

SPACE WARPS: II. New lens candidates from the CFHTLS discovered through citizen science

Anupreeta More,^{1*} Aprajita Verma,² Phil Marshall,^{2,3} Amit Kapadia,⁴ Michael Parrish,⁴ Chris Snyder,⁴ Julianne Wilcox, Elisabeth Baeten, Christine Macmillan, Claude Cornen, Surhud More,¹ Raphael Gavazzi,⁶ Chris Lintott,¹ Robert Simpson,¹ David Miller,³ Arfon Smith,³ Edward Paget,³ Prasenjit Saha,⁴ Rafael Kueng,⁴ Kelly Borden,³ Tom Collett, Thomas Jennings, Matthias Tecza,¹ Layne Wright and possibly others

¹*Kavli IPMU (WPI), University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa 277-8583, Japan*

²*Dept. of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK*

³*Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94035, USA*

⁴*Adler Planetarium, Chicago, IL, USA*

⁵*Department of Physics, University of Zurich, Switzerland*

⁶*Institut d'Astrophysique de Paris, UMR7095 CNRS Université Pierre et Marie Curie, 98bis bd Arago, 75014 Paris, France*

to be submitted to MNRAS

ABSTRACT

The CFHT Legacy Survey has been searched for strong lenses with semi-automated algorithms both at galaxy and groups scales. With the aim of improving these lens finding robots, we carry out a blind lens search in the complete CFHTLS WIDE survey with SPACE WARPS. We describe the training sample used for both training the citizen scientists that participated in SPACE WARPS and calibrating their performance. We generate realistic looking simulated samples of lenses both at galaxy and group-scales as part of this training sample. We present 61 new strong gravitational lens candidates discovered from the SPACE WARPS-CFHTLS search out of which 40 candidates are promising. Furthermore, we compile a sample of false positives which can be used for robot testing, first of its kind for the lensing community. The simulated sample is also used to make a comparison between human performance against the robots. We find that the XX per cent of the known lens sample is recovered and the new lens sample has XX completeness with respect to the simulated sample. We make available the following data products: simulated sample, simulation code and the false positives sample at XXX.

Key words: gravitational lensing – methods: statistical – methods: citizen science

1 INTRODUCTION

The last few decades have seen a rise in the discoveries of strong gravitational lenses owing to the plethora of interesting applications they have in astrophysics and cosmology. Strong lenses are routinely used to probe the dark matter distribution from galaxy (ref) to cluster scales (ref), to study distant young galaxies by using the lensing magnification as

a natural telescope (ref), to test the cosmological model by constraining cosmological parameters such as the Hubble constant (ref) and dark energy (ref) and many more. Even though strong lenses are rare, since a foreground massive object needs to be sufficiently aligned with a distant background source to produce multiple images, systematic lens searches have led to discovery of over 500 lenses till date (XXX add mld url).

Rarity of lenses implies searching for them is a painstaking task. Efficient automated methods are thus imperative

* anupreeta.more@ipmu.jp

to finding a reasonably complete and pure sample of strong lenses.

Since the inception of the first citizen science project, Galaxy Zoo, to classify galaxy morphologies (Lintott et al. 2008), several astronomy and non-astronomy projects have been launched by the Zooniverse leading to many interesting project. For example, PlanetHunters has discovered XXX transiting exoplanets (Schwamb et al. 2012), Supernova XXX add ref and XXX add ref

In Paper I, we describe SPACE WARPS, an online system that enables crowd-sourced detection of gravitational lenses. In this paper (referred to as Paper II), we describe our first lens search with SPACE WARPS in the CFHTLS data.

This paper is organised as follows. In Section ?? we give a brief overview of the SPACE WARPS system, focusing on the aspects most relevant to the interpretation of the results of this first lens search. In Section 2 we introduce the CFHTLS imaging data and the known lens samples from the CFHTLS. In Section ??, we explain the training sample generated to aid the SPACE WARPS users in the lens search. In Section 5, we present the new lens candidates and our findings. We discuss the implications of our results for future lens searches in Section 6 and draw conclusions in Section 7.

2 DATA

2.1 The CFHT Legacy Survey

The Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) is a photometric survey in five optical bands ($u^*g'r'i'z'$) carried out with the wide-field imager MegaPrime which has a 1 deg^2 field-of-view and a pixel size of $0.186''$. The CFHTLS WIDE covers a total area of 171 deg^2 on the sky and it consists of four fields W1, W2, W3 and W4. The field W1 has the largest sky coverage of 63.65 deg^2 . The fields W2 and W4 have similar sky coverages of 20.32 deg^2 and 20.02 deg^2 , respectively¹. The field W3 has a sky coverage of 42.87 deg^2 and is more than twice as large as W2 and W4.

The CFHTLS imaging is very homogeneous and has great image quality. Most of the lensed arcs are much brighter in the g band thus, deep imaging in this band is desirable. The limiting magnitude is 25.47 for the g band which goes the deepest among all of the five bands. The mean seeing in the g band is $0.78''$. The zero point to convert flux to AB magnitude for all bands is 30. These characteristics make CFHTLS ideal to do visual inspection for finding lenses. We use the data from the final T0007 release taken from the Terapix website² for this work.

We note that the CFHTLS is a niche survey with a unique combination of wide imaging with deep sensitivity. It is a precursor to the ongoing wide imaging surveys such as the DES, KIDS and HSC and planned surveys such as the LSST. Searching for lenses with SPACE WARPS in the CFHTLS will teach us important lessons and help prepare us for these larger imaging surveys.

2.2 Image Presentation in SPACE WARPS

Our aim is to do a blind lens search over the entire CFHTLS WIDE. We use the g , r and i -band color imaging which are most useful for visual identification of lenses (and provide access to the u and z band imaging through dashboard with additional tools XXX).

We make color composite images using publically-available code³ following the prescription of Lupton et al. (2004). Specifically, we first rescale the pixel values of each channel image into flux units, and then apply an arcsinh stretch. The stretch parameters are chosen using a small random sample of images, to ensure that the background noise is just visible, and that the centres of bright, intermediate redshift galaxies are not saturated. The color scales are chosen to maximize the contrast between faint extended objects. These parameters are then fixed during the production of all the tiles, in order to allow straightforward comparison between one image and another, and for intuition to be built up about the appearance of stars and galaxies across the survey.

We extract contiguous cutouts of size $81.84''$ (440 pixels) with an overlapping region of $10''$ (54 pixels) between the neighbouring cutouts resulting into $\sim 430\,000$ cutouts for the entire CFHTLS WIDE. The size of the individual cutout is determined by optimising factors such as the typical angular scales of gravitational lenses, the number of objects seen in a single cutout and the total number of image cutouts in the survey. It is possible that a lens candidate happens to be too close to the edge of a cutout, the overlap between neighbouring cutout allows the inspector to get a clearer view of the same candidate in at least one of the cutouts. We note that since the images are shown randomly, an inspector may not necessarily come across the neighbouring cutout unless the inspector classifies a large number of images. This is not a problem since our user base is extremely large and we receive multiple classifications of the same cutout.

2.3 Existing CFHTLS lens samples

The CFHTLS data has been searched for lenses using semi-automated algorithms, primarily, in the g band where the predominantly faint blue source galaxies are bright compared to the predominantly red deflector galaxies. Here, we briefly mention the lens samples which were known to the authors prior to the lens search with SPACE WARPS.

At galaxy-scales, we focus on two primary lens searches. The RINGFINDER (Gavazzi et al. 2014) was used for finding compact rings or arcs around centers of isolated and massive early-type galaxies. By subtracting the PSF-matched i -band images from the g -band images, the algorithm looks for excess flux in the bluer g -band. An object detector measures the properties of these residual blue features, and candidates meeting length-width ratio and tangential alignment criteria are then visually inspected to form the final sample. Gavazzi et al. (2014) first selected some 638,000 targets as either photometrically-classified early type galaxies, or objects selected to have red centers and blue outer parts, from the T06

¹ These numbers are estimated from http://terapix.iap.fr/cplt/table_syn_T0006.html

² <http://terapix.iap.fr/cplt/T0006-doc.pdf>

³ The open source color image composition code used in this work is available from <http://github.com/drphilmarshall/HumVI>

CFHTLS data release catalogs. 14370 were found to show detectable blue residuals, and 2524 were visually inspected, having passed the automatic feature selection process. This led to a sample of 42 good quality (`q_flag` = 3) and 288 medium quality (`q_flag` = 2) lens candidates. In addition to this well defined sample, Gavazzi et al. (2014) reported a further 71 serendipitously detected lens candidates. From this sample of “RINGFINDER candidates,” the SL2S team found, during their follow-up campaign, 39 confirmed lenses (and 17 promising candidates). We use this sample of 39 “confirmed RINGFINDER lenses” in our completeness analysis.

The second galaxy-scale lens search was to find edge-on galaxy lenses in the CFHTLS (Sygnet et al. 2010) by selecting the galaxy’s profile from the output of SExtractor in the *i* band and with low inclination angle. This sample has about 3 promising and a total of 18 lens candidates.

On the other hand, the ARCFINDER More et al. (2012) was used for finding blue arc-like features in the complete CFHTLS data without any pre-selection on the type of the lensing object. The search was carried out in the *g*-band which is the most efficient wavelength to find typical lensing galaxies. This sample, called the SARCS, has 55 promising and a total of 127 lens candidates and consists of both galaxy and groups/cluster scale lens candidates. Arc finding is better suited for lensed images or arcs with larger image separations i.e. more massive systems like groups and clusters. Thus, the SARCS sample has mostly groups/cluster-scale lenses and a few galaxy-scale lenses.

For the purposes of transparency and to help a little with their training, the volunteers participating in SPACE WARPS-CFHTLS lens search were made aware of these known lens samples. Images containing the systems from the above samples were labelled as “Known Lens Candidates” in the Talk forum, where volunteers have the opportunity to discuss their findings with other volunteers and the science team.

3 TRAINING SAMPLE: SIMULATED LENSES

State the importance

3.1 Methodology

We create two main types of simulated lens sample a) galaxy-scale lenses and b) group or cluster-scale lenses. The galaxy-scale lenses are further divided into two types based on the nature of the background sources, namely, galaxies and quasars. Below, we describe how each type of the lens sample was generated.

3.1.1 Galaxy-scale lenses

The N_{src} behind a lens is then calculated by doing the following integral,

$$N_{\text{src}} = N_{\text{src}}(> L_s, z_s) \int_{z_l}^{\infty} \sigma_{\text{lens}}(\sigma, z_l, z_s, q) D_s^2 (1+z_s)^2 \frac{d\chi}{dz_s} dz_s \quad (1)$$

$$N_{\text{src}}(> L_s, z_s) = \int_{L_{\text{min}}}^{\infty} \Phi(L_s, z_s) dL_s \quad (2)$$

where $\Phi(L_s, z_s)$ is the source luminosity function per unit comoving volume, q is the projected axis ratio of the lens ellipticity, χ and D_s are the comoving and angular diameter distances to the source,

We use the elliptical galaxy (LRG) catalog from the CFHTLS XXX to select all the foreground galaxies (e.g. $z < 1$) that are potential lenses for the simulated sample. We exclude all those galaxies whose positions match with the lensing galaxies from the known CFHTLS SL2S lens samples More et al. (2012) within 2 arcsec (XXX check).

