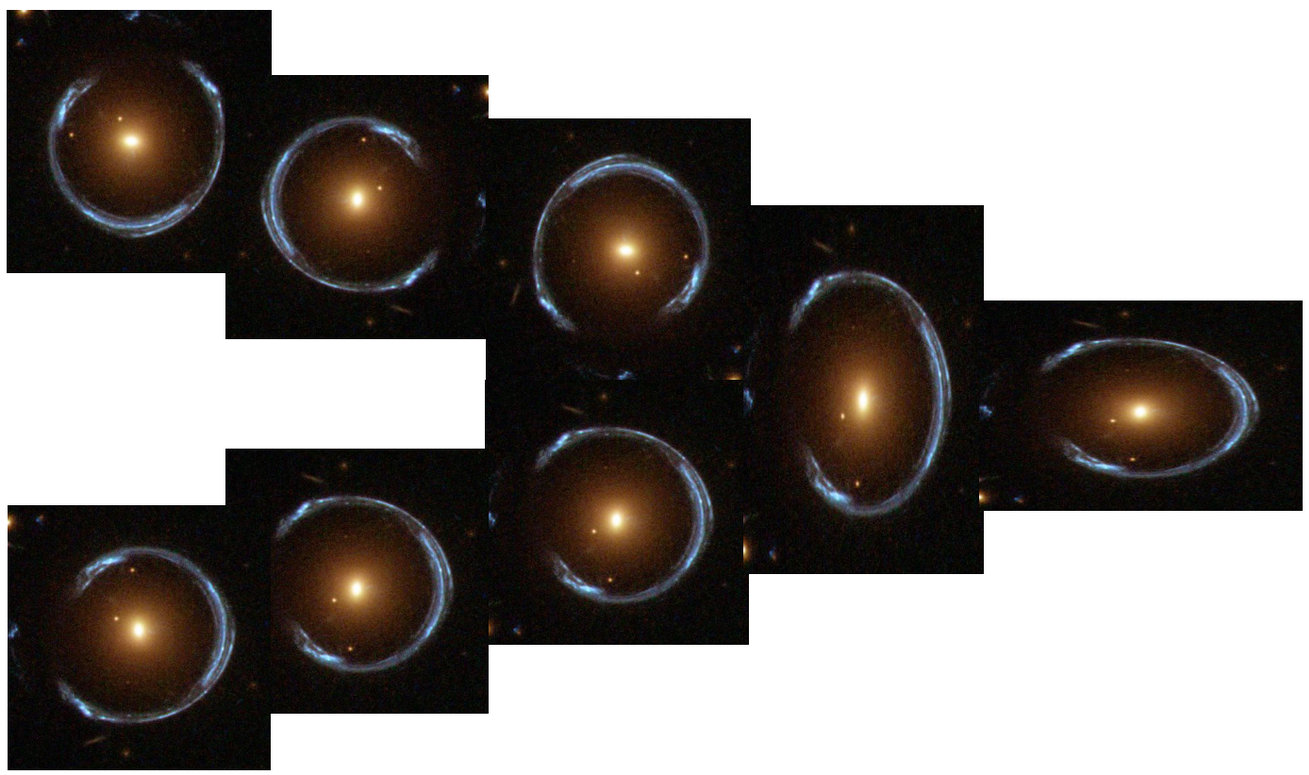
Galaxy-scale Strong Lensing in DES, UNIONS, Euclid



F. Courbin (EPFL, Switzerland)

K. Rojas, E. Savary, B. Clément, *R. Gavazzi, M. Maus, J. Chan, C. Lemon,*
G. Vernardos, S. Suyu, R. Cañameras, S. Schuldt, R. Joseph, A. Galan,
J.-C. Cuillandre, S.Fabbro, S. Gwyn, M. Hudson, M. Kilbinger, D. Scott, C. Stone

FNSNF
FONDS NATIONAL SUISSE
DE LA RECHERCHE SCIENTIFIQUE

European Research Council
Established by the European Commission

EPFL Laboratoire d'astrophysique http://lastro.epfl.ch
Ecole Polytechnique Fédérale de Lausanne (EPFL)

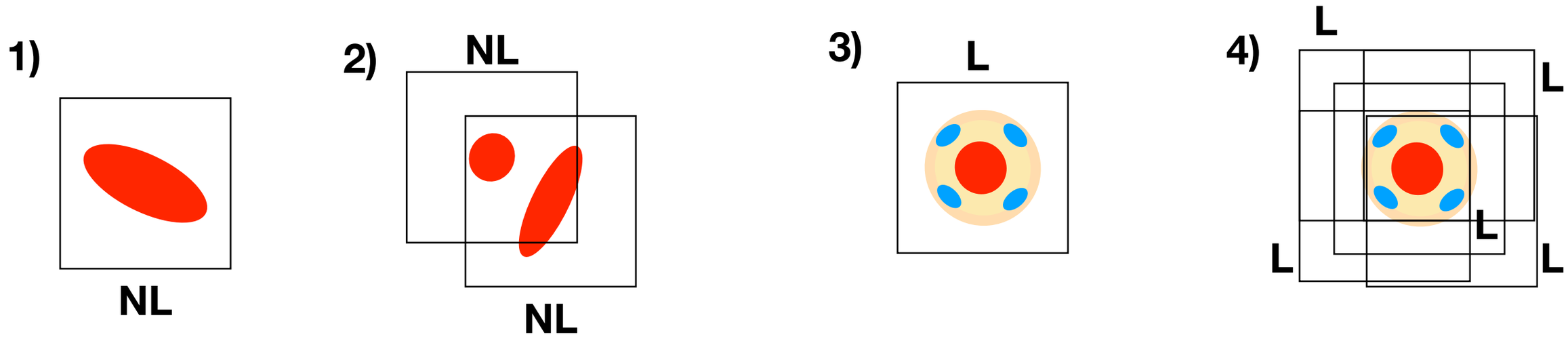
Paris/IAP - Oct 2021

Motivations, challenges and questions

- Use **real data as a test bench** for strong-lens finding in Euclid
- Explore **diversity** of different lenses
- What preselection?
- **Color or resolution?** Is one better than the other ?

