022)

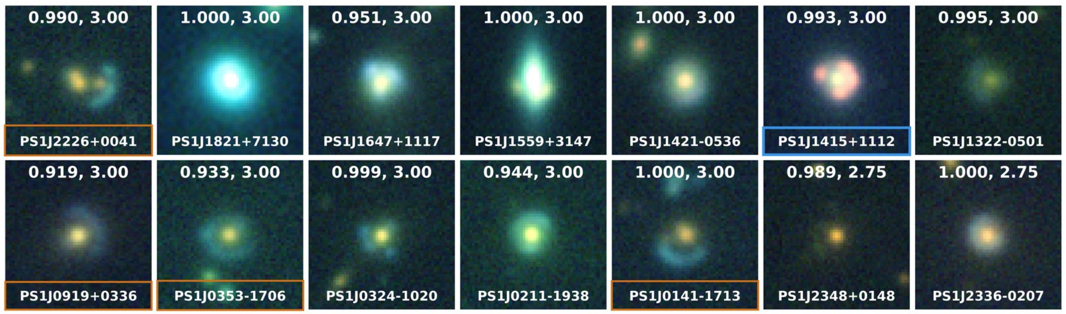
Method for all-sky classification works + Applicable to the future Rubin LSST stacks

Lens candidates from HSC (Shu et al. 2022)

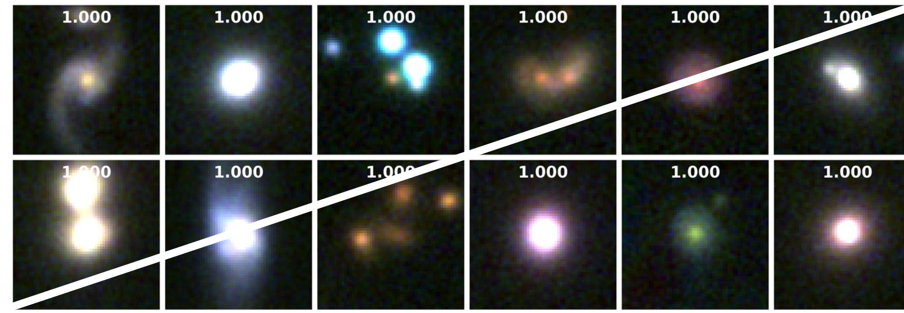
All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163; Cañameras et al. 2021, A&A 6, L6





330 new high-quality lens candidates from Pan-STARRS, follow-up on-going (Taubenberger et al., in prep.)



BUT ALSO several false positives!

Also successful on ~4 mag deeper imaging from HSC Wide survey (Cañameras et al. 2021, Shu et al. 2022)

Method for all-sky classification works + Applicable to the future Rubin LSST stacks

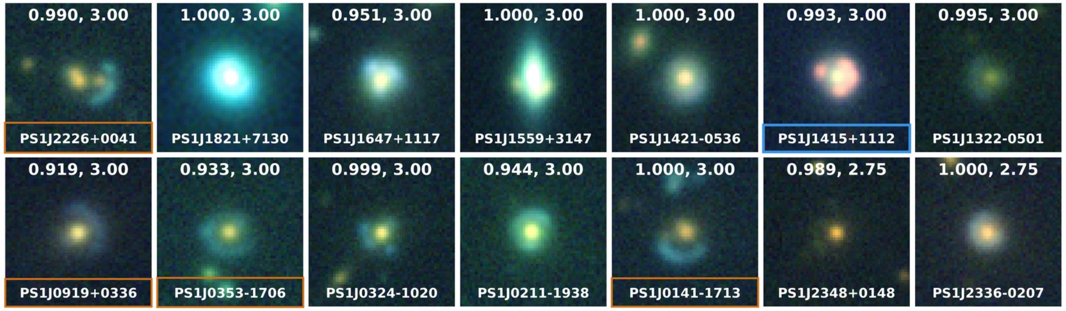
Not fully automated due to number of contaminants

Inspection of few 10^4 to 10^6 cutouts depending on survey depth

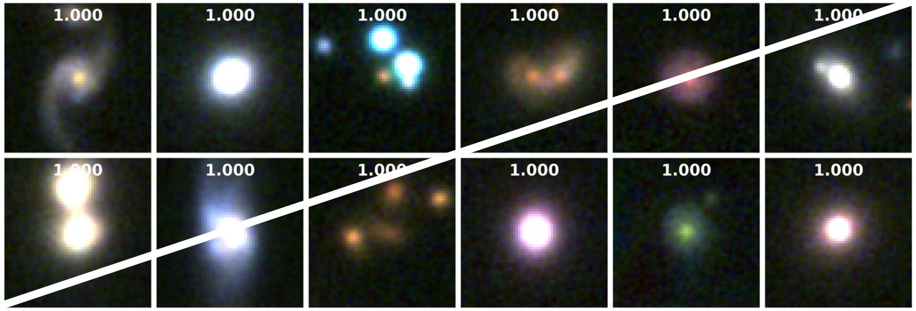
After few hours of “expert” visual inspection of lens candidates

All-sky classification of static images

Cañameras et al. 2020, A&A 644, 163; Cañameras et al. 2021, A&A 6, L6



330 new high-quality lens candidates from Pan-STARRS, follow-up on-going (Taubenberger et al., in prep.)



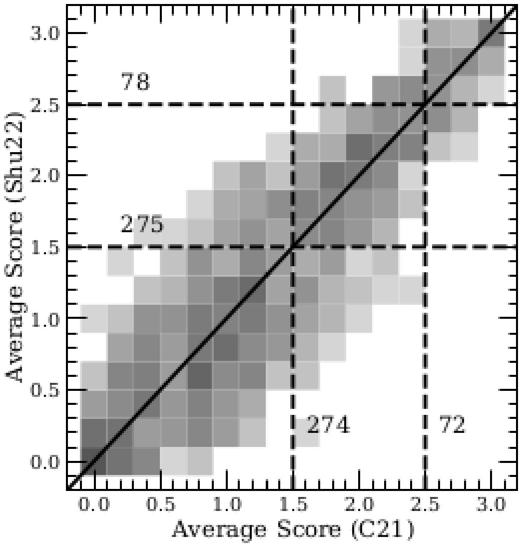
BUT ALSO several false positives!

Also successful on ~4 mag deeper imaging from HSC Wide survey (Cañameras et al. 2021, Shu et al. 2022)

Method for all-sky classification works + Applicable to the future Rubin LSST stacks

Not fully automated due to number of contaminants

Inspection of few 10^4 to 10^6 cutouts depending on survey depth ➜ **Strong biases?**



Grading and regrading (Shu et al. 2022)

Evaluation of supervised neural networks

How can we reduce human input?

Cañameras et al. 2023, arXiv:2306.03136

➜ Build **test set from real survey data** (e.g. HSC Wide)

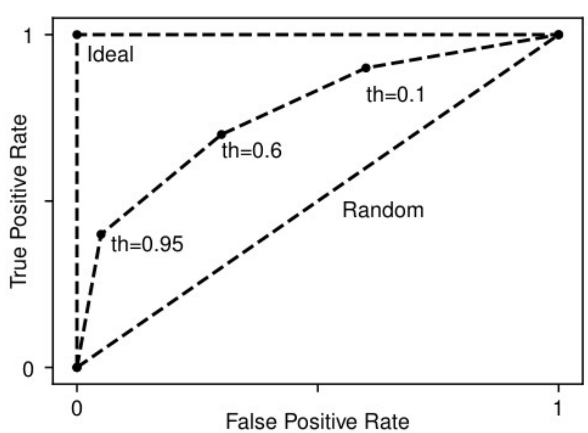
- 220 known galaxy-scale lenses (SuGOHI project)
  ➜ Test completeness for different configurations
- 70,000 non-lenses in COSMOS
  ➜ Measure correct number of false positives

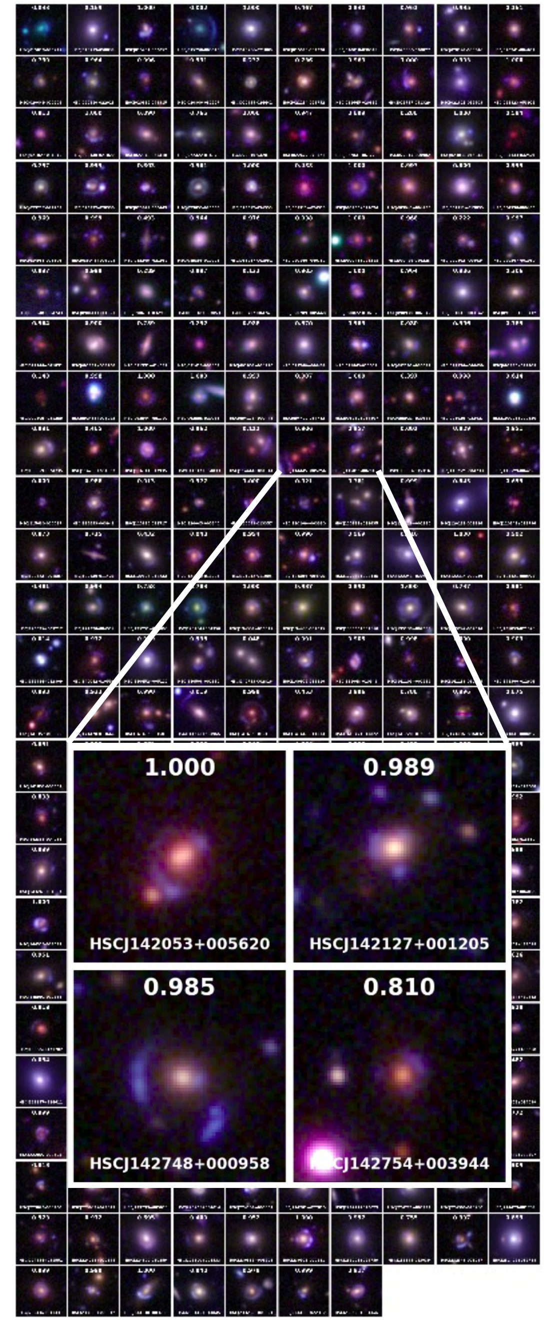
Metrics

➜ area under ROC

➜ TPR_0 and TPR_{10}

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}; \ \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}$$





Evaluation of supervised neural networks

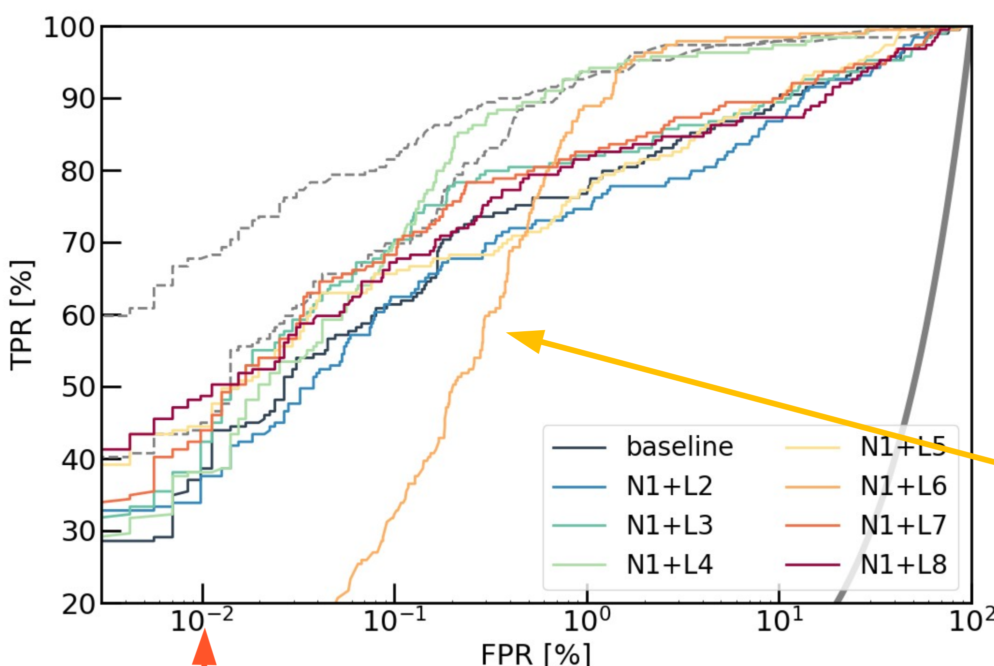
Cañameras et al. 2023, arXiv:2306.03136

The design of the ground-truth data set is key to improve performance

Interactive machine learning ➜ here by modifying the training sample iteratively

Test multiple **combinations of positive examples** (simulated lenses), fixing everything else

➜ Balanced data sets of about 10^5 examples



	CNN		
Data set	AUROC	TPR_0	TPR_{10}
baseline (L1+N1)	0.9557	0.0	44.4
N1 + ...			
L2	0.9527	7.4	41.8
L3	0.9581	0.0	49.8
L4	0.9919	0.0	43.4
L5	0.9686	8.5	49.7
L6	0.9891	0.0	0.0
L7	0.9624	0.0	50.3
L8	0.9531	0.0	50.3