First, we calculate the luminosity and velocity dispersion of each potential lensing galaxy using the CFHT Megacam *g* and *r* band magnitudes along with the photometric redshift (z_l) from the LRG catalog. The Megacam magnitudes are converted to SDSS magnitudes⁴ and are further *k*-corrected to redshift $z = 0.1$ (Frei & Gunn 1994). We assume that the evolution of galaxy luminosities is similar to that determined by (Faber et al. 2007), that is, a decline of 1.5 in the m_{r*} from redshift $z = 1$ to $z = 0$ (see Eq. 4).

$$\frac{L}{L_*} = 10.0^{-0.4 (m_{r\text{SDSS}} - m_{r*})} \quad (3)$$

where m_{r*} is

$$m_{r*} = -20.44 + 1.5 (z_l - 0.1). \quad (4)$$

We use the $L - \sigma$ relation from (Parker et al. 2005) to get the velocity dispersion as given in Eq. 5.

$$\sigma = 142 \left(\frac{L}{L_*} \right)^{1/3} \quad (5)$$

Next, in order to decide whether a galaxy is likely to act as a strong lens, we calculate the lens cross-section (σ_{lens}) and the number of sources (N_{src}) that are in the background. Following (?), the lens cross-section is calculated analytically for an isothermal model and is given by

$$\sigma_{\text{lens}} = b_I^2 \int_0^{2\pi} 0.5r^2(\theta) d\theta \quad (6)$$

where b_I is

$$b_I = b_{\text{SIS}} \epsilon_3 / \sin^{-1}(\epsilon_3), \quad (7)$$

the eccentricity (ϵ_3) is

$$\epsilon_3 = (1 - q_3^2)^{1/2} \quad (8)$$

and the projected axis ratio is given by

$$q_k = \sqrt{q_3^2 \sin^2 i_e + \cos^2 i_e}. \quad (9)$$

In the above equations, q_3 is the 3d axis ratio of the ellipsoid and i_e is the inclination angle. Also, $b_{\text{SIS}} = 4\pi \frac{c^2}{\sigma^2} \frac{D_s}{D_{ls} D_l}$ and is referred to as the Einstein radius where

⁴ <http://www3.cadc-ccda.hia-ih.nrc-cnrc.gc.ca/megapipe/docs/filters.html>



Figure 1. Examples of the three types of simulated lenses.

D_s, D_l and D_{ls} are angular diameter distances to the source, the lens and between the lens and source, respectively.

Next, if the foreground galaxy can act as a lens and has at least one source in the background, then we determine a redshift (z_s) and i -band magnitude of the background source(s). We assume two types of background sources namely, galaxies and quasars. For each source, the redshift and magnitude are generated by drawing randomly from the following redshift and luminosity distributions. For galaxies, we assume the redshift distribution is

$$p_s = \frac{\beta z_s^2 \exp(\frac{z_s}{z_0(m_{\text{lim}})})}{\Gamma(3/\beta) z_0^3 (m_{\text{lim}})} \quad (10)$$

where $\beta = 3/2$ and $z_0(m_{\text{lim}}) = 0.13m_{\text{lim}} - 2.2$ and the luminosity function is

$$n_s = \int_{-\infty}^{m_{\text{lim}}} \frac{n_0 dm}{\sqrt{10^{2a(m_1-m)} + 10^{2b(m_1-m)}}} \quad (11)$$

with parameters $a = 0.30$, $b = 0.56$, $m_1 = 20$ and $n_0 = 3 \times 10^3 \text{ deg}^{-2}$ as given in (Faure et al. 2009) and references therein. For quasars, we calculate the luminosity function by following the prescription of (?) and use k -corrections by (Richards et al. 2006).

The luminosity function is expressed as

$$\frac{d\Phi}{dM} = \frac{\Phi_*}{10^{0.4(\alpha+1)(M_{\text{abs}}-M_*)} + 10^{0.4(\beta+1)(M_{\text{abs}}-M_*)}} \quad (12)$$

where the normalization, $\phi_* = 5.34 \times 10^{-6} h^3 \text{ Mpc}^{-3}$ and break magnitude, $M_* = -20.90 + 5 \log h - 2.5 \log f(z)$. The redshift dependent factor in M_* is given by

$$f(z) = \frac{e^{\zeta z_s} (1 + e^{\xi z_*})}{(\sqrt{e^{\xi z_s}} + \sqrt{e^{\xi z_*}})^2}. \quad (13)$$

We adopt the best-fit values $\zeta = 2.98$, $\xi = 4.05$, $z_* = 1.60$ (Oguri & Marshall 2010). For the faint end slope, we use $\beta = -1.45$ whereas for the bright end slope, we use $\alpha = -3.31$ when $z_s < 3$ and $\alpha = -2.58$ at higher redshifts, as prescribed by (Oguri & Marshall 2010). respectively. We note that when calculating N_{src} , the source number density is artificially boosted by a factor (see Table 1) to increase the occurrence of simulated lenses. This helps in creating a large enough sample to carry out various performance tests.

Next, we determine properties of the background source for every lens. We follow similar procedures for both background galaxies and quasars. For simplicity, we simulate a

single background source behind every lens. In order to select one background source from the N_{src} per lens, we do ray-tracing for all of the N_{src} sources with GRAVLENS (Keeton et al. 2000) and choose sources that satisfy criteria as given below. We determine fluxes of the lensed images and the total magnification of each of the lensed source. We draw a random source for which the flux of the second brightest lensed image and the total magnification of all lensed images are above the thresholds given in Table 1.

Since we want to produce realistic looking lens systems, we simulate lenses in each of the five CFHTLS filters. The colors of the background galaxies are drawn randomly from the photometric CFHTLenS catalog (Hildebrandt et al. 2012; Erben et al. 2013). Similarly, we use a quasar catalog from the SDSS Data Release 9 (P  ris et al. 2012) from which colors are drawn to simulate quasar lenses. Next, we assume deVaucouleur's profile to account for the size and shape of the galaxies. The ellipticity and the position angle are drawn randomly between the range given in Table 1. The effective radius of the galaxy is estimated from the Luminosity–size relation (Bernardi et al. 2003) given by

$$R_{\text{eff}} = 10^{0.52} \frac{L_r^{2/3}}{(1 + z_s)^2} \quad (14)$$

where $L_r = L_s/10^{10.2}$. On the other hand, quasars are assumed to follow a Gaussian profile where the σ is equated to that of the median seeing for every filter. The median seeing values are taken from Table 4 of the official Terapix T0007 release explanatory document ⁵.

Once all the parameters are determined for the lens and source models, GRAVLENS is used to simulate lensed images. After accounting for the shot noise in the lensed images and convolving them with the median seeing in each of the filters, the simulated image is added to the real CFHTLS image centered on the lensing galaxy. Note that we ensure that the lensed galaxies and lensed quasars do not have the same lensing galaxy in the foreground. Similarly, the lensing galaxies from the galaxy-scale lenses are distinct from the central galaxies of groups-scale lenses which are described in the following section.

⁵ <http://terapix.iap.fr/cpl/T0007/doc/T0007-doc.pdf>

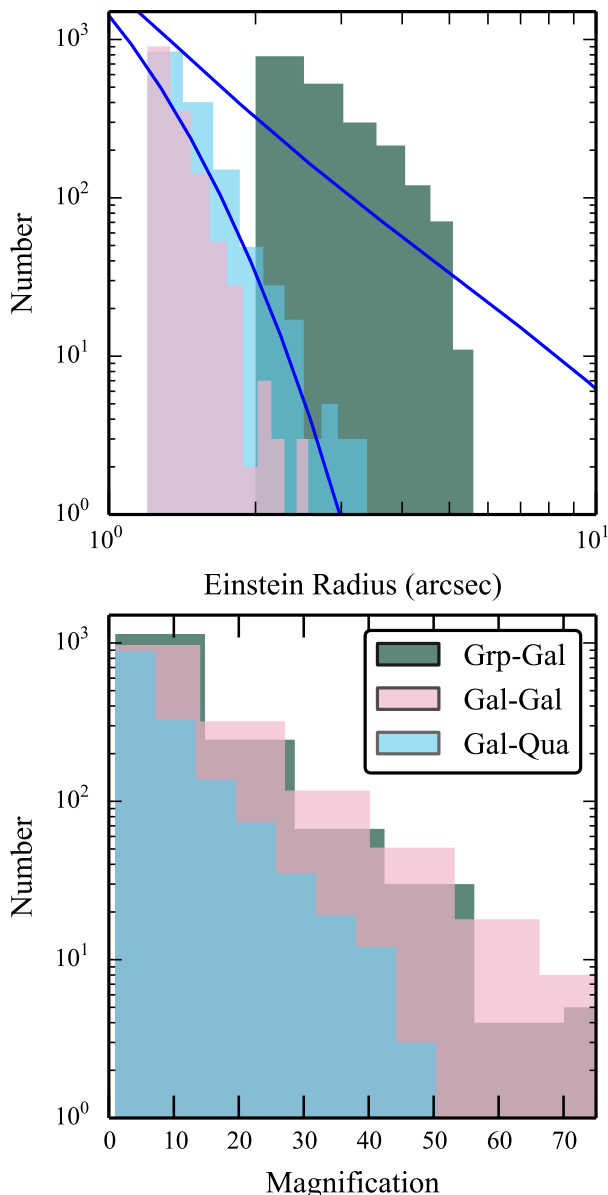


Figure 2. Einstein radius distribution for all types of lenses. The dashed-dotted (blue) curves show the theoretical prediction assuming an SIS model at galaxy-scales and a total (NFW+Hernquist) model at groups-scales taken from (More et al. 2012).

3.1.2 Groups-scale lenses

At group or cluster-scales, the brightest group galaxy (BGG) at the center alone does not cause strong lensing. We need to account for the extra convergence arising from the dark matter component as well as satellite galaxies, at least, in the inner regions which are typically responsible for the multiple lensed images (Oguri 2006). Owing to the lack of an appropriate group catalog for our purposes, we create a basic group catalog based on the magnitudes and photometric redshifts available for the CFHTLS. We select all galaxies with $10^{10.8} M_{\odot}$ as the BGGs. We select the member galaxies

Table 1. Thresholds used in the selection of the simulated lenses.

Name	gal		qso	
	min	max	min	max
Source Redshift	1.0	4.0	1.0	5.9
Source Magnitude	21.0	25.5	21.0	25.5
boost factor	100 (40†)		1200	
Second brightest image	23		23	
Total magnification	19		20	
Lens shear strength	0.001	0.02	0.001	0.02
Lens shear pa	0	180	0	180
Source ellipticity	0.1	0.6		
Source PA	0	180		

† – corresponds to the factor used for Groups scale lenses.

such that their photometric redshifts are within $\delta z = 0.01$ of the BGG and within an aperture of 250 Kpc.

