Q1. what image they used in Dieleman 2015

A1. use GZ2 images (424 by 424 JPEF, RGB)

424x424 color -> 69x69 color ->69x69x3

Q2. what b and s in Fig 13. represent?

A2. b is median offset (average) and s is average of scatters(dispersions) for each class

Q3. what value CNN returns in 5.1?

A3. -3~6 (context infers)

Q4. what they did in 5.2?

A4. -3~0 positive(1) -5 negative(0) p_{S0} = 0~1

Table 1
T-Type Classification Schemes

Class	c0	E0	E+	S0-	S0	S0+	S0/a	Sa	Sab	Sb	Sbc	Sc	Scd	Sd	Sdm	Sm	Im	?
RC3	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10	:
Fukugita	0	0	0	1	1	1	1	2	2	3	3	4	4	5	5	6	6	−1
P.N. (this work)	−5	−5	−5	−3	−2	−2	0	1	2	3	4	5	6	7	8	9	10	99

B4 Seminar #2
26.7.2018
Nao Sakai

Finding Strong Gravitational Lenses in the Kilo Degree Survey with Convolutional Neural Networks

C.E.Petrillo et al. 2017

Abstract

The volume of data that will be produced by new-generation surveys requires automatic classification methods to select and analyze sources. Indeed, this is the case for the search for strong gravitational lenses, where the population of the detectable lensed sources is only a very small fraction of the full source population. We apply for the first time a morphological classification method based on a Convolutional Neural Network (CNN) for recognizing strong gravitational lenses in 255 square degrees of the Kilo Degree Survey (KiDS), one of the current-generation optical wide surveys. The CNN is currently optimized to recognize lenses with Einstein radii $\gtrsim$1.4 arcsec, about twice the r-band seeing in KiDS. In a sample of 21789 colour-magnitude selected Luminous Red Galaxies (LRG), of which three are known lenses, the CNN retrieves 761 strong-lens candidates and correctly classifies two out of three of the known lenses. The misclassified lens has an Einstein radius below the range on which the algorithm is trained. We down-select the most reliable 56 candidates by a joint visual inspection. This final sample is presented and discussed. A conservative estimate based on our results shows that with our proposed method it should be possible to find $\sim$ 100 massive LRG-galaxy lenses at $z \lesssim 0.4$ in KiDS when completed. In the most optimistic scenario this number can grow considerably (to maximally $\sim$2400 lenses), when widening the colour-magnitude selection and training the CNN to recognize smaller image-separation lens systems.

- a search for strong gravitational lenses
- detectable lensed sources are very few compared with full sources

  ➡ a morphological classification based on a CNN (first time)
- 761/21789 are selected, correctly classifies two out of three
- the most reliable 56 candidates are selected by visual inspection

Introduction

- Gravitational lens

  total mass, IMF, Hubble's constant and so on…

- New Survey

  data increase, impossible by visual inspection

- Machine Learning and Deep Learning

  suitable for image recognition task

★ In This Paper

a CNN on KiDS

KiDS is suitable for finding strong lenses

~seeing, pixel scale and large sky coverage~

Fundamental Information of Images

- the Kilo-Degree Survey
  - VLT (Parental Observatory ,Chile, ESO)
  - OmegaCAM wide-field imager
  - KiDS ESO-DR3

Luminous Red Galaxies (LRGs)

parameters: S-Extractor

magnitude: MAG_AUTO

$$r < 20$$
$$|c_{\mathrm{perp}}| < 0.2$$
$$r < 14 + c_{\mathrm{par}}/0.3$$

where

$$c_{\mathrm{par}} = 0.7(g-r) + 1.2[(r-i) - 0.18)]$$
$$c_{\mathrm{perp}} = (r-i) - (g-r)/4.0 - 0.18$$

LRGs are more likely to be lensing galaxies

- Software used
  - Astro-WISE: data handling, analysis
  - S-Extractor: source extraction, photometry

Fundamental Information of CNN

■ class

lens (positive, 1) / non-lens (negative, 0)

■ CNN output

p = 0~1 p>0.5 lens / p<0.5 non-lens

■ training set

▲ images

- mock gravitational lensed sources
  simulation, discussed in 3.1.1
- r-band KiDS images
  real, single band, discussed in 3.1.2

▲ number

- three million for lens / non-lens respectively

0.5 である妥当性

mock sourcesが結果を左右する

Images of Training Set

- Real Galaxy Sample
  - LRGs: 6326
  - contaminants (LRGs): 218
  - false positives: 990
- Mock Lensed-Source Sample

