FERMILAB-PUB-14-578-AE

arXiv:1409.2885v1 [astro-ph.IM] 9 Sep 2014

Crowded Cluster Cores: Algorithms for Deblending in Dark Energy Survey Images

Yuanyuan Zhang[1], Timothy A. McKay[1,2,3,⋆], Emmanuel Bertin[4], Tesla Jeltema[5], Christopher J. Miller[1,2], Eli Rykoff[6], Jeeseon Song[1]

ABSTRACT

Deep optical images are often crowded with overlapping objects. This is especially true in the cores of galaxy clusters, where images of dozens of galaxies may lie atop one another. Accurate measurements of cluster properties require deblending algorithms designed to automatically extract a list of individual objects and decide what fraction of the light in each pixel comes from each object. We present new software called the Gradient And INterpolation based deblender (GAIN) as a secondary deblender to improve deblending the images of cluster cores. This software relies on using image intensity gradient and using an image interpolation technique usually used to correct flawed terrestrial digital images. We test this software on Dark Energy Survey coadd images. GAIN helps extracting unbiased photometry measurement for blended sources. It also helps improving detection completeness while introducing only a modest amount of spurious detections. For example, when applied to deep images simulated with high level of deblending difficulties, this software improves detection completeness from 91% to 97% for sources above the 10σ limiting magnitude at 25.3 mag. We expect this software to be a useful tool for cluster population measurements.

Subject headings: surveys: catalogs: techniques: image processing

1. Introduction

Deep, wide-field, ground based optical imaging surveys have an important role to play in the near-term future of astronomy. Existing projects like the Pan-STARRS[1] and the Dark Energy Survey (DES)[2] are mapping thousands square degrees of the sky to 24th magnitude and beyond, while a future project like the Large Synoptic Survey Telescope (LSST)[3] plans to image 20,000 square degrees to 28th magnitude. These surveys support an enormous range of science goals, from identifying Near-Earth Objects to measuring the expansion history of the universe. One important goal for all of these surveys is to measure the properties of galaxy clusters. Clusters are extreme objects, marking the upper limit of structure formation. As such, properties of their population are extremely sensitive to basic cosmological parameters like the expansion history and cosmic mass density. Detecting galaxy clusters and measuring their properties precisely is an important task.

While deep images greatly facilitate the detection and measurement of clusters, they present

[1]Physics Department, University of Michigan, Ann Arbor, MI, USA

[2]Department of Astronomy, University of Michigan, Ann Arbor, MI, USA

[3]Michigan Center For Theoretical Physics, University of Michigan, Ann Arbor, MI, USA

[4]Institut d'Astrophysique de Paris, Univ. Pierre et Marie Curie & CNRS UMR7095, F-75014 Paris, France

[5]Department of Physics and Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA, USA

[6]SLAC National Accelerator Laboratory, Menlo Park, CA, USA

⋆Following authors listed alphabetically.

[1]http://pan-starrs.ifa.hawaii.edu/

[2]http://www.darkenergeysurvey.org/

[3]http://www.lsst.org/

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

some data processing challenges as well. One of the most challenging is deblending. Image crowding is a problem in many astronomical applications, and a wide variety of approaches to solving it have been proposed. Many were created for special purposes, for example deblending images of objects crowded with point sources, like globular clusters (Federici et al. 1983; Stetson 1987; Diolaiti et al. 2000; Savage and Oliver 2007). In this application, the intrinsic shape of every object can be more or less accurately estimated. Deep images of galaxy clusters present more complexity. Cluster cores are dense with galaxies, each with its own unknown shape, brightness, and size. Their images often overlap, making the identification of individual galaxies and measurements of their brightness a challenge.

In this manuscript, we describe a package designed for deblending of cluster cores. This software automatically identifies blended sources, separates them, and measures the photometry of each individual source. We describe tests of this software on both real and simulated images. This software is primarily designed for DES, which has started its 5-year survey since August 2013 (Diehl et al. 2014). DES obtained a full depth scientific verification data set in 2012 with the new DECam camera (Flaugher et al. 2012) mounted on the Blanco telescope, allowing this package to be directly tested on DES images with DES nominal depth. However, the algorithms may also be of use for other data sets like HST data and SDSS data.

Several software packages for astronomical image processing exist. Each attempts to automatically detect astronomical sources and measure their properties. One of the most widely used packages is SExtractor (Bertin and Arnouts 1996), which can extract sources in a wide field in a short amount of time. Before SExtractor, a few packages (Jarvis and Tyson 1981; Beard et al. 1990; Maddox et al. 1990; Yee 1991) existed, but are not up to date for application with modern wide field surveys. Since the emergence of SExtractor, scientists have been primarily focused on developing software for specific instruments (York et al. 2000; Lupton 2001) or to do precise photometry and shape measurement without attempting to do astronomical source detection (Simard et al. 2002; Peng et al. 2002, 2010).