High AUROC but low TPR_{10}

Our goal

Evaluation of supervised neural networks

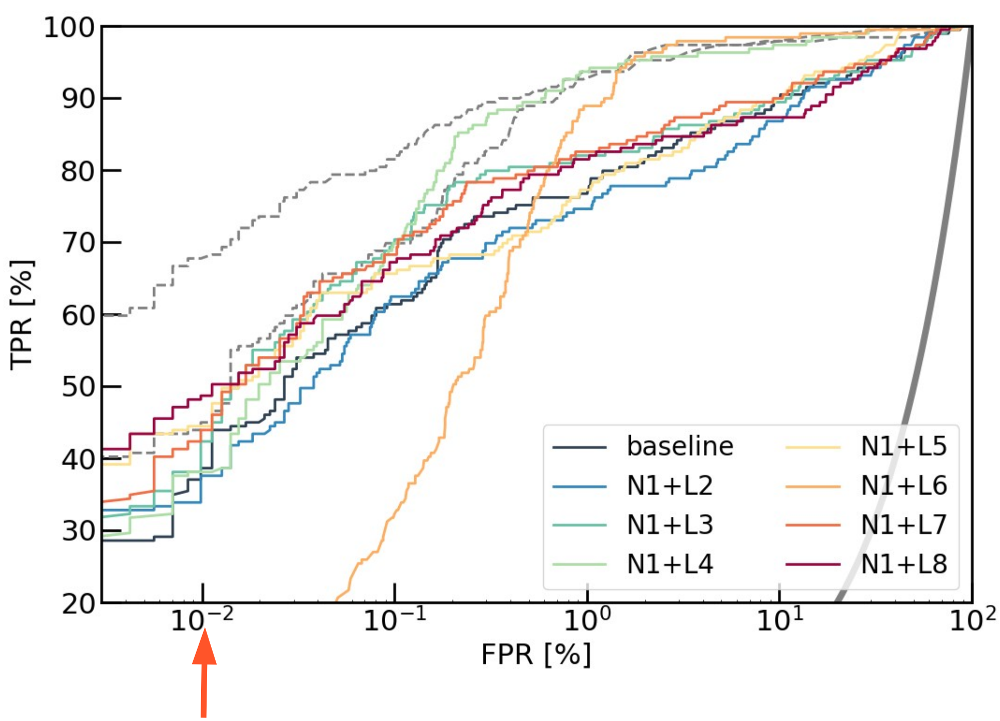
Cañameras et al. 2023, arXiv:2306.03136

The design of the ground-truth data set is key to improve performance

Interactive machine learning ➜ here by modifying the training sample iteratively

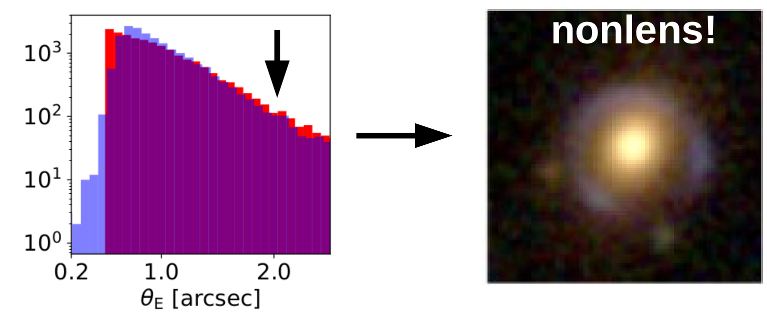
Test multiple **combinations of positive examples** (simulated lenses), fixing everything else

➜ Balanced data sets of about 10^5 examples

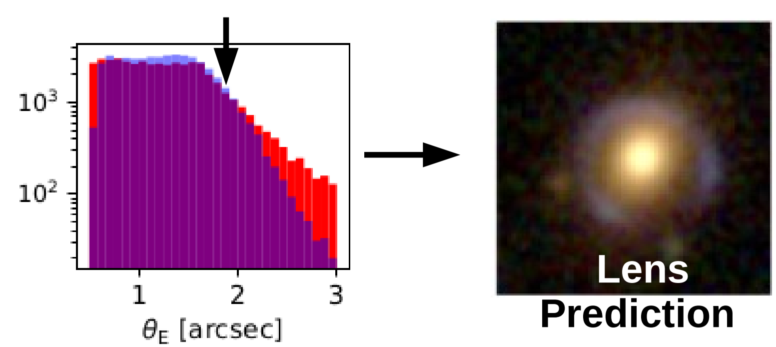


Our goal

Prediction



Parameter distributions chosen in simulations play a major role (no need to follow nature)



Prediction

Evaluation of supervised neural networks

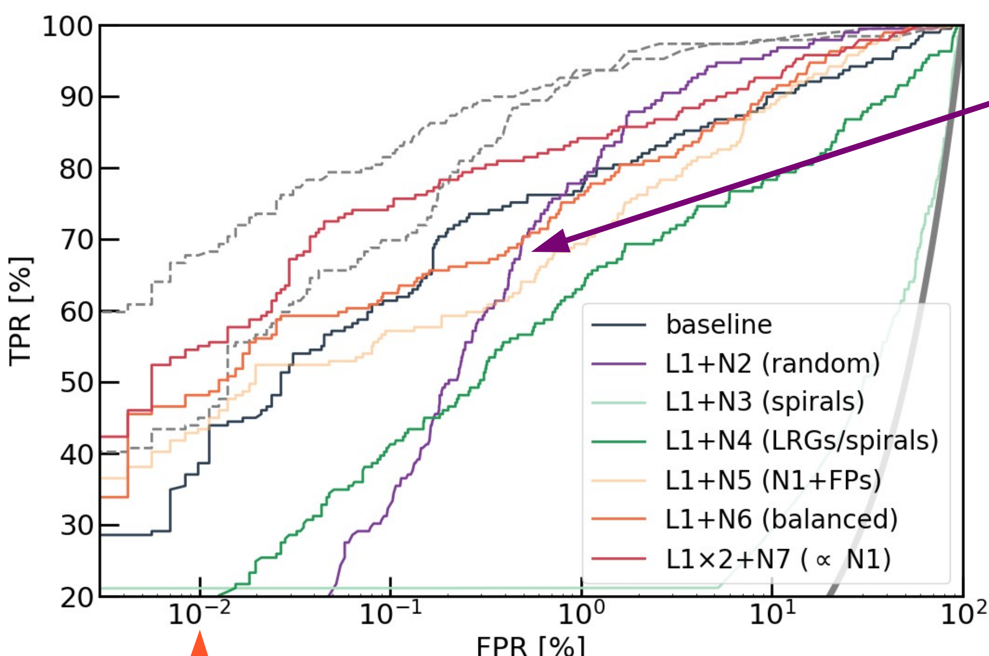
Cañameras et al. 2023, arXiv:2306.03136

The design of the ground-truth data set is key to improve performance

Interactive machine learning ➜ here by modifying the training sample iteratively

Test multiple **combinations of negative examples** (non-lenses), fixing everything else

➜ Balanced data sets of about 10^5 examples

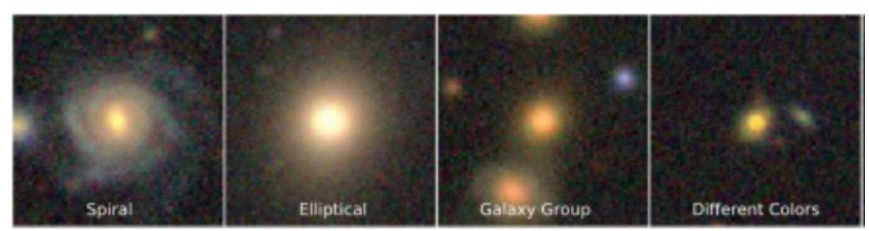


Our goal

Drawing random non-lenses does not work

➜ **Need to boost fractions of usual contaminants (spirals, rings, groups, etc)**

➜ Use external citizen science projects or unsupervised ML classifications

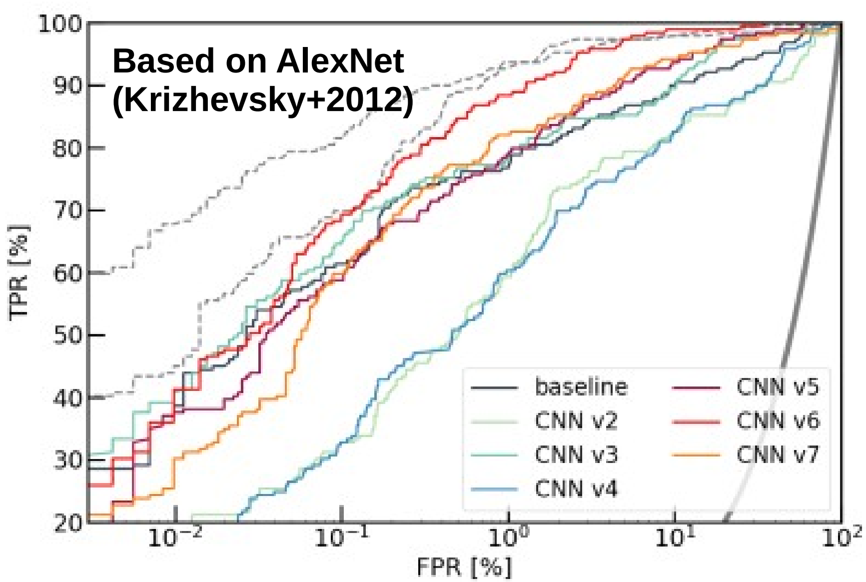


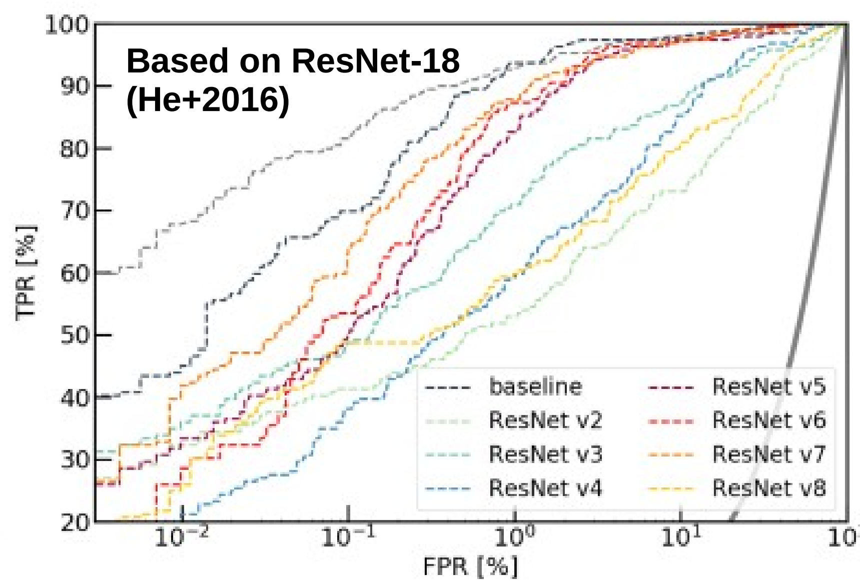
Evaluation of supervised neural networks

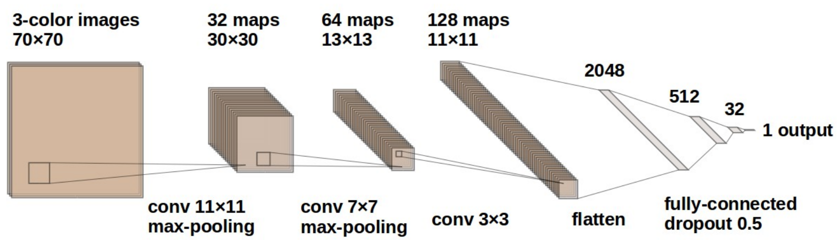
Cañameras et al. 2023, arXiv:2306.03136

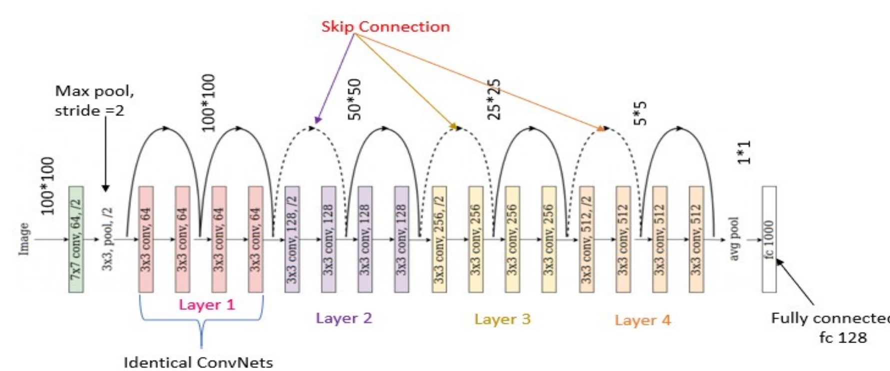
Fine-tuning the network architecture also plays a major role