10^6 simulated lensed images

101 by 101 pixels 20 by 20 arctic

the same spatial resolution of KiDS (0.21 arctic per pixel)

include galaxy-galaxy group-galaxy lenses

about twice FWHM r-band KiDS

lenses are typically early-type galaxies

Parameter	Range	Unit
Lens (SIE)		
Einstein radius	1.4 - 5.0	arcsec
Axis ratio	0.3 - 1.0	-
Major-axis angle	0.0 - 180	degree
External shear	0.0 - 0.05	-
External-shear angle	0.0 - 180	degree
Source (Sérsic)		
Effective radius	0.2 - 0.6	arcsec
Axis ratio	0.3 - 1.0	-
Major-axis angle	0.0 - 180	degree
Sérsic index	0.5 - 5.0	-

╳ z > 0.5 small size small Service index

spiral galaxies increase

↓

source redshift < 0.5

exclude

╳ spiral galaxy source

╳ very elliptical ones

Building Training Set

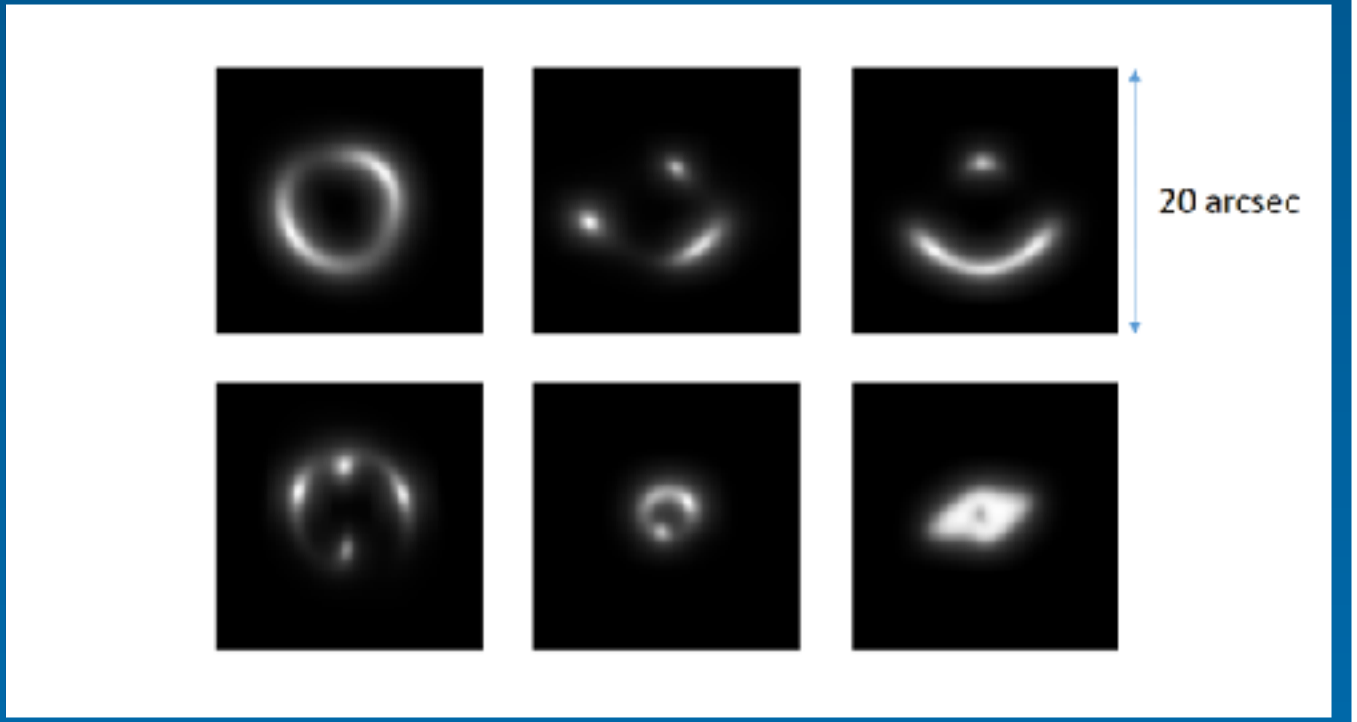
- Positive Sample
  1. choose a mock lensed source and a LRG randomly
  2. perturb both of them randomly see Augmentation
  3. rescale the source peak brightness
  4. add the two images
  5. clip negative values to zero/a square-root stretch
  6. normalize by the peak brightness