The Dark Energy Survey has adopted an advanced version of SExtractor in the data processing pipeline for catalog production. To handle deblending, SExtractor has a procedure that decides if a detected object should be further separated as several branching components. A major parameter that controls this procedure is DEBLEND_MINCONT. The DEBLEND_MINCONT parameter is a threshold on the intensity ratio between a branching component and the original detected object. However, upon carefully tweaking SExtractor's deblending procedure, especially the DEBLEND_MINCONT parameter, it is found that high levels of deblending are accompanied by increasing numbers of spurious detections and also deteriorating photometry. At an optimum setting of SExtractor to achieve the balance between real and spurious detections for DES, additional deblending is still needed, especially in cluster cores. We note that a future version of SExtractor featuring an improved deblender is in the works (Bertin, private communication). As of now, it is difficult for a general purpose suite of software to also handle deblending well. We suspect that this deblending dilemma plagues other reduction packages. The data processing pipeline of the Sloan Digital Sky Survey is very different from that of the Dark Energy Survey. However, there are also reports about suppressed completeness around bright objects (Adelman-McCarthy et al. 2006; Mandelbaum et al. 2005), implying that deblending was also an issue in that analysis. The particular challenges of accurately deblending galaxy cluster cores in both surveys led to this project.

In this paper, we describe new software that aids blended source detection and photometry measurement. It operates on clean, already processed astronomical images. We describe steps in the image analysis in Section 2, explain the algorithms in Section 3, and present our analysis on the effectiveness of this approach in Section 4. We summarize the algorithms and the testing in Section 5.

2. Application

This package, which we call the Gradient And INterpolation based deblender (GAIN deblender), is written in c++ and IDL and can be acquired on-

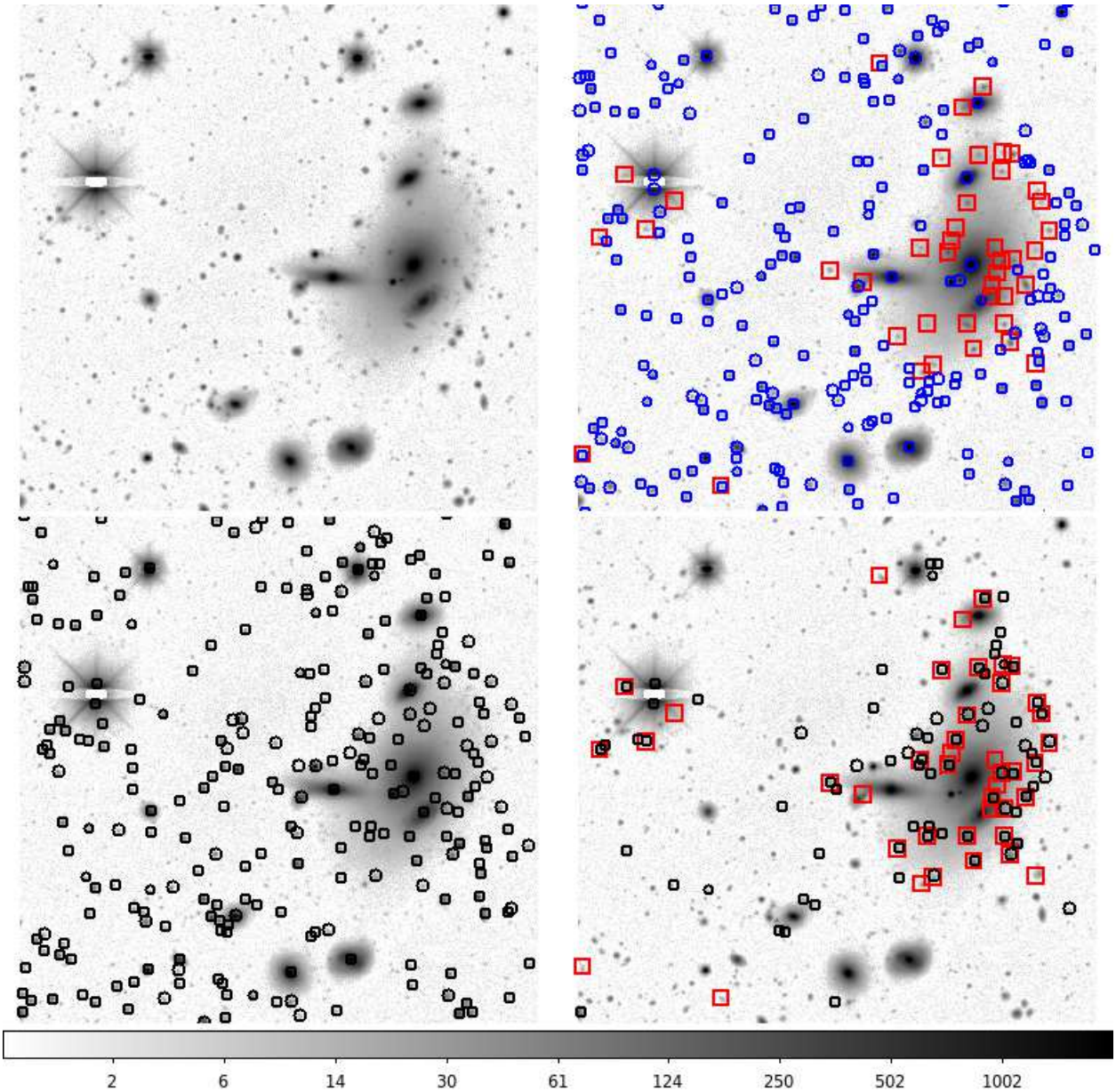

Fig. 1.— Upper Left: A DES image, which is the linear combination of the r, i and z coadd images. There is a bright star on the left and a brightest cluster galaxy on the right. Upper Right: Same image with detected sources marked. Blue circles mark SExtractor detections with the optimum DES deblending setting. Red squares mark additional detections found through GAIN. Lower left: The circles mark SExtractor detections with a very low SExtractor DEBLEND_MINCONT setting (1×10^{-6}). Lower Right: Black circles mark the differences between using different SExtractor settings. For comparison, GAIN detections marked by red squares in the upper right panel are shown again in this figure. This figure illustrates that adjusting the SExtractor DEBLEND_MINCONT parameter is insufficient for finding all the blended sources. In these figures, we only show sources brighter than 24.0 mag in i.

