INITIAL ANALYSIS OF CROWDSOURCED MORPHOLOGICAL CLASSIFICATIONS OF DECALS IMAGES

K.W. WILLETT[1] ET AL.
[1]University of Minnesota
Draft version March 3, 2016

ABSTRACT

Abstract.
Keywords: keywords

1. SCIENCE CASE

The morphology of galaxies has long been a critical parameter in studying their formation and evolution over cosmic time. This includes basic divisions between early-type, late-type, and merging galaxies as well as more detailed measurements of bulge/disk ratio, spiral arm angle and orientation, and fainter features such as tidal tails or dust lanes. The former parameters measure the large scale integrated dynamical history of the galaxy, as well as its history of star formation (Schawinski et al. 2014), while the latter categories typically probe secular processes acting on longer timescales.

Automated measurements of galaxy morphology, however, are not yet sufficiently accurate to be used on large scale surveys, especially for faint targets and features with small angular sizes. The Galaxy Zoo project has provided reliable morphologies for several large surveys (including SDSS, UKIDSS, COSMOS, and CANDELS) by leveraging crowdsourced visual classifications into quantitative probabilities. This technique shows each image to dozens of people and seeks consensus answers, using our knowledge of user reliability to weight their responses, and has been proven (Lintott et al. 2011; Willett et al. 2013) to be as reliable (and in many cases better) that state-of-the art automated methods. By using the DECaLS images in the existing Galaxy Zoo platform, the team will provide morphologies that complement the other globally measured host galaxy properties such as colour and luminosity. Galaxy Zoo has a strong track record in publication from Galaxy Zoo morphologies, with 50 publications which collectively have ~ 2000 citations. Some of our science highlights include the discovery of new types of galaxies (e.g. green peas; Cardamone et al. 2009), as well as the investigation of red spirals (Bamford et al. 2009; Masters et al. 2010) and blue ellipticals (Schawinski et al. 2009).

The advantage of the DECaLS images over existing GZ measurements is primarily driven by two aspects: namely, the deeper imaging (especially in g- and z-bands, but also r-band), as well as improved spatial resolution compared to the SDSS which has provided the bulk of the low-redshift GZ images to date.

In addition, the Dark Energy Spectroscopic Instrument (DESI; Levi et al. 2013) is planning to use DECaLS for targets as part of major new galaxy redshift surveys covert 14000 deg^2 of the northern sky. The low redshift part of this is the Bright Galaxy Sample (BGS) which will use bright time to observe 10 million nearby galaxies (median redshift of $z = 0.2$) from a simple magnitude-limited sample of $r < 19.5$. Galaxy Zoo morphologies of BGS galaxies will provide several scientific opportunities, especially studying the clustering (real and redshift space) of galaxies as a function of their detailed morphologies. Such measurements will be significant improvements on existing SDSS-based GZ measurements (see Skibba et al. 2009, 2012) as it will provide greater statistics, larger scales and finer morphologies (e.g. better bar and bulge classifications). Moreover, morphologies may offer a new and unique way of helping to characterise the halo occupation statistics of galaxies thus assisting the detailed modeling of the DESI galaxies as a function of galaxy properties e.g., bright, massive, ellipticals are usually central halo galaxies. Moreover, morphologies may offer a new and unique way to help characterize the halo occupation statistics of galaxies, thus assisting the detailed modelling of the DESI galaxies as a function of galaxy properties, since bright massive ellipticals are usually central halo galaxies. In addition, they will enable analyses of the morphology-halo mass relation of galaxies with better precision than previous surveys.

We list below four example science cases that can be addressed using the combination of DECaLS imaging/metadata and GZ crowdsourced morphologies.

1. **Statistical studies of galaxy evolution with morphology** *(Lead GZ scientists: Lintott, Masters, Skibba, Schawinski, Willett).* The complexity of galaxy evolution, along with the development of large surveys, like SDSS has driven forward the technique of statistical galaxy evolution in recent years, revolutionising our understanding of the galaxy population, and revealing the complex interdependence of galaxy properties such as mass, environment, dynamics, morphology and their star formation histories. The multi-dimensional correlations common in galaxy evolution (e.g. between stellar mass, colour, environment and morphology) makes larger samples vital to disentangle the primary variables. Galaxy Zoo has enable the addition of quantitative visual morphology to these techniques, and the involvement of a large number of Galaxy Zoo team members in this project underscores how this covers the core GZ:DECaLS science). Our previous work with GZ:SDSS has found that samples of $N \simeq 10^4$ may be needed for even a detection of some effects (Skibba et al. 2012). GZ:DECaLS data (especially when combined with GZ:DES, a proposal for which is under review within DES) will reduce the statistical uncertainty in such measurements. Questions we will address are:
   - How does star formation history correlate with morphology as a function of stellar mass and environment? Schawinski will provide significant expertise on constraining galaxy star formation histories and their links to morphology, also see Smethurst et al. (2015).

- Are bars triggered or destroyed by galaxy interactions in different environments? Follows on previous work by Skibba, Masters to greater statistical significance (Skibba et al. 2012; Masters et al. 2011, 2012).
- What is the contribution of bar dynamics to the feeding of AGN? e.g. an extension of Galloway et al. (2015) to greater statistical significance.

2. **Discovery of rare objects** *(Lead GZ Scientists: Keel, Maksym, Simmons and Lintott).* The Galaxy Zoo experience with the SDSS data showed the value of having classifiers able to note objects which fell outside the normal range of galaxy structure, both as regards color and morphology; the greater surface brightness sensitivity, in particular, the DECaLS data promises to similarly reward broad inspection. Galaxy Zoo is perhaps best well known by the public for its discovery of a variety of unusual objects with very low spatial densities. The most famous is "Hanny's Voorwerp" (an AGN-ionized gas cloud) (Lintott et al. 2009), but GZ has also discovered other classes of rare objects, including smaller versions of the Voorwerp (Keel et al. 2012), the "green peas" (compact star forming galaxies; Cardamone et al. 2009), bulgeless galaxies with AGN (Simmons et al. 2013), and overlapping galaxies for studying dust content (Keel et al. 2013). We expect that new classes of objects that may be revealed by deeper imaging, and intend to not only use catalogues on known galaxies in GZ:DECaLS, but also include Tractor identification of extended objects which are not otherwise catalogued.

3. **Galaxy morphology at low luminosities** *(Lead GZ scientists: Bamford and Willett).* Visual morphologies for large, low-redshift samples have so far been limited by the depth and resolution of SDSS and UKIDSS. Studies that have pushed down to low luminosities have therefore been restricted to small volumes and correspondingly small samples. For example, the GAMA survey study by Kelvin et al. (2014) studies only 3727 galaxies down to $M_r < -17.4$ mag, as it is limited to redshifts below 0.06. GAMA will soon gain greater depth with VST-KiDS imaging, and has complete spectroscopic coverage, but covers only 250 deg^2. DECaLS will reach similar depths, but over a much larger area. GZ:DECaLS data will therefore greatly improve upon both the quality and number of classifications at fainter luminosities. Combining results from these complementary surveys will enable us to make a definitive morphological census of the "faint end of the Hubble sequence", where the galaxy population is dominated by very late-type disks, irregulars and dwarf spheroidals.

4. **Exploring the role of minor mergers** *(Lead GZ Scientist: Kaviraj).* Minor merging is a fundamental process that is thought to drive around half of the local star formation budget ($\sim 40\%$ in local spirals) (Kaviraj 2014). However, this process is poorly explored because minor mergers produce only faint tidal features that are invisible in typical imaging surveys (e.g. SDSS). We will use the deep DECaLS images to select minor merger remnants and study this process in detail in the local and intermediate redshift Universe ($z < 0.5$). Minor merger remnants will be selected by identifying spiral galaxies (i.e. galaxies which cannot have had a major merger) that are morphologically disturbed (so must have had a minor merger). This sample will be used to answer the following questions:

    - What is the minor merger rate at $z = 0$ and does it show an evolution to intermediate redshift?
    - Do minor mergers trigger AGN activity? What role do they play in the low and high excitation modes in active galaxies?
    - How does the minor merger rate depend on stellar mass and environment?
    - What is the star formation enhancement due to minor merging in the host galaxies? Does this enhancement show any evolution with redshift?
    - What fraction of the star formation budget is driven by minor merging in the nearby Universe?

    The methodology for this project has already been developed in Kaviraj (2014) who performed an identical study using Stripe 82 (see Figure 1 in that paper for examples of deep images that will allow us to select minor merger remnants).