Test multiple **network architectures**, fixing everything else









ResNets not always better than classical CNNs for small image sizes

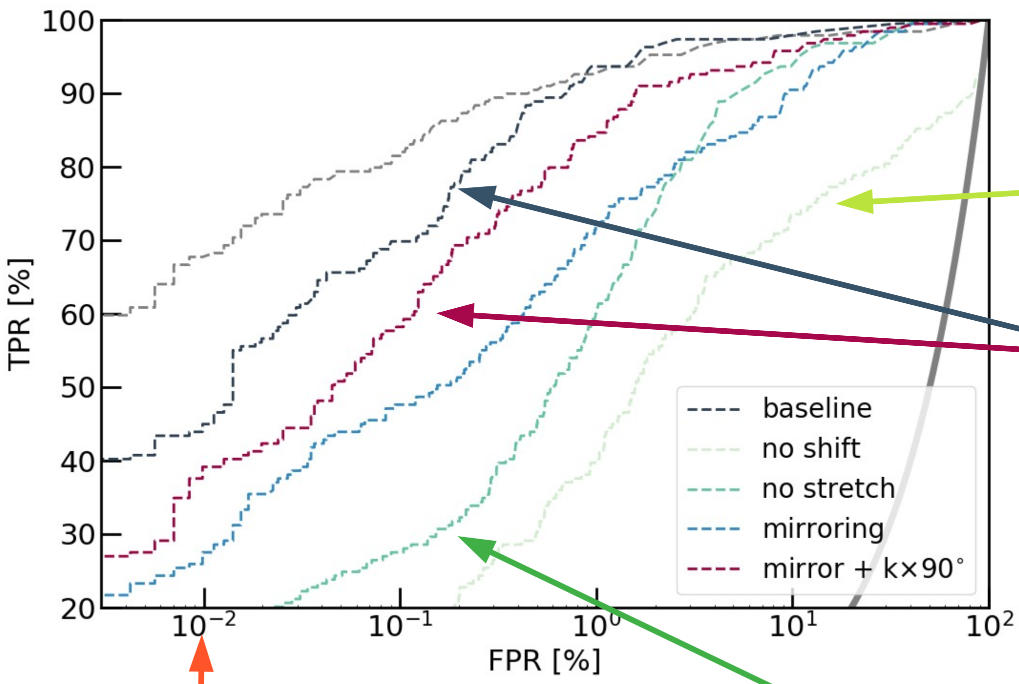
Evaluation of supervised neural networks

Cañameras et al. 2023, arXiv:2306.03136

The processing of the ground-truth data set is important

Test multiple **data augmentation techniques**, fixing everything else

➜ Balanced data sets and baseline ResNet architecture



NN can learn when galaxies are centered “too perfectly”

Random rotations in data loader are sufficient to achieve rotation invariance

Square root stretches to the pixel values are helpful

Our goal

Evaluation of supervised neural networks - Summary

Cañameras et al. 2023, arXiv:2306.03136

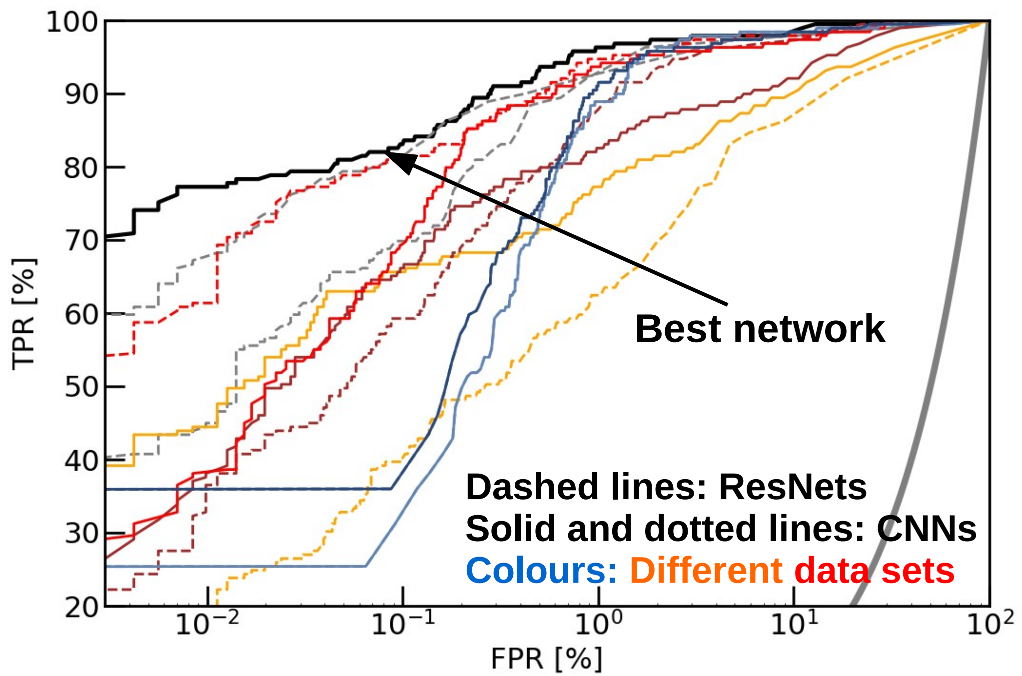
- Major improvements for specific networks and data sets ➜ **FPRs from 1% to 0.01%!**
- Performance directly indicate behaviour on real survey data
- Recall at low contamination has significantly increased (see ROC curves)

Given class imbalance, still o(10^4) galaxies to eyeball ➜ How can we improve?

Other tests completed

Architecture level: group-invariant network architectures (e.g., Cohen+2016, Schaefer+2018) - networks pre-trained on ImageNet - multiclass classification (e.g., Teimoorinioa+2020)

Data set level: Influence of the number of observing bands - Lens-light subtraction - Masking of neighbouring galaxies - Denoizing image cutouts - Deconvolving image cutouts - combining classification + modeling networks

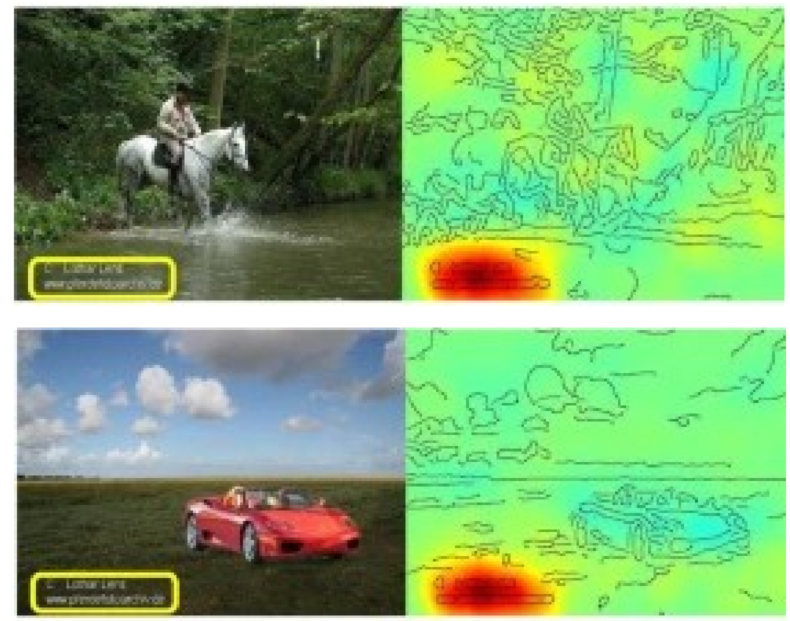


Receiver Operating Characteristic (ROC) curves using observed HSC lenses and non-lenses.

Robustness of neural network classification

Cañameras et al. 2023, arXiv:2306.03136

Has the model based its decision on a spurious correlation in the training data ?



Two pictures labelled as “horse”
(Credit: Lapuschkin et al. 2019)

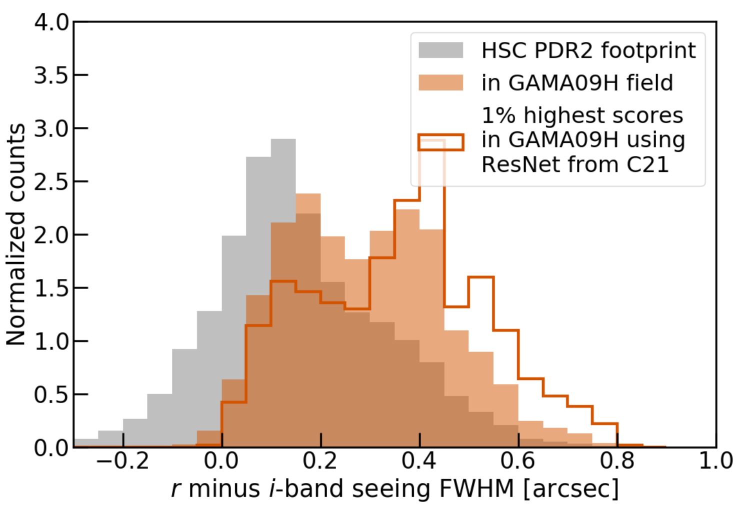
Robustness of neural network classification

Cañameras et al. 2023, arXiv:2306.03136

Has the model based its decision on a spurious correlation in the training data ?

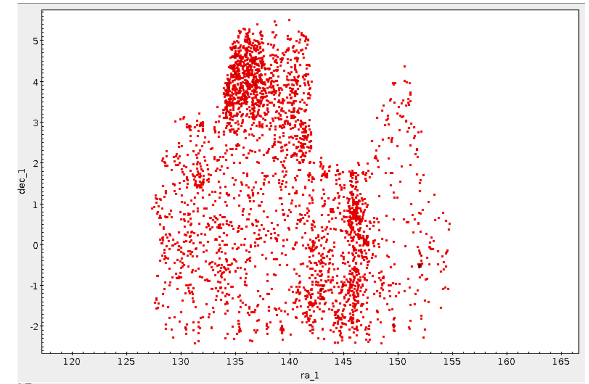
➜ Problems due to PSF mismatches between bands ?

➜ Some networks show systematic dependence with variations in seeing FWHM



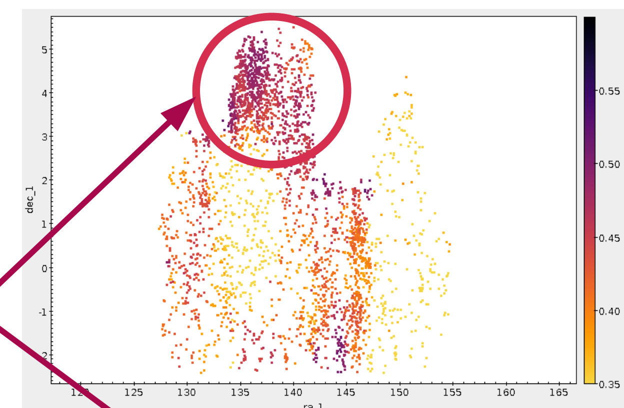
➜ Add PSF frames as input with science frames

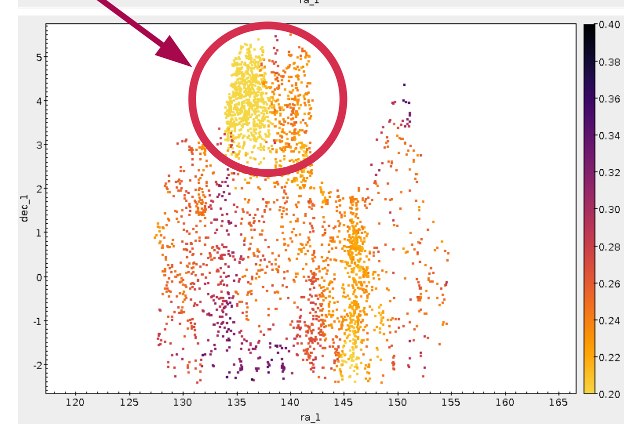
CNN candidates in GAMA09H



Opposite trends in r- and i-band seeing

➜ color gradients



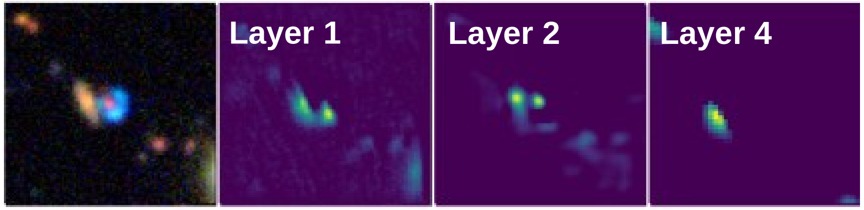


Interpretability of strong-lens finding neural networks
Examples of local methods

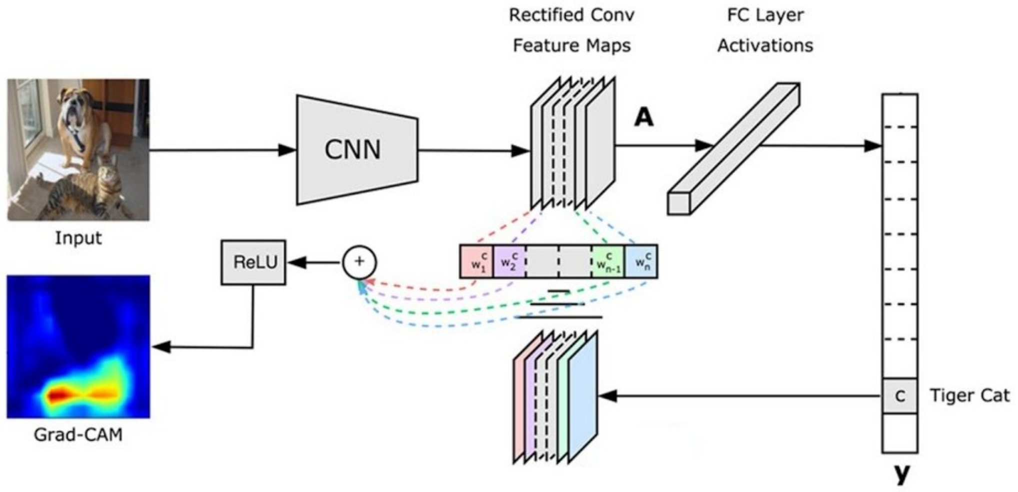
To constrain selection functions, improve performance, and identify biases ?