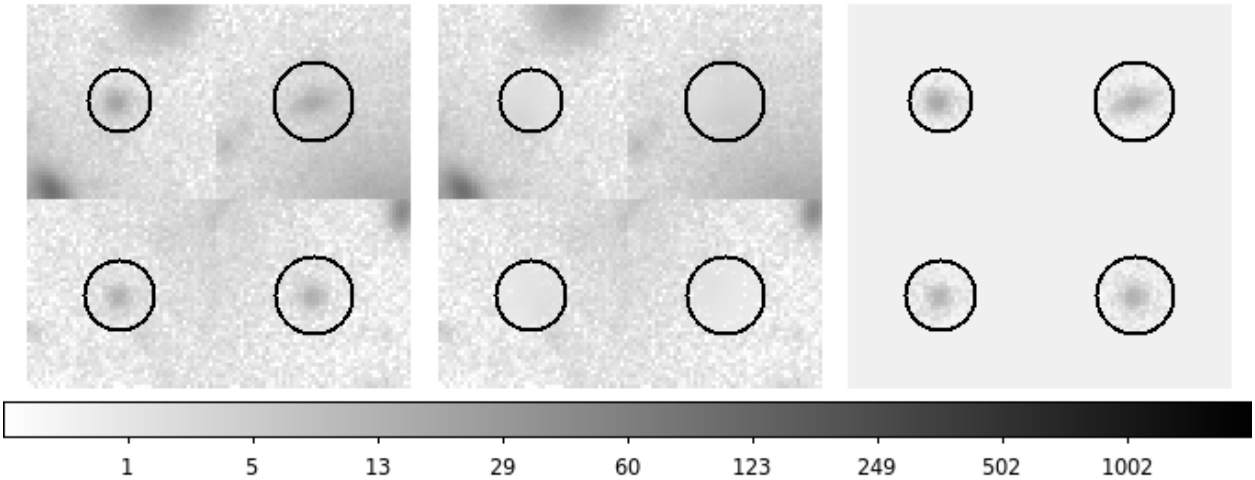

Fig. 2.— Left: r band DES coadd images of four blended sources. The circles indicate the regions to be deblended and are centered on the blended sources. Middle: Images that contain light from neighbors of the blended sources. Right: the residual between the left and middle panels, which shows the light from the blended sources within the circle.

line[4]. It operates on processed images that have been flat corrected, background subtracted, cosmic rays removed etc. We assume that the users have already done one round of source extraction by other means, and are using this package for finding blended sources that are missed in the previous procedure or to improve the photometry of blended sources.

The first stage in our analysis is the identification of blended sources. It begins by independently identifying all sources in the image, then matching its results to the user supplied object list. Sources not matched to the user supplied input catalog are kept as new sources to be extracted in the blended image. This procedure is illustrated in Figure 1. The upper left panel in Figure 1 shows a DES image containing a bright star and a brightest cluster galaxy. The upper right panel in Figure 1 shows sources detected by SExtractor (the blue circles) and some additional sources identified by GAIN (the red boxes). Note that this package is sensitive to image imperfections as is the case around the saturated star in the left half of this image. It has the beauty of finding very faint sources but also the peril of finding false sources around image artifacts like satellite trails or saturated star streaks. Qualitatively, the GAIN algorithm can find blended sources without introducing as many false detections as SExtractor with a low DEBLEND_MINCONT setting, as illustrated in Figure 1.

The second stage in our analysis aims to correctly assign the light in each pixel to the blended sources which contribute to it. This procedure is illustrated in Figure 2. Given a detected blended source and a region (this region can be computed from our package) to be deblended, as shown in the left panel, the package interpolates for the light that comes from "background sources" as shown in the middle panel. The residual between the original image and the "background sources" interpolation is the light from the blended source alone (right panel).

We find GAIN to be useful for two scenarios. In many applications, GAIN would be used both to detect blended sources using the first component and to extract light for those sources using the second component. It may also be used by skipping the first step, and using the second component to extract light for user supplied blended sources. Subsequent photometry measurement can be done on light separated images. This step is recommended if the user has blended sources in their first round of source extraction and wants to ob-

[4] https://github.com/yyzhang/gain_deblend

tain consistent photometry for all blended sources.

The speed of running this package is fast enough to apply with wide field optical surveys. For one image, the running time of GAIN is only a small fraction of the running time of SExtractor. On a workstation computer equipped with Intel Xeon Processor E5645, running on a $10,000 \times 10,000$ pixels image without parallelization, the source detection takes approximately 300 seconds to identify every 10,000 sources (without matching to a user supplied input catalog). For light separation, the light separation module takes approximately 10 seconds for every 1,000 sources. The memory usage of this package depends on the size of the input image. It is generally more than twice the size of the image. For a $10,000 \times 10,000$ pixels image which has the size of $\sim$ 3.5GB, the peak memory usage can be as intensive as $>$ 8GB.