2. SELECTION CRITERIA

Galaxy Zoo 2 — main spectroscopic sample

- Galaxy in MGS or Stripe 82
- spectroscopic redshift available
- $0.0005 < z_{spec} < 0.25$
- $m_r < 17.0$
- `petro90_r` $> 3''$
- flag is not SATURATED, BRIGHT, or BLENDED

Galaxy Zoo 2 — Stripe 82 coadd

- version 1 of coadded data
- same as MGS with exception of $m_r < 17.77$

Dark Energy Camera Legacy Survey (DECaLS)

- Galaxies in NASA-Sloan (Extended) Atlas
- Good-quality measurements in g, r, z bands as of Jun 2015
- Note: attempted to do size cutoffs using the Tractor catalog data for $\theta > 10''$, but couldn't find robust matches for galaxies despite searching in multiple bricks.

So the Sloan and DECaLS samples definitely were not selected in the same manner. However, all the relevant parameters are stored in the NASA-Sloan Atlas if we wanted to cut on that. More specifically, we can do a simple match (ideally through previously matched catalogs, but if necessary through a tight positional match) to directly compare morphological measurements for *the same galaxies* (Table 2).

DECaLS images are selected from the extended NASA-Sloan Atlas (v1.0.0; M. Blanton, priv. comm.). The JPG images are sized at 424×424 pixels; the pixel scale is set using both the 50% R_{50} and 90% R_{90} r-band Petrosian radii in the NSA catalog. The pixel scale used is the smaller of ($R_{50} * 0.04$) or ($R_{90} * 0.02$) arcsec/pixel; an absolute minimum of 0.10 arcsec/pixel is imposed to prevent images from being zoomed-in to the point where obvious pixellation is an issue.

Note: while we had originally planned to select only galaxies of $r > 10''$, the Tractor measurements weren't reliable and KWW ultimately didn't apply the cuts to the NSA catalog. As a result, there is no minumum angular size for the GZ-DECaLS DR1 sample. This means that a few galaxies down to $r_{Petro} = 0$. (in the catalog) are included, and a significant number with $r_{Petro} < 2''$. The original limit on GZ2 with SDSS imaging was $r_{Petro} > 3''$; the seeing at CTIO should be an improvement, but there's still no real reason to visually classifiy galaxies that aren't at least several resolution elements across. Visual inspection of the DR1 images show that the majority of stellar artifacts in the sample have $r_{Petro} = 1''.02072$ (this apparently gives the DECaLS PSF, and these images should clearly be removed). Very few galaxies with $r_{Petro} > 3''$ have visible disc features. There is also a strong correlation for the overall classification as a function of angular size. By plurality vote, 56% of galaxies with $r_{Petro} > 10''$ are feature/disk, while only 33% of galaxies with $5'' < r_{Petro} < 10''$ and 7% of galaxies with $r_{Petro} < 5''$ are feature/disk.

This means: a) analysis of DECaLS galaxies should put a lower limit on Petrosian radius for reliable sample selection, likely at least $3''$, and b) the same limit should be applied to sample selection for GZ-DECaLS DR2 and beyond. We really don't want to waste users' time by putting in galaxies too small for reliable classification.

The images are created using the g-, r-, and z-band exposures. RGB images are created using a modified arcsinh stretch (Lupton et al. 2004) for each band:

$$f[x|\text{band}] = \frac{\text{asinh}(xQ/s_{\text{band}})}{\sqrt{Q}}, \tag{1}$$

where x is the raw pixel value in each band. We adopt a value of $Q = 1.0$ and scaling factors of [$s_g = 0.0066$, $s_r = 0.01385$, $s_z = 0.025$]. The resulting RGB values are then clipped at extremes and mildly smoothed in order to "desaturate" outlier pixels which can strongly skew the color gradient of the entire image. Scaling values were chosen by visually examining images made with a range of Q, s and choosing those which emphasized low-surface brightness features while balancing contrast of brighter regions and color balance. The same stretch is applied to every image in the sample.

3. RESULTS

3.1. *Project design*

GZ-DECaLS uses a hierarchical decision tree (Figure 2) to assign morphological classifications, similar to all previous iterations of Galaxy Zoo. The design of the tree and morphological categories are similar to those in Galaxy Zoo 2 (Willett et al. 2013), but with several key differences. GZ-DECaLS reduces the options for the task measuring bulge prominence in disk galaxies from four to three, eliminating the "just noticeable" option. This is justified based on GZ2 data showing that the "dominant" option in the original GZ2 tree was almost never selected; only eight galaxies in the entire main spectroscopic sample ($< 0.0003\%$) met the `clean` threshold 80% for this question. An odd number of responses also eliminates the requirement for users to implicitly decide whether the bulge strength is objectively weak or strong (Clark & Watson 1995), increasing the discriminating power of the lowest and highest responses.

GZ-DECaLS also eliminated the "can't tell" option for the number of spiral arms, requiring users instead to give their best estimate on the arm multiplicity. Hart et al. (in prep) show that requiring users to select a discrete arm multiplicity increased the overall confidence without adding to the noise.

For the task listing any "odd features" that users see in the image, GZ-DECaLS allows users to select more than one question (since, for example, a galaxy might have both a dust lane and a merger). GZ2 required users to select only a single option, leading to much lower levels of consensus for answers that are not necessarily mutually exclusive. GZ-DECaLS eliminated the options for "disturbed" (dominated by pixel trails in GZ2; Willett et al. 2013) and added on option to explicitly identify galaxies with line-of-sight overlaps that are not physically merging. Such overlapping galaxies were annotated and separately collected by users in the project Forum for Keel et al. (2013, 2014), but are now an explicit part of the interface.

Finally, GZ-DECaLS adds a question for explicitly identifying the effect of mergers and tidal debris. This question is answered for all images not labeled as "star or artifact"; users are instructed to use separate labels for merging galaxies (where the nuclei are still physically distinct) and galaxies with tidal debris, the effects of which can range from spectacular tails hundreds of kpc in length to very faint irregularities or extensions of the galaxy's light profile (Kaviraj 2010; Kaviraj et al. 2012). A similar question was first introduced for the GZ:CANDELS project, but this is the first case in Galaxy Zoo where it is applied to low-redshift galaxies.

3.2. *Statistics*

There were a total of 32,429 images classified in DR1 of DECaLS. The final data set consists of 1,290,465 classifications by 15,990 registered users. The most prolific single user classified 20,272 galaxies (63% of the sample).

The main period of classifications for DR1 ran from 21 Sep 2015 to 15 Feb 2016; images were retired once they had been classified 40 times apiece. After reaching this limit for the full sample, the retirement threshold was raised to 60 in order to improve statistics on disk/featured galaxies and provide a bridge to the DR2 phase.

Vote fractions were collated and combined using equal weighting for each users. A more advanced weighting scheme based on group agreement is intended before the final data release.

4. MATCHING DECALS AND SDSS

There is some overlap between the existing morphological classifications from Galaxy Zoo 2 (Willett et al. 2013). Most of the difference comes from the fact that SDSS was located at Apache Point Observatory in the Northern Hemisphere (latitude 32.780278°), while the DECaLS camera is mounted on the Blanco 4-m telescope at CTIO in the Southern Hemisphere (−30.169661°). They can cover a significant fraction of the same portion of the sky, but DECaLS will be limited at high northerly latitudes. Figure 1 shows the overlap between the NASA-Sloan ATLAS[1] (derived from SDSS) and

[1] http://www.nsatlas.org/

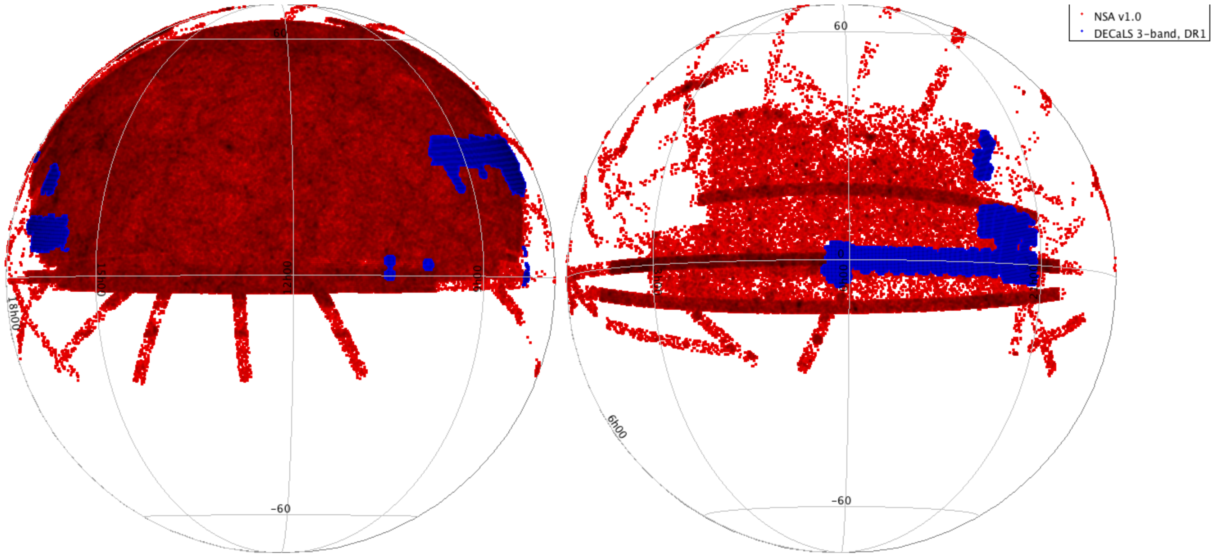

Figure 1. Overlap between galaxies in the NASA-Sloan Atlas (red) and selected targets for Galaxy Zoo from DECaLS DR1 (blue).