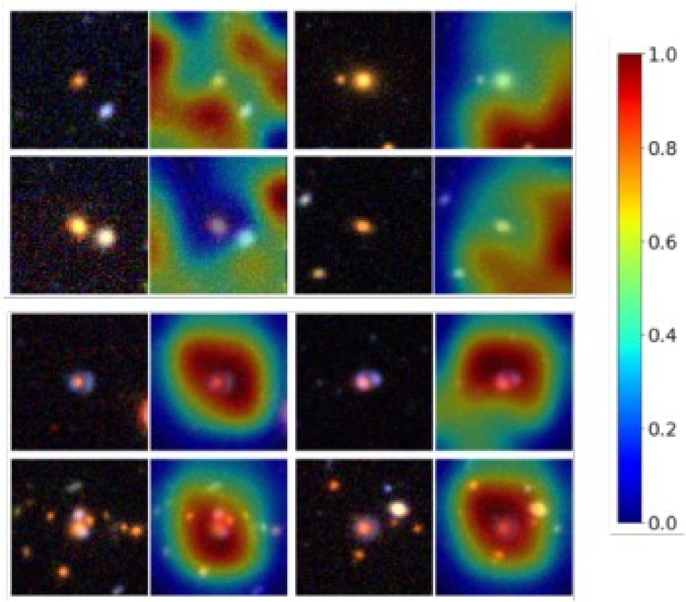
➜ **Where** is the most useful information for lens/nonlens classifications ?

➜ Visual inspection of feature maps: Initial vs. later layers (Jacobs+2022)



➜ Saliency mapping: Gradient-weighted Class Activation Mapping (Selvaraju+2017)





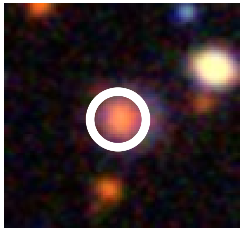
Interpretability of strong-lens finding neural networks
Examples of local methods

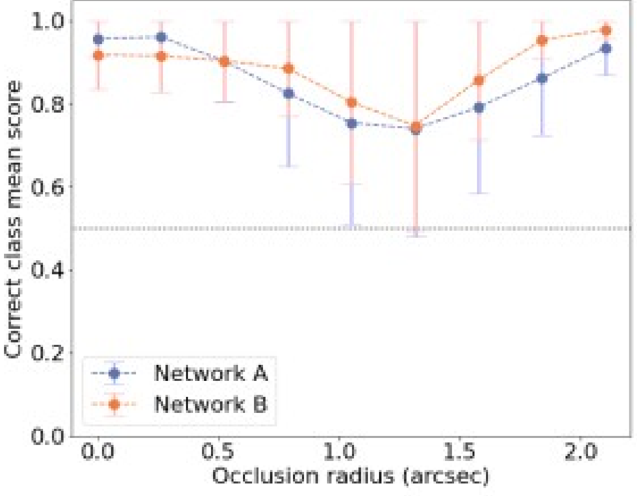
To constrain selection functions, improve performance, and identify biases ?

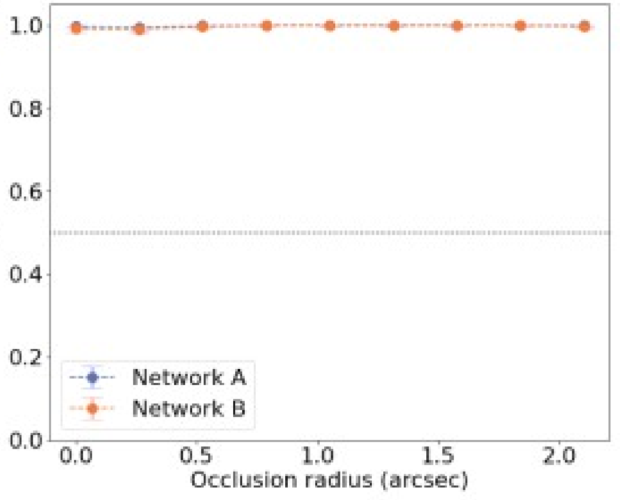
➜ **Where** is the most useful information for lens/nonlens classifications ?

➜ Sensitivity probes are easy to implement ➜ Occlusion mapping (Zeiler & Fergus 2014)

- Annular masks centered on the central galaxy
- CNNs “seem to” use relevant information (lensed arcs and multiple images)





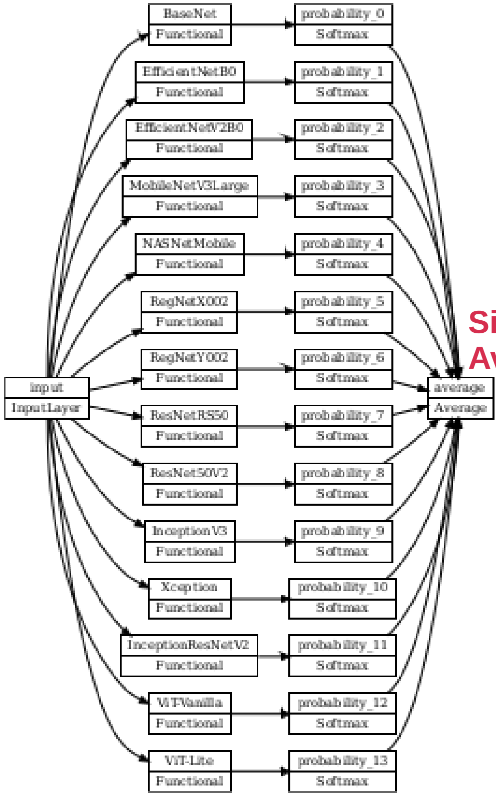


Finding rare strong lenses in large data sets

Towards automated selections with deep, wide-scale surveys

Network ensembles are very efficient in decreasing false positive rates (Hansen+1990)

- ➜ Individual neural networks learn different representations
- ➜ Ensembles mitigate stochasticity of learning process, and lower the variance in output scores
- ➜ Combine networks, e.g. **with different architectures** (fixed training data, fixed augmentation)



Architecture	Parameters	CPU hours	TPR	FPR	AUROC	Candidates	Lens	Model reference
BaseNet	581 828	104	0.906	0.097	0.973	62 326	13	Andika et al. (2023)
RegNetX002	2 338 692	1202	0.924	0.072	0.983	160 852	17	Radosavovic et al. (2020)
RegNetY002	2 816 896	1317	0.899	0.078	0.977	192 797	13	Radosavovic et al. (2020)
MobileNetV3Large	3 000 484	148	0.938	0.046	0.989	190 808	16	Howard et al. (2019)
EfficientNetB0	4 055 275	314	0.945	0.042	0.991	147 705	19	Tan & Le (2019)
NASNetMobile	4 274 520	434	0.948	0.048	0.991	169 309	17	Zoph et al. (2018)
EfficientNetV2B0	5 925 012	243	0.958	0.030	0.994	151 319	15	Tan & Le (2021)
ViT-Vanilla	9 208 772	1185	0.949	0.050	0.990	24 642	16	Dosovitskiy et al. (2020)
ViT-Lite	9 230 060	3499	0.960	0.029	0.994	12 872	16	Lee et al. (2021)
Xception	20 870 252	1041	0.970	0.017	0.997	216 620	17	Szegedy et al. (2016)
InceptionV3	21 811 556	549	0.967	0.020	0.996	186 190	17	Szegedy et al. (2015)
ResNet50V2	23 579 268	1094	0.967	0.021	0.996	164 722	18	He et al. (2016)
ResNetRS50	33 705 060	1245	0.966	0.021	0.996	155 262	16	Bello et al. (2021)
InceptionResNetV2	54 343 460	1630	0.966	0.014	0.996	150 435	18	Chollet (2016)
Ensemble	195 741 135	···	0.963	0.016	0.996	3080	16	This work

Ensemble classifier for lensed quasar searches in HSC Wide images (Andika+2023)

Finding rare strong lenses in large data sets
Towards automated selections with deep, wide-scale surveys

Network ensembles are very efficient in decreasing false positive rates (Hansen+1990)

→ Individual neural networks learn different representations

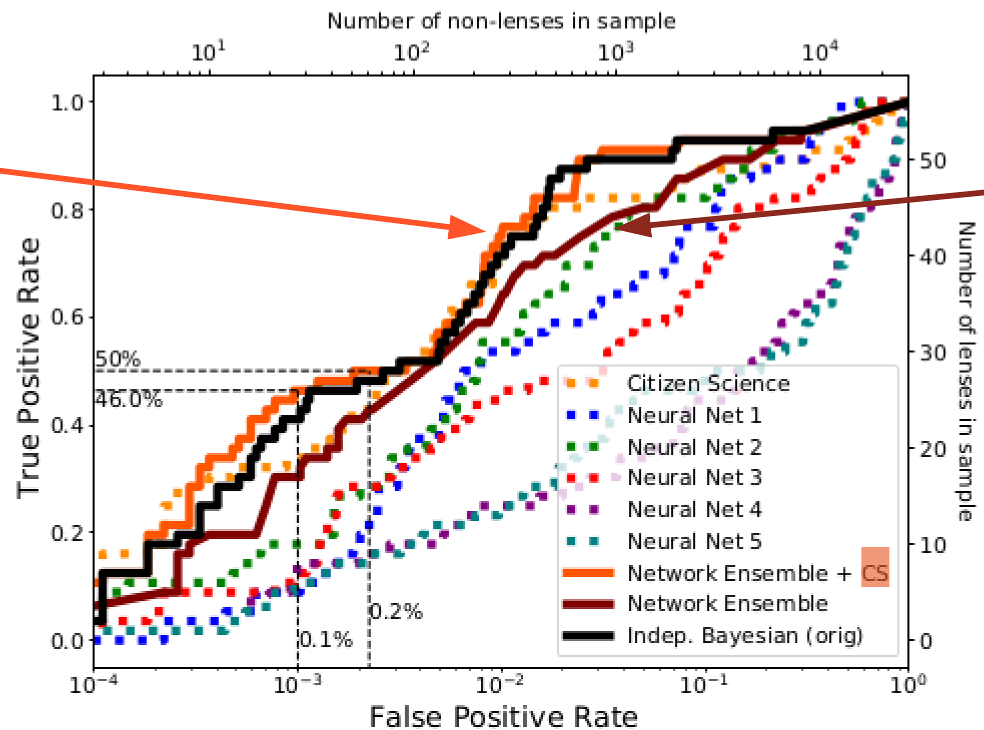
→ Ensembles mitigate stochasticity of learning process, and lower the variance in output scores

→ Combine networks with **different architectures, different ground-truth, data augmentation**

  → Independent networks identify different populations of contaminants

Combine neural networks with Citizen Science

→ Very different selection functions (Knabel et al. 2020)