We adopt an isothermal ellipsoid for the BGG and members whenever the ellipticities are available else we use an isothermal sphere. On the other hand, we adopt an NFW profile for the underlying dark matter halo. Assuming a constant mass-to-light ratio of $3 \times 0.7 h M_{*}/L_{*}$, we use the BGG luminosity to estimate the stellar mass. The stellar mass–halo mass relation (?), including random scatter, is then used to calculate the halo mass for the lens. Given the halo mass, other key parameters such as the scale radius (r_s) and the density at the scale radius (ρ_s) can be determined for an NFW profile.

As described in Section 3.1.1, we calculate the luminosity and velocity dispersion for the BGG and each of the member galaxies. Next, we calculate the lens cross-section for each potential lensing group. The complexity in the lens models makes it analytically intractable to calculate the size of the caustics⁶. Hence, we use GRAVLENS to determine the area covered by the caustics. We consider only galaxies as our background source population since group or cluster-scale quasar lenses are not expected to be found in the CFHTLS (check XXX). Following the same procedure as described in Section 3.1.1, we calculate the number of galaxies expected to lie behind every potential lensing group (see Eq. 1). As before, for each background galaxy within the lens cross-section, a redshift and an i -band magnitude is determined by drawing galaxies randomly from the respective distributions (see Eqs. 10-11).

All those groups that are found to have no background galaxies within the cross-sectional area are rejected and the rest are included as potential lenses. As mentioned earlier, we artificially boost the total number of sources behind every lens but ensure that (check XXX) the statistical properties such as the profile of the image separation distribution are not affected (see Figure 2). We follow the same procedure and apply the same thresholds to determine properties of the lensed galaxies for every lens as are described for galaxy-galaxy lenses in the previous section. The simulated images

⁶ The lens mass distribution determines size and shape of the caustics. Any source located within the caustics will form multiple lensed images which is the criteria for strong lensing. To further understand caustics, see XXX.

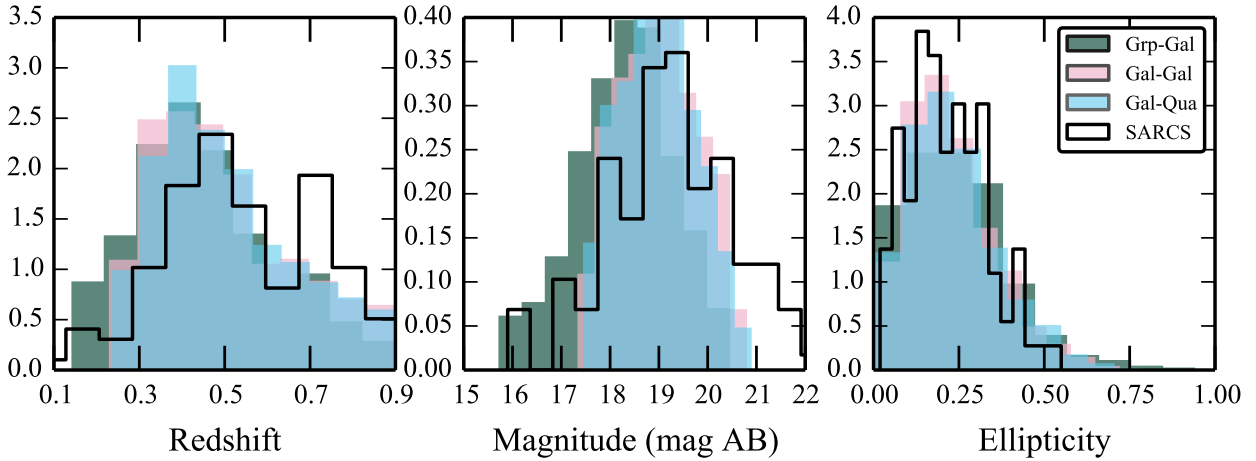


Figure 3. Distributions of properties of the lensing galaxies of the simulated sample compared to the known lens sample SARCS XXXX check?

are added to the real CFHTLS images with the BGGs as the center.

3.2 Simulated Lens Sample and Catalog Description

In this section, we describe some of the properties of our simulated sample for each of the three types of lens samples.

The Figure 2 shows Einstein radius distribution for the galaxy-scale (dashed for background quasars and dotted for background galaxies) and groups-scale simulated lenses. For comparison, we show the expected distributions (blue dashed-dotted curves) for an SIS-like density profile at galaxy-scales and an NFW+Hernquist profile at groups-scales. The theoretical curves are taken from (More et al. 2012) wherein the models are explained in detail. We note that the model we adopt at groups-scale also includes SIS or SIE components for the group members unlike the theoretical prediction. The theoretical curves have arbitrary normalizations.

We show the redshift and magnitude distributions of the lensing galaxies in the left and right panels of Figure 3 respectively. Furthermore, we overplot the distributions of respective properties of the SARCS lenses from (More et al. 2012) for comparison with arbitrary normalizations. We note qualitative similarities between the simulated and the real lens samples.

Similarly, we show the ellipticity and the position angle of the simulated lens galaxy population extracted from the T0007 release of the CFHTLS catalogs in the left and right panels of Figure 3, respectively. As before, the dashed-dotted (blue) curves show the same distribution for the SARCS lens population with arbitrary normalization (More et al. 2012).

We produce catalogs with lens and source properties for each of the three types of lenses. These catalogs are available XXX. The catalogs typically have lens position, redshift, magnitudes, Einstein radius, ellipticity (whenever available) and shear (for galaxy-scale lenses only). For the background sources, we provide the offset from the lens center, redshift, magnitudes, total magnification, number of lensed images.

Additionally, ellipticity and effective radius when the background sources are galaxies.

3.3 Limitations of the simulated sample

The simulated sample is not perfect because our understanding of various phenomena in the Universe is not complete and because of the uncertainties in the parameters of our model. Here, we describe some of the cases or aspects in which the simulations are known to have failed or seem unrealistic.

The parameters required by various scaling relations and the models primarily depend on the photometry of the galaxies, groups and quasars detected in the survey. For galaxy-scale lenses, the lensing galaxies at higher redshifts or which are fainter have poor photometric measurements. This causes relatively larger uncertainties in its luminosity and velocity dispersion and leads to simulated lenses which look implausible. For example, due to a larger uncertainty in the velocity dispersion of the lensing galaxy, the lensed images may have larger image separation than what is expected given the visual priors from the galaxy.

At group-scales, the photometric and redshift estimates are used when defining the group membership. Therefore, errors in redshift estimates generate galaxy groups with BGG or member galaxies with dissimilar properties. In some cases, low redshift spiral galaxies are incorrectly assigned high redshift. Spiral galaxies are typically less massive and low redshift spiral galaxies are unlikely to act as gravitational lenses. Hence, the resulting simulated lenses are not convincing.

We use single Sersic component to describe the light profiles of background galaxies. This is clearly not the most accurate description for galaxies, especially, star-forming galaxies which form a significant fraction of the lensed galaxy population. Star-forming galaxies have complex structures such as star forming knots, spiral arms, bars and disks.

4 TRAINING SAMPLE: DUDS AND FALSE POSITIVES

A good training sample consists of a representative set of objects that one wants to find and another set of objects which appear to be from the former set but are of different origin in reality and which one can learn to discard efficiently. Indeed, we wanted to have a good training sample for the SPACE WARPS users so that they can correctly identify the true lens candidates. Hence, in addition to the simulated lenses, we added a sample of duds and false positives to the training sample. Duds are images which have been visually inspected by experts and confirmed to contain no lenses. False positives (FPs) are systems which look like lenses but are not, for example, spiral galaxies, starforming galaxies, chance alignments of features arranged in a lensing configuration and stars.

We selected a sample of 450 duds for the Stage I classification in SPACE WARPS and a sample of 500 false positives for the Stage II inspection. The sample of false positives was selected from the candidates which passed the Stage I of SPACE WARPS. We note that this is the first time, we have a systematically compiled sample of visually inspected false positives by the SPACE WARPS users and categorized by the science team. Such a sample is tremendously helpful for training and understanding performances of various lens finding algorithms (Chan et al. 2014). **TBA data products; make the duds and FP sample available**

5 RESULTS

5.1 SPACE WARPS Lens Sample

The SPACE WARPS works as a single unified neural network which uses the method of visual inspection to find gravitational lenses. The citizens are shown images at two stages. At stage I, citizens focus on rapid inspection and selection of lens candidates ranging from possible lenses to almost certain lenses. At stage II, citizens carefully inspect the candidates from stage I and select promising lens candidates only. The science team consisting of lens experts then categorize and grade the lens candidates further purifying the sample of lens candidates. In Table 2, we list the total number of detections of the known lens candidates, known confirmed lenses and the new lens candidates at both stages I and II. The known sample corresponds to the combined Ring Finder and Arc Finder samples. We also show their fractions with respect to the total number of lens candidates found at these stages. We note that all the confirmed lenses found at stage I are also recovered at stage II whereas a small fraction of the known lens candidates are rejected at stage II. Also, the new lens candidates increases the known lens candidates sample by over 50%.

Below, we describe the steps taken after stage II for selecting the final sample of lens candidates.

5.1.1 Selection of the SPACE WARPS lens candidates

The images inspected in SPACE WARPS are assigned probabilities (P) through the SPACE WARPS Analysis Pipeline (SWAP), built in the SPACE WARPS system (see Paper I). In an ideal situation, at the end of stage II, all images containing lenses

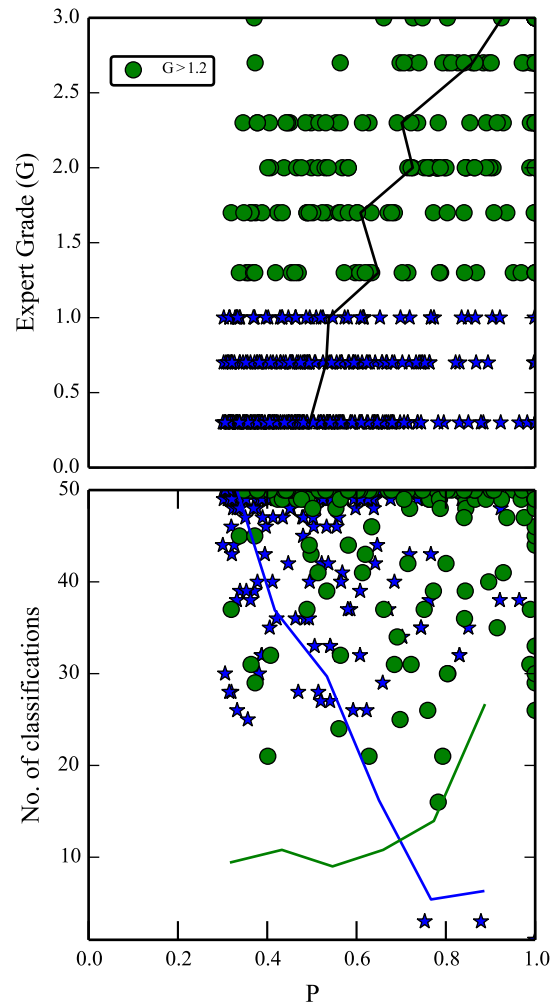


Figure 4. Comparison of the $P(\text{lens})$ with the expert grades and number of classification for each subject.

would have high P values and those without lenses will have low P values. However, practically, a fraction of the real lenses will be assigned low P value decreasing the completeness and a fraction of non-lenses will be assigned high P value decreasing the purity. Our goal is to find a P value which will result in acceptable levels of completeness and purity of the final sample of lens candidates.