DECaLS DR1.

Morphologies for this analysis are taken from the published GZ2 tables in Willett et al. (2013) for SDSS. The DECaLS morphologies have been collated, but not systematically debiased to account for changes in morphological fraction as a function of apparent size and brightness. Therefore, we only compare the *weighted vote fractions* ($f_{w,morph}$) from GZ2 to the *raw vote fractions* ($f_{r,morph}$) from DECaLS.

In the GZ2 main spectroscopic sample (243,500 galaxies), we matched galaxies within a $3''.0$ radius and find 9,281 subjects appearing in both catalogs. We match a further 5,800 subjects using the same radius for the Stripe 82 data. There is overlap of 2,814 bright Stripe 82 galaxies that are included in both. The unique total of 12,267 subjects is only 38% of the catalog, despite the fact that the original SDSS Legacy sky coverage (Strauss et al. 2002) overlaps with all of the current DECaLS bricks.

Part of the mismatch comes from the limited spatial coverage of the Stripe 82 region in SDSS, which only covered a declination range of $-1.26° < \delta < +1.26°$ (Annis et al. 2014). The DECaLS imaging bricks have NSA targets in a larger area, extending between roughly $-2.5° < \delta < +2.5°$. These are presumbly targets imaged in SDSS DR8 or later, since otherwise they would have been included in the original GZ2 selection.

However, there are many DECaLS galaxies in the imaging area covered by the main Legacy survey. Galaxies in DECaLS but *not* GZ2 have $\langle m_r \rangle = 17.27$ mag, $\langle r_{petro} \rangle = 6''.64$, $\langle z \rangle = 0.093$. Galaxies in *both* DECaLS and GZ2 are on average brighter ($\langle m_r \rangle = 16.29$ mag), larger ($\langle r_{petro} \rangle = 7''.99$), and lower-redshift ($\langle z \rangle = 0.080$). The vast majority of the DECaLS images with no GZ2 counterpart are galaxies with $17.0 < m_r < 17.77$ — the fainter magnitude limit is that set by the GZ2 main sample, and the brighter was the spectroscopic targeting limit for SDSS (required for a redshift and inclusion in the NSA; Figure 3). The few remaining galaxies brighter than 17.0 mag but not in GZ2 may be the result of positional matching errors, very low-redshift ($z < 0.0005$) galaxies or targets with a large angular size that were shredded in the initial SDSS pipeline.

Summary: roughly 40% of the DECaLS galaxies have morphological measurements from GZ2, and can be used for direct comparison. We believe we understand the reasons for the remainder of DECaLS images that are not matched to GZ2; these will be valuable scientific additions as new targets, and can serve as independent checks on the accuracy of the classifications.

5. ANALYSIS

5.1. *Science cases*

1. Statistical studies of galaxy evolution with morphology

   Nothing yet.

2. Discovery of rare objects

   The majority of rare objects discovered in previous phases of Galaxy Zoo took place through interactions on the Galaxy Zoo Forum, a bulletin-board style website where users could tag individual objects and discuss them with other users and scientists. The replacement for the now-retired forum is Talk[2]. Talk has two features that are designed to aid in identifying unusual and interesting objects: hashtags and collections.

   Hashtags are discrete labels applied to an image by a user, and are not specified by any aspect of the interface. The most common hashtags applied to DECaLS subjects were identical or nearly identical to the morphological categories in the interface, including "spiral" (1,381), "merger" (1,013), "disturbed" (790), and "ring" (637). Common hashtags *not* associated with categories in the decision tree include "starforming" (535), "agn" (309), and "oiii" (168). The latter two hashtags specifically refer most often to cases where users examine the associated spectrum of the galaxy (typically from SDSS, which has a direct link from the GZ Examine page) and posting additional information, since neither AGN nor [OIII] emission can be reliably identified from visual morphology alone in these bands (but see Cardamone et al. 2009, for an example of emission-line galaxies using different bands).

[2] http://talk.galaxyzoo.org

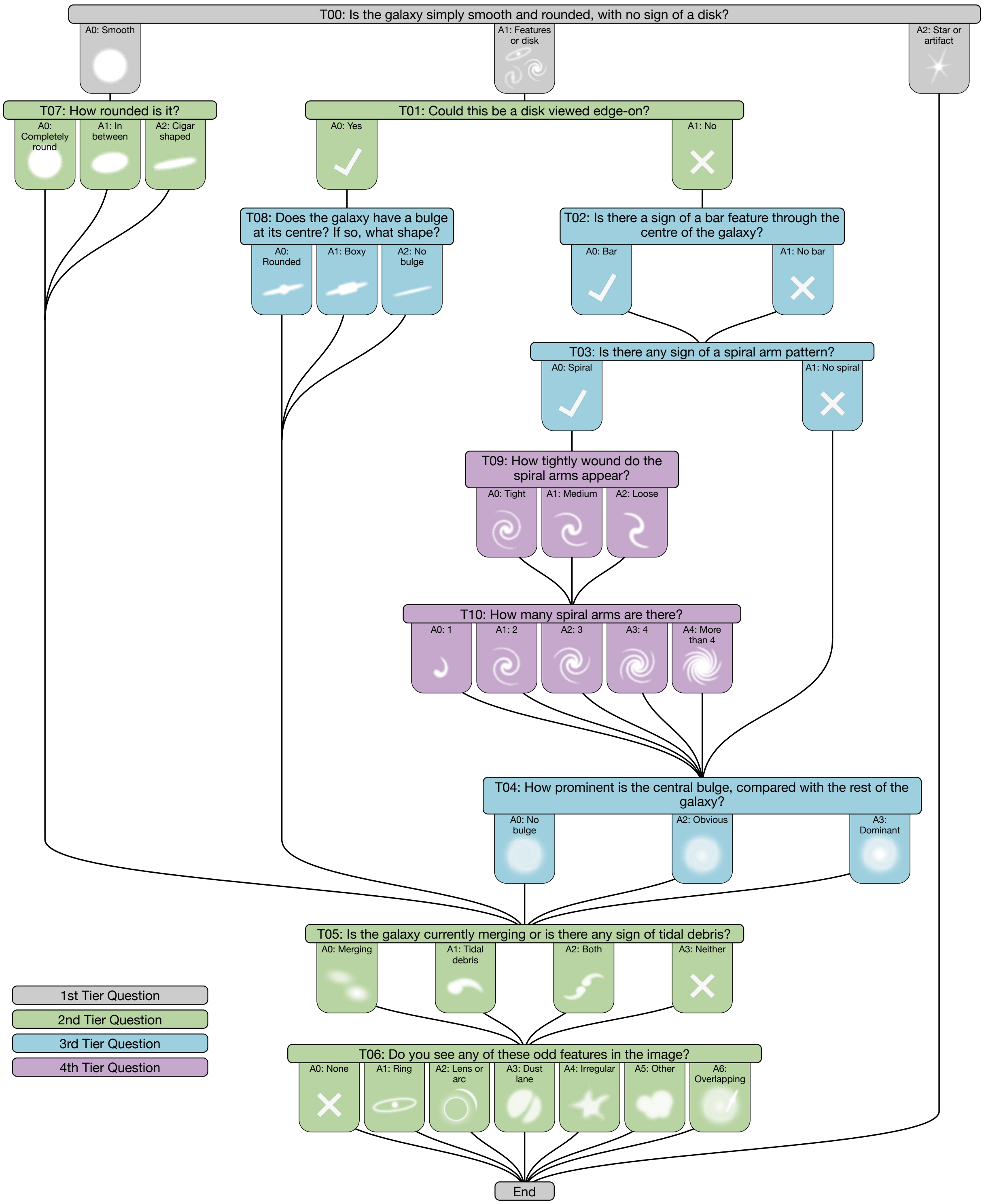

Figure 2. Hierarchical decision tree for the morphology annotations in GZ-DECaLS. A comparison of all Galaxy Zoo morphological schemes is available at `http://data.galaxyzoo.org/gz_trees/gz_trees.html`.