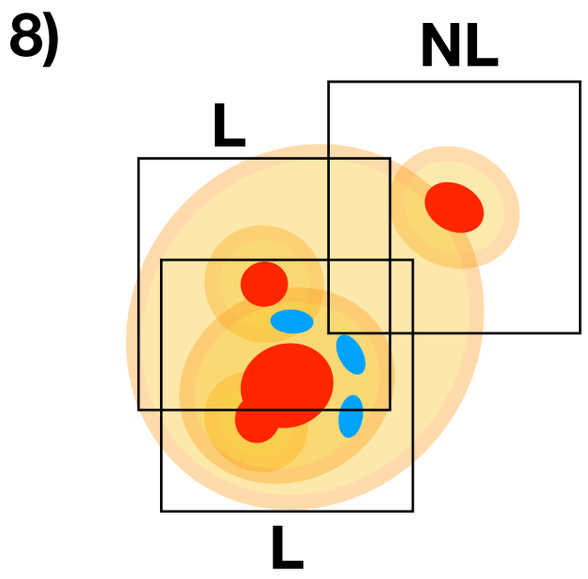
- Build catalogues of **contaminants** on real data
- Find lenses in large numbers to **enable follow-up** already
- Test lens **modeling tools** on real lenses
- Test lens/source **deblending** methods
- Eventually **use real Euclid images** as a training set or to do « self-supervised learning »