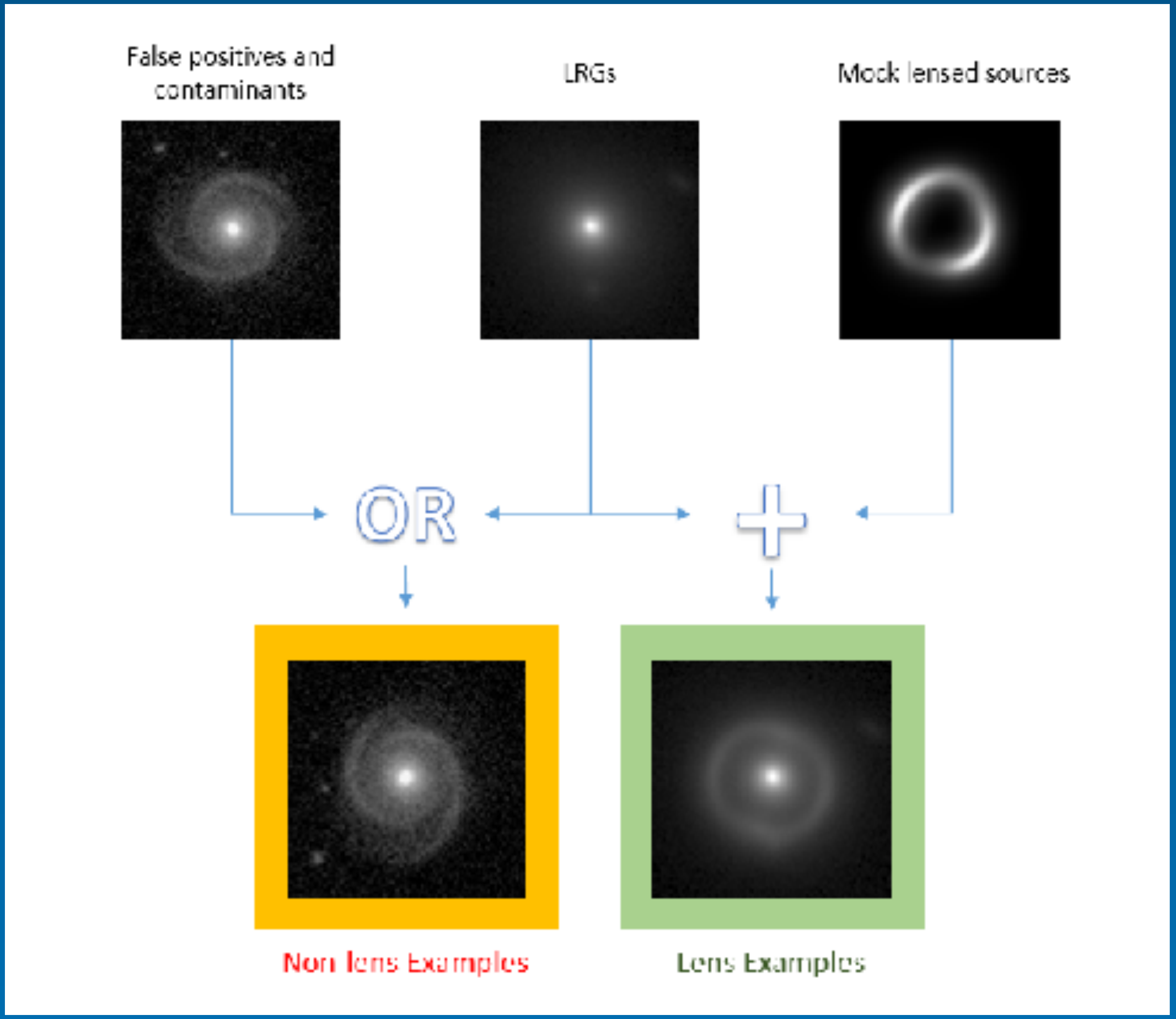
2-20% of the peak brightness of the LRG
typical lower magnitude of the lensing feature

emphasize lower luminosity features

- Negative Sample
  1. choose a galaxy 60% LRG 40% contaminant or false positive
  2. perturb it randomly
  3. a square-root stretch
  4. normalize by the peak brightness

Building Training Set





- Augmentation
  1. random rotation 0~2π
  2. random shift x, y -4~+4
  3. horizontal flip 50%
  4. rescale 1/1.1~1.1 log uniformly

on 101 by 101 pixels images
then cut out into 60 by 60 pixels

categorizeのスコアはこれでいい？
恣意性

Candid ... ion

21789 LRGs (full)

CNN

★ forecasted ~50 (LENSEPOP)