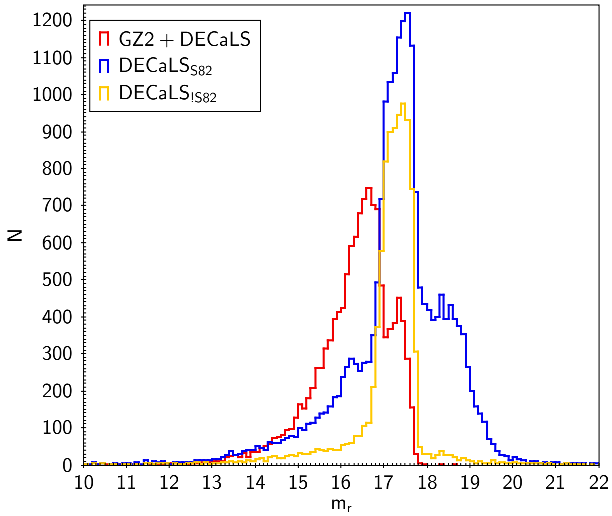

Figure 3. Histogram of the apparent r-band magnitude distribution for various subsamples, binned in $\Delta m_r = 0.1$. The red histogram shows galaxies classified in both GZ2 and GZ-DECaLS; the blue represents GZ-DECaLS galaxies from the Stripe 82 region but not detected in GZ2, and the yellow represents GZ-DECaLS galaxies from outside the Stripe 82 region but not detected in GZ2.

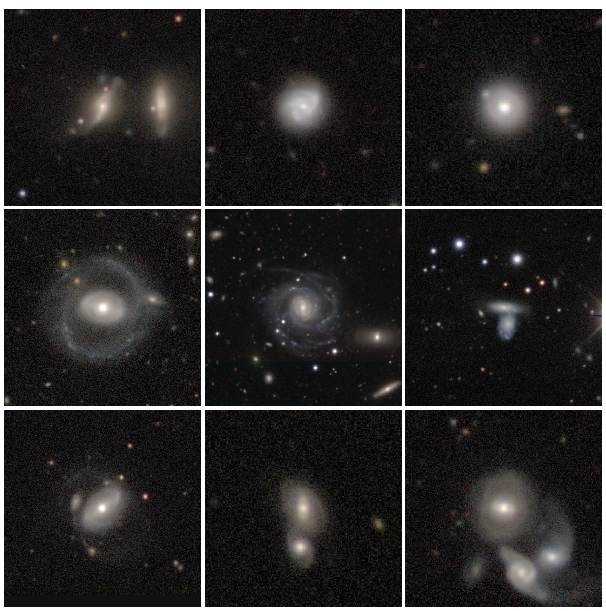

Figure 4. The most commonly collected images from GZ-DECaLS in the Talk interface. Images are arranged in order of the number of their collections they appear in, from 23 (upper left) to 19 (lower right).

Collections are discrete sets of objects assembled by volunteers; they often (although are not required to) revolve around a theme, such as mergers or galaxies with a particular visual color. One indication of unusual objects could be the tendency of users to place it in a collection. Among the 10 most-commonly collected galaxies in GZ-DECaLS (each selected by between 19 and 23 users), three were merging systems (including one multi-merger), three were spectacular examples of grand-design spirals, two visually striking overlaps, one a distorted edge-on disk, and one ring galaxy (Figure 4). One of the spirals (`AGZ000be3p`) had an imaging artifact in one of the DECaLS bands that may have caused additional interest.

Discussions on Talk can also take place that are not associated with any particular object, although they can be referenced. Only seven galaxies in GZ-DECaLS are mentioned in more than two discussion threads; these constitute a variety of mergers, artifacts, and grand design spirals. One potentially interesting set are galaxies with a visible point source that could be detections of supernovae (such as `AGZ000atnd; see` Frieman et al. 2008).

Identifying the most common tags and collections, however, will not necessarily identify truly unusual galaxies (and in many ways will select directly *against* that). For scientists with particular interests, the searchable Talk archives offer a rich database for particular features or combinations. Objects worthy of further study need some mechanism to filter up to scientists who can offer context and the means to follow-up observations, if necessary.

3. Galaxy morphology at low luminosities

   Nothing yet.

4. Exploring the role of minor mergers

   Both mergers and tidal debris are explicitly identified as Task 09 of the decision tree (Figure 2), which is answered for every image not designated as a star or artifact. While major mergers (with mass ratios of $\sim 3:1$ or less) with tidal tails are relatively easy to identify (Darg et al. 2010), the tidal debris from either a minor or significantly relaxed merger can have a very low surface brightness. GZ-DECaLS increased the number of example images for this particular question, including annotations pointing out faint features and a longer blog post[3] demonstrating the various types of tidal debris that were expected to be seen in DECaLS.

   The challenging role for minor mergers and tidal debris is establishing a reliable threshold on vote fraction for which feature identification is considered reliable. Figure 5 shows the number of potential merger candidates as a function of the total number of people who answered Task 09. This is $N_{\rm T09} = 40 - N_{\rm star/artifact}$, where the mean value of unweighted $f_{\rm star/artifact}$ is 0.18 for the entire GZ-DECaLS sample. This is significantly higher than GZ2, which had an average $f_{\rm star/artifact}$ of only 0.04. Among the GZ-DECaLS images, the number of candidates decreases sharply as the number of people who answered the question with a positive answer (either "merger", "tidal debris", or "both") is increased.

   At the canonical "clean" threshold of 80% adopted in the original Galaxy Zoo by Lintott et al. (2008), there are 782 potential merger/tidal debris candidates, or 0.02% of the sample. Of those, 574 were identified (from the plurality vote) as mergers, 111 as tidal debris, and 97 as having both a merger and tidal debris. This is a small enough sample to permit expert scientists to do individual follow-up of the images. Note that there is no reason that a specific prior of 80% should apply to this task, necessarily, and examination of the purity vs. completeness for a lower threshold is critical for science results.

[3] http://blog.galaxyzoo.org/2015/11/05/searching-for-tidal-debris-in-decals-images/

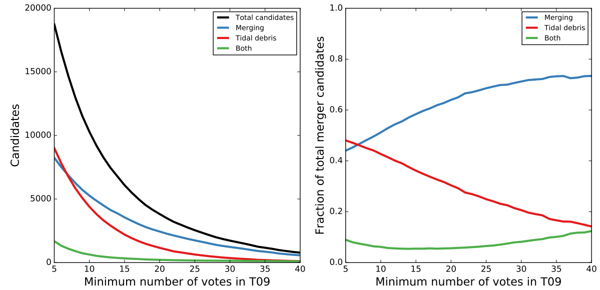

Figure 5. *Left*: Number of candidate minor mergers in GZ-DECaLS as a function of the number of people who gave a positive identification to Task 09 (Figure 2). *Right*: Minor merger candidates split by their type according to plurality vote.

Figure 5 also shows that the relative fraction of galaxies identified as merger vs. tidal debris has a strong dependence on the overall number of votes for the task. As the number of total positive votes for the question increases, the likelihood of "tidal debris" being the plurality answer drops from roughly 50% to 15%, with the likelihood of "merger" correspondingly increasing. The likelihood of identifying *both* tidal debris and a merger, on the other hand, is relatively constant at $\sim 10-15\%$ over the entire range of votes.

5.2. *DECaLS — SDSS comparison*

Figure 7 shows the comparison between the GZ2 SDSS morphological classifications and those in the deeper DECaLS imaging. Direct comparison between tasks is possible for almost all morphological categories in Tables 1 and 2, since the DECaLS tree was designed to closely mimic the GZ2 tree. The exceptions are for the task measuring bulge prominence in disk galaxies; DECaLS had only three options for ranking the bulge prominence (compared to four in GZ2), eliminating the "just noticeable" option. We justify this by noting that the "dominant" option in the original GZ2 tree was almost never selected; only eight galaxies in the entire main spectroscopic sample ($< 0.0003\%$) met the `clean` threshold 80% for this question. For Figure 7, we map bulge prominence for GZ2→DECaLS as 'no_bulge' → 'no_bulge','just_noticeable' → 'obvious', and 'obvious' → 'dominant'. The other exceptions are the elimination of the "can't tell" option for the number of spiral arms, and the difference in options and ability to select multiple categories in the "odd features" question. Given the very small numbers of galaxies for the latter option, we suggest that any potential targets be individually inspected.