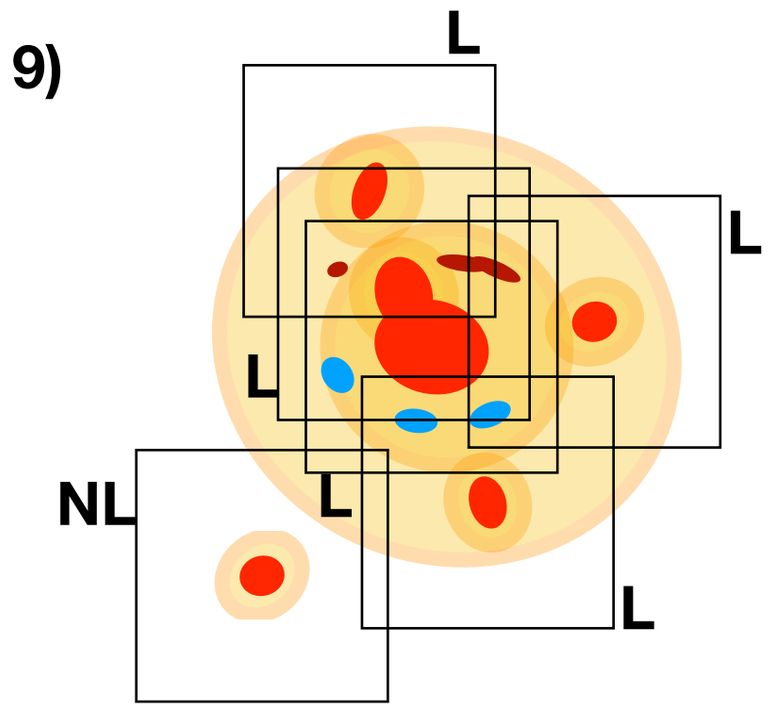
Challenge 1: Variety in lens properties

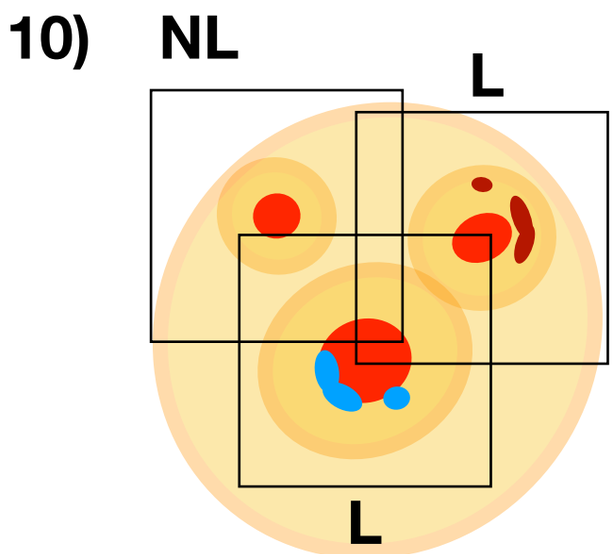


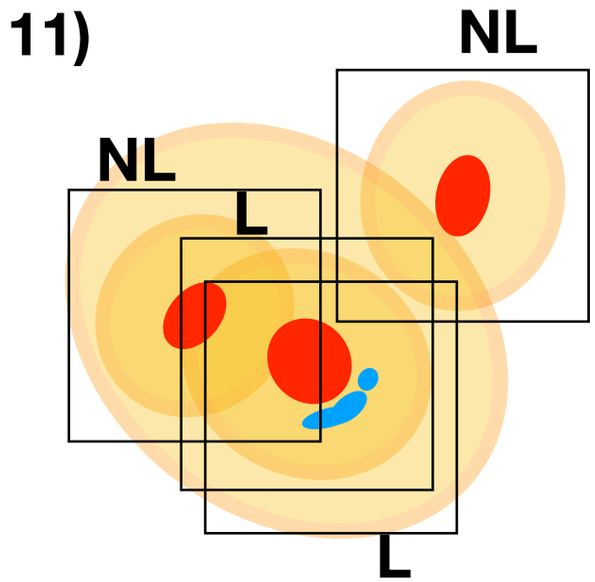
Courtesy: B. Clément

Challenge 1: Variety in lens properties









Galaxy Light

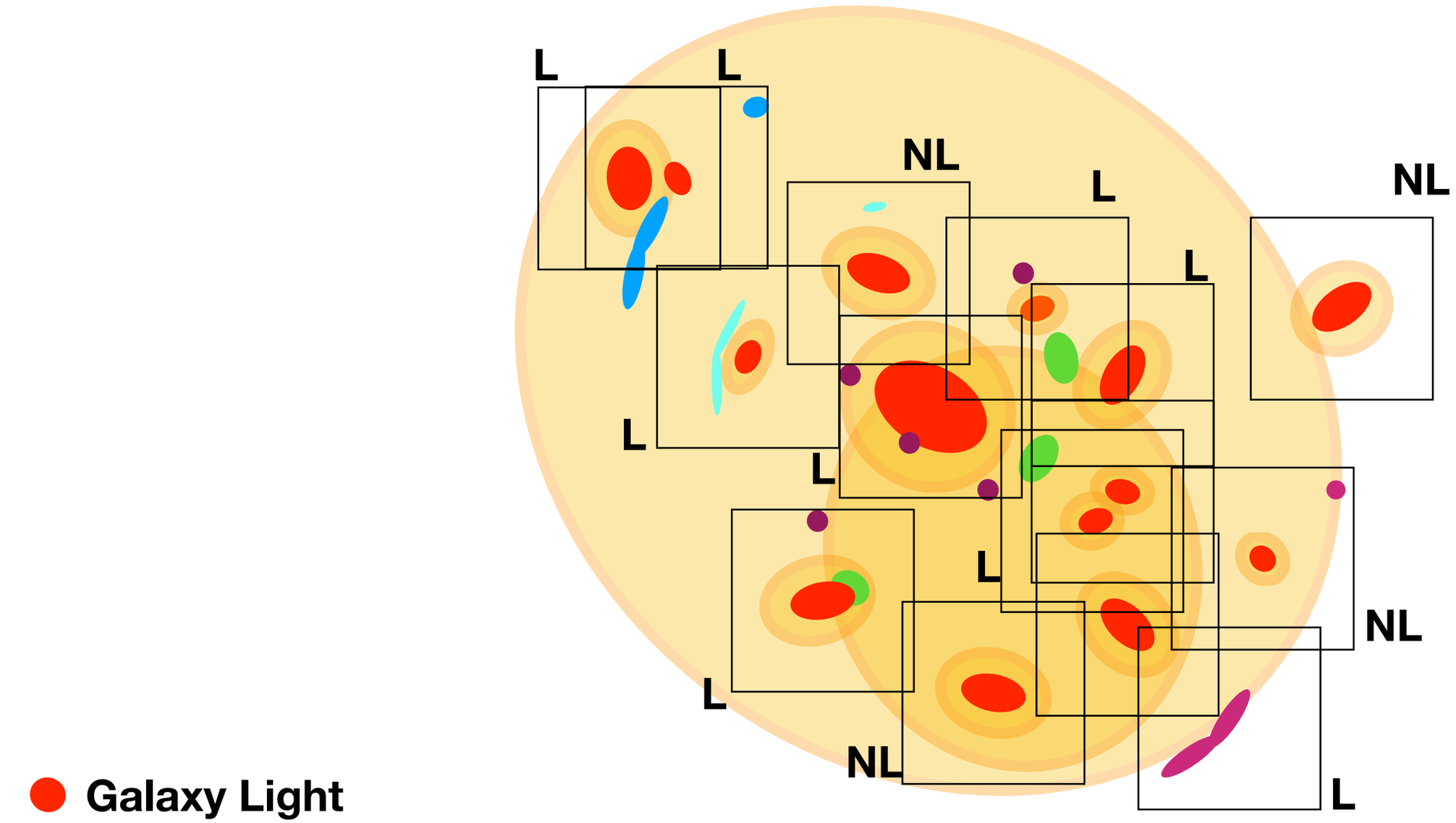
Mass

Arc

10”x10” cutout

Courtesy: B. Clément

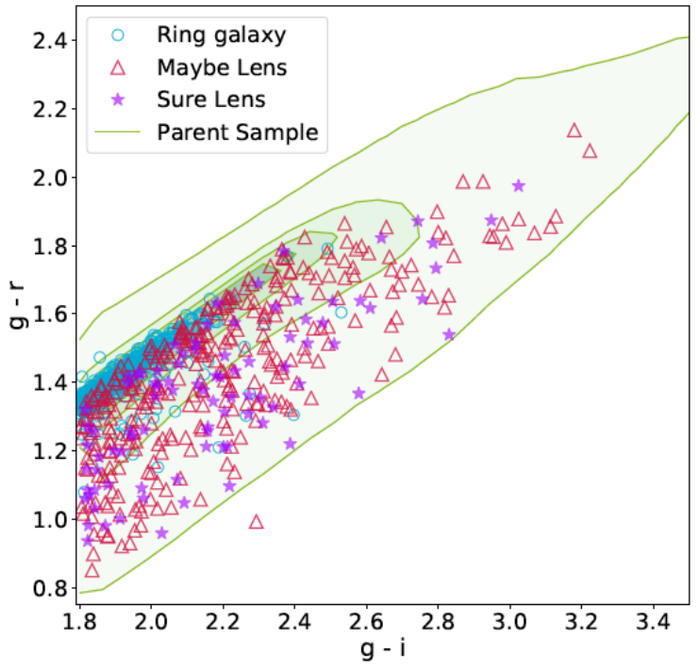
Challenge 1: Variety in lens properties



Courtesy: B. Clément

Challenge 2: Preselection for training and lens search

DES



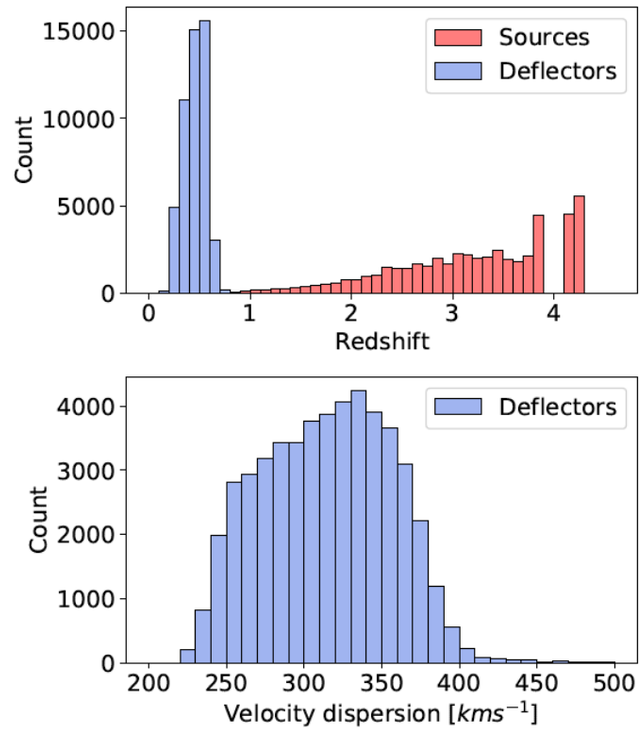
$$
\begin{aligned}
&1.8 < g - i < 5,\\
&0.6 < g - r < 3,\\
&18 < r < 22.5,\\