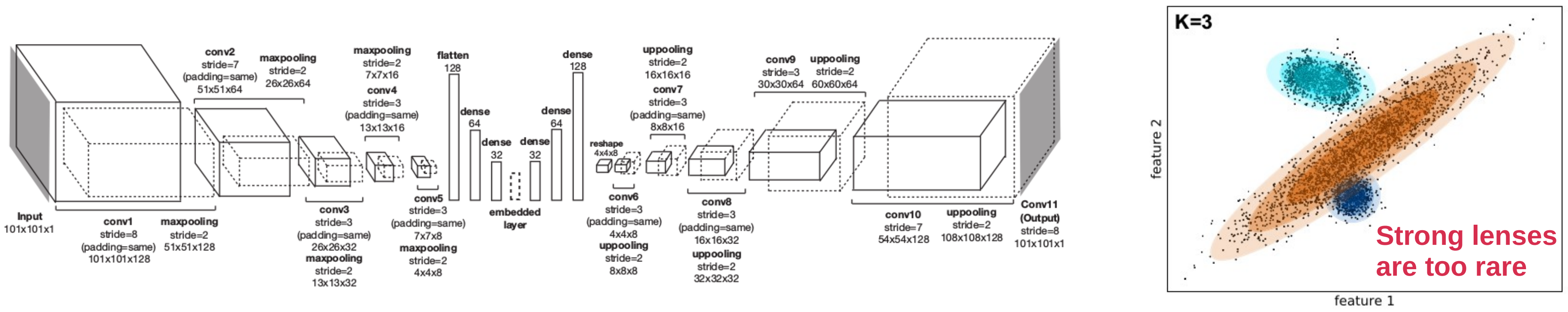
Simple Average

Ensemble classifier for HSC Wide images (Holloway+2023)

Finding rare strong lenses in large data sets
Towards automated selections for deep, wide-scale surveys

Unsupervised learning algorithms

➜ For direct, fully-automated classification of strong lenses (Cheng et al. 2020)



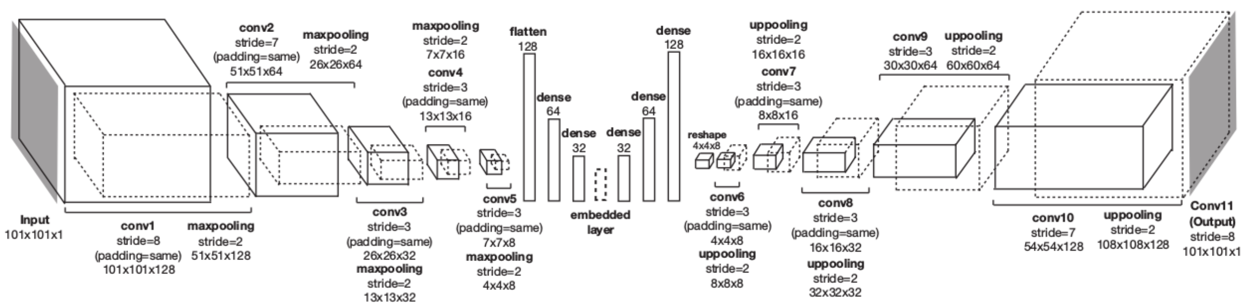
Lens finding with (1) a convolutional autoencoder, and (2) a Bayesian Gaussian mixture model (Cheng et al. 2020)

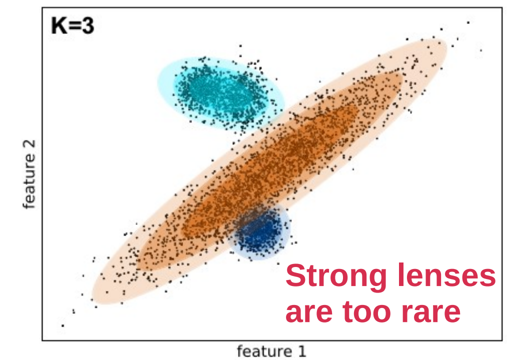
➜ **Result in elevated contamination rates ...**

Finding rare strong lenses in large data sets
Towards automated selections for deep, wide-scale surveys

Unsupervised learning algorithms

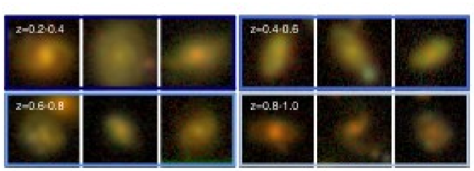
➜ For direct, fully-automated classification of strong lenses (Cheng et al. 2020)



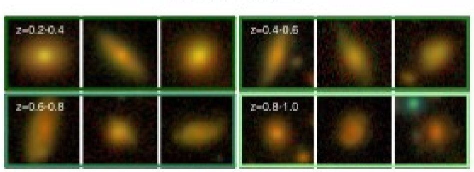


Lens finding with (1) a convolutional autoencoder, and (2) a Bayesian Gaussian mixture model (Cheng et al. 2020)

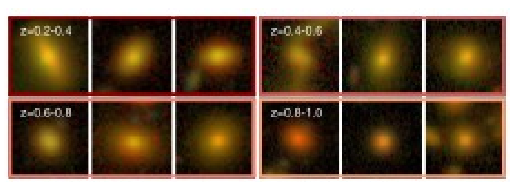
➜ For creating large catalogs of non-lens galaxies + retrain a supervised learning algorithm



(a) *Spiral galaxies.*



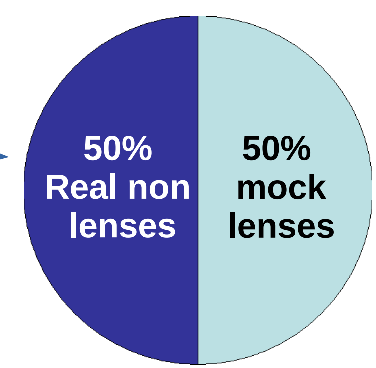
(b) *S0/Sa galaxies.*



(c) *Elliptical galaxies.*

Morphological groups from unsupervised classification of HSC Udeep (Martin et al. 2020)

Millions of objects for updated training set



Finding rare strong lenses in large data sets
Towards automated selections for deep, wide-scale surveys

Unsupervised learning algorithms

➜ for deblending image components (Savary+2022), or conducting image denoising (Cheng+2020) ?

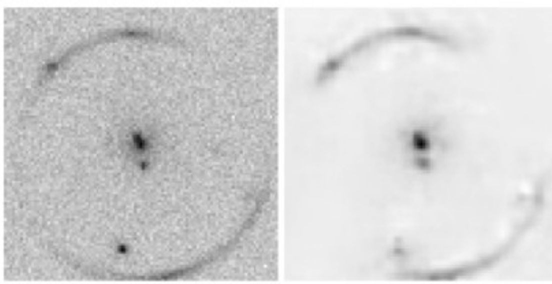
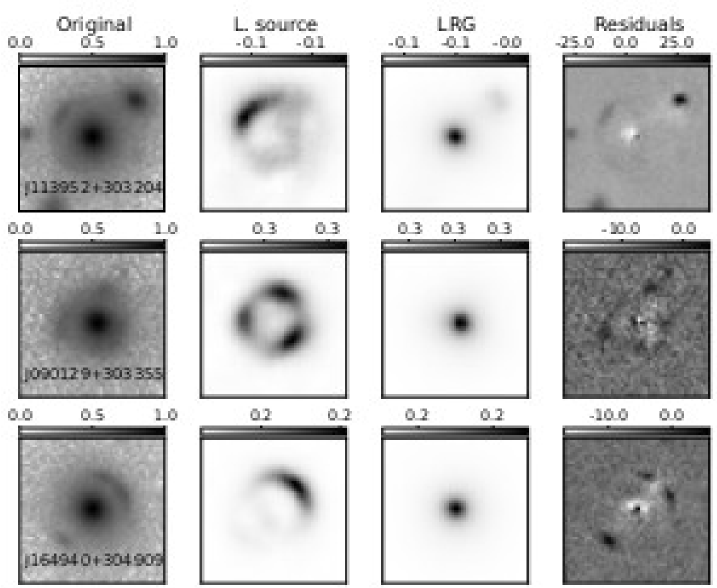
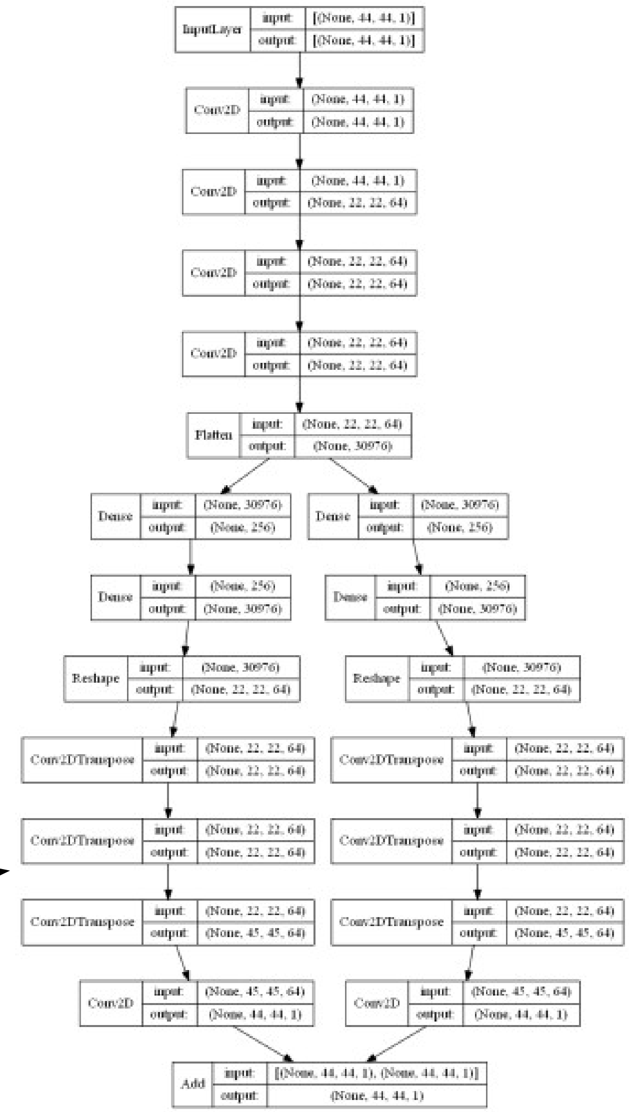


Image denoizing with a convolutional autoencoder (Cheng+2020)



Lens/source deblending with a convolutional autoencoder (Savary+2022)



Efficient strong lens modeling with deep learning

Schuldt et al. 2021a, A&A 646, A126; Schuldt et al. 2023a, A&A 671, A147

Estimate mass profile parameters

➜ Singular Isothermal Ellipsoid + external shear and uncertainties

Regression convolutional neural network

(see also, e.g., Hezaveh+2017, Perreault-Levasseur+2017, Madireddy+2019, Park+2020, Pearson+2019,+2021)

Efficient strong lens modeling with deep learning

Schuldt et al. 2021a, A&A 646, A126; Schuldt et al. 2023a, A&A 671, A147

Estimate mass profile parameters

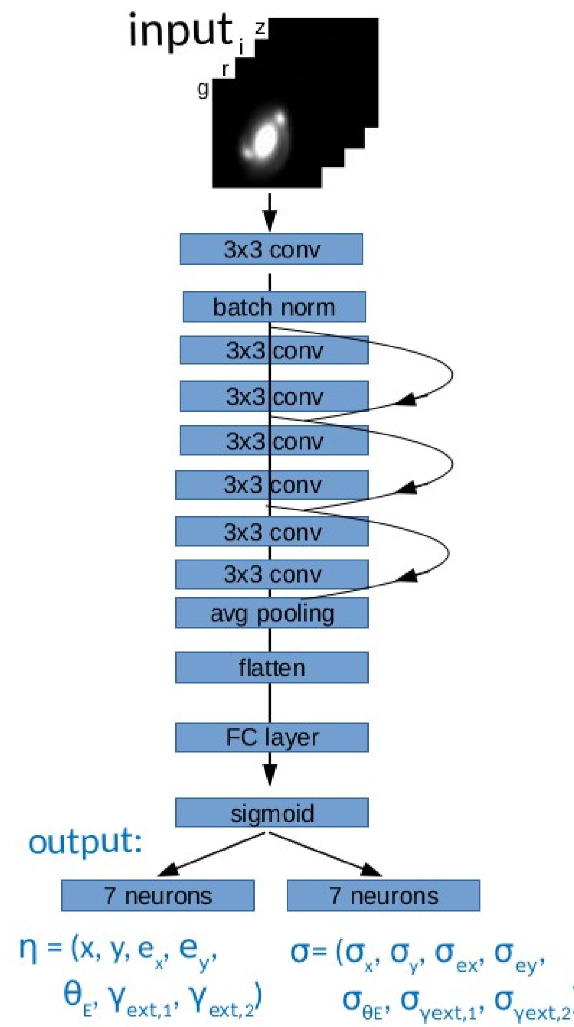
Regression convolutional neural network

(see also, e.g., Hezaveh+2017, Perreault-Levasseur+2017, Madireddy+2019, Park+2020, Pearson+2019,+2021)

- **Realistic lens simulations** ➜ Train and test on HSC Wide
- **log-probability loss with a regularisation term**

$$L = \sum_{k=0}^{N} \sum_{l=0}^{p} \left[-w_l \times P\left(\eta_{k,l}^{\text{pred}}, \eta_{k,l}^{\text{tr}}, \sigma_{k,l}\right) + \epsilon_l \times \log\left(\sigma_{k,l}^2\right) \right]$$

$$P(\eta_{k,l}^{\text{pred}}, \eta_{k,l}^{\text{tr}}, \sigma_{k,l}) = -\frac{\left(\eta_{k,l}^{\text{tr}} - \eta_{k,l}^{\text{pred}}\right)^2}{2\sigma_{k,l}^2} - \ln(\sigma_{k,l}) - \ln(\sqrt{2\pi}).$$