The top level question — distinguishing between smooth (early-type) and feature/disk (late-type) galaxies, as well as flagging any stars or artifacts — is strongly correlated for the GZ2 and DECaLS classifications. This task is in essence a binary choice for the sample, since it must be answered for every image and the number of stars/artifacts is very small (only 2%). The results confirm the initial expectations that the deeper DECaLS imaging decreases the likelihood of a galaxy being classified as smooth; the effect is strongest in galaxies previously robustly classified as smooth ($f_{smooth,GZ2} > 0.80$) lying significantly below the one-to-one line in the top left panel of Figure 7. More than 33% of the galaxies have $(f_{smooth,GZ2} - f_{smooth,DECaLS}) > 0.25$, indicating a significant difference in the likelihood of being classified as early type. On the other side, only 90 galaxies ($< 1\%$ of the sample) increased their $f_{smooth,DECaLS}$ by the same amount. Specific numbers depend on the threshold being applied, but strongly favor the discovery of faint features (although not necessarily disks) due to the deeper imaging.

Stars/artifacts are not well-correlated, although the majority of all images in the overlapping sample have $f_{star} < 0.1$ in both datasets. This is to be expected - while the presence of a star incorrectly identified as a galaxy would show up in surveys, artifacts are specific to the imaging process and are expected to be almost entirely uncorrelated (unless caused by the presence of a permanent feature on the sky, such as chip bleed for a galaxy near a bright star). The DECaLS images do show a significantly higher rate of objects labeled as stars or artifacts than GZ2; tags in the Talk[4] platform indicate that these are largely driven by images that are too zoomed-out for accurate classification, caused by faulty measurements of the Petrosian radii in the NSA atlas. These images are good candidates to have the apparent image size corrected (manually, if necessary) and re-inserted into the Galaxy Zoo interface for classification.

The more detailed morphological categories in the features/disk branch are more tightly correlated than the top-level question. Edge-on galaxies are concentrated at the extremes; users tend to strongly agree whether a galaxy is edge-on or not, with very few intermediate examples. Galaxies designated as edge-on primarily have rounded bulges. Boxy bulges were almost never selected with high confidence, while bulgeless galaxies were rare, but well-correlated in both sets of data.

The bar fraction (splitting by majority vote) for disk galaxies actually dropped from 24% in GZ2 to 19% in DECaLS, although the vote fractions themselves are overall well-correlated, albeit with significant scatter. This is somewhat surprising, since the overall number is significantly below the fraction measured in both GZ2 (Masters et al. 2011) and other studies using optical images in volume-limited samples (eg, Barazza et al. 2008; Aguerri et al. 2009; Lee et al. 2012). An explanation for both the difference in GZ2 and DECaLS plus the overall difference with other studies may simply be the use of majority vote (eg, 50%) as a threshold for separating barred from unbarred galaxies. Several GZ2 papers (Hoyle et al. 2011; Willett et al. 2013) show that the crowdsourced classifications have a much higher completeness for strong bars; Galloway et al. (2015) show that a lower limit of $p_{bar} = 0.3$ identifies a more complete sample of bars while maintaining a low false positive rate.

Galaxies with spiral structure show no significant differences between DECaLS and GZ2 classifications, with spirals outnumbering lenticulars or other features by roughly 4 : 1 in both samples. There is no immediate sign that the ability of users to reliably identify lenticulars has changed, but this merits more careful follow-up by matching to expert samples (eg, Nair & Abraham 2010). In the spiral structure itself, users are more likely to identify tightly-wound spiral arms in the DECaLS images, primarily at the expense of medium-wound arms. The number of loosely-wound spirals (which correlates strongly with the presence of a merger; see Casteels et al. 2013) does not change. Interestingly, users also identified more galaxies with high-multiplicity spiral arms (3, 4, and 5+). This would be a natural outcome from improved imaging; better S/N would result in seeing fainter arms near the edge of a galactic disk, and improved seeing could im-

[4] `http://talk.galaxyzoo.org`

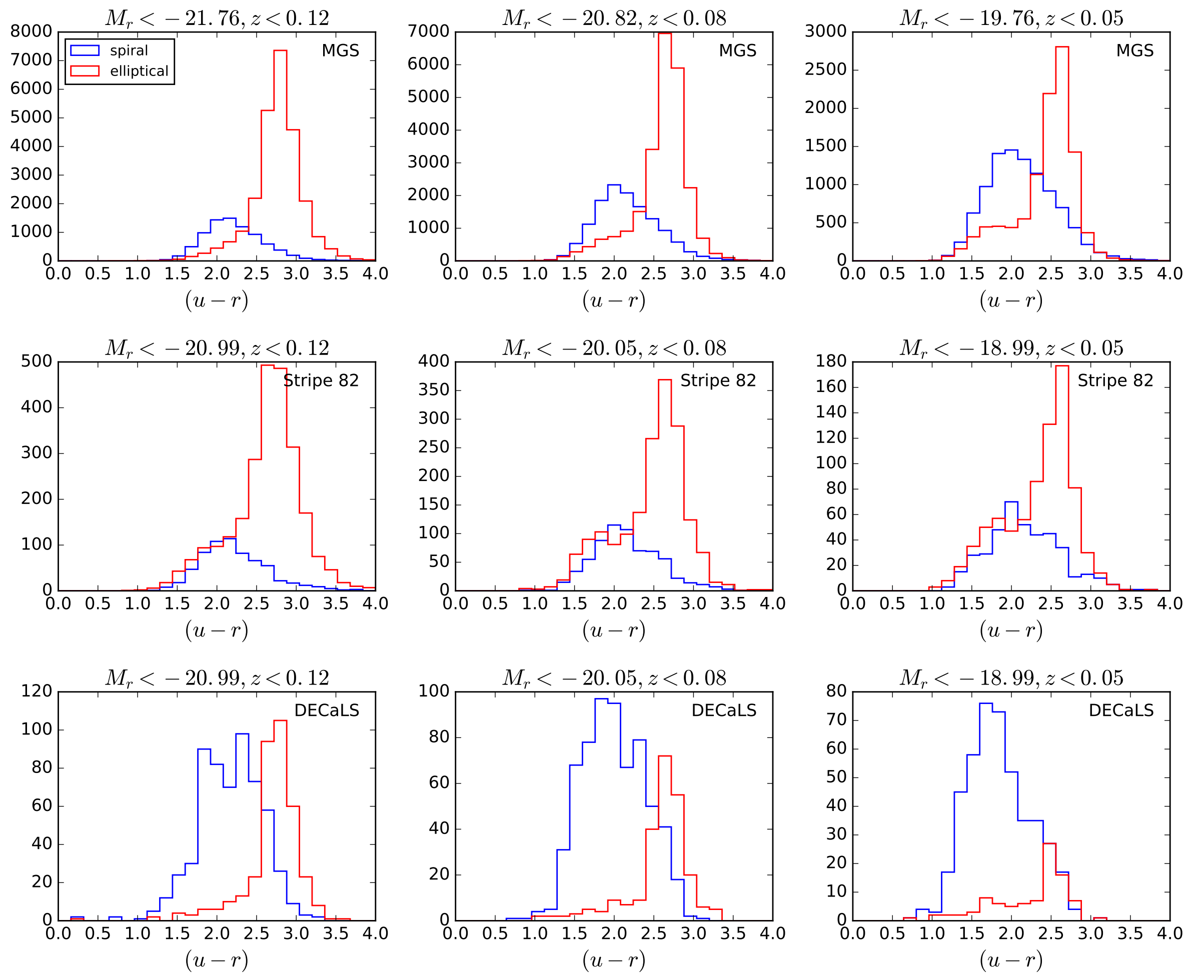

Figure 6. Histograms of the optical $(u-r)$ color distribution for various volume-limited and sample choices, separated by highly-confident ($p >= 0.8$) morphological classifications into spiral and ellipticals. Galaxies with intermediate morphologies ($0.2 < p < 0.8$) are not shown. **Top row**: GZ2 main spectroscopic sample. **Middle row**: Stripe 82 coadded. **Bottom row**: DECaLS.

prove the contrast in the interarm regions of the disk to reveal tightly-wound arms. Further analysis should also specifically investigate the distribution of DECaLS spiral arm votes for galaxies originally classified as "can't tell" in GZ2.[5]

Bulge prominence does show a significant difference between the GZ2 and DECaLS data. The likelihood of a "medium" bulge (designated as "just noticeable" in GZ2 and "obvious" in DECaLS) has a higher average vote fraction in DECaLS, and is roughly flat above $f_{medium\ bulge,GZ2} > 0.5$. This may be a strong function of the elimination of the "dominant" category, however; one of the priorities of the data reduction should be to better map bulge prominence between the two trees, and ideally transform it to a continuous variable that approximates the bulge-to-total ratio (eg, Lackner & Gunn 2012).