3. Methods

3.1. Source Detection

The source detection component of GAIN aims to identify local maxima in images which are associated with real objects. Our approach to this is inspired by the crowded field stellar photometry software DAOPHOT (Stetson 1987). The presence of a separable astronomical object usually causes a local image intensity maximum. Unfortunately, many, even most, local intensity maxima are generated by image noise rather than real astronomical objects. Procedures to eliminate these noise peaks are therefore necessary.

One approach to reducing the impact of noise is to smooth the whole image. While smoothing is very effective at eliminating noise, it also washes out the saddle points which separate close pairs of real astronomical sources, exacerbating the problem of blending. When close pairs are separated by distances around twice the FWHM of the seeing, and one source is brighter than the other, even very modest smoothing merges the two. To do the best job of deblending, we would like to avoid smoothing. Instead of finding maxima in smoothed images, GAIN uses the image segmentation procedure and the image Laplacian map to reduce the impact of noise.

The source detection of GAIN begins with identifying sources on raw, unsmoothed images, then purges the identifications assigned with low pixel area in the segmentation map. To further eliminate spurious detections, GAIN cross matches the remaining identifications to sources identified in a "weighted Laplacian" map (which we explain in Section 3.1.2). The software performs two other rounds of segmentation area purging during the "weighted Laplacian" step and during the "cross matching" step.

3.1.1. Segmentation of Images

Purging of noise peaks can be aided by segmenting an image into separate regions, each associated with one element in the seed list of maxima. This seed list may include **all** local maxima, or may be produced by some other means, for example it could be a list of objects identified by SExtractor. We use a simplified version of Meyer's watershed flooding algorithm (Meyer 1992) for segmentation. To prime the segmentation procedure, we give each of the pixels in the seed list a unique region label. The goal of segmentation is then to label every pixel in the image as either belonging to one of the regions or residing in a boundary between the regions. The process begins by ranking all pixels in the image in descending order of intensity. As we move down the list, we apply the following logic to each pixel.

1. If this pixel is already labeled (because it was in the seed list), its label remains unchanged.

2. If this pixel is unlabeled, and all of the neighboring pixels are unlabeled, this pixel is marked as a *boundary*. This should be true only for local maxima not found in the seed list.

3. If this pixel is unlabeled, and some or all of the neighboring pixels are already labeled, and their labels (except those labeled as *boundary*) are **not** all the same, this pixel is labeled as a *boundary*.

4. If this pixel is unlabeled, and some or all of the neighboring pixels are already labeled, and their labels (except those labeled as *boundary*) are all the same, this pixel is given that label. If all the labeled neighbors are labeled as *boundary*, this pixel will also be a *boundary*.