As this is the first SPACE WARPS lens search, we explore a large range in the P value to understand the balance between the real lens candidates and false positives. In Figure 4, we plot the distribution of false positives and real lens candidates categorized by the lens experts as a function of their P value. As expected, the fraction of real lens candidates is higher for higher P values and it decreases for lower P values. We note that from $P \sim 0.7$ onwards to lower P values, the fraction of false positives starts to exceed the fraction of real lens candidates. This could be a good threshold to choose to maximize the purity of the final sample. However, we choose $P > 0.3$ as our threshold to increase the completeness of our sample. We achieve further purity by including the grades of the lens experts.

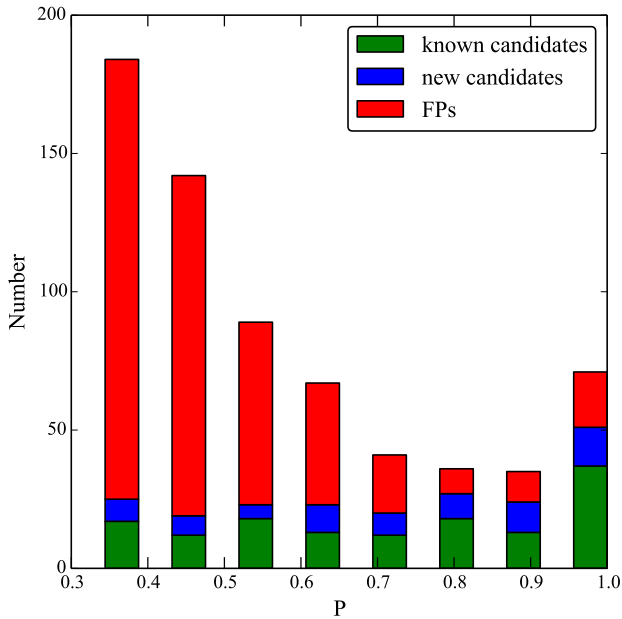


Figure 5. Total number of detections of the known candidates, new candidates and the FPs shown relatively for binned values of P .

Table 2. Statistics of detections in SPACE WARPS

	Stage I		Stage II			
	KC	KL	KC	KL	NC	AC
Number	128	34	107	34	61	168
%	29	58	25	58	12	34
P_{thresh}	0.95	0.95	0.3	0.3	0.3	0.3

KC– Known lens candidates

KL– Known confirmed lenses

NC– New lens candidates

AC– All (known and new) lens candidates

P_{thresh} – systems with P above this threshold are selected

Note about %: For KC and KL, fractions are with respect to the known population whereas for NC and AC, fractions are with respect to the total population of lens candidates.

5.1.2 New lens candidates from SPACE WARPS

All images with $P > 0.3$, at the end of stage II, are inspected by three lens experts (AM, AV, PJM) and are assigned integral grades from 0,1,2 or 3 which imply that the candidate is almost certainly not a lens, possibly a lens, probably a lens and almost certainly a lens, respectively. This process led to a final sample of 61 new lens candidates with averaged grade ≥ 1.3 (see Table 3) and are shown in Figure 6. For this final sample, we give a SPACE WARPS ID and Name of the lens system, Ra, Dec, photometric redshift (z_{phot}), i band magnitude of the lensing galaxy, averaged grade from the lens experts, zoo ID (identifier used in TALK⁷, the discussion forum for SPACE WARPS), P value at stage II and a visual categorization of type of lensed images and the lensing galaxy in

the Comments column in Table 3. Whenever available the lens properties are taken from the CFHTLS photometric catalog (Coupon et al. 2009) otherwise the reported lens galaxy positions are measured visually. The visual categorization of the lens type is only suggestive and the explanation of the notations in the Comments column is given at the bottom of the table.

As the first lens search was a blind search with no pre-selection of candidates using any algorithm, we expected to find various types of lenses and this is indeed what we find. The final sample consists of both galaxy and groups-scale lens candidates. There are detections of elongated arcs and point-like quasar lensed images. Most of them are brighter in bluer g band but some candidates brighter in the redder i band are also found.

CHECK FOR STAGE 1 – COMPLETENESS OF SIMS AND KNOWN LENSES and P value comparison, what about the duplicate lenses near the borders - stats on those? check the detection plot of arc mag / Reinst and find out why some of the bins have unexpected results

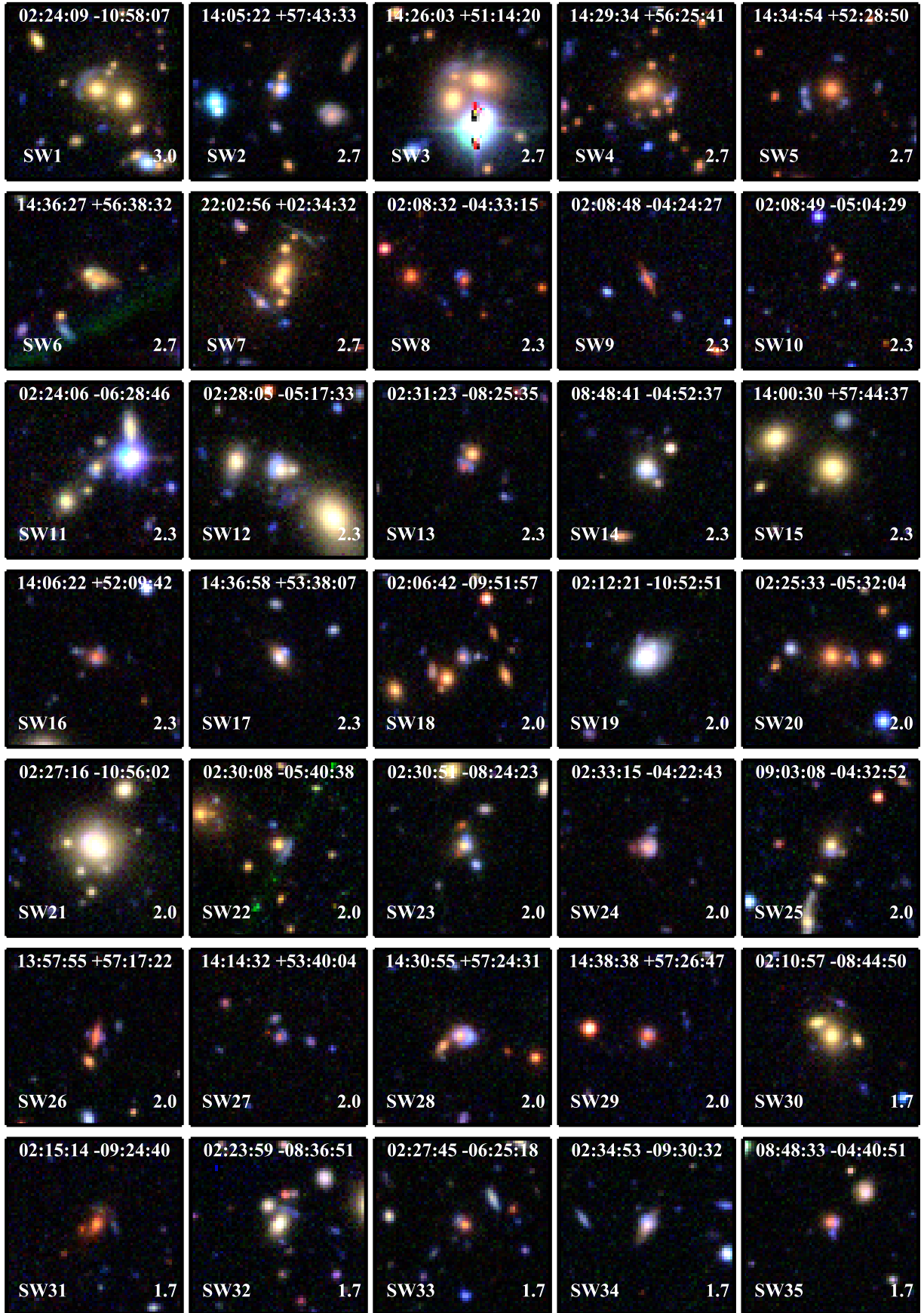
5.2 Measurements of lens and arc properties

In the subsequent sections, we make comparison of various properties of the lens candidates. Here, we describe how we measured those properties, namely, the lens redshift, the Einstein radii and the total flux of the arcs.

We use the publicly available redshifts for the lens galaxy from the CFHTLS photometric catalogs (Coupon et al. 2009). The definition of the Einstein radius is different in different cases. For the galaxy-scale lenses in the simulated sample, we use the value of the input lens model parameter for the R_E . For groups-scale lenses, since the lens model is multi-component, we need to determine the R_E from the image positions. We use half of the averaged image separations of the lensed image counterparts with minimum and maximum image separations. For the RINGFINDER sample, the arcs are detected in the scaled difference image between g and i bands where the lensing galaxy is subtracted (Gavazzi et al. 2014, for details, see). We use the peak position of the lensed images measured by running SEXTRACTOR on this difference image and identified visually. We calculate the image separation from the lens center as an estimate of the R_E . For the ARCFINDER (SARCS) sample, we use the same definition as above except that the peak position is identified either by the ARCFINDER or visually. For the SPACE WARPS lens sample, the same definition is used where the peak positions are identified either with ARCFINDER or SEXTRACTOR.

The total flux of the arc is measured in the g band. For the simulated sample, we multiply the magnification of the second brightest image with the source magnitude. For the RINGFINDER sample, we use the flux of the lensed images measured by SEXTRACTOR from the scaled difference image, that is, $g - \alpha i$ and convert it to the g band flux using mean colors of the foreground and background population. For the ARCFINDER and the SPACE WARPS sample, we integrate the flux in the image pixels identified by ARCFINDER or SEXTRACTOR.