There are few galaxies with robust classifications in the catch-all "odd" categories. Ring galaxies are the only category with confident classifications by the user (eg, examples at both ends with few vote fractions at intermediate values). Users were slightly more confident in identifying rings in GZ2. There were no robust identifications of lens- or arc-shaped features in any GZ2/DECaLS pair of images. This is unsurprising, given the very low likelihood of detecting a true strong gravitational lens ($\sim 2 \times 10^{-4}$; Marshall et al. 2016) and the fact that the stretch and SDSS band selection does not optimize the color contrast for detecting lensed objects. Irregular objects are present, although with only a handful at high confidence in either survey. There are a large number of galaxies with high GZ2 vote fractions for "other", but this was not a commonly-selected option in DECaLS. The opposite effect is seen in the galaxies with a dust lane, for which many galaxies had essentially zero votes in GZ2 but a range of high confidences in DECaLS. It will be intriguing to investigate whether this seems to be driven by the improved DECaLS imaging, or whether it is a reflection of a more experienced user base with several years of identifying galaxies with dust lanes, and thus are moving away from the relatively vague "other" option in the decision tree.

5.3. *Other*

Describe overall demographics in Figure 8.

Figure 9 shows the change in average elliptical-spiral ratio. What does that mean?

Correlation between the various morphological tasks (Figure 10).

[5] R. Hart et al. (in prep) are working on a new method for debiasing the spiral arm data in the Galaxy Zoo decision tree.

Table 1
Galaxy Zoo morphological demographics for low-z optical imaging — all galaxies

Task	SDSS main sample			Stripe 82 coadd			DECaLS		
	N_{tot}	f_{tot}	$f_{prevtask}$	N_{tot}	f_{tot}	$f_{prevtask}$	N_{tot}	f_{tot}	$f_{prevtask}$
smooth	179153	0.74	0.74	16209	0.82	0.82	23292	0.72	0.72
features/disk	64067	0.26	0.26	3346	0.17	0.17	7967	0.25	0.25
star/artifact	280	0.00	0.00	210	0.01	0.01	1170	0.04	0.04
edge-on	9932	0.04	0.16	624	0.03	0.19	1726	0.05	0.22
not edge-on	54135	0.22	0.84	2722	0.14	0.81	6241	0.19	0.78
barred disk	14366	0.06	0.26	801	0.04	0.29	1174	0.03	0.19
no bar	39887	0.16	0.74	1932	0.10	0.71	5167	0.15	0.81
spiral	45462	0.19	0.84	2520	0.13	0.92	4973	0.15	0.80
no spiral	8791	0.04	0.16	213	0.01	0.08	1368	0.04	0.20
tight spiral arms	17322	0.07	0.39	1113	0.06	0.45	2279	0.07	0.46
medium spiral arms	20691	0.08	0.46	981	0.05	0.40	1637	0.05	0.33
loose spiral arms	6821	0.03	0.15	388	0.02	0.16	871	0.02	0.18
1 spiral arm	1879	0.01	0.04	119	0.01	0.05	237	0.00	0.05
2 spiral arms	26413	0.11	0.59	1602	0.08	0.65	3566	0.10	0.72
3 spiral arms	3025	0.01	0.07	188	0.01	0.08	625	0.01	0.13
4 spiral arms	837	0.00	0.02	51	0.00	0.02	192	0.00	0.04
5+ spiral arms	758	0.00	0.02	44	0.00	0.02	167	0.00	0.03
?? spiral arms	11922	0.05	0.27	478	0.02	0.19	—	—	—
no bulge	3962	0.02	0.07	103	0.01	0.04	593	0.01	0.10
noticeable bulge	34139	0.14	0.63	1139	0.06	0.42	—	—	—
obvious bulge	15791	0.06	0.29	1321	0.07	0.48	5316	0.16	0.85
dominant bulge	361	0.00	0.01	170	0.01	0.06	432	0.01	0.07
round edge-on bulge	6506	0.03	0.66	524	0.03	0.85	1244	0.03	0.76
boxy edge-on bulge	173	0.00	0.02	5	0.00	0.01	53	0.00	0.03
no edge-on bulge	3135	0.01	0.32	84	0.00	0.14	329	0.01	0.20
round elliptical	62308	0.26	0.35	6092	0.31	0.38	9279	0.28	0.39
in-between elliptical	91284	0.37	0.51	8331	0.42	0.51	11369	0.35	0.48
cigar-shaped elliptical	25561	0.10	0.14	1786	0.09	0.11	2644	0.08	0.11
odd feature	23795	0.10	0.10	1713	0.09	0.09	—	—	—
no odd features	219425	0.90	0.90	17842	0.90	0.91	—	—	—
ring	4099	0.02	0.18	178	0.01	0.11	317	0.01	0.01
lens/arc	155	0.00	0.01	17	0.00	0.01	4	0.00	0.00
disturbed	720	0.00	0.03	47	0.00	0.03	—	—	—
irregular	5761	0.02	0.25	113	0.01	0.07	44	0.00	0.00
other	4919	0.02	0.21	589	0.03	0.38	14	0.00	0.00
merger	7018	0.03	0.31	599	0.03	0.39	—	—	—
dust lane	220	0.00	0.01	6	0.00	0.00	141	0.00	0.00
overlapping	—	—	—	—	—	—	52	0.00	0.00
nothing	—	—	—	—	—	—	979	0.03	0.03
merger	—	—	—	—	—	—	1543	0.04	0.04
tidal debris	—	—	—	—	—	—	343	0.01	0.01
merger and tidal debris	—	—	—	—	—	—	141	0.00	0.00
neither	—	—	—	—	—	—	29232	0.93	0.93
total	243500			19765			32429		

Note. — A value of 0.00 is rounded down, indicating that the fraction of galaxies in this category was $f < 0.01$.

REFERENCES

Aguerri, J. A. L., Méndez-Abreu, J., & Corsini, E. M. 2009, A&A, 495, 491
Annis, J., Soares-Santos, M., Strauss, M. A., et al. 2014, ApJ, 794, 120
Bamford, S. P., Nichol, R. C., Baldry, I. K., et al. 2009, MNRAS, 393, 1324
Barazza, F. D., Jogee, S., & Marinova, I. 2008, ApJ, 675, 1194
Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, MNRAS, 399, 1191
Casteels, K. R. V., Bamford, S. P., Skibba, R. A., et al. 2013, MNRAS, 429, 1051
Clark, L. A., & Watson, D. 1995, Psychological Assessment, 7, 309
Darg, D. W., Kaviraj, S., Lintott, C. J., et al. 2010, MNRAS, 401, 1043
Frieman, J. A., Bassett, B., Becker, A., et al. 2008, AJ, 135, 338
Galloway, M. A., Willett, K. W., Fortson, L. F., et al. 2015, MNRAS, 448, 3442
Hoyle, B., Masters, K. L., Nichol, R. C., et al. 2011, MNRAS, 415, 3627
Kaviraj, S. 2010, MNRAS, 406, 382
—. 2014, MNRAS, 437, L41
Kaviraj, S., Darg, D., Lintott, C., Schawinski, K., & Silk, J. 2012, MNRAS, 419, 70
Keel, W. C., Manning, A. M., Holwerda, B. W., Lintott, C. J., & Schawinski, K. 2014, AJ, 147, 44
Keel, W. C., Manning, A. M., Holwerda, B. W., et al. 2013, PASP, 125, 2
Keel, W. C., Lintott, C. J., Schawinski, K., et al. 2012, AJ, 144, 66
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, MNRAS, 444, 1647
Lackner, C. N., & Gunn, J. E. 2012, MNRAS, 421, 2277
Lee, G.-H., Park, C., Lee, M. G., & Choi, Y.-Y. 2012, ApJ, 745, 125
Levi, M., Bebek, C., Beers, T., et al. 2013, ArXiv e-prints, arXiv:1308.0847 [astro-ph.CO]
Lintott, C., Schawinski, K., Bamford, S., et al. 2011, MNRAS, 410, 166
Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, MNRAS, 389, 1179
Lintott, C. J., Schawinski, K., Keel, W., et al. 2009, MNRAS, 399, 129
Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, PASP, 116, 133
Marshall, P. J., Verma, A., More, A., et al. 2016, MNRAS, 455, 1171
Masters, K. L., Mosleh, M., Romer, A. K., et al. 2010, MNRAS, 405, 783
Masters, K. L., Nichol, R. C., Hoyle, B., et al. 2011, MNRAS, 411, 2026
Masters, K. L., Nichol, R. C., Haynes, M. P., et al. 2012, MNRAS, 424, 2180
Nair, P. B., & Abraham, R. G. 2010, ApJS, 186, 427
Schawinski, K., Lintott, C., Thomas, D., et al. 2009, MNRAS, 396, 818

Table 2
Galaxy Zoo morphological demographics for low-z optical imaging — GZ2/DECaLS overlaps only