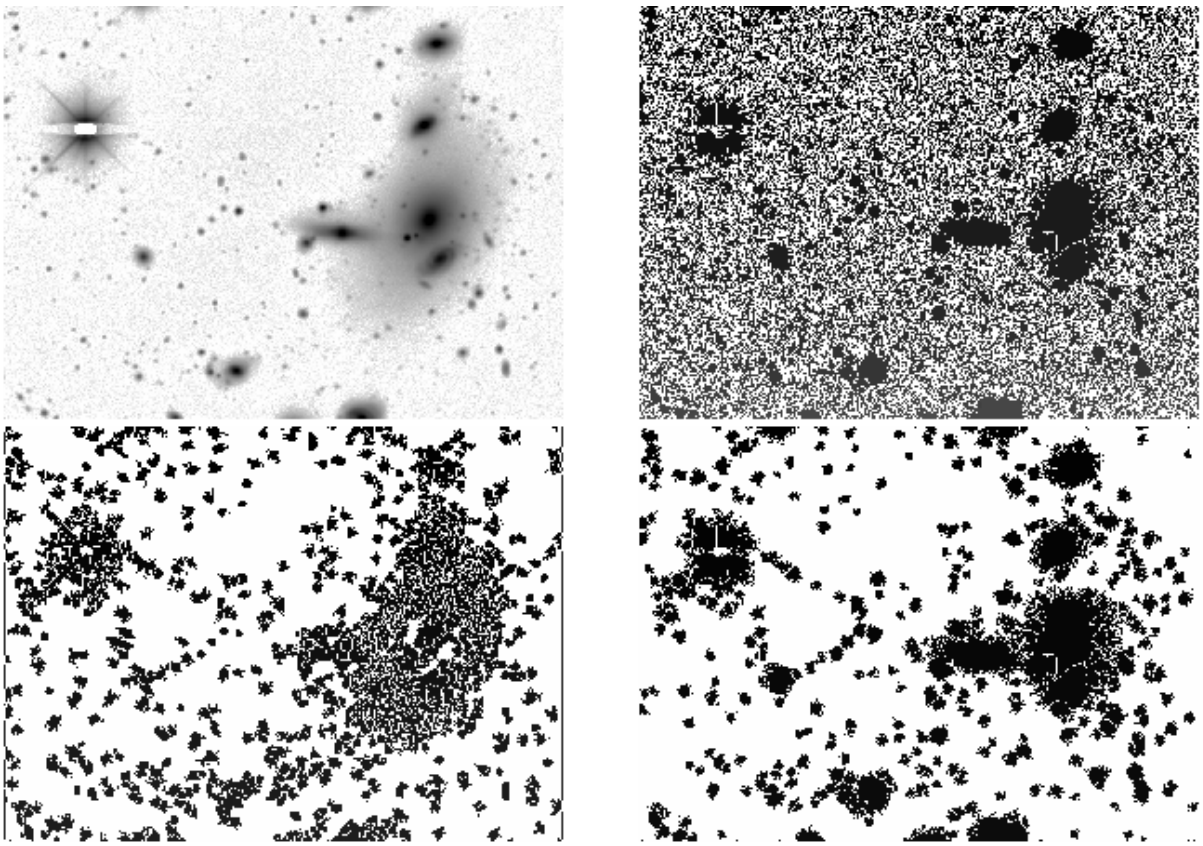

Fig. 3.— Upper left: Original DES coadd image, same as the one shown in Figure 1. The image is slightly zoomed in to best illustrate the segmentation procedure. Upper right: The segmentation map derived using all the local intensity maxima. Lower left: The segmentation map derived using all the local maxima of the smoothed and weighted Laplacian map. Lower right: The segmentation map derived using the cross matched local maxima from the original image and from the smoothed and weighted Laplacian image. In the three segmentation maps, the pixels in white belong to the "boundary" (see Section 3.1.1 for definition) and the pixels in black are associated with the seeds (local maxima from the intensity map, from the smoothed and weighted laplacian map, and the cross matching between the two).

Note that this procedure differs from the original implementation of Meyer's algorithm. When the original seed list includes **all** of the local maxima (as shown in the upper right panel of Figure 3), the procedure above gives the same result as Meyer's, though it is much faster. When the seed list does not include all of the local maxima, as will be the case in our GAIN application, this algorithm may yield thick boundaries (shown in the lower panels of Figure 3). This may not be desirable for computer vision applications like those for which Meyer's watershed algorithm was invented, but it is perfectly acceptable for our purposes.

After the segmentation is complete, we count the number of pixels labeled as belonging to each element of the original seed list. This number reflects the total area associated with each seed by the algorithm. We then purge false detections by eliminating seeds with segmentation areas smaller than some threshold value (for example, as purge all seed maxima associated with fewer than 27 pixels in DES coadd images).

3.1.2. Cross Matching to Laplacian Maxima

To populate an effective list of sources for deblending, GAIN generates a Laplacian map (3×3 pixels laplacian) of the original image. This is a useful approach because the contrast between real objects and their local background can be greatly enhanced in a Laplacian map. In terrestrial image processing applications, the Laplacian of Gaussian method is often used for edge detection (Lindeberg 1993). In these applications, an image is first smoothed on some scale with a Gaussian kernel, and then the Laplacian of the resulting image is calculated. However, in astronomical images, real objects already have had their spatial extent smoothed by an instrumental PSF, so that finding features in the Laplacian is an effective approach to object detection.

One limitation of this approach is that extended, low contrast sources, are not prominent in the Laplacian map. Since many of the sources indeed have low surface-brightness, we ameliorate this problem by weighting the Laplacian map with a transformed version of the original image intensity map. This weight map, denoted as $\mathrm{I}_w(x, y)$, is computed from the original image intensity map $\mathrm{I}(x, y)$ as,

$$\mathrm{I}'(x, y) = \mathrm{I}(x, y) - \min(\mathrm{I})$$
$$\mathrm{I}_w(x, y) = \begin{cases} \mathrm{I}'(x, y), & \text{if } \mathrm{I}'(x, y) \leq \mathrm{mean}(\mathrm{I}') \\ \log \frac{\mathrm{I}'(x,y)}{\mathrm{mean}(\mathrm{I}')} + \mathrm{mean}(\mathrm{I}'), & \text{otherwise.} \end{cases} \tag{1}$$

We use the logarithmic values for high intensity pixels to suppress their weight. The weighted Laplacian map is then derived from the raw Laplacian map $\mathrm{L}(x, y)$ and the image weight map $\mathrm{I}_w(x, y)$, as

$$\mathrm{L}_w(x, y) = \mathrm{L}(x, y)\mathrm{I}_w(x, y). \tag{2}$$

Unlike the original image intensity map, where pixel values span range across many orders of magnitude, the pixel values of $\mathrm{I}_w(x, y)$ span a much narrower range. Finally, we smooth $\mathrm{L}_w(x, y)$ with a Gaussian function ($\sigma = 1$ pixel for application to DES coadd images). Local maxima identified in this smoothed, weighted Laplacian image then become the seeds for the segmentation and purging step described in Section 3.1.1. For illustration, we show the smoothed, weighted Laplacian image of a DES coadd image in Figure 4. We also show the segmentation result using local maxima from this smoothed, weighted Laplacian image in the lower left panel of Figure 3.