&g > 20,\\
&i > 18.2,
\end{aligned}
$$

Parent sample: 18 745 029 LRGs
5186 sq degree of DES DR1
13” x 13” stamps (50 pix)
g, r, i bands used
No redshift and no velocity dispersion

Proxy to allocate a velocity dispersion to each galaxy

Rojas, Savary, Clément, et al. (2021, arXiv2109.00014)

CFIS



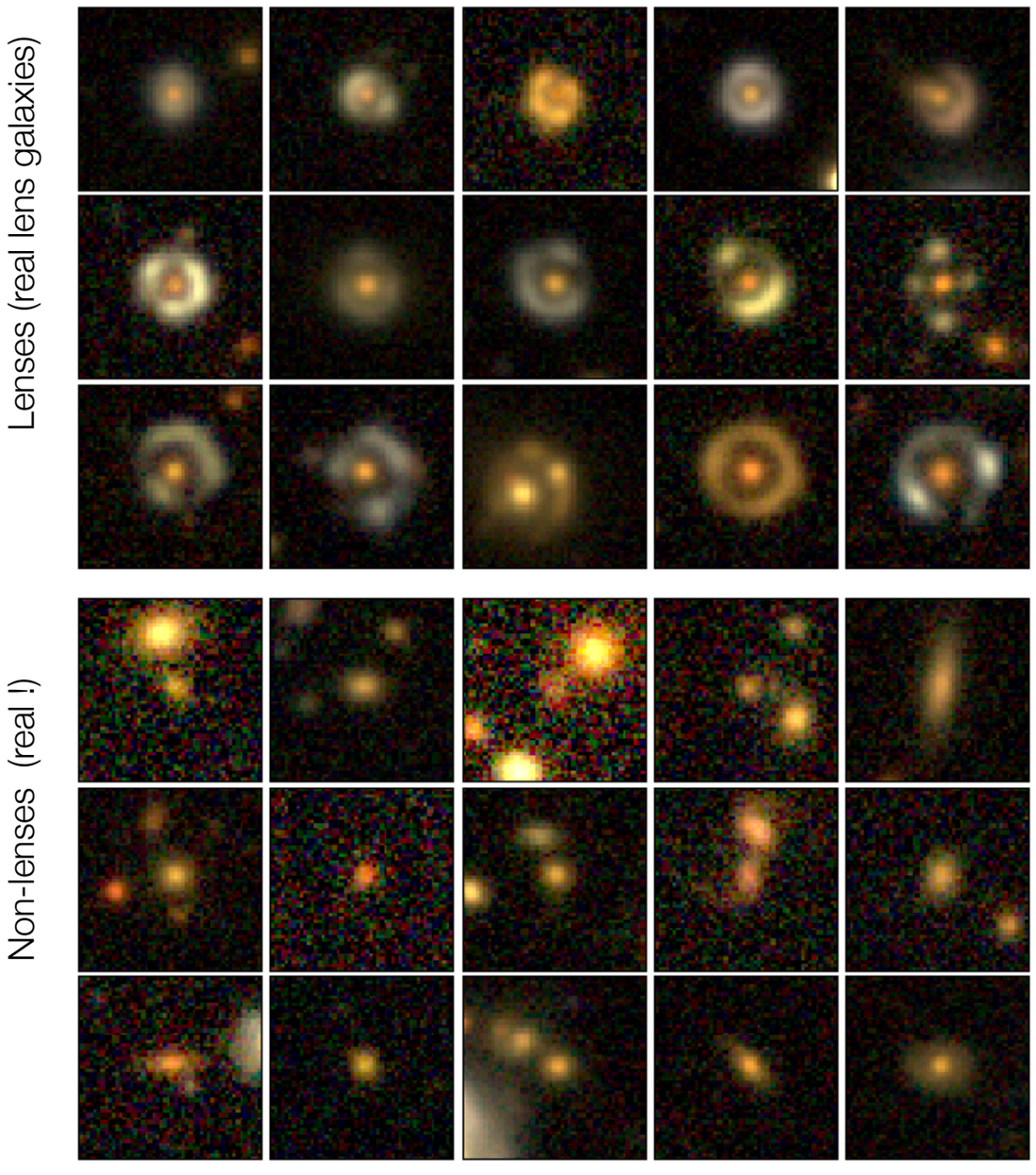
Parent sample: 2 344 002 LRGs
2500 sq degree of CFIS imaging
8” x 8” stamps (44 pix)
Only r-band but color from PANSTARRS
Redshift and vel. disp. for the training set

Color selection as in Cañameras et al. 2020, A&A 644, A45

Savary, Rojas, Clément, et al. (2021, arXiv0000.00000)