Task	SDSS main sample			Stripe 82 coadd			DECaLS		
	N_{tot}	f_{tot}	$f_{prevtask}$	N_{tot}	f_{tot}	$f_{prevtask}$	N_{tot}	f_{tot}	$f_{prevtask}$
smooth	6715	0.72	0.72	4705	0.81	0.81	7583	0.62	0.62
features/disk	2554	0.28	0.28	1054	0.18	0.18	4482	0.37	0.37
star/artifact	12	0.00	0.00	41	0.01	0.01	202	0.02	0.02
edge-on	395	0.04	0.15	211	0.04	0.20	781	0.06	0.17
not edge-on	2159	0.23	0.85	843	0.15	0.80	3701	0.30	0.83
barred disk	528	0.06	0.24	257	0.04	0.30	706	0.06	0.19
no bar	1636	0.18	0.76	589	0.10	0.70	3049	0.25	0.81
spiral	1842	0.20	0.85	785	0.14	0.93	3044	0.25	0.81
no spiral	322	0.03	0.15	61	0.01	0.07	711	0.06	0.19
tight spiral arms	654	0.07	0.36	355	0.06	0.46	1545	0.13	0.52
medium spiral arms	848	0.09	0.47	300	0.05	0.39	974	0.08	0.33
loose spiral arms	315	0.03	0.17	118	0.02	0.15	428	0.03	0.15
1 spiral arm	73	0.01	0.04	34	0.01	0.04	125	0.01	0.04
2 spiral arms	1055	0.11	0.58	488	0.08	0.63	2136	0.17	0.72
3 spiral arms	120	0.01	0.07	60	0.01	0.08	442	0.04	0.15
4 spiral arms	35	0.00	0.02	18	0.00	0.02	142	0.01	0.05
5+ spiral arms	24	0.00	0.01	14	0.00	0.02	102	0.01	0.03
?? spiral arms	510	0.05	0.28	159	0.03	0.21	—	—	—
no bulge	152	0.02	0.07	36	0.01	0.04	191	0.02	0.05
noticeable bulge	1410	0.15	0.65	374	0.06	0.44	—	—	—
obvious bulge	590	0.06	0.27	382	0.07	0.45	3293	0.27	0.88
dominant bulge	12	0.00	0.01	54	0.01	0.06	271	0.02	0.07
round edge-on bulge	266	0.03	0.68	183	0.03	0.88	566	0.05	0.78
boxy edge-on bulge	7	0.00	0.02	2	0.00	0.01	23	0.00	0.03
no edge-on bulge	117	0.01	0.30	23	0.00	0.11	138	0.01	0.19
round elliptical	2198	0.24	0.33	1581	0.27	0.34	3160	0.26	0.42
in-between elliptical	3482	0.38	0.52	2513	0.43	0.53	3583	0.29	0.47
cigar-shaped elliptical	1035	0.11	0.15	611	0.11	0.13	840	0.07	0.11
odd feature	763	0.08	0.08	407	0.07	0.07	—	—	—
no odd features	8506	0.92	0.92	5352	0.92	0.93	—	—	—
ring	142	0.02	0.20	51	0.01	0.14	200	0.02	0.34
lens/arc	4	0.00	0.01	6	0.00	0.02	3	0.00	0.01
disturbed	30	0.00	0.04	15	0.00	0.04	—	—	—
irregular	223	0.02	0.31	40	0.01	0.11	12	0.00	0.02
other	126	0.01	0.17	107	0.02	0.29	9	0.00	0.02
merger	188	0.02	0.26	143	0.02	0.39	—	—	—
dust lane	9	0.00	0.01	1	0.00	0.00	46	0.00	0.08
overlapping	—	—	—	—	—	—	14	0.00	0.02
nothing	—	—	—	—	—	—	296	0.02	0.51
merger	—	—	—	—	—	—	458	0.04	0.04
tidal debris	—	—	—	—	—	—	167	0.01	0.01
merger and tidal debris	—	—	—	—	—	—	64	0.01	0.01
neither	—	—	—	—	—	—	11376	0.93	0.94
total	9281			5800			12267		

Note. — A value of 0.00 is rounded down, indicating that the fraction of galaxies in this category was $f < 0.01$.

Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, MNRAS, 440, 889
Simmons, B. D., Lintott, C., Schawinski, K., et al. 2013, MNRAS, 429, 2199
Skibba, R. A., Bamford, S. P., Nichol, R. C., et al. 2009, MNRAS, 399, 966
Skibba, R. A., Masters, K. L., Nichol, R. C., et al. 2012, MNRAS, 423, 1485
Smethurst, R. J., Lintott, C. J., Simmons, B. D., et al. 2015, MNRAS, 450, 435
Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810
Willett, K. W., Lintott, C. J., Bamford, S. P., et al. 2013, MNRAS, 435, 2835

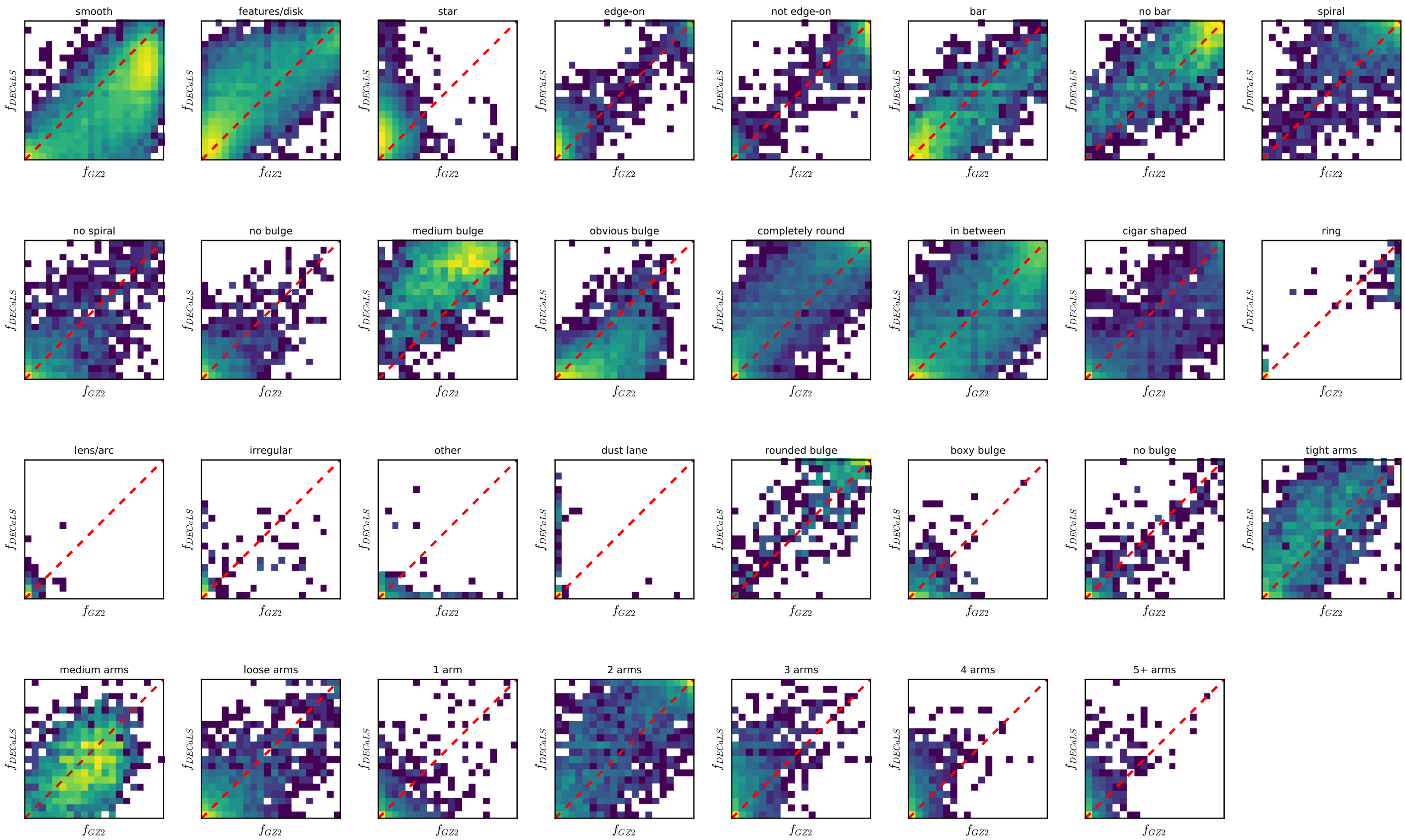

Figure 7. Comparison of the unweighted morphological vote fractions for the 12,267 galaxies in both DECaLS (x-axis) and GZ2 (y-axis). Colors are a log-histogram of the number of galaxies in each bin, normalized to the peak bin for each category. Each morphological category plots only galaxies for which the question was well-sampled (determined by plurality vote through the decision tree). The red dashed line shows the one-to-one correspondence.