761 LRGs

(include many contaminants)

contaminants examples

include LRGs used for training set

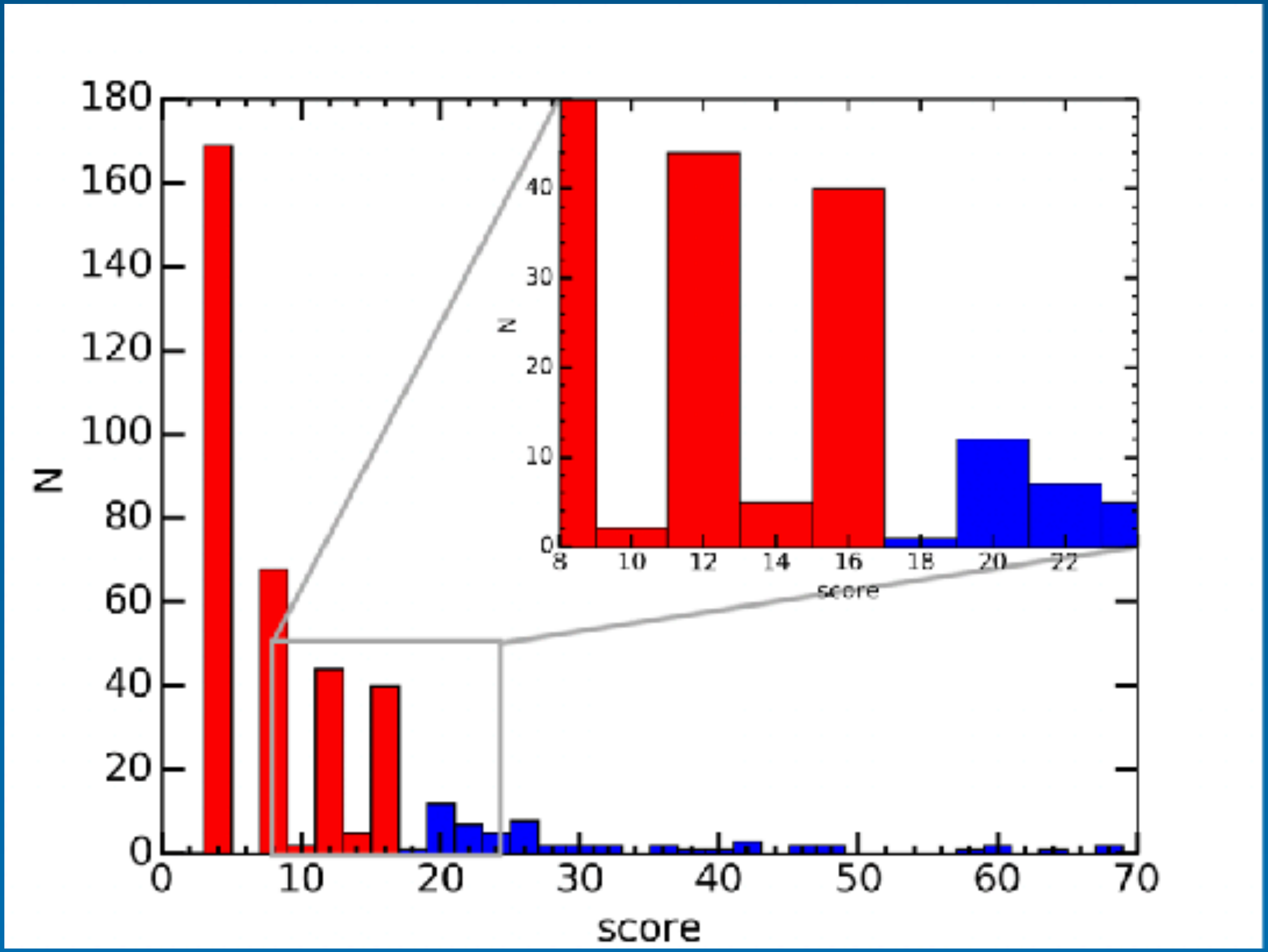
- Remove Contaminants

  seven of authors checked them all

  u, g, r, i + RGB(g, r ,i) composite image (by STIFF)

  → categorize into ‘Sure’(10), ‘Maybe’(4), ‘No’(0)

Candidates Selected



★ forecasted ~50 (LENSEPOP)