Challenge 2: Preselection for training and lens search

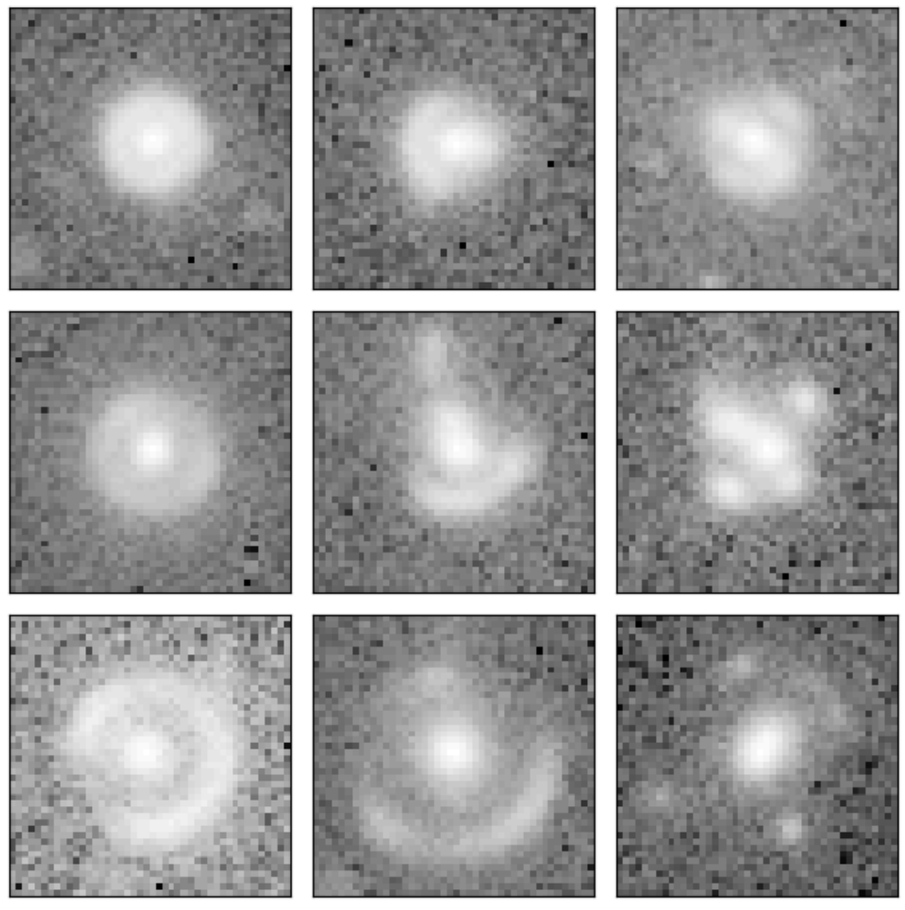
DES simulation



100 000 simulated lenses with 1.2” < R_E < 1.8”

Rojas, Savary, Clément, et al. (2021, arXiv2109.00014)

CFIS simulation



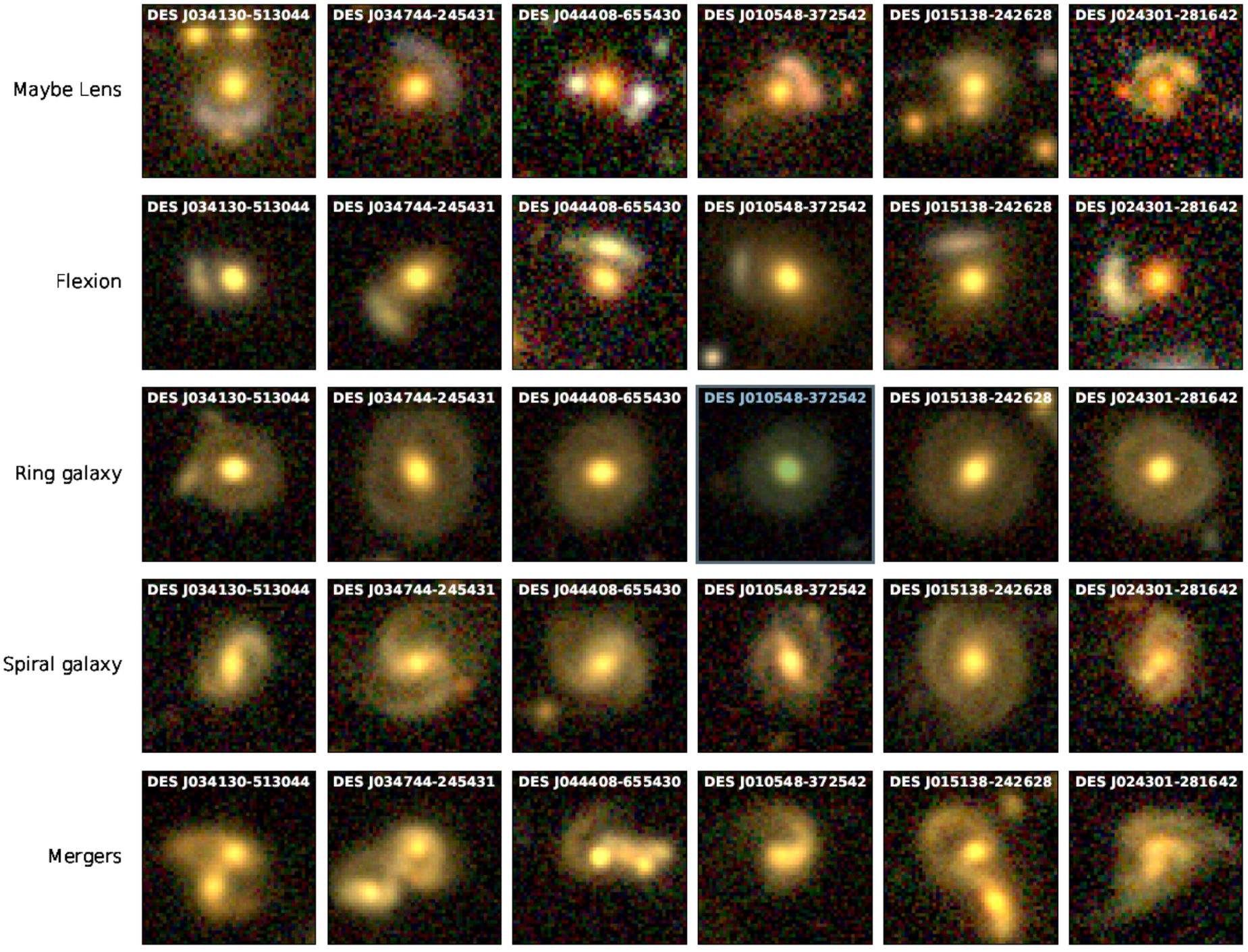
10 600 simulated lenses with 0.8” < R_E < 3.0”