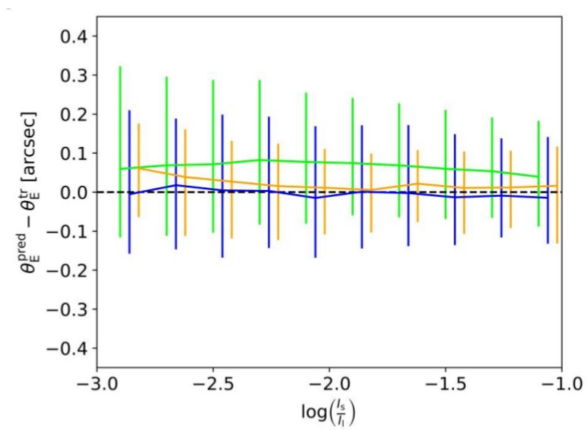
Efficient strong lens modeling with deep learning

Schuldt et al. 2021a, A&A 646, A126; Schuldt et al. 2023a, A&A 671, A147

Estimate mass profile parameters

Regression convolutional neural network

- **Lens mass profile parameters are recovered**
- Results are stable, e.g. for fainter lensed sources
- Translates into accurate predictions of image positions and time delays

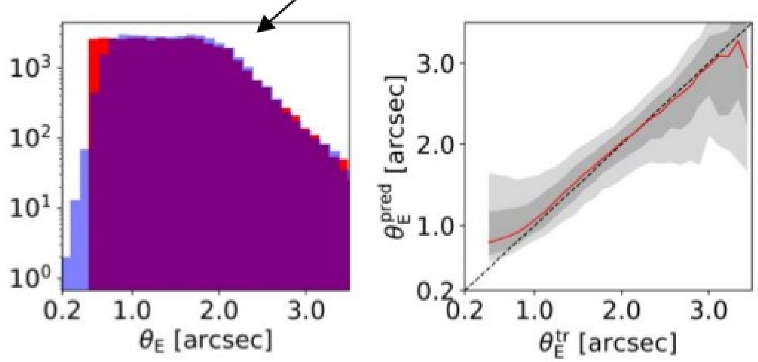


Einstein radius

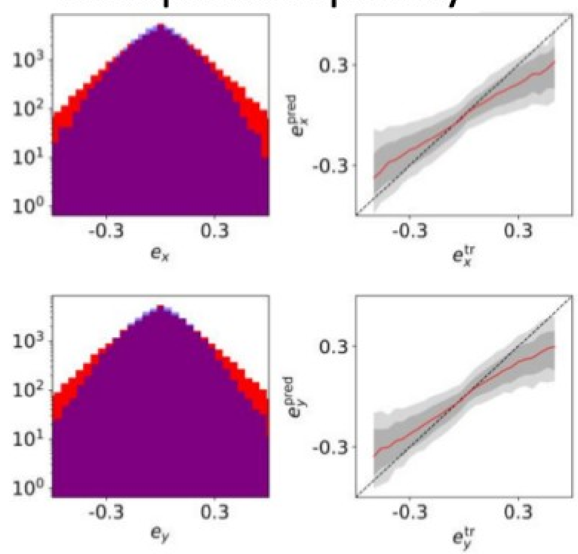
Flat up to ~2"

Input

Output



Complex ellipticity



Efficient strong lens modeling with deep learning

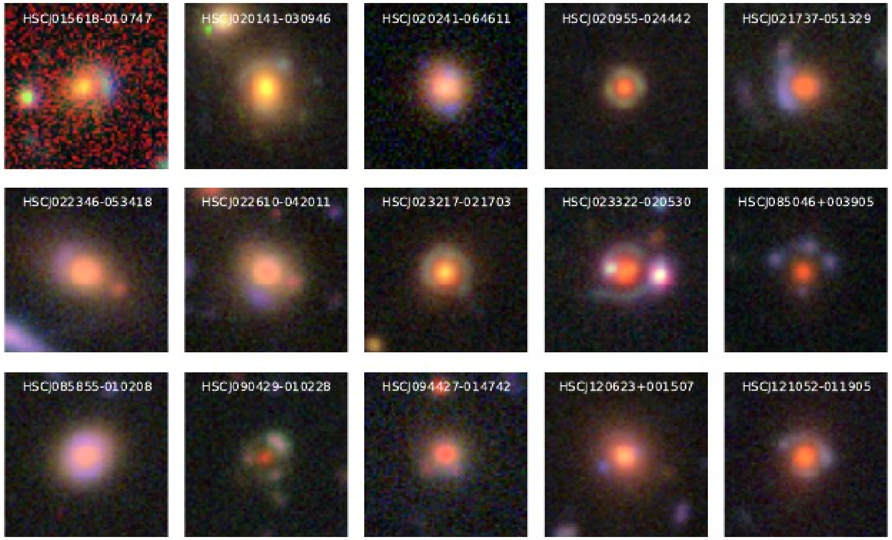
Schuldt et al. 2023b, A&A 673, A33

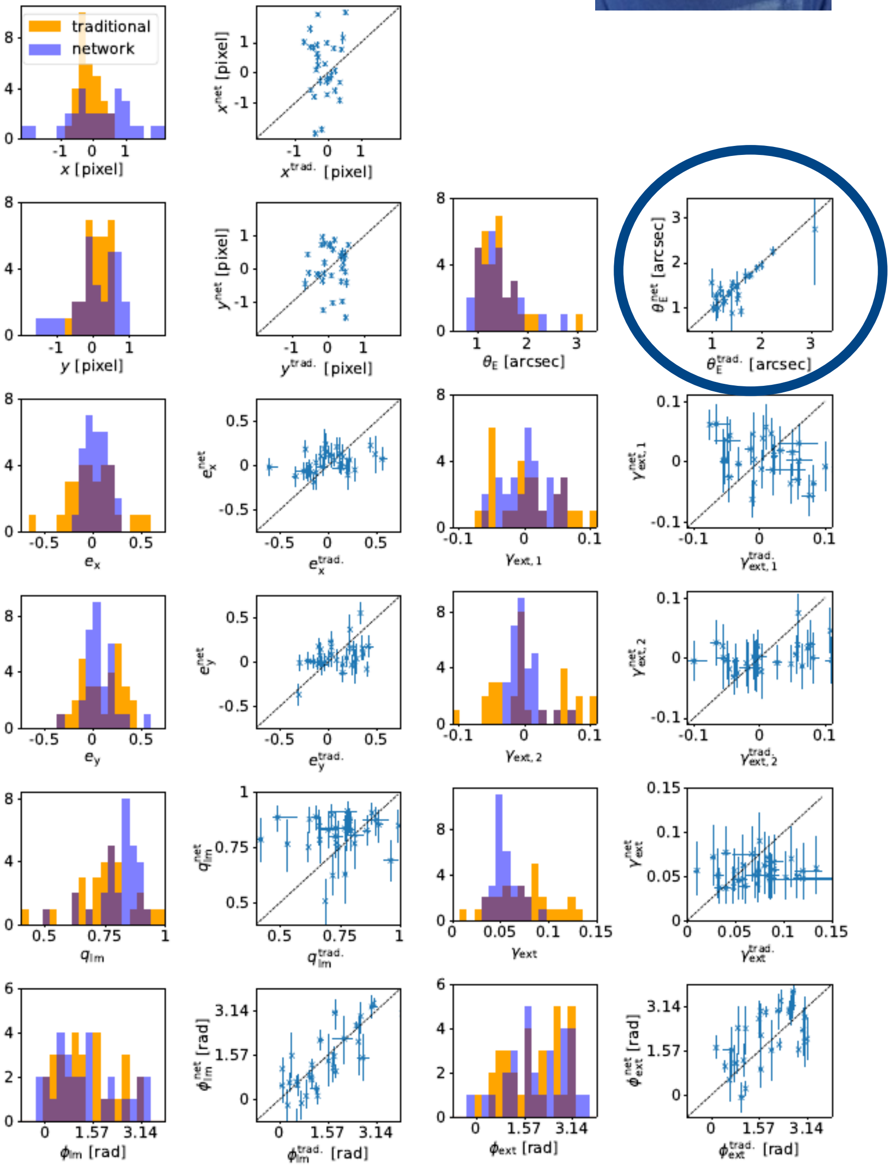
Neural network vs traditional modeling

- Use galaxy-scale strong lenses from HSC Wide (Sonnenfeld+2018, Wong+2018) to compare

  (1) CNN-based modeling (Schuldt+2023a)

  (2) Traditional MCMC sampling-based models with a semi-automated pipeline





Efficient strong lens modeling with deep learning

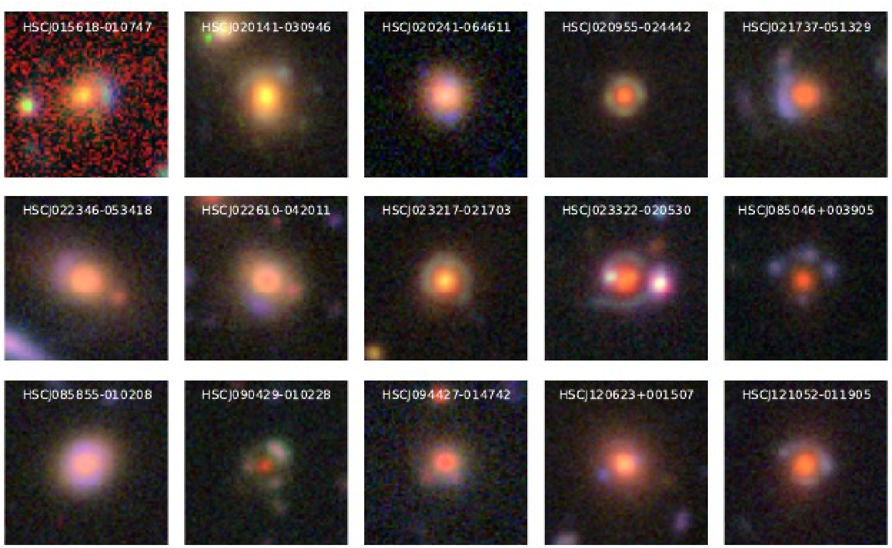
Schuldt et al. 2023b, A&A 673, A33

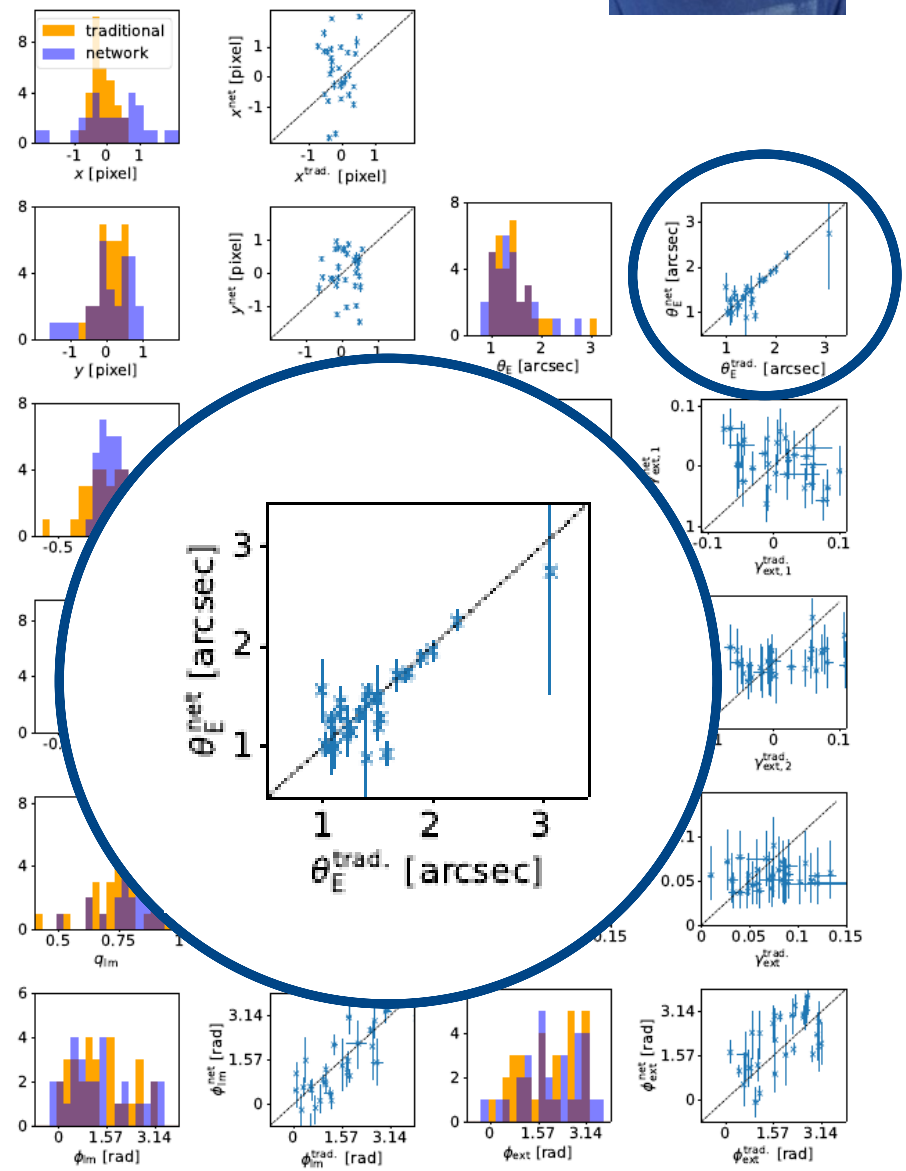
Neural network vs traditional modeling

- Use galaxy-scale strong lenses from HSC Wide (Sonnenfeld+2018, Wong+2018) to compare