384 candidates
classified in ‘Sure’ or ‘Maybe’
at least by one classifier

threshold 17
‘five “Maybe”s at least’
(blue bar)

only two candidates
soccer 70

‘Maybe’ -> 6 points and relocate threshold appropriately
gave no big difference

Final Samples

- Examples
  - successfully classified as lenses

J085446-012137

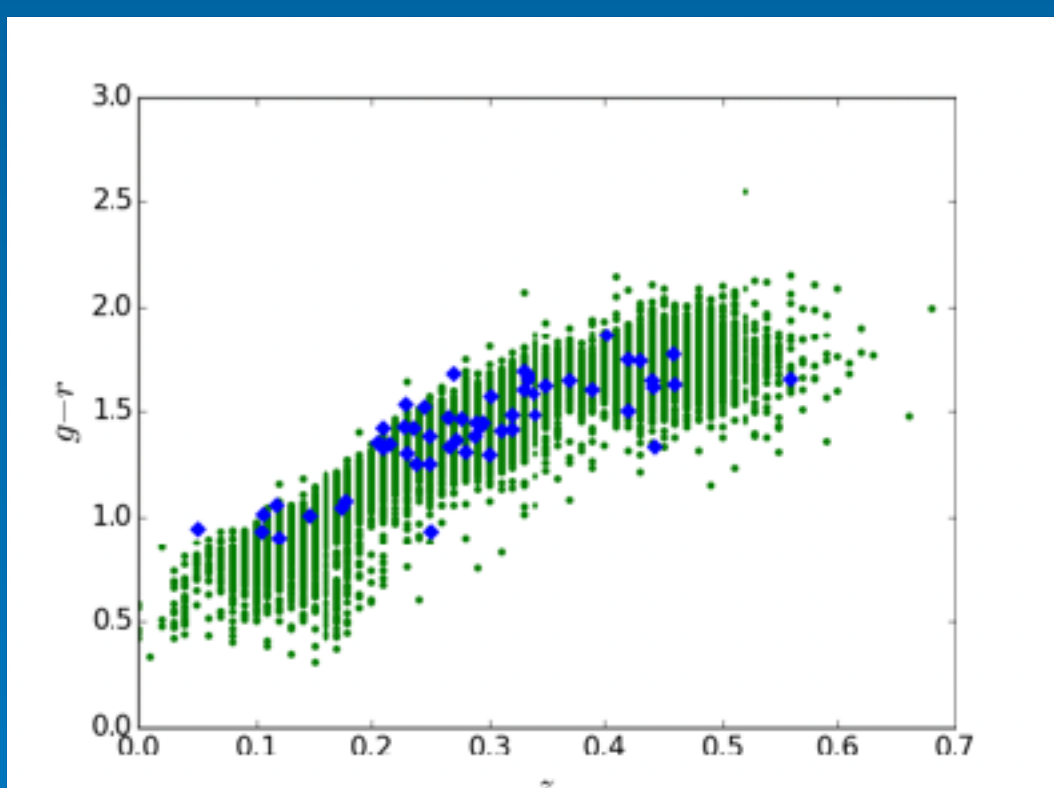
J114330-014427

  - misclassified as non-lenses

J1403+0006

Einstein radius ~0.83 < 1.4
(the CNN was trained over 1.4)

- Other Information
  - z and g-r color relation



56 candidates and full LRGs

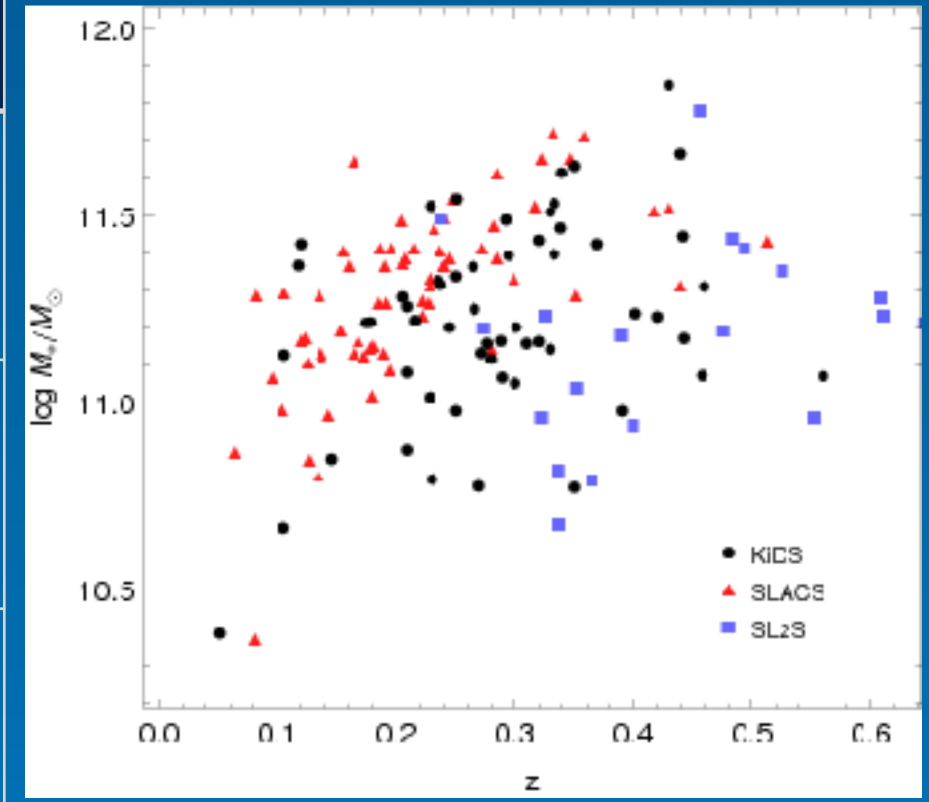
- RGB images (Figure 11)
- r-band images (Appendix C)
- z, magnitudes and so on (Table 2)