We do NOT attempt to follow realistic distributions in R_E but sample uniformly (both in DES and CFIS)

Lenses sources come from HST but with HSC colors

Savary, Rojas, Clément, et al. (2021, arXiv0000.00000)

Challenge 3: Visual inspection (takes a few days)

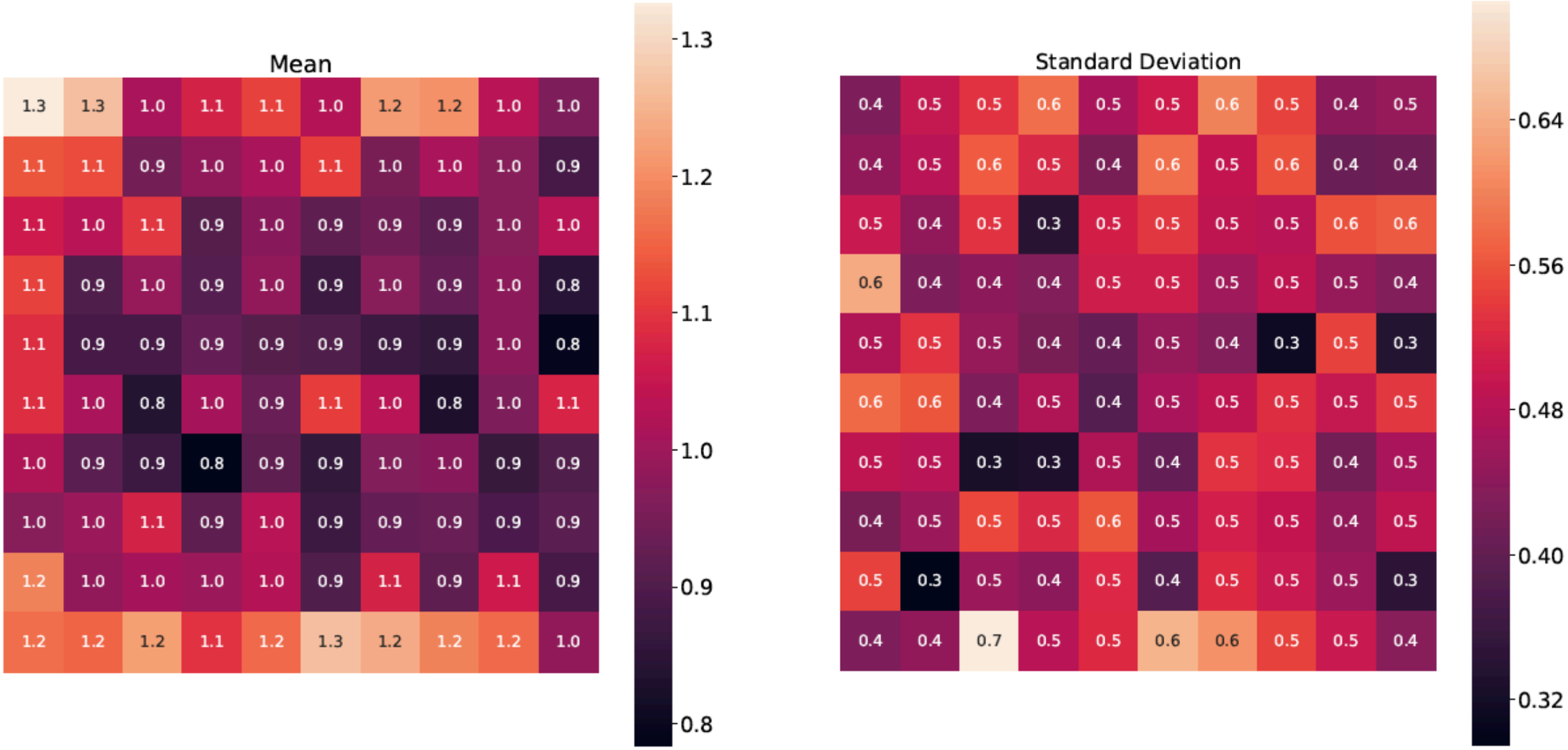


Rojas, Savary, Clément, et al. (2021, arXiv2109.00014)

Challenge 3: Visual inspection (takes a full week)

76 000 objects with CNN score > 0.9 in DES
9 460 objects with CNN score > 0.5 in CFIS
7 « inspectors » among the authors

6 different users inspect mosaics of 100 x 100 objects (Step 1): Lenses & ring galaxies
Detailed classification (Step 2): sure lens, maybe lens, flexion, ring, spiral, merger