⁷ <http://talk.spacewarps.org/>



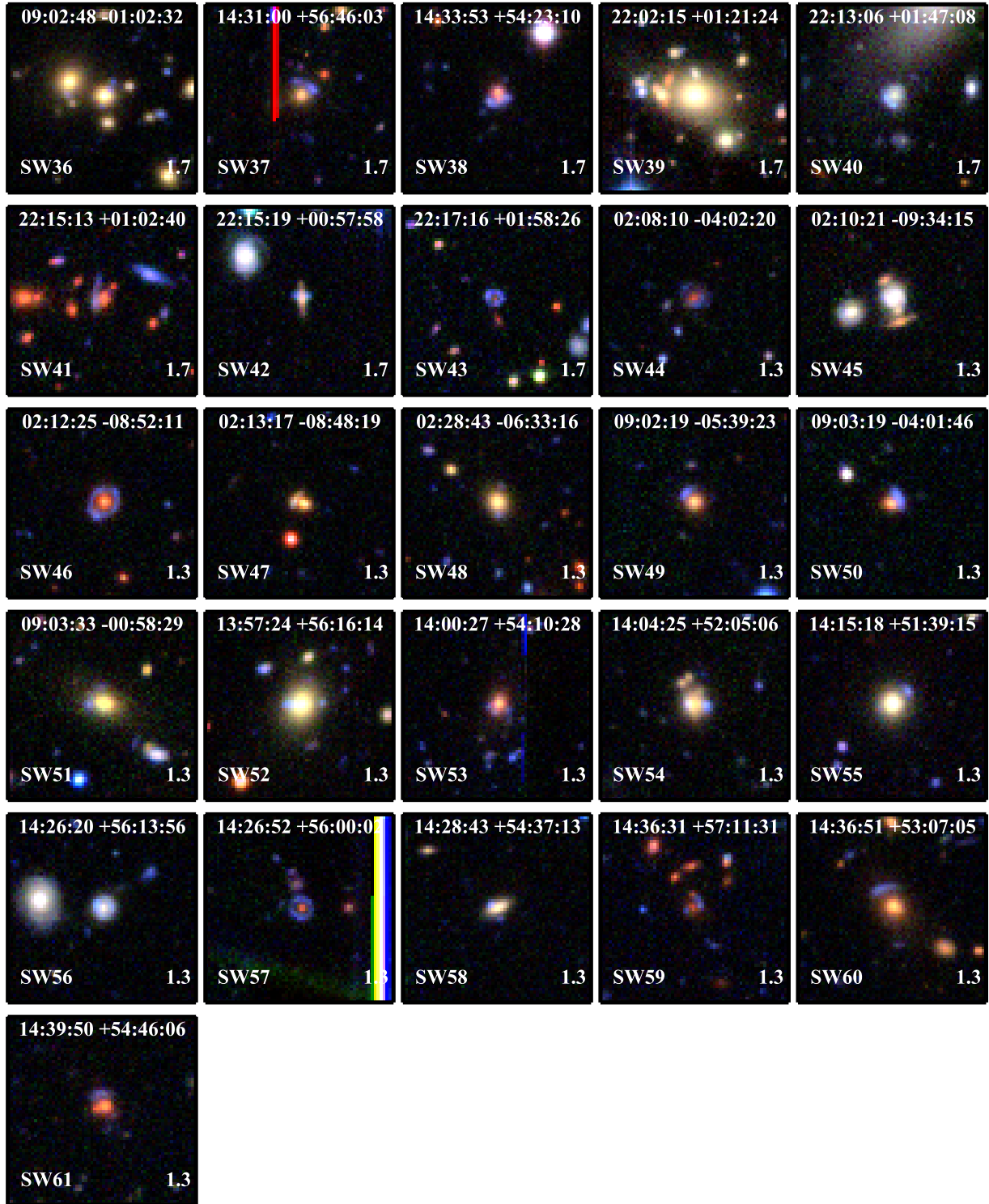


Figure 6. Sample of lens candidates with $G \geq 1.3$.

5.3 Recovery of known CFHTLS lenses with SPACE WARPS

We determine what fraction of the known sample of lenses are recovered by SPACE WARPS. In Table 2, we show that

$\sim 25\%$ - 30% of the known candidates and $\sim 60\%$ of the known confirmed lenses are found at stages I and II in SPACE WARPS. The left and the middle panels of Figure 7 show the fraction of detections as a function of arc magnitude and the Einstein radius of the lens systems.

As expected, we find that systems with brighter images and/or with larger Einstein radii are detected more often in SPACE WARPS.

We find that most of the confirmed lenses and candidates that are missed by SPACE WARPS are systems with fainter arcs and smaller Einstein radii and they come from the RINGFINDER sample. The main reason why RINGFINDER found such candidates is because their team used lensing galaxy-subtracted images to detect the presence of the lensed images both during the automated object-finding phase and during the visual inspection and classification of their candidates. This approach naturally improves the detection efficiency at smaller Einstein radii and for fainter systems. The SPACE WARPS volunteers were not shown any galaxy-subtracted images. However, in light of the improved detection efficiency, this might be a better strategy to adopt for future lens searches at galaxy-scales with SPACE WARPS. In the Discussion Section 6.2, we further explore and discuss why the confirmed lenses may have been missed in SPACE WARPS.

5.4 Image separation distribution

The distribution of image separations (i.e. twice the Einstein radius) can be used to probe the average density profile of the lens population (Oguri 2006; More et al. 2012). However, the lens sample found by the Arc Finder may have incompleteness as a function of the image separation. Thus, the lack of understanding of the selection function of the lens sample may affect the constraints on the density profile. A blind lens search done by visual inspection alone e.g. by SPACE WARPS may find lenses missed by the Arc Finder search and thereby, improve completeness. Furthermore, the completeness of the lens sample found by SPACE WARPS is characterised with the help of the simulated sample. For example, in the right most panel of Figure 7, we plot the completeness of the simulated sample found at stage I of SPACE WARPS as a function of the arc magnitude and Einstein radius. We find that the sample is more than 80% complete between image separations 2''-10''.

In Figure 10, we plot theoretical curves for the image separation distribution corresponding to three density profiles, namely, isothermal sphere (IS), NFW (Navarro et al. 1997) and Total profile which has NFW and Hernquist profiles combined with an adiabatically contracting model for dark matter component (Gnedin et al. 2004) taken from More et al. (2012).

6 DISCUSSION

Finding gravitational lenses is a difficult and complex task. No one method is perfect. Each method has some advantages over the other. It may be the case that a single method may not be the best means for optimising completeness and purity. Visual inspection will likely be required for pruning candidates at some stage of lens candidate selection even in the future. Therefore, we would like to understand how best we should combine the strengths of robots and humans to optimize the lens finding method.

In this section, we attempt to understand and com-

pare the lens (candidates) that are found by the lens finding robots and missed by humans and vice versa.

6.1 Why SPACE WARPS candidates were missed by lens finding robots?

We rerun the RINGFINDER and ARCFINDER on images centered on the new SPACE WARPS candidates to trace and understand at what stage the algorithm failed to detect them.

First, we rerun RINGFINDER on the new SPACE WARPS sample. At the beginning, a galaxy catalog is generated based on magnitude, redshift and SED type (Gavazzi et al. 2014, see) to select galaxies which are most likely to act as lenses. We find that about 40% of the new SPACE WARPS candidates fail to meet this initial selection criteria e.g. SW1, SW13, SW19, SW22, SW27 and SW30. All of the lensing galaxies are bright enough to satisfy the $i < 22$ criterion. However, some of them have a bright companion galaxy, some of them do not look like E/S0 type galaxies and some are edge on galaxies. **there are a few which look regular galaxies so which criteria are they failing??**

In the following steps, the flux from the galaxy is subtracted from the scaled difference image to enhance the visibility of the faint blue lensed features. An object finder is run on this image to quantify the lensed image properties. About 50% of the SPACE WARPS candidates could not be detected by the object finder because properties such as the image area, axis ratio, magnitude/color and alignment with respect to the lensing galaxy are not satisfied. Some of the candidates missed at this stage are e.g. SW4, SW5, SW6, SW26, SW36, SW39 and SW47.

Next, we rerun the ARCFINDER on the same SPACE WARPS sample of new candidates. The ARCFINDER is directly run on images to look for elongated arc like objects and does not require a list of targets to begin with. Objects are identified by placing thresholds on the noise level in the images. Thus, ARCFINDER detections are sensitive to changes in the noise levels.

Originally, the ARCFINDER was run on a large image with an area of $\sim 19350 \times 19350$ pixels². For the rerun, we work with much smaller images because this is faster but this alters the measured noise and hence, affects the number and type of arc detections. We find that about 30% of the new candidates are detected without changing any of the thresholds in the code because of the change in the noise level. **does this make sense?**

The ARCFINDER code calculates second order brightness moments around every pixel to decide if the distribution of flux is elongated in some direction in order to detect elongated arc-like objects. An elongation estimator is assigned to every pixel. All pixels with a value of the estimator above a certain threshold are connected to form the arc feature. This is building of the arc candidate. Subsequently, arc properties such as the area, mean flux, length and curvature are determined. We relax the threshold at the building stage and also relax thresholds mainly on the area of the arc which leads to detection of about 75% of the new SPACE WARPS candidates. We find that relaxing thresholds on other arc properties do not improve the detection rate significantly and thus are left unchanged.

Typically, the candidates are missed from the ARCFINDER sample if a) they are brighter in the redder band

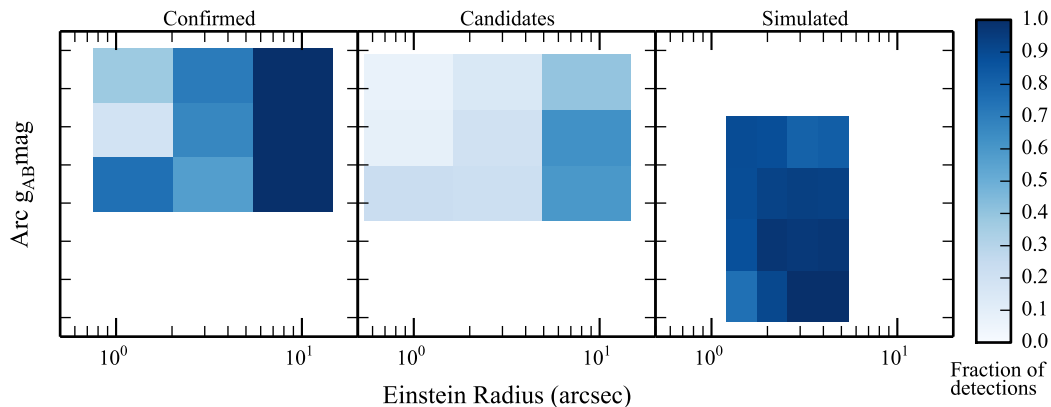


Figure 7. Fraction of lens candidates detected by SPACE WARPS as a function of the arc magnitude (g band) and the Einstein radius for three lens samples, namely, simulated lenses, known confirmed lenses and the lens candidates sample.

rather than the bluer g band used for arc detection because then the arc feature will be at the noise level and will fail detection b) the flux of the arc and the galaxy are blended in the g band such that the ARCFINDER mistakenly connects part of the galaxy to the pixels belonging to the arc candidate which then results in a candidate with no morphological characteristics of an arc or c) if the candidate is not elongated which means most of the lensed quasars with circular looking lensed images will be missed.

Relaxing the thresholds obviously increases the total number of candidate detections which includes a large sample of false positives. For example, the number of arc candidate detections increase by a factor of ~ 2 when we relaxed the thresholds in the reruns described above whereas the number of false positives increase by a factor of ~ 5 . Thus, it is not recommended to relax these thresholds alone. A better alternative is to relax the thresholds to increase the completeness and cross-correlate the arc candidate positions with a galaxy catalog to discard those candidates which are not close to a putative lensing galaxy within typical radius in order to reduce the false positives.