3.1.3. Further Purging

As a final step of cleaning, we cross match the complete list of local maxima from the original image with the list of local maxima from the smoothed and weighted Laplacian image. We retain those local maxima which have corresponding peaks (within 2 pixels or 0.53") in the smoothed and weighted Laplacian image. This list of matched local maxima then becomes the seed list for the segmentation process (illustrated in the lower right panel of Figure 3). Local maxima associated with sufficiently large pixel lists in the segmentation map then form the final list of GAIN detected objects.

3.1.4. Matching to User Supplied Catalog

The usual application of GAIN follows an initial round of source extraction using tools like SExtractor. GAIN aims to search for additional blended sources missed by these applications. For

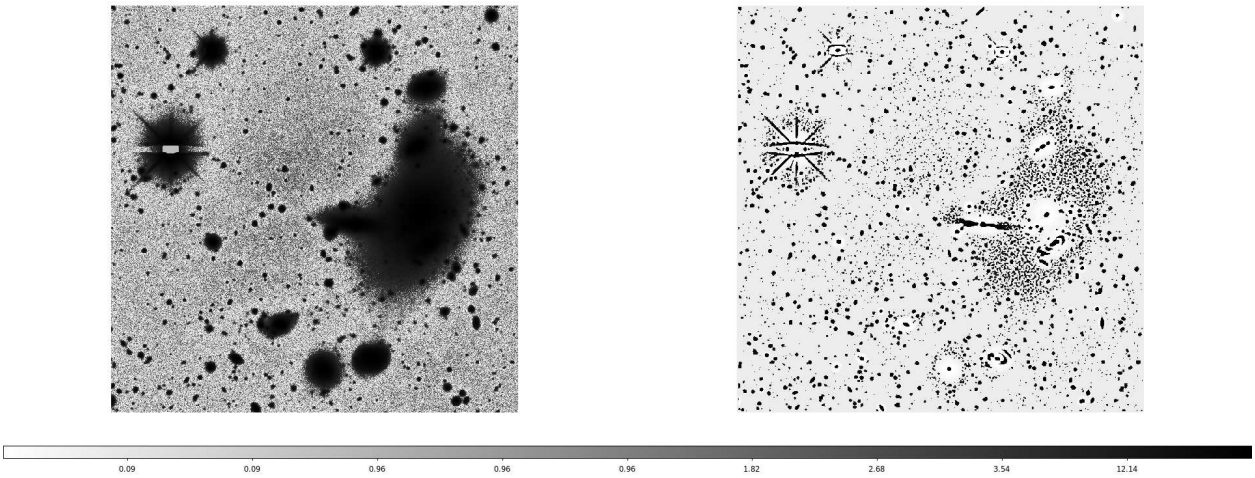

Fig. 4.— Left: Original DES coadd image, same as the one shown in Figure 1. The image is slightly zoomed in and shown at the "Histogram" scale of DS9. Right: The weighted Laplacian map of the same image derived in Section 3.1.2. The right image is also shown at the "Histogram" scale of DS9. Note that the substructure feature of the original image, like the spirals of the spiral galaxies at the bottom right, is enhanced in a weighted Laplacian map.

this reason, we match the GAIN source list to the user supplied source list, and single out those sources not identified by the original reduction as a list of newly identified, deblended sources.

For this matching procedure, the (x, y) coordinates of the user supplied sources are taken as input. Then in descending order of the intensity value at the user supplied sources' coordinates, we search for each source's nearest match in our source list from Section 3.1. All matches with a separation less than a threshold value (set to 10 pixels or 2.7" for our application to the DES data) are considered valid matches. The matched source is then removed from our list and the matching process continues. After removing all GAIN sources which match the user supplied list, the remaining sources constitute our list of newly identified, deblended sources.

3.2. Photometry for Blended Sources

3.2.1. Separating the Light

The most difficult aspect of blended object photometry is untangling the relative contributions of light from multiple sources. It is not uncommon for a faint source near a bright object to sit atop a background with photon count equal to – or in extreme cases several times higher than – the photon count of the source. To measure photometry for blended objects, we must meticulously account for the light contributed by their neighbors. For this purpose, we use an image inpainting technique which is originally developed in the computer vision field. In this field, many techniques have been developed to recover parts of an image that is deteriorated, or remove components that are unwanted. Our method is inspired by the Telea (2004) technique which is used to "inpaint" an image, i.e., to recover the texture of a small patch of an image from its surrounding pixels. It cannot re-create new patterns in the to-be-filled region, but rather fills them with a smooth background through interpolation. The problem this technique tries to solve is similar to our light separation problem. It allows us to estimate the light contribution from the more extended source in the blended pixels. We choose the Telea (2004) technique over a variety of other available approaches (Bertalmio et al. 2000; Criminisi et al. 2004) because it is computationally efficient and has been extensively studied.

The Telea (2004) technique works as follows. Given an intensity map, an unknown pixel can be inpainted with a value approximated from its

known neighbors. Telea (2004) developed an efficient way to prioritize pixels in an unknown patch and determine the order in which they are inpainted, starting from the pixels nearest to the boundary and progressing inward (Sethian 1996). This technique explicitly maintains a *narrow band* of pixels to be filled in as one of its features. Our implementation is adapted from the Telea (2004) technique, with a few modifications. A brief description follows.

1. Identify the regions to be inpainted: for each deblended source, this region is defined as a circle centered on the object, with an area equal to the segmentation area from Section 3.1.1. Pixels in this region are labeled as *unknown*, and the rest as *known*.