  (1) CNN-based modeling (Schuldt+2023a)

  (2) Traditional MCMC sampling-based models with a semi-automated pipeline



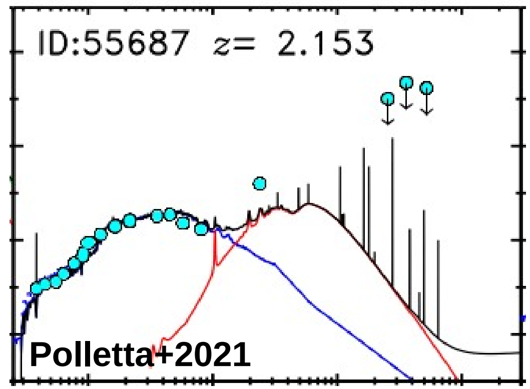


Photometric redshift estimation

Schuldt et al. 2021b, A&A 651, A55

Traditional photo-z codes

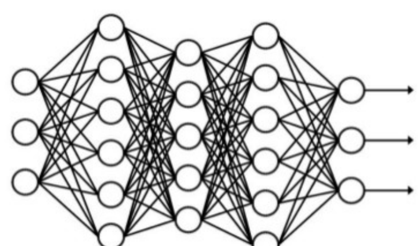
- Deblended photometry and fitting of the spectral energy distributions



Redshift *z*

Directly predict galaxy redshifts from multiband images

- Regression convolutional neural network (d'Isanto+2018, Pasquet+2019, Treyer+2023)
- **More systematic pipeline** ➜ Train and test on HSC Wide *grizy* to prepare for LSST



Redshift z

Photometric redshift estimation

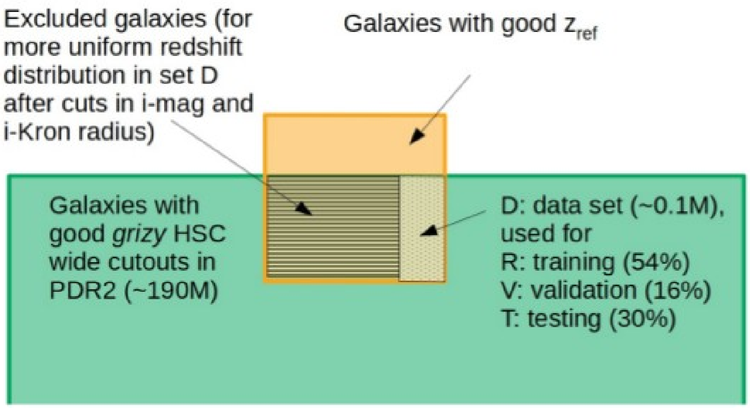
Schuldt et al. 2021b, A&A 651, A55

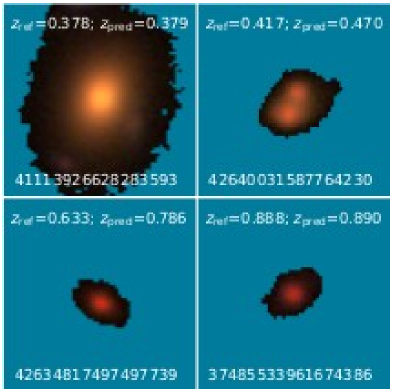
Predict galaxy redshifts from images

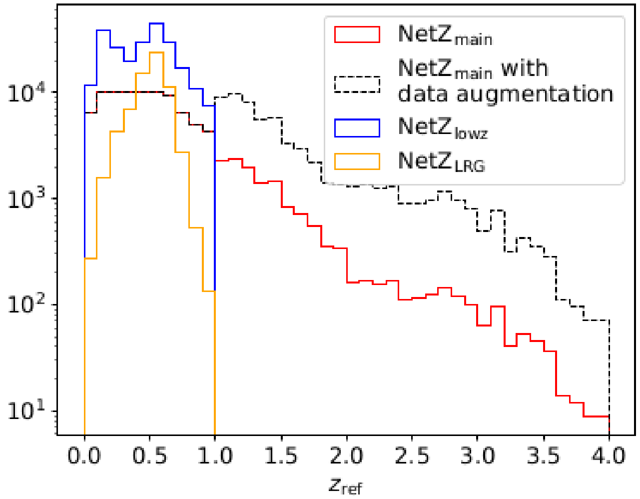
– Regression convolutional neural network (d'Isanto+2018, Pasquet+2019, Treyer+2023)

- 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" + masking + **augmented data set**
  ➜ 10^5 examples for training a simple CNN







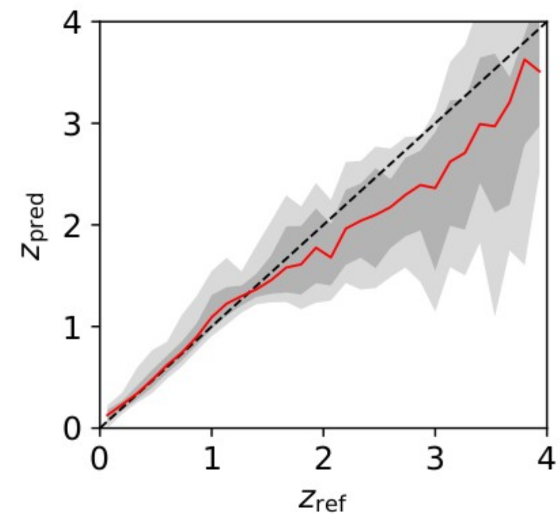
Photometric redshift estimation

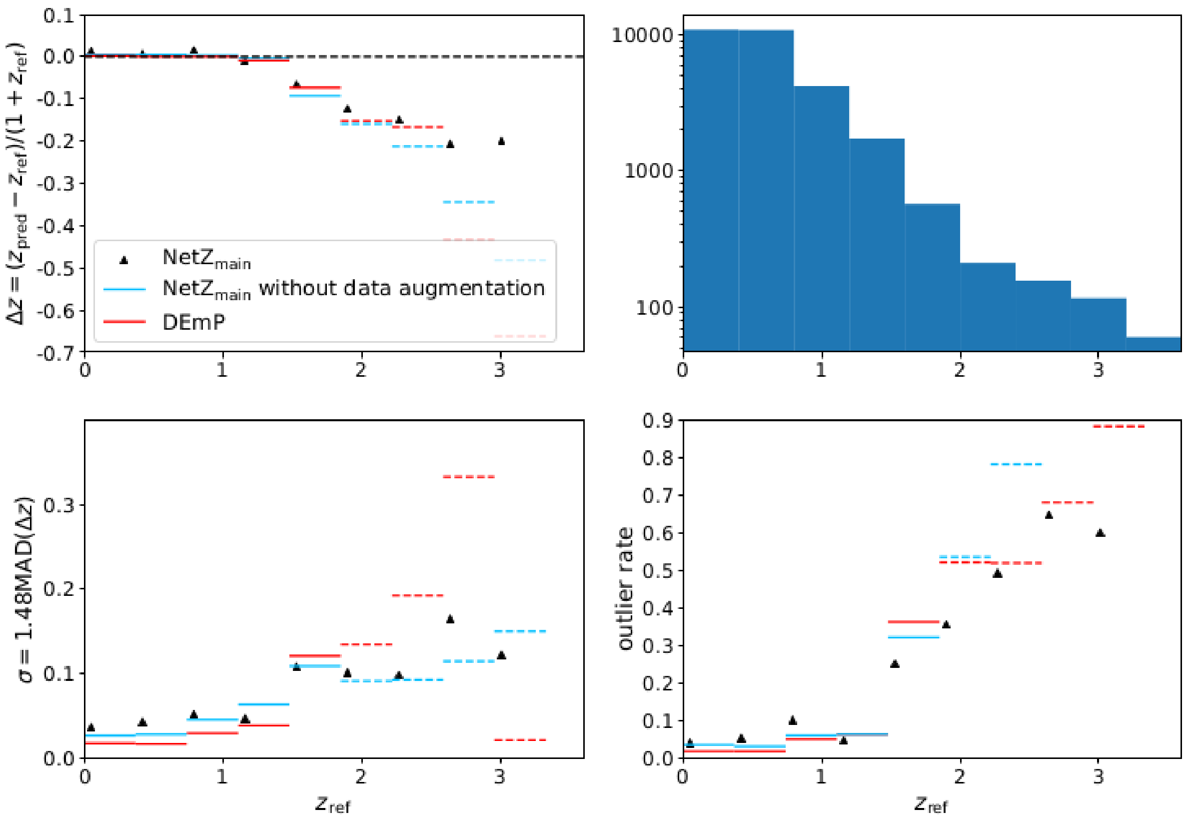
Schuldt et al. 2021b, A&A 651, A55

Predict galaxy redshifts from images

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

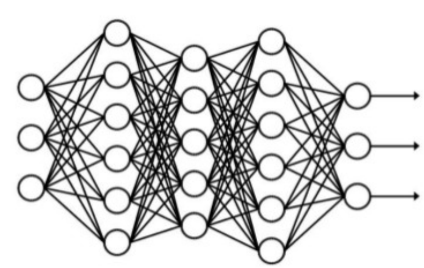
Dispersion $\sigma = 1.48 \times \mathrm{MAD}(\Delta z_i) = $

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

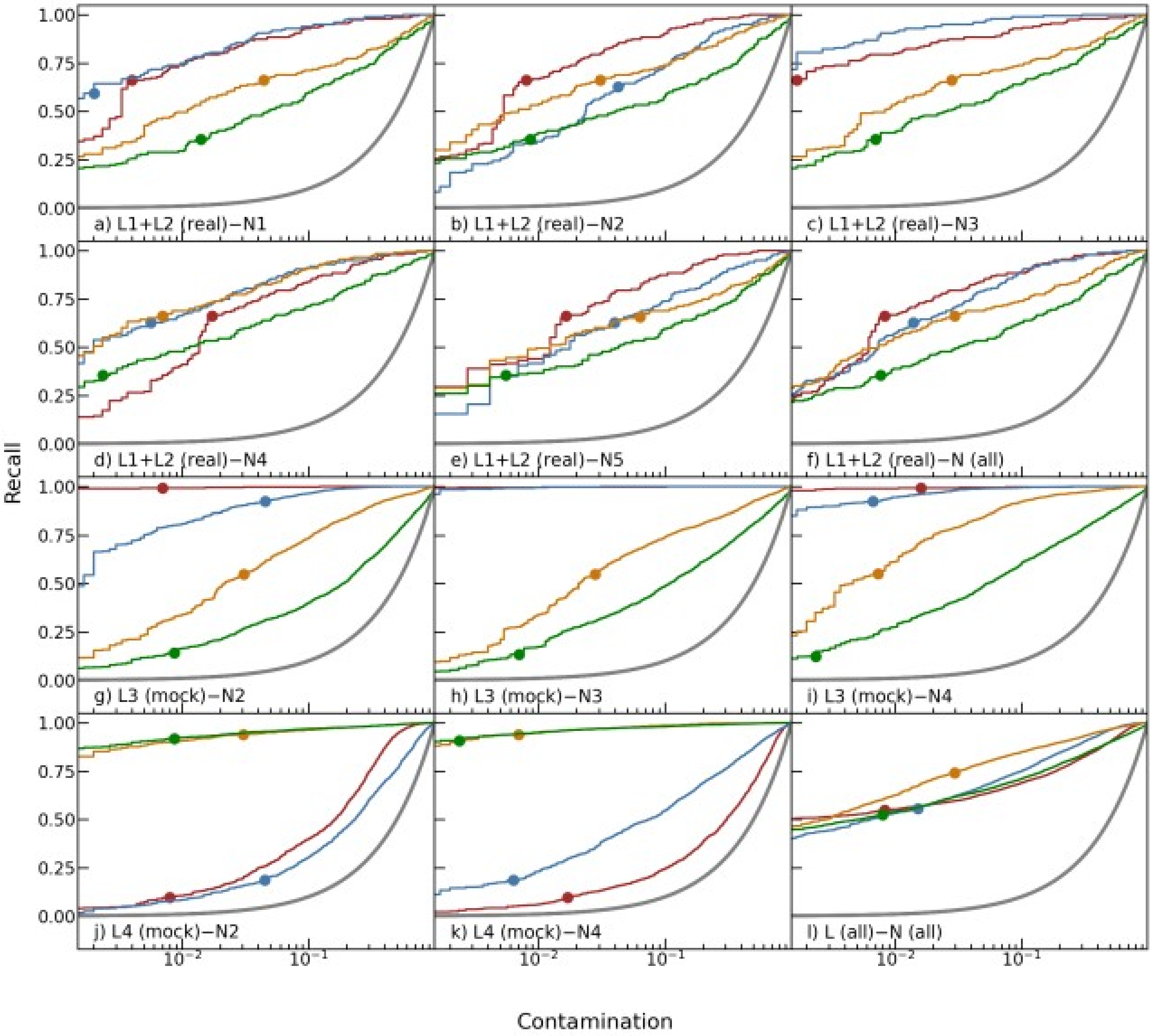
Outlier rate $f_{\mathrm{outlier}} = \frac{N(|\Delta z_i| > 0.15)}{N_{\mathrm{bin}}}$