TBD: comment on any lensed quasars, on any exotic lenses; of course, there must be lenses which cannot be detected currently by either robots/humans check if some candidates were detected because they were hidden underneath the sims ie. from the D11;

6.2 False negatives: known lenses missed by SPACE WARPS

Like any lens finding method, the SPACE WARPS system can fail to detect certain kinds of lenses. Here, we explore and discuss why some of the known lenses may have been missed.

We find that about 40% of the known confirmed sample of lenses are missed (see Table 2). While this seems like a fraction, we note that many of the missed lenses are from the RINGFINDER sample with small Einstein radii and faint lensed images (see Figure 7). Among the confirmed lenses from the RINGFINDER, about 50% are missed. Out of the missed sample of 18 lenses, about half of them are visually difficult to detect and the other half appear to have faint

blue smudges around galaxies which should have been easier to identify. Similarly, if we consider the ARCFINDER lens sample, $\sim 23\%$ are missed by SPACE WARPS. This is a relatively small sample of ~ 5 systems and visual inspection suggests that, by and large, either the lensed features are faint or they have odd properties which makes them difficult to identify correctly. For further tests, we combine the RINGFINDER and ARCFINDER sample.

For a lens finding method which uses collective skill, experience and knowledge of a group of citizens, it may be difficult to find a single factor with certainty which causes a lens candidate to be missed. We attempt to understand whether there is indeed a single dominant factor that is resulting in the loss of these lenses or the lenses are being missed due to a combination of multiple reasons. Below, we consider some of the factors that could affect the efficiency of finding lenses.

First, we check if the Nclass is different for the detected and the missed sample. Surprisingly, most of the lenses in the known sample have few classifications ($N_{\text{class}} < 10$) which includes both the detected and missed lenses. Many of the remaining have between 10 to 20 classifications. And, only a few subjects have $N_{\text{class}} > 20$ which continue to remain for long in the database because these candidates may have been difficult to identify.

6.2.1 Blind lens search

Efficiency of a visual search can vary in different sections of an image. Our eyes tend to focus usually at the center of an image and lens candidates close to the borders could go undetected. Therefore, it is essential to test and understand if SPACE WARPS tends to miss certain lenses because they happen to be close to the borders of the image cutouts.

From the SWAP, the image cutouts inspected by the SPACE WARPS volunteers receive a status of detected (if $P > P_{\text{accthresh}}$), rejected (if $P < P_{\text{rejthresh}}$) and undecided (if $P_{\text{rejthresh}} < P < P_{\text{accthresh}}$). In Figure 11, we compare the positions of simulated lenses which are detected (red), undecided (green) and rejected (blue). We also make the same comparison for the known lens candidates in the right panel. We note that for some cases randomly selected subsamples

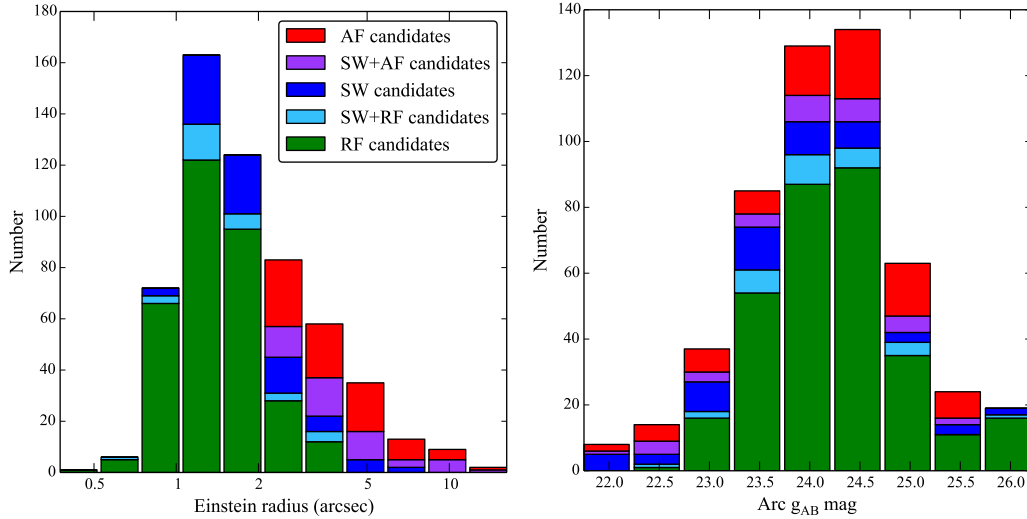


Figure 9. Candidate detections by the RINGFINDER, SPACE WARPS and the ARCFINDER as a function of the Einstein radius and g band magnitude of the lensed images. (XXXX check how different are the estimates from the different measurement techniques for a given lens candidates.)

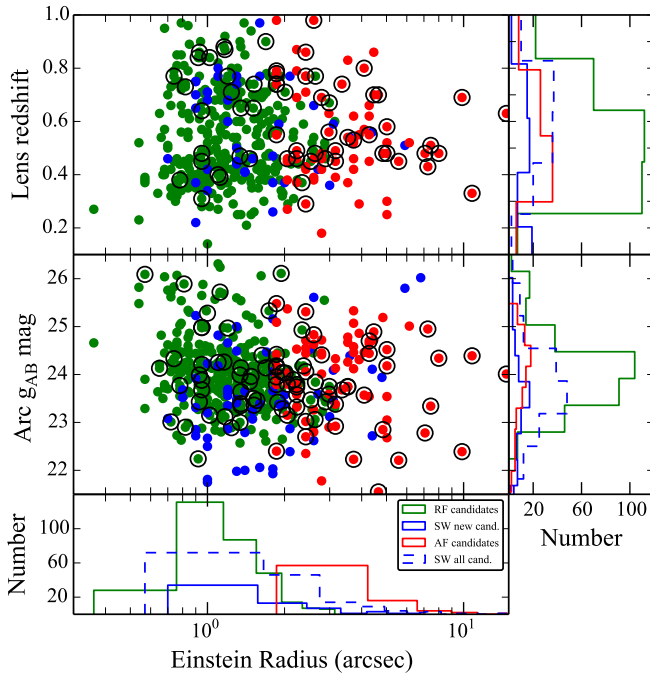


Figure 8. Comparison of the lens redshift and of the arc magnitude with the Einstein radius for all of the three lens samples, namely, from the RINGFINDER (green dots), SPACE WARPS (all candidates black circles and new candidates only in blue dots) and ARCFINDER (red dots).

are shown for the ease of visual comparison. For example, the absolute number of simulated lenses detected is too high compared to the rejected sample. We do not find any strong visual correlation in the rate of detections as a function of the position for both the simulated and the known lens sample. Thus, the completeness of the lens sample does not seem

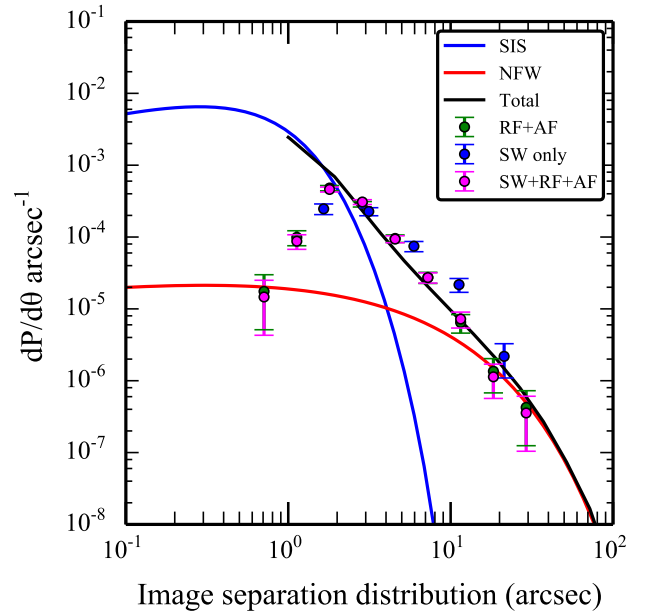


Figure 10. Image separation distribution. Comparing theoretical predictions (solid curves) with the CFHTLS known lens samples (REF, green points) and the same CFHTLS sample after combining with the incremental lens sample from SPACE WARPS (magenta points). The SPACE WARPS only sample including new and known lens candidates are shown with blue points.

to be significantly affected by whether a lens is located close to the border or well within the center.

6.2.2 Volunteer profile

Here, we investigate if the skill⁸ of volunteers is systematically different between the sample of detected and missed lenses.

We check how the posterior probability P (see XXXX Paper I) of an image or a subject to contain a lens changes as the image receives more classifications from multiple users. In Figure 12, we show the trajectory plots of a few examples of detected lenses (top left panels) and missed lenses (bottom left panels) by SPACE WARPS at stage I. The number of classifications (Nclass) for a subject increase from top to bottom. Also, every subject is assigned a prior probability $P_0 = 2 \times 10^{-4}$ (grey dashed line) and starts at the middle of the trajectory plot. The P value of a subject is updated with every classification from the volunteer. If a volunteer identifies a lens candidate, the trajectory moves to the right otherwise moves to the left. A subject is accepted if it crosses the blue-dashed line marking the ($P_{\text{accthresh}} = 0.95$) on the right and is rejected, if it crosses the red-dashed line marking the $P_{\text{rejthresh}} = 10^{-7}$ on the left.

By how much the P value will change depends on how well the volunteers are performing on the training sample. Thus, more skilled volunteers will change the P by a large factor compared to the less skilled volunteers. This is evident in the trajectory plots as large and small distances in between the consequent points which we refer to as kicks. Comparison of the kick sizes between the detected and the missed lenses suggests that the missed lenses do not have as many volunteers giving large kicks. We also note that most of the large kicks seen in the trajectories of the missed lenses seem to be moving the subjects to the right. In other words, the more skilled volunteers are mostly classifying them as subjects with lens candidates. Further qualitative comparison and discussion of the trajectory plots is given in Appendix A.

For a quantitative comparison of the large and small kicks for the entire samples of detected and missed known lenses, we show a plot of histogram on the right of Figure 12. Qualitatively, there are four types of volunteers making classifications - those causing positive large kicks (correct classifications by more skilled volunteers, those causing positive small kicks (correct classifications by less skilled volunteers), those causing negative small kicks (incorrect classifications by less skilled volunteers) and those causing negative large kicks (incorrect classifications by more skilled volunteers). The four histograms in the figure correspond to these four types of volunteers for each sample (that is, detected or missed). The kick size is small if $\Delta \log(P) = \log(P_{\text{current}}) - \log(P_{\text{previous}}) < \Delta \log(P)_{\text{cut}}$ (chosen to be 1.2) and is large if greater than $\Delta \log(P)_{\text{cut}}$.