of 56 final candidates
(on the handout)

Sample Comparison

- spectroscopic redshift if available, photometric if not
- estimate stellar mass with software LE PHARE
- some of candidates have no velocity dispersion data

	KiDS	SLACS	SL2S
Redshift	$0.28^{+0.12}_{-0.08}$	$0.20^{+0.09}_{-0.07}$	$0.48^{+0.23}_{-0.16}$
Mass (solar mass (scatter))	11.2 (0.2)	11.3 (0.2)	11.2 (0.25)
Velocity Dispersion (km/s)	232^{+46}_{-20}	243^{+47}_{-33}	258^{+42}_{-53}



redshift, mass: median velocity dispersion: average

Estimate Einstein's Radius

SIS model

$$\theta_{\mathrm{E}} = 4\pi \left(\frac{\sigma_{\mathrm{SIS}}}{c} \right)^2 \frac{D_{ls}}{D_s}$$

θ_{E}: Einstein radius (rac

σ_{SIS}: velocity dispersion

D_{ls}: angular diameter d

the lens and th

D_s: angular diameter distance between

the observer and the source

Dls, Dsはどの値を使ったか書かれていない。
zを用いて求めたか。

Jeans equation
the motion of collecting stars in gravitational field

- Dynamical Estimation ➡ calculate velocity dispersion
  - $\sigma_{\mathrm{SIS}} = \sigma_*$
  - Jeans dynamical analysis
  - stellar mass and velocity dispersion relation