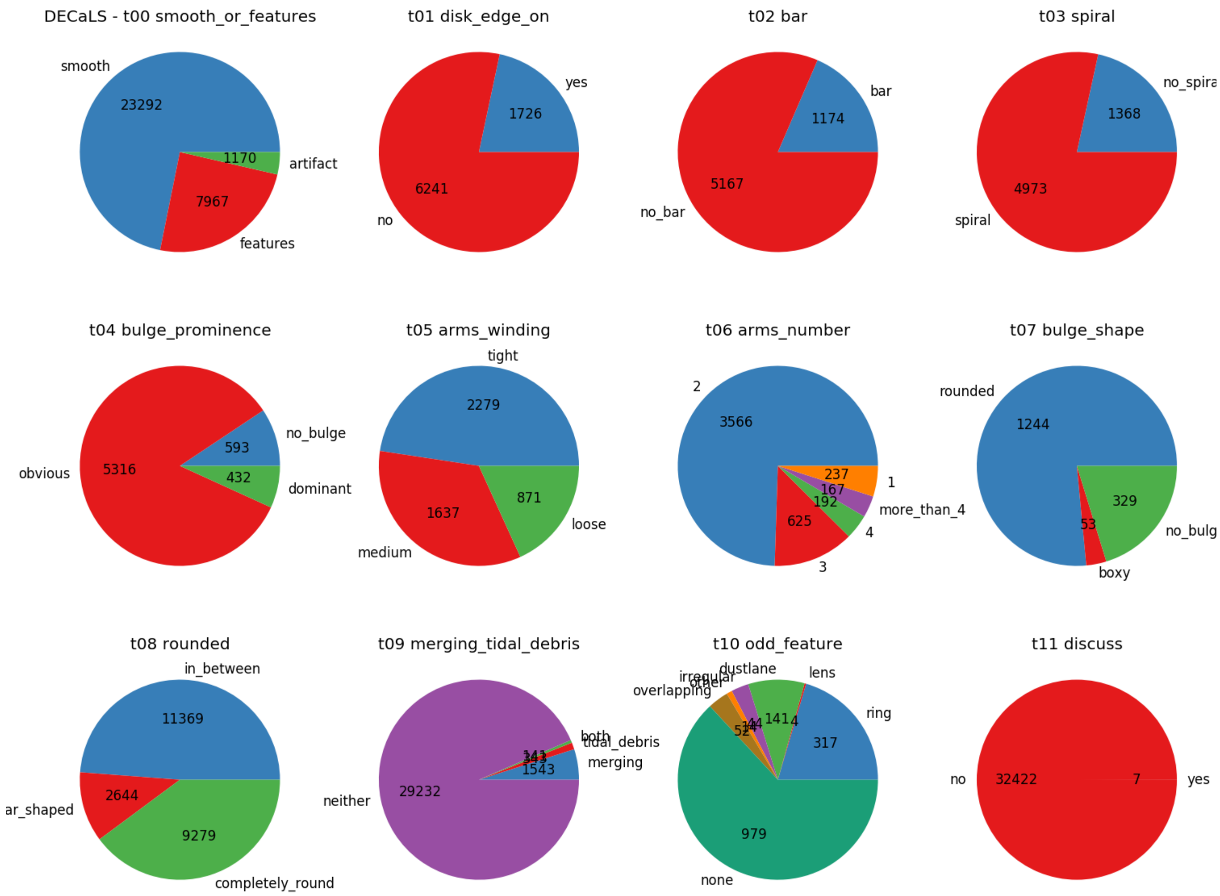

Figure 8. Distribution of the plurality morphological category for each task in GZ:DECaLS. Labels show the count of the number of galaxies with each morphology, including only those for which the question was reached in the hierarchical tree.

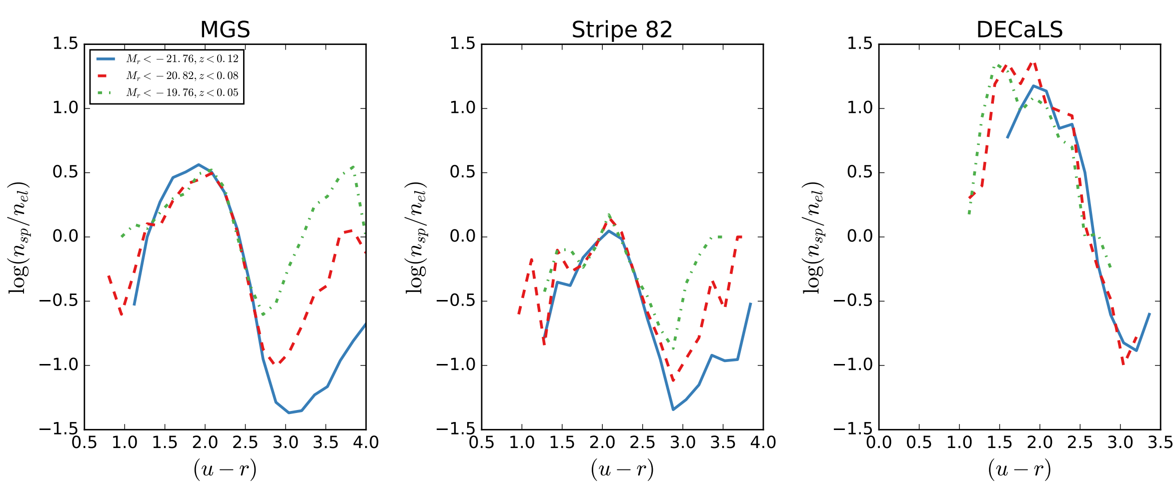

Figure 9. Average spiral-elliptical ratio for various galaxy samples as a function of optical $(u-r)$ color. From left to right, curves are for the GZ2 main spectroscopic sample, the deeper Stripe 82 coadded images, and the DECaLS images. Colors/linestyles show different volume/absolute magnitude limits for each sample.

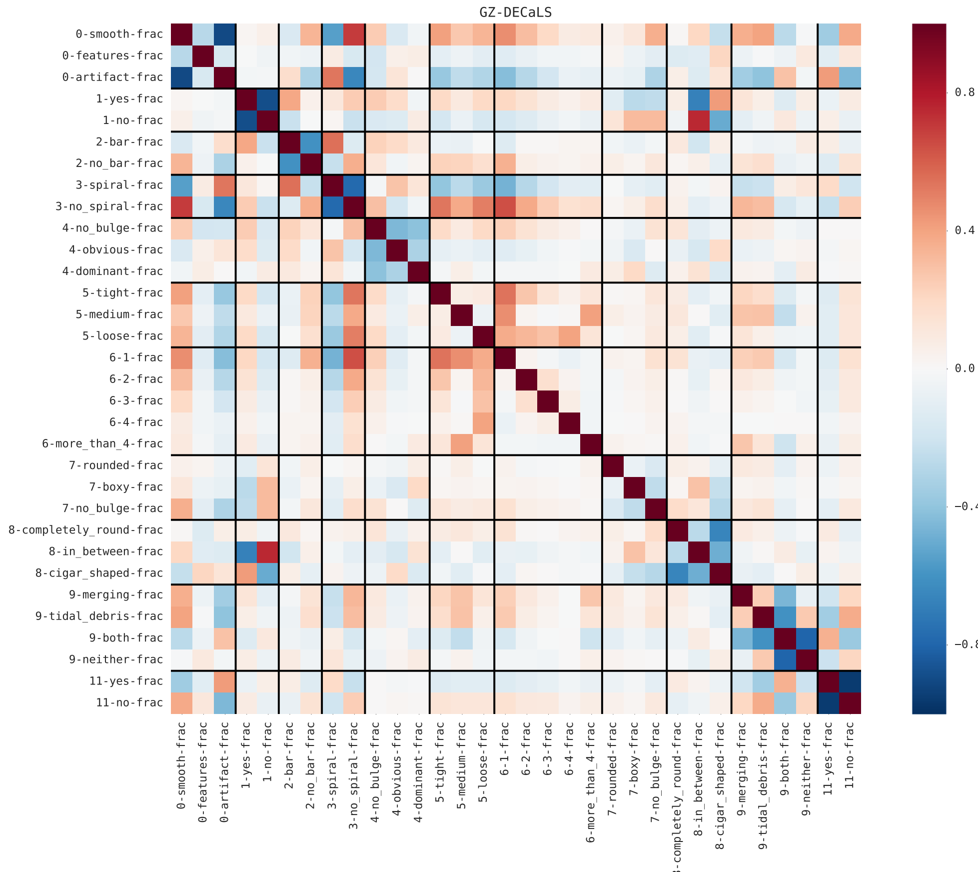

Figure 10. Correlation matrix between the various tasks in the GZ-DECaLS decision tree. Lower left and upper right corners show the same (mirrored) sets of data.