Some of the key differences are i) the ratio of positive large kicks to positive small kicks for the detected sample is higher than the missed sample suggesting that the fraction of more skilled volunteers is large for the detected sample ii) in terms of Nclass, the missed sample is dominated by volunteers causing negative small kicks whereas for the detected sample there is a comparable contribution from three

types of volunteers - positive small and large kicks and negative small kicks. iii) the Nclass from volunteers with negative large kicks are lower both for the detected and the missed samples. this is consistent with our expectation that more skilled volunteers should not be making incorrect classifications. Therefore, lack of more skilled volunteers classifying the missed lenses seems to be one of the major factors.

We rerun SWAP for stage I where we use classifications from volunteers who produce $\Delta \log(P) > 1.2$ only. This obviously means reducing the total Nclass per subject by a large fraction. As a result, we also need to change the $P_{\text{accthresh}}$ which is chosen to be 0.1 and we find that about a third of the missed lenses are detected while all the previously detected lenses remain detected too. The others simply do not have enough classifications from volunteers producing large positive kicks. Thus, it may be possible to detect more promising candidates if we preferentially show the rejected systems to a few more skilled volunteers. We understand that changing the rejection and acceptance thresholds will likely increase the rate of FPs along with improved completeness but most probably the increase in FPs will not be too significant and that it will still be worth using the new strategy.

7 SUMMARY AND CONCLUSIONS

We report the discovery of new gravitational lens candidates from the the first lens search through SPACE WARPS. In this search, volunteers are shown g-r-i color images of random regions of the sky taken by the CFHT Legacy Survey. The aim of this lens search is to find lenses that have been missed by previous searches done on the CFHTLS with lens finding algorithms. In this search, we use a training sample to train the volunteers and calibrate their performance which helps us in more efficient pruning of the lens candidates. The training sample has simulated lenses, duds and FPs. In this paper, we describe the details of the training sample generated and tuned for the CFHTLS data. Volunteers receive feedback messages when they click or fail to click on the training images during live classification of real images which helps them remain focused and improve their performance over time.

The blind lens search in the CFHTLS is done in two stages: stage I and stage II. The aim of stage I is to do fast inspection of a large sample of $\sim 430\,000$ images and reduce them to a smaller sample of ~ 3000 images with interesting lens candidates. The aim of stage II is to carefully reinspect the smaller pool of images to exclude as many FPs as possible and reduce the sample to ~ 500 candidates. In the next step, the images are graded by three lens experts (AM, AV and PJM) leaving in systems that are mildly interesting to very promising lens candidates.

We draw the following conclusions:

- Crowd-sourced gravitational lens detection works. We present a new sample of 61 gravitational lens candidates with average expert grade of 1.3 and above where 1-mildly interesting, 2-probably, a lens and 3-certainly, a lens.
- Compared to the sample of robotically (the RINGFINDER and ARCFINDER) detected lens candidates, the SPACE WARPS sample finds lens systems with

⁸ see the Appendix B for the choice of this term and how it is different from that used in paper I

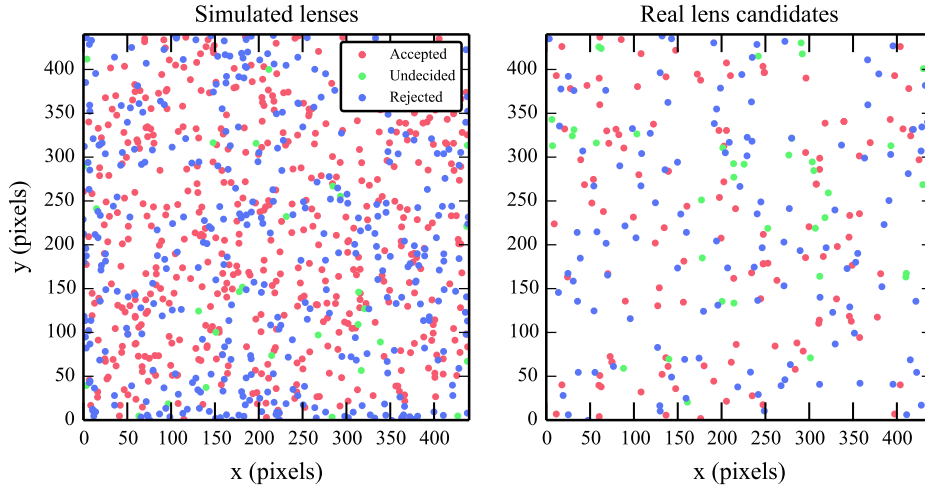


Figure 11. Completeness as a function of position of lens systems. Simulated lenses (left) and real lens candidates (right) are shown. Irrespective of the status of the lenses, that is, detected, undecided or rejected, there is no strong dependency on where the lenses are located both for the simulated and the real sample of lens candidates.

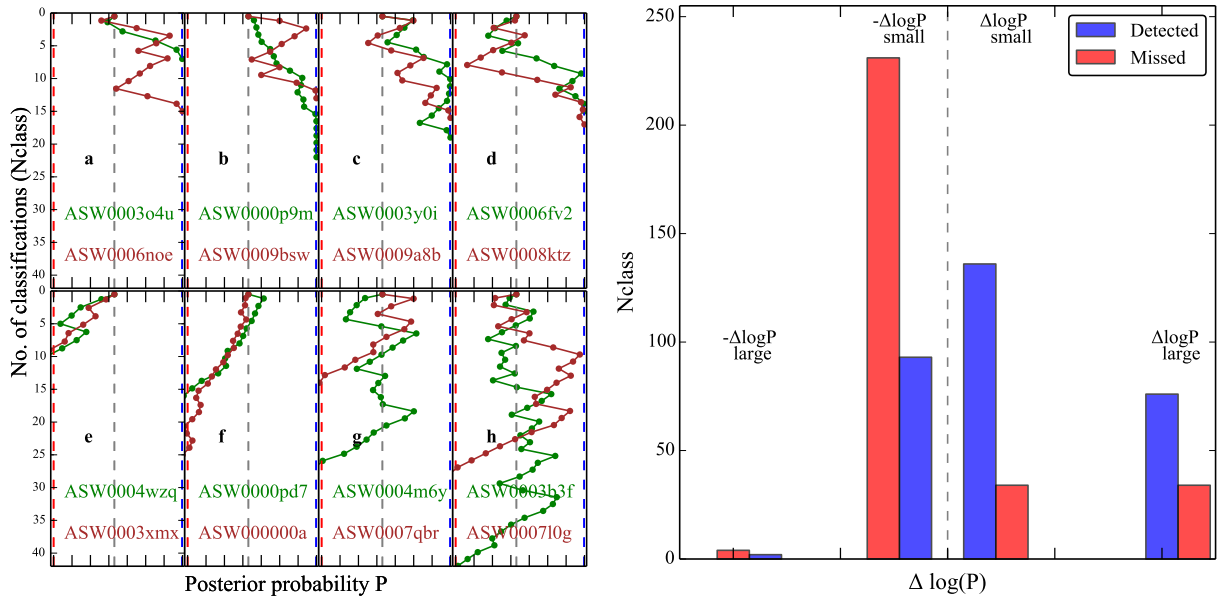


Figure 12. Examples of trajectories of some known lenses (top-left: detected and bottom-left: missed) at stage I and histogram of the known lenses for different types of volunteers (right).

similar range in the lens redshift and the flux and the Einstein radii of the lensed images. However, the SPACE WARPS search being a purely visual one cannot find some of the RINGFINDER-identified-lensed images with sub-arcsecond Einstein radii because the flux of the typically faint lensed images is obscured by the flux from bright lensing galaxies. The RINGFINDER finds these candidates because it works with lensing galaxy-subtracted images.

- Qualitatively, the SPACE WARPS sample finds lens systems with different types of lensing galaxies, for example, elliptical, spiral (face on and edge on) and small red galaxies unlike those found from robotic searches. Similarly, the lensed images too have different colors, morphologies and

sizes which are again typically missed by any given algorithm.

-

Table 3: Sample of the SPACE WARPS new lens candidates.