$$\log\sigma_* = -0.1 + 0.22\log M_*/M_{\mathrm{sun}}$$

for candidates without velocity dispersion

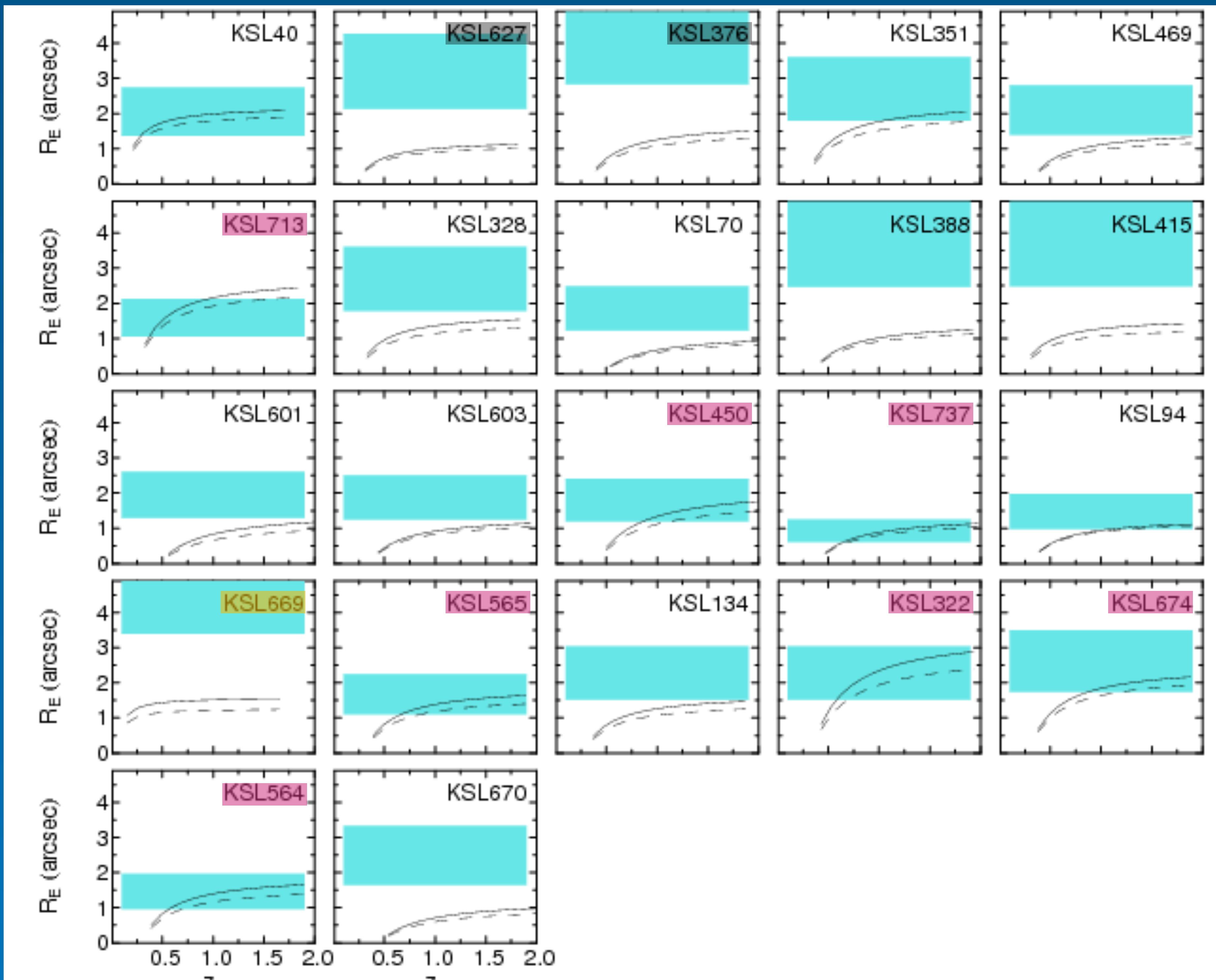
median fit, SLACS ETG lenses

- Observational Estimation

Einstein's radius = $0.5R_E$ ~ R_E (based on Koran et al. 1994)

Observational and Dynamical Radius

~candidates with velocity dispersion values~



solid line: Jeans analysis   dashed line: $\sigma_{SIS} = \sigma_*$ cyan region: $0.5R_E \sim R_E$

lenses

excellent agreement

a marger event

~$2\times10^{10}M_{sun}$

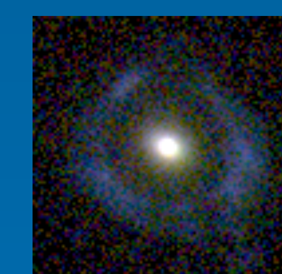
z = 0.05

too small

KSL 669 (26)

ring galaxies

KSL 627 (60)



KSL 376 (48)

~4.3 arctic   ~5.7 arctic

15 kpc   25 kpc

radius do not match for galaxy lenses

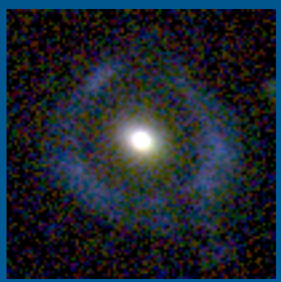
Necessity of Additional Observation

ring galaxies

KSL 627 (60)

~4.3 arctic

15 kpc



KSL 376 (48)

~5.7 arctic

25 kpc

too large

they may be parts of a group of galaxies

e.g. J085446-012137 (KSL317)

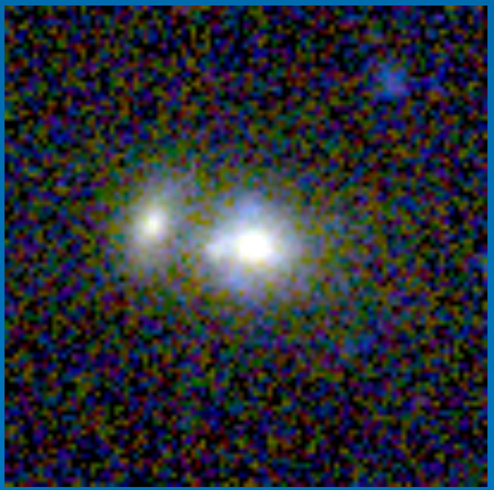
Limousin et al. 2010

spectroscopic validation for the arcs is needed

Examples of Lenses

lenses

KSL 322 (20)



KSL 737 (26)