2. Initiate the *narrow band*: the *narrow band* is a list of pixels originally identified as *unknown*, that have at least one neighbor labeled as *known*. Pixels in the *narrow band* are prioritized in the ascending order of their original intensity value.

3. Begin inpainting: select the highest priority pixel from the *narrow band* and inpaint it. We explain how this is done in the next item. After inpainting this pixel, label it as *known*, and check if it has any *unknown* neighbors. If there are any, add them to the *narrow band* list, re-prioritize the *narrow band*, and repeat this step until the *narrow band* is empty.

4. Inpaint a pixel: to inpaint one pixel, we fill it with the zeroth order approximation value from its known neighbors,

$$\mathrm{I}(q) = \frac{\Sigma_p w_p \times \mathrm{I}(p)}{\Sigma_p w_p} \tag{3}$$

   with w_p being the weighting for each neighboring pixel.

At step 2, we prioritize the *narrow band* pixels according to their image intensity values. This is fundamentally different from the implementation of Telea (2004) algorithm in that Telea (2004) prioritize the *narrow band* in the order of pixels' distance to the original *known* and *unknown* region boundary, while the distance to the boundary is calculated from the fast marching solution to the Eikonal Equation (Sethian 1996). This is designed to mimic the practice in manual inpainting that the pixels closest to the *known* region are filled first (Bertalmio et al. 2000, 2001). However, we prioritize pixels for inpainting using their intensity value rather than their distance to the *known* region. This is because pixels with lower intensity are less affected by the astronomical object we are removing, and filling them in first allows for more reliable background reconstruction. Also, we use Equation 3 for step 4 with zeroth order approximation rather than the first order approximation used in Telea (2004) because astronomical images are noisy, and the derivatives at pixel scale are unreliable.

After using the above method for "background" interpolation for one object, the interpolated image contains light from the object's neighbors. The residual between the original image and the interpolated image contains the extracted light for this deblended object.

3.2.2. From Image to Catalog

To produce a useful catalog of deblended sources, we need to measure magnitudes and shapes, as well as to classify each as a star or a galaxy. Many packages capable of doing this are available (Bertin and Arnouts 1996; Simard et al. 2002; Peng et al. 2002, 2010), and a user might choose their favorite. For our application to the DES data, we use SExtractor (Bertin and Arnouts 1996). The tests described in the following sections are also based on the application of SExtractor. We provide the wrapper code for such an application in our package. Applications of other software, like GALFIT(Peng et al. 2002, 2010), are possible. When choosing software for cataloging, we advise people to consider a few things:

1. The light extracted image of one object is usually smaller than the area that contains **all** of its light. One should consider how to reconstruct/account for the light of the object outside this region.

2. Because of the above constraint, photometry from model fitting is probably more appropriate. With our application of SExtractor, we find that Kron (Kron 1980) and Petrosian (Petrosian 1976) magnitudes can provide reliable photometry when proper corrections are made for reconstructing total light.

3. Some star/galaxy separation methods may not work on light extracted images. When we use SExtractor on such images, we find that the class_star quantity fails most of the time because of the small area used for light separation. Star/galaxy separation using SExtractor spread_model quantity is still effective.

4. Methods Verification

We verify the performance of the following aspects of GAIN : photometry measurement, source detection completeness, and source detection purity. Because of the importance of star/galaxy separation in extra-galactic science, we also include a modest test of it in the photometry measurement test.

We want to test this package on optical images with complex deblending challenges, while all the sources in these images are known and already reliably measured. Our principal goal is to improve deblending around brightest cluster galaxies, so we designed a test to simulate this challenge by adding simulated BCGs to real deep optical images. The images we use are coadded images taken with DECam on the Victor M. Blanco telescope by the Dark Energy Survey. Cleaning of individual exposures, coaddition of the images, and initial extraction of sources are all done using standard data processing pipelines of the DES collaboration (Mohr et al. 2012). We then select regions of these images with no bright stars or real BCGs, so that deblending is not an important issue before the addition of a simulated BCG. Object catalogs extracted from these images are then used as "truth tables" in our testing process. When we add simulated BCGs to these images, some objects which are initially isolated and clean become blended, giving us a well understood deblending challenge to test against.

In Figure 5, we show an image before and after adding simulated BCGs. For the results presented in this section, we make simulated galaxy images of Sersic profile with Sersic index n = 4 and Sersic radius Re = $10''$ at ~ 19.0 magnitude (exact values vary depending on the metrics). We convolve these profiles with the PSF function of one coadd image and add them into that coadd image. We also ran the test with simulated galaxies of different magnitudes and different Sersic parameters. Since the results are weakly sensitive to these changes, we present results for just the model described above. In this test, we combine GAIN with SExtractor for photometry measurement. We compare the performance of this application to application of solely using SExtractor when possible. Throughout the test, the major parameter of SExtractor's deblending procedure, DEBLEND_MINCONT is set at 0.001, which is found to be optimum for processing Dark Energy Survey early data.

We note that the algorithm we describe in this paper is designed for deblending between a satellite object and its much brighter neighbors. It may also be desirable to deblend closely spaced pairs and triples of astronomical objects. GAIN can indeed help this kind of deblending problem, but the performance of GAIN on this aspect remains un-verified. Because pairs or triplets do not always cause local maxima, it is hard to distinguish them from extended sources without using the image PSF (which is being implemented in a future version of SExtractor). GAIN is not designed for this kind of deblending.