SW ID	Name	RA (deg)	Dec (deg)	z_{phot}	m_i (mag)	R_A (")	G	ZooID	P	Comments
SW1	CFHTLS J022409-105807	36.0398	-10.9688	0.0	0.0	4.8	3.0	ASW0004dv8	1.0	A,G
SW2	CFHTLS J140522+574333	211.3426	57.7259	0.7	19.7	1.0	2.7	ASW000619d	0.7	A,R
SW3	CFHTLS J142603+511420	216.5149	51.2390	0.0	0.0	4.4	2.7	ASW0006mea	0.7	A,G
SW4	CFHTLS J142934+562541	217.3926	56.4281	0.5	19.0	5.9	2.7	ASW0009cjs	0.8	A,G
SW5	CFHTLS J143454+522850	218.7270	52.4808	0.6	19.4	4.4	2.7	ASW0007k4r	0.4	Q,G/R
SW6	CFHTLS J143627+563832	219.1164	56.6425	0.5	19.4	1.5	2.7	ASW0008swn	0.9	A,D
SW7	CFHTLS J220256+023432	330.7369	2.5758	0.0	0.0	6.8	2.7	ASW0007e08	0.8	A,G/C
SW8	CFHTLS J020832-043315	32.1340	-4.5543	1.0	21.0	1.6	2.3	ASW0002asp	1.0	A,R
SW9	CFHTLS J020848-042427	32.2011	-4.4075	0.8	20.5	1.1	2.3	ASW0002bmc	0.9	D,D
SW10	CFHTLS J020849-050429	32.2078	-5.0749	0.8	20.6	0.9	2.3	ASW0002qtn	1.0	A,R
SW11	CFHTLS J022406-062846	36.0256	-6.4796	0.4	19.6	0.9	2.3	ASW0003wsu	0.7	A,E
SW12	CFHTLS J022805-051733	37.0236	-5.2927	0.4	18.8	1.4	2.3	ASW0009ans	1.0	Q,E
SW13	CFHTLS J023123-082535	37.8468	-8.4266	0.0	0.0	1.2	2.3	ASW0004xjk	0.3	A,R
SW14	CFHTLS J084841-045237	132.1708	-4.8772	0.3	19.0	1.0	2.3	ASW0004nan	1.0	A,E
SW15	CFHTLS J140030+574437	210.1260	57.7437	0.4	18.2	2.0	2.3	ASW0009bp2	0.6	A,E
SW16	CFHTLS J140622+520942	211.5958	52.1617	0.7	20.3	1.2	2.3	ASW0005rnb	0.7	A,R
SW17	CFHTLS J143658+533807	219.2425	53.6355	0.7	19.6	0.9	2.3	ASW0007hu2	0.6	D,D
SW18	CFHTLS J020642-095157	31.6750	-9.8658	0.2	20.8	0.9	2.0	ASW0001ld7	0.8	A,R
SW19	CFHTLS J021221-105251	33.0881	-10.8811	0.3	17.9	1.8	2.0	ASW0002dx7	0.8	D,E/S
SW20	CFHTLS J022533-053204	36.3888	-5.5346	0.5	19.4	3.6	2.0	ASW0004m3x	0.4	A,R/G
SW21	CFHTLS J022716-105602	36.8186	-10.9341	0.4	17.2	1.8	2.0	ASW0009ab8	0.7	A,E/G
SW22	CFHTLS J023008-054038	37.5359	-5.6774	0.6	19.7	1.9	2.0	ASW0003r61	0.5	A,E
SW23	CFHTLS J023051-082423	37.7141	-8.4064	0.0	0.0	0.8	2.0	ASW000412m	0.4	A,E
SW24	CFHTLS J023315-042243	38.3133	-4.3789	0.7	19.7	1.0	2.0	ASW00050sk	0.8	A,R
SW25	CFHTLS J090308-043252	135.7844	-4.5479	0.0	0.0	1.2	2.0	ASW00007mq	0.6	A,E
SW26	CFHTLS J135755+571722	209.4827	57.2897	0.8	20.2	1.3	2.0	ASW0005ma2	0.8	D,D
SW27	CFHTLS J141432+534004	213.6372	53.6679	0.7	21.4	0.9	2.0	ASW0006jh5	0.8	A,R
SW28	CFHTLS J143055+572431	217.7333	57.4088	0.7	19.3	1.0	2.0	ASW0007wfi	0.9	A,R
SW29	CFHTLS J143838+572647	219.6589	57.4464	0.8	20.2	1.1	2.0	ASW0008qsm	0.9	A,R
SW30	CFHTLS J021057-084450	32.7414	-8.7474	0.0	0.0	2.5	1.7	ASW0002p8y	0.4	A,G
SW31	CFHTLS J021514-092440	33.8109	-9.4111	0.7	19.9	2.6	1.7	ASW00021r0	0.4	A,R/G
SW32	CFHTLS J022359-083651	35.9995	-8.6143	0.0	0.0	3.1	1.7	ASW0004iye	0.4	A,E
SW33	CFHTLS J022745-062518	36.9387	-6.4218	0.6	20.5	1.2	1.7	ASW0003s0m	0.5	A,R
SW34	CFHTLS J023453-093032	38.7232	-9.5089	0.5	19.8	0.7	1.7	ASW00051ld	0.3	A,D
SW35	CFHTLS J084833-044051	132.1385	-4.6809	0.8	20.2	0.9	1.7	ASW0004wgd	0.7	A,R
SW36	CFHTLS J090248-010232	135.7020	-1.0424	0.4	19.1	1.4	1.7	ASW000096t	0.6	D,E
SW37	CFHTLS J143100+564603	217.7511	56.7675	0.0	0.0	1.8	1.7	ASW00086xq	0.8	A,E
SW38	CFHTLS J143353+542310	218.4736	54.3862	0.8	19.8	1.6	1.7	ASW0009cox	0.6	A,R/G
SW39	CFHTLS J220215+012124	330.5635	1.3567	0.3	17.4	4.6	1.7	ASW0005qiz	0.5	rA,G
SW40	CFHTLS J221306+014708	333.2758	1.7856	0.0	17.1	1.4	1.7	ASW0008wmr	0.9	A,S
SW41	CFHTLS J221513+010240	333.8056	1.0446	0.0	0.0	0.8	1.7	ASW0008dxh	0.3	A,R/G
SW42	CFHTLS J221519+005758	333.8321	0.9661	0.4	20.2	1.0	1.7	ASW0008xbu	0.8	A,D
SW43	CFHTLS J221716+015826	334.3189	1.9739	0.1	21.6	1.0	1.7	ASW00096rm	1.0	A/R,R
SW44	CFHTLS J020810-040220	32.0450	-4.0389	1.0	20.8	1.8	1.3	ASW0001c3j	0.7	A,R
SW45	CFHTLS J021021-093415	32.5898	-9.5711	0.4	18.4	2.7	1.3	ASW0002k40	0.4	D,S
SW46	CFHTLS J021225-085211	33.1051	-8.8697	0.8	19.5	2.1	1.3	ASW00024id	1.0	R,R
SW47	CFHTLS J021317-084819	33.3234	-8.8055	0.5	19.8	1.3	1.3	ASW00024q6	0.4	A,R/E
SW48	CFHTLS J022843-063316	37.1794	-6.5547	0.5	19.1	1.8	1.3	ASW0003r6c	0.3	D/A,E
SW49	CFHTLS J090219-053923	135.5794	-5.6566	0.0	0.0	2.0	1.3	ASW0000g95	1.0	A,R/E
SW40	CFHTLS J090319-040146	135.8311	-4.0297	0.0	19.8	1.2	1.3	ASW00007ls	0.5	A,R/E
SW51	CFHTLS J090333-005829	135.8886	-0.9749	0.0	0.0	2.1	1.3	ASW00008a0	1.0	A/D,E/G
SW52	CFHTLS J135724+561614	209.3536	56.2707	0.0	0.0	2.6	1.3	ASW0006e0o	0.9	D,E
SW53	CFHTLS J140027+541028	210.1164	54.1746	0.0	0.0	1.2	1.3	ASW0006a07	0.6	Q,R/E
SW54	CFHTLS J140425+520506	211.1062	52.0850	0.4	18.9	1.4	1.3	ASW0005o0w	0.6	D,E
SW55	CFHTLS J141518+513915	213.8290	51.6542	0.4	18.3	3.0	1.3	ASW00070vl	0.8	D,E
SW56	CFHTLS J142620+561356	216.5870	56.2323	0.5	19.5	1.3	1.3	ASW0007sez	0.8	A/R,S
SW57	CFHTLS J142652+560002	216.7201	56.0006	0.0	0.0	1.5	1.3	ASW0007t5y	1.0	R,R
SW58	CFHTLS J142843+543713	217.1815	54.6204	0.4	19.7	1.3	1.3	ASW0007pga	0.6	D,D

SW ID	Name	RA (deg)	Dec (deg)	z_{phot}	m_i (mag)	R_A (")	G	ZooID	P	Comments
SW59	CFHTLS J143631+571131	219.1315	57.1922	0.7	20.9	1.3	1.3	ASW0008pag	0.6	D/A,R
SW60	CFHTLS J143651+530705	219.2150	53.1183	0.6	19.2	3.1	1.3	ASW0007h27	1.0	A,E/G
SW61	CFHTLS J143950+544606	219.9609	54.7686	0.0	0.0	1.7	1.3	ASW00085cp	0.4	A,G/R

The column Comments has two type of notes. The first is about the lens image configuration where the symbols mean the following A: Arc, D: Double, Q: Quad, R: Ring. The second is a comment on the type of lens assessed visually. Note that this classification is not based on colors or spectral analysis. The symbols are E: Elliptical, S: (face on) Spiral, G: Group-scale, D: Edge on disk, R: Red starforming galaxy. This galaxy falls within the masked region as per the catalog from which the magnitudes and the redshift are extracted.

ACKNOWLEDGEMENTS

We thank all XXXmembers of the SPACE WARPS community for their contributions to the project so far. A complete list of collaborators is given at... In particular we would like to recognise the efforts of XXX, YYY etc in moderating the discussion.

We are also grateful to Brooke Simmons, David Hogg, XXX and YYY for many useful conversations about citizen science and gravitational lens detection. PJM was given support by the Royal Society, in the form of a research fellowship, and by the U.S. Department of Energy under contract number DE-AC02-76SF00515. AV acknowledges support from the Leverhulme Trust in the form of a research fellowship. The work of AM and SM was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. The work of AM was also supported in part by National Science Foundation Grant No. PHYS-1066293 and the hospitality of the Aspen Center for Physics. The SPACE WARPS project is open source. The web app was developed at <https://github.com/Zooniverse/Lens-Zoo> while the SWAP analysis software was developed at <https://github.com/drphilmarshall/SpaceWarps>. This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program.

APPENDIX A: QUALITATIVE COMPARISON OF DETECTED AND MISSED KNOWN LENSES

APPENDIX B: SKILL

In paper I, the term skill is given by the expectation value of the information gain which is related to the P_L , the probability to correctly identify a lens as a lens and P_D , the probability to correctly identify a dud as a dud. The skill does not have simple dependency on P_L and P_D . One can get the same value of skill for different combinations of P_L and P_D .

As described in the appendix of paper I, the posterior probability P is related to the volunteers skill (or information gain). Also, a volunteer with a given skill, will have different positive ($\Delta \log(P)$) and negative ($-\Delta \log(P)$) kicks on the trajectory plot. Thus, it is not useful to compare classifications of high and low skill volunteers to understand if there is a difference in their performance and if the high skill volunteers can increase the completeness of the lenses. For example, a volunteer who is very good at rejecting duds and not so good at identifying lenses will still have a high skill. But such a volunteer will fail to detect many of the interesting lenses. Therefore, we do not compare high skill vs. low skill volunteers to address this issue.

For lack of a better word, we also use the same term skill which has a similar meaning to that described in paper I but not identical. We use a more qualitative definition implying the ability of the volunteers to make correct lens identification.

XXXXX

REFERENCES

- Bernardi, M., et al. 2003, *AJ*, 125, 1866
 Chan, J. H. H., Suyu, S. H., Chiueh, T., More, A., Marshall, P. J., Coupon, J., Oguri, M., & Price, P. 2014, *ArXiv e-prints*
 Coupon, J., et al. 2009, *A&A*, 500, 981
 Erben, T., et al. 2013, *MNRAS*, 433, 2545
 Faber, S. M., et al. 2007, *ApJ*, 665, 265
 Faure, C., et al. 2009, *ApJ*, 695, 1233
 Frei, Z., & Gunn, J. E. 1994, *AJ*, 108, 1476
 Gavazzi, R., Marshall, P. J., Treu, T., & Sonnenfeld, A. 2014, *ApJ*, 785, 144
 Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, *ApJ*, 616, 16
 Hildebrandt, H., et al. 2012, *MNRAS*, 421, 2355
 Keeton, C. R., Christlein, D., & Zabludoff, A. I. 2000, *ApJ*, 545, 129
 Lintott, C. J., et al. 2008, *MNRAS*, 389, 1179
 Lupton, R., Blanton, M. R., Fekete, G., Hogg, D. W., O'Mullane, W., Szalay, A., & Wherry, N. 2004, *PASP*, 116, 133
 More, A., Cabanac, R., More, S., Alard, C., Limousin, M., Kneib, J.-P., Gavazzi, R., & Motta, V. 2012, *ApJ*, 749, 38
 Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
 Oguri, M. 2006, *MNRAS*, 367, 1241
 Oguri, M., & Marshall, P. J. 2010, *MNRAS*, 405, 2579
 Pâris, I., et al. 2012, *A&A*, 548, A66
 Parker, L. C., Hudson, M. J., Carlberg, R. G., & Hoekstra, H. 2005, *ApJ*, 634, 806
 Richards, G. T., et al. 2006, *AJ*, 131, 2766
 Schwamb, M. E., et al. 2012, *ApJ*, 754, 129
 Sygnet, J. F., Tu, H., Fort, B., & Gavazzi, R. 2010, *A&A*, 517, A25

This paper has been typeset from a \LaTeX file prepared by the author.