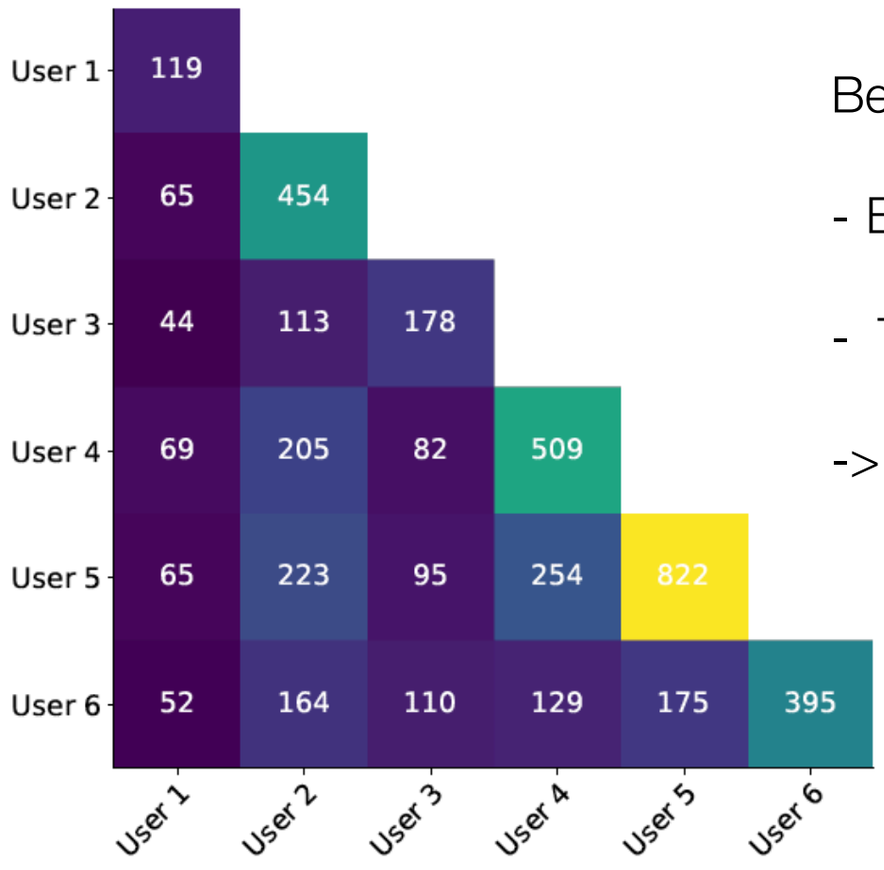
Rojas, Savary, Clément, et al. (2021, arXiv2109.00014)

Challenge 3: Visual inspection (takes a few days)

Comparison of « inspectors » in the CFIS data

Classification	User 1	User 2	User 3	User 4	User 5	User 6
NL	4492	4151	4389	3941	3398	4012
FL	17	23	61	178	408	221
ML	96	423	144	474	775	357
SL	23	31	34	35	47	38

Non-lens
Flexion
Maybe lens
Sure lens



Before the mosaic classification and detailed classification:

- Establish a joint classification scheme

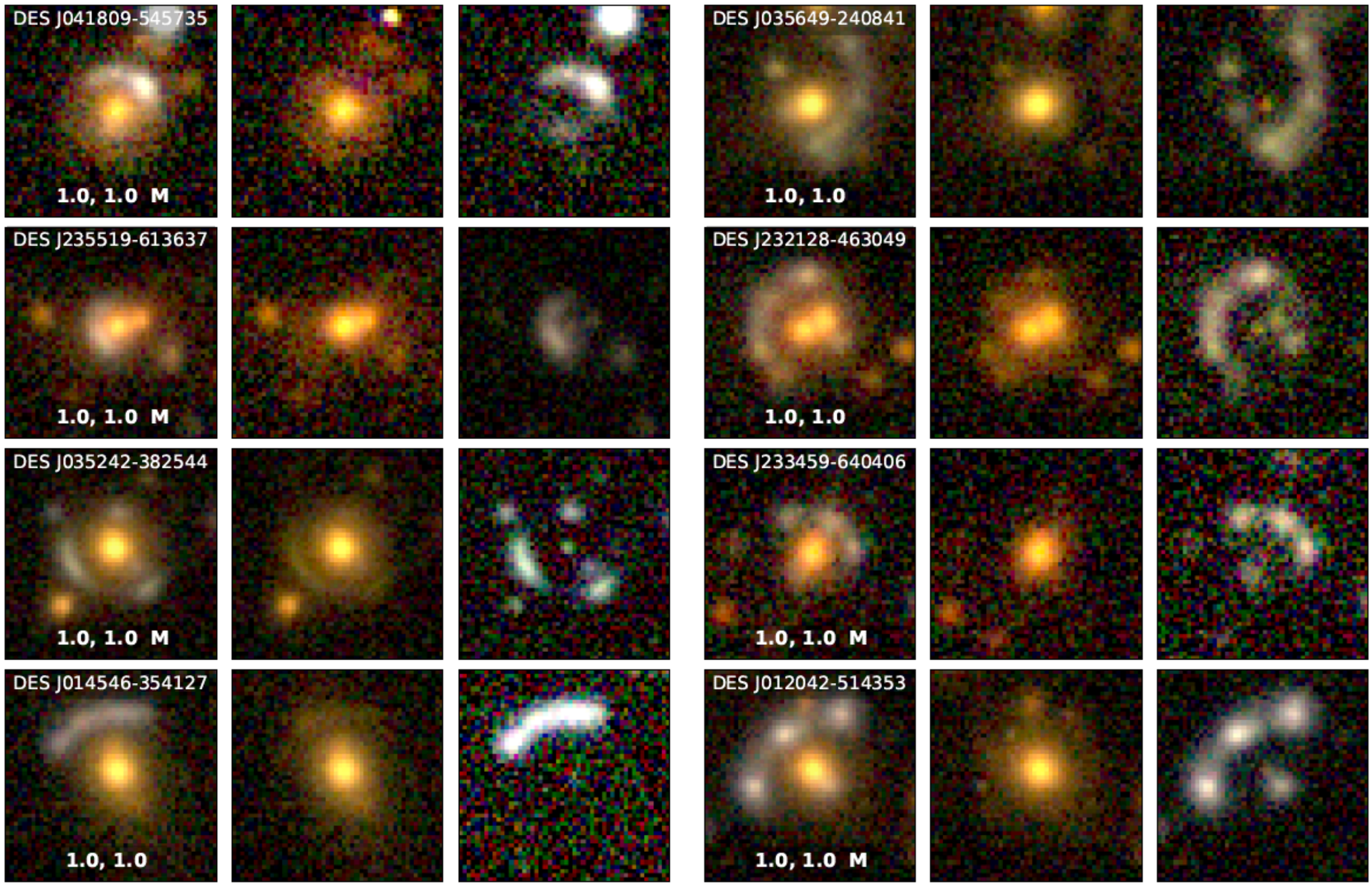
- Train on a few dozens of objects

-> still **little overlap** between the classification of the 6 inspectors

Savary, Rojas, Clément, et al. (2021, arXiv0000.0000)