- **Supervised machine learning is key to identify rare objects, e.g. strong lenses**
- All-sky classification pipeline works, but many contaminants + long visual inspection
- Full automation need systematic network evaluation on external, realistic test sets

  ➜ Major improvements for specific networks and data sets ➜ **FPRs from 1% to 0.01%!**
- Some networks learn spurious correlation in the training data ➜ Need interpretability
- **Ensembles of neural networks leverage diversity of individual models**
- Unsupervised machine learning not ready for rare object identification
- ResNet for automated lens modeling + parameter uncertainties ➜ Performance are promising + validated on real strong lens systems
- CNN for automated redshift estimates ➜ Competitive approach with broad applications

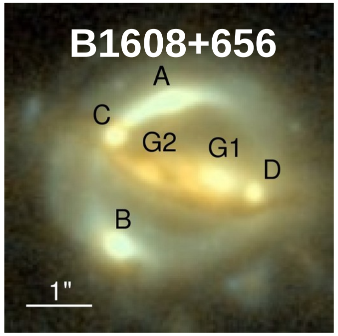


Bonus slides

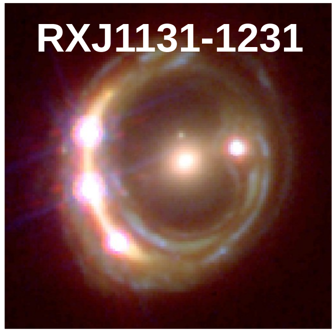


Recall
Contamination
1.00
0.75
0.50
0.25
0.00
10^{-2}
10^{-1}
a) L1+L2 (real)−N1
b) L1+L2 (real)−N2
c) L1+L2 (real)−N3
d) L1+L2 (real)−N4
e) L1+L2 (real)−N5
f) L1+L2 (real)−N (all)
g) L3 (mock)−N2
h) L3 (mock)−N3
i) L3 (mock)−N4
j) L4 (mock)−N2
k) L4 (mock)−N4
l) L (all)−N (all)

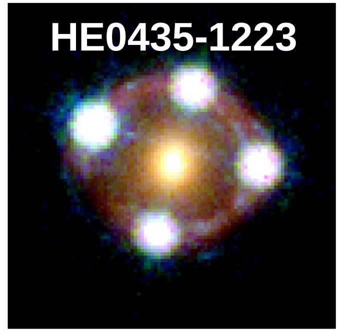
Cosmology with 6 lensed quasars: H0LiCOW project



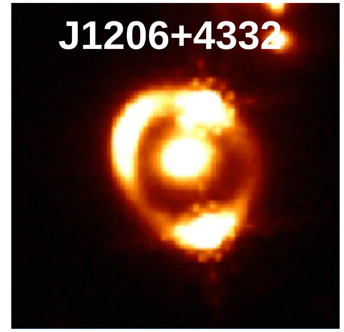
Suyu et al. 2010



Suyu et al. 2013, 2014; Tewes et al. 2013

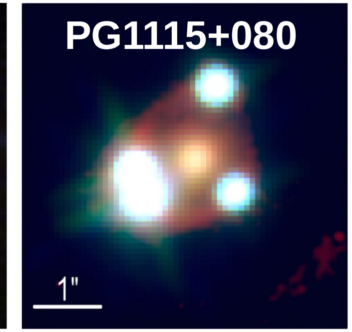


Wong et al. 2017; Rusu et al. 2017; Sluse et al. 2017; Bonvin et al. 2017



Birrer et al. 2019

Bonvin et al. 2019; Sluse et al. 2019; Rusu et al. 2019



Chen et al. 2019

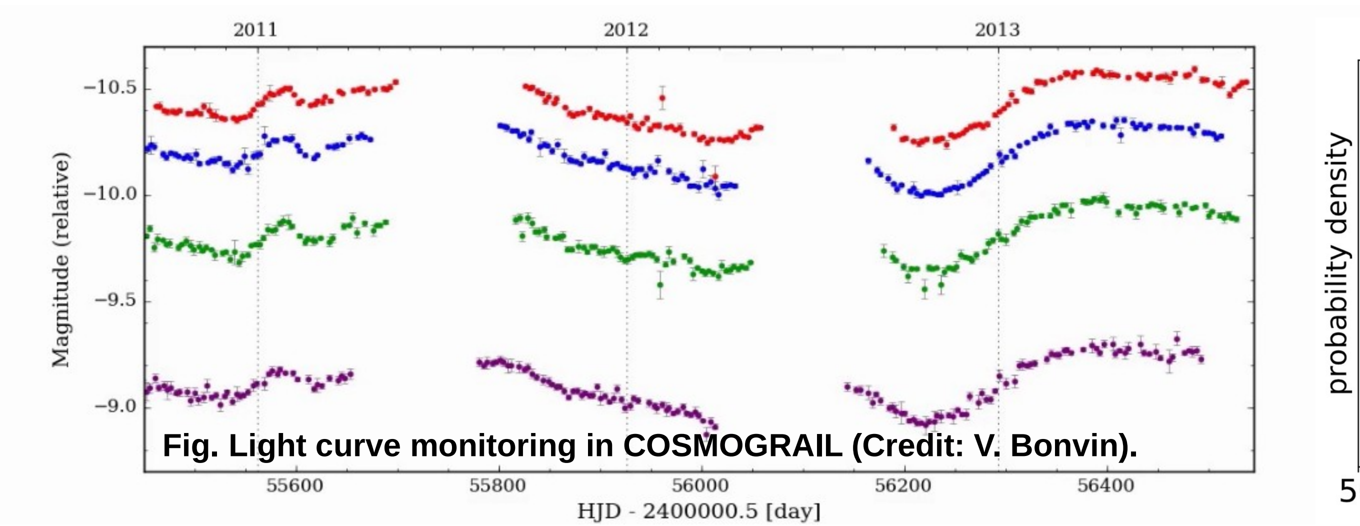
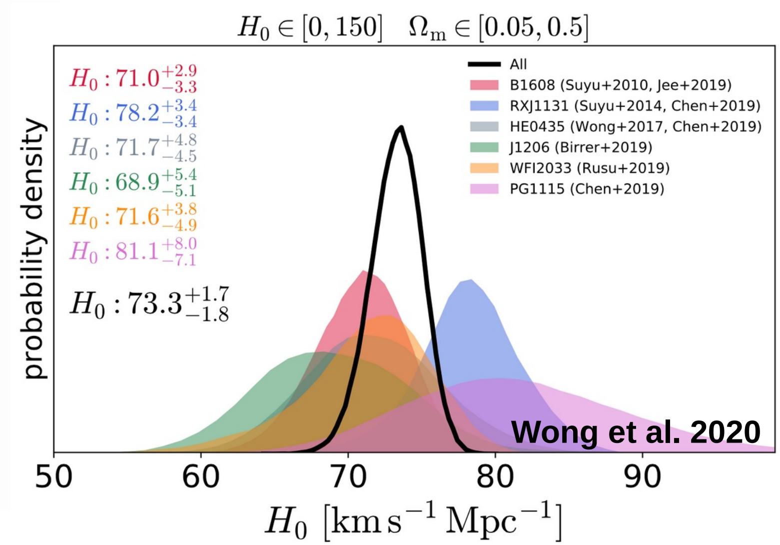


Fig. Light curve monitoring in COSMOGRAIL (Credit: V. Bonvin).



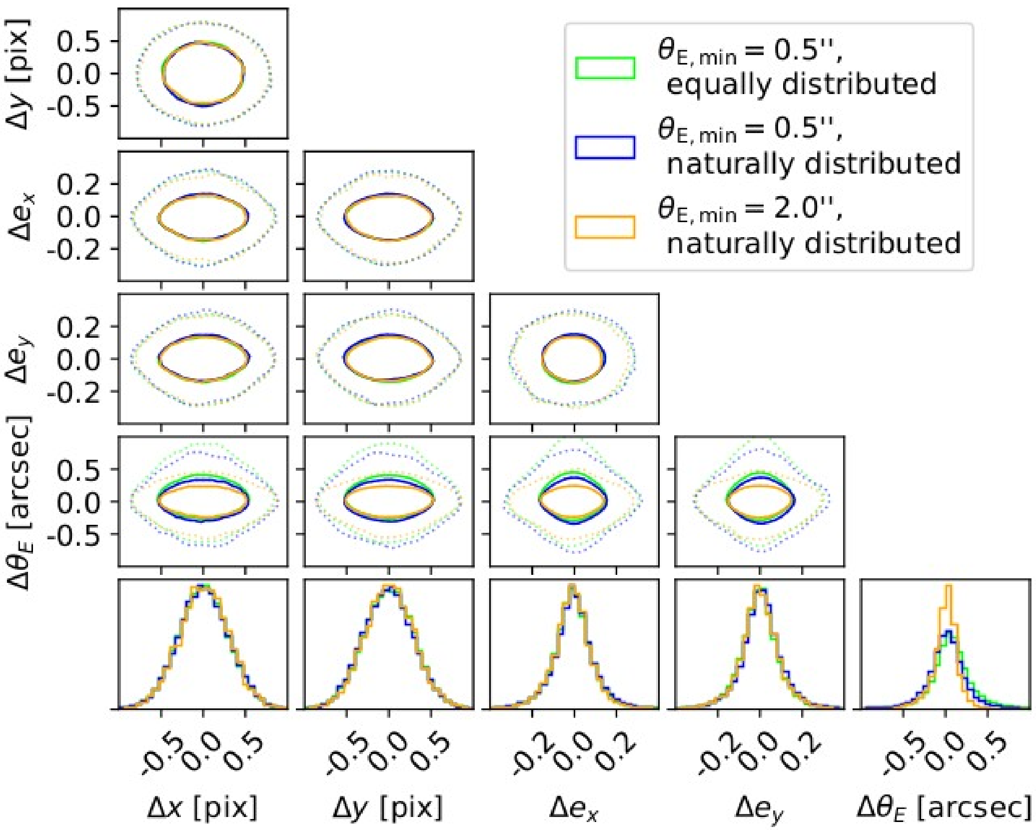
Time delays from COSMOGRAIL + Lens modeling + Line-of-sight mass modeling

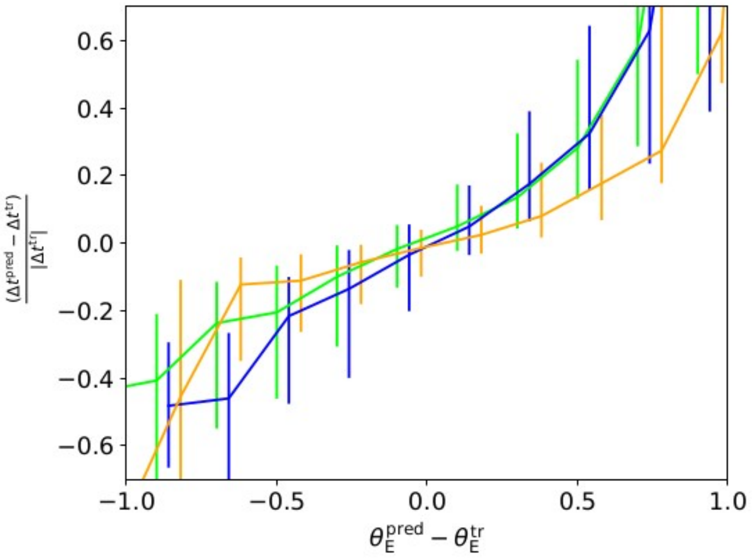
➜ H0 with 2.4% precision in flat ΛCDM (blind analysis)

Efficient strong lens modeling

Schuldt et al. 2021a, A&A 646, A126

Predict lens mass profile parameters





Photometric redshift estimation

Schuldt et al. 2021b, A&A 651, A55

Predict photometric redshifts

- Morphological information helps
  - New CNN trained on the same set ➜ replacing cutouts with point-like sources
  - Larger scatter + larger bias at higher z