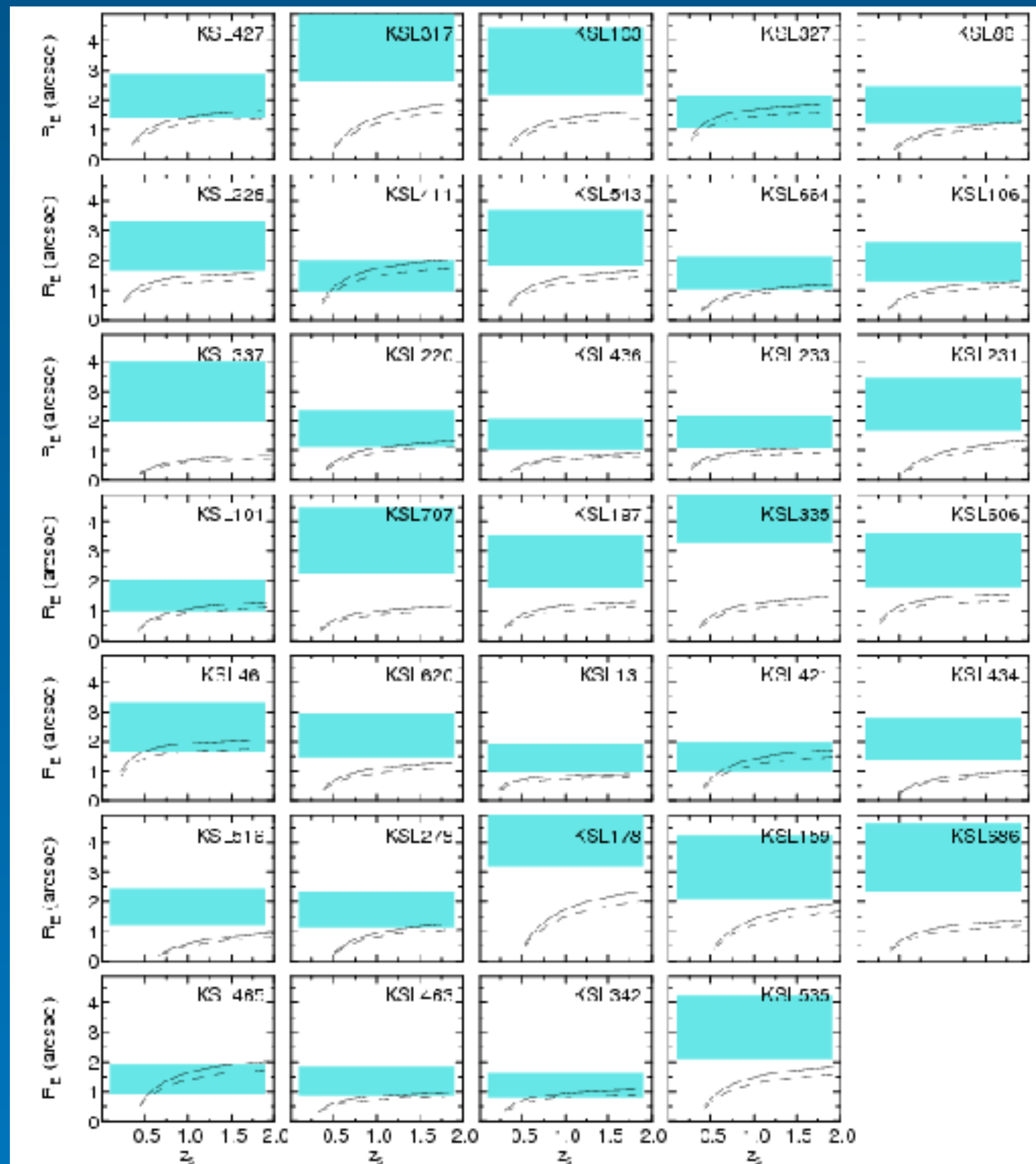
KSL 564 (20)

KSL 565 (24)

4 RESULTS

Observational and Dynamical Radius

~candidates without velocity dispersion values~



solid line: Jeans analysis (ave. σ_{SIS}/σ_* is assumed) dashed line: $\sigma_{SIS} = \sigma_*$ cyan region: $0.5R_E \sim R_E$

COUCLUSIONS

- Candidate Selection

trained CNN on KiDS DR3

		CNN		visual selection		velocity dispersion	
LRGs	21789	→	761	→	56	→	22

- Expected Quality of the CNN
  - ~50 this work
  - ~100 full KiDS
  - ~2400 loosen restriction
- Improve the CNN
  - additional color information
  - training the CNN on the false galaxies and the true galaxies
  - model averaging i.e. different structure and parameters for the same task averaging output
  - subtract galaxy (especially small radii and bright galaxies)