Challenge 4: deblending the lens and source light

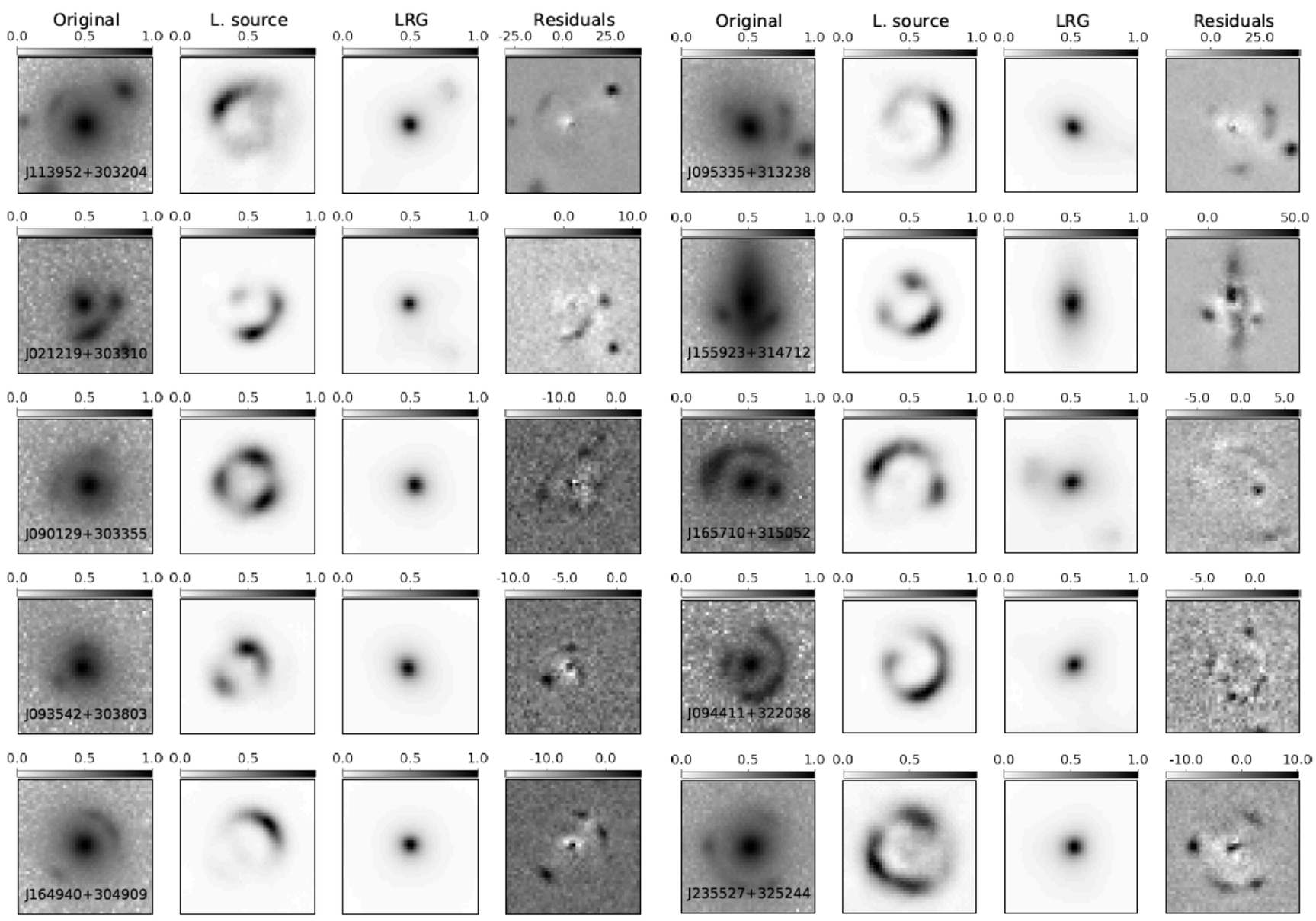
If **color information available** (DES) then use Multiband morpho-Spectral Component Analysis Deblending Tool (MuSCADeT; Joseph, Courbin, Starck, 2016 A&A 589, A81)



Rojas, Savary, Clément, et al. (2021, arXiv2109.00014)

Challenge 4: deblending the lens and source light

If **no color information available** (CFIS) then use auto-encoders (here shown for some of the CFIS « Sure Lenses »)
Use mapping between inputs and labels
Training on 10000 simulated lenses and 5000 real LRGs with no lensing



Savary, Rojas, Clément, et al. (2021, arXiv0000.0000)

Quick word about Euclid lens finding

- Reduction pipeline produces **tiles** of 32’ x 32’
- 15000 deg^2 ~ 50000 tiles
- To Euclid depth in the wide survey 1000 objects have VIS < 22
- Collett (2015) expects 170000 lenses in the full survey
- This is 3 lenses per tile
- Euclid does 20 tiles per day
- 60 lenses expected per day, among 20000 objects to test
- 0.5% false positive rate (what we **measure** in CFIS and DES) means 100 objects to check per day -> feasible?

Summary - expectations for Euclid

- 405 Sure or maybe lenses in DES DR1, starting from 18 million objects
- 133 Sure or maybe lenses in CFIS starting from 2.3 million objects
- Overall numbers compatibles with other searches
- Quality of training set prevails over « quality of network »
- Color information less important than high resolution

- For Euclid, use single-band: VIS is the best but complemented by NIR
- No color selection, to enable lensing by anything and make pipeline simpler
- Far too many false positives so far (0.5%)
- Towards a fully data-driven training set? After 1 year of mission?
- Citizen science to clean from false positives?
- Modeling step with standard models: 2h per lens