Finally, to thoroughly evaluate the deblending problem, the performance of GAIN and more importantly the detection and photometry measurement of cluster galaxies in DES, one may wish to compare a data set with much higher resolution to DES data. A project of comparing HST and DES data to study the DES data processing performance and its scientific effects has been undertaken within the DES collaboration (Palmese, Lahav, et al., private communication).

4.1. Photometry Measurement

To test whether GAIN can improve photometry measurement for blended sources, we use it to measure sources that become blended in the altered coadd image. We run SExtractor on the unaltered image as well as the image with simulated BCGs. We select sources that were clean in the original image (SExtractor flag = 0) but become blended (SExtractor flag = 3) upon the addition of simulated BCGs for photometry measurement comparison. In the original image, the selection standard ensures that these sources were free from blending. Therefore, their photometry measurement from the original image can be used as a truth table. In the images altered with sim-

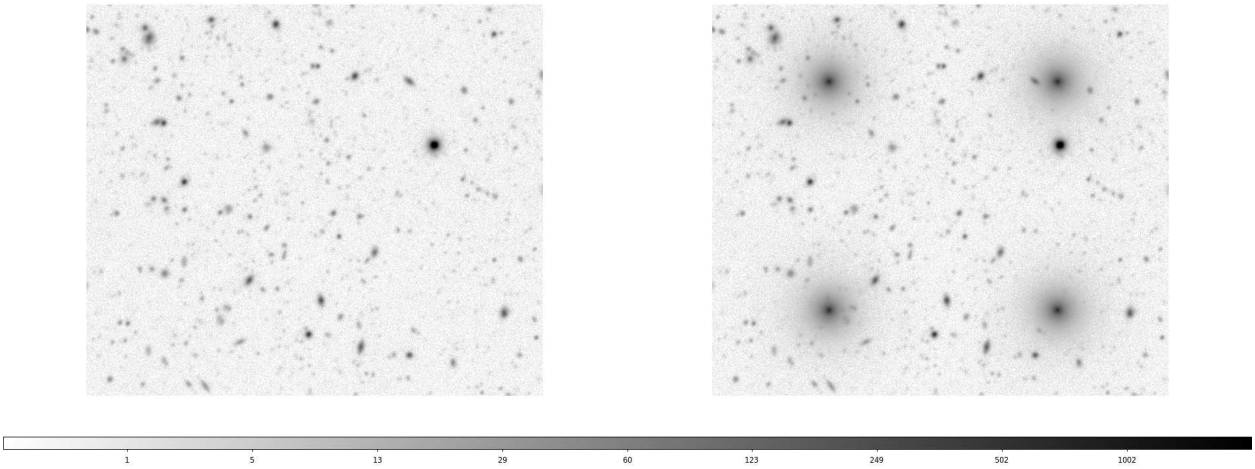

Fig. 5.— Left: An r band DES coadd image. The 10 sigma limiting magnitude of this image is 25.3 mag as measured from the SExtractor mag_auto uncertainty. Right: The same image after adding simulated BCGs. The size of these two images is approximately $3.15' \times 3.15'$. The apparent magnitude of the simulated BCGs is approximately ~ 19 mag (exact values vary depending on the metrics).

ulated BCGs, we compare measurements of these sources from SExtractor directly and from SExtractor with GAIN implementation to their "truth table" values. The result is shown in Figure 6 (a)(b).

Because light from simulated BCGs in the altered image is not completely accounted for in the basic SExtractor reductions, the blended sources typically have their brightness significantly overestimated, often by as much as much as 0.5 mag. Comparing to model magnitudes (Bertin 2011; Desai et al. 2012) in the truth table, Kron magnitudes (mag_auto) from the blended image are subject to more bias than model magnitudes. When GAIN is implemented, results become much better: either model magnitudes or Kron magnitude for these "artificially" blended sources are almost unbiased. Potential bias, if there is any, is below 0.05 mag.

The photometry measurement from SExtractor in Figure 6 are obtained with SExtractor "global background" setting. One might think that the photometry measurement biases in blended images will go away with SExtractor "local background" setting, but our tests show that this is not the case. In Figure 7, we show comparisons adopting SExtractor local background setting. A local background setting does diminish the biases, but is not sufficient to eliminate them. In addition, the scatter of photometry measurement is much larger in local background setting.

In addition to magnitude measurements, we also include a modest comparison of star/galaxy separation parameters for these sources. This is shown in Figure 6 (c)(d) and Figure 7 (c)(d). We find that the class_star quantity has become ineffective for star/galaxy separation as discussed in Section 3.2.2. Another star/galaxy separation parameter spread_model seems to remain effective, but has large scatter associated with it. Users who require precise star/galaxy separation may wish to consider other approaches.

4.2. Purity and Completeness

To test for completeness (did we recover all real objects) and purity (are all the new deblended objects real), we again utilize images containing simulated BCGs.

After introducing a deblending challenge into optical images with simulated BCGs, issues associated with deblending procedures emerge: object detection becomes incomplete and spurious detections appear. We use GAIN to improve deblending, then verify the completeness and purity. We point out here that the face values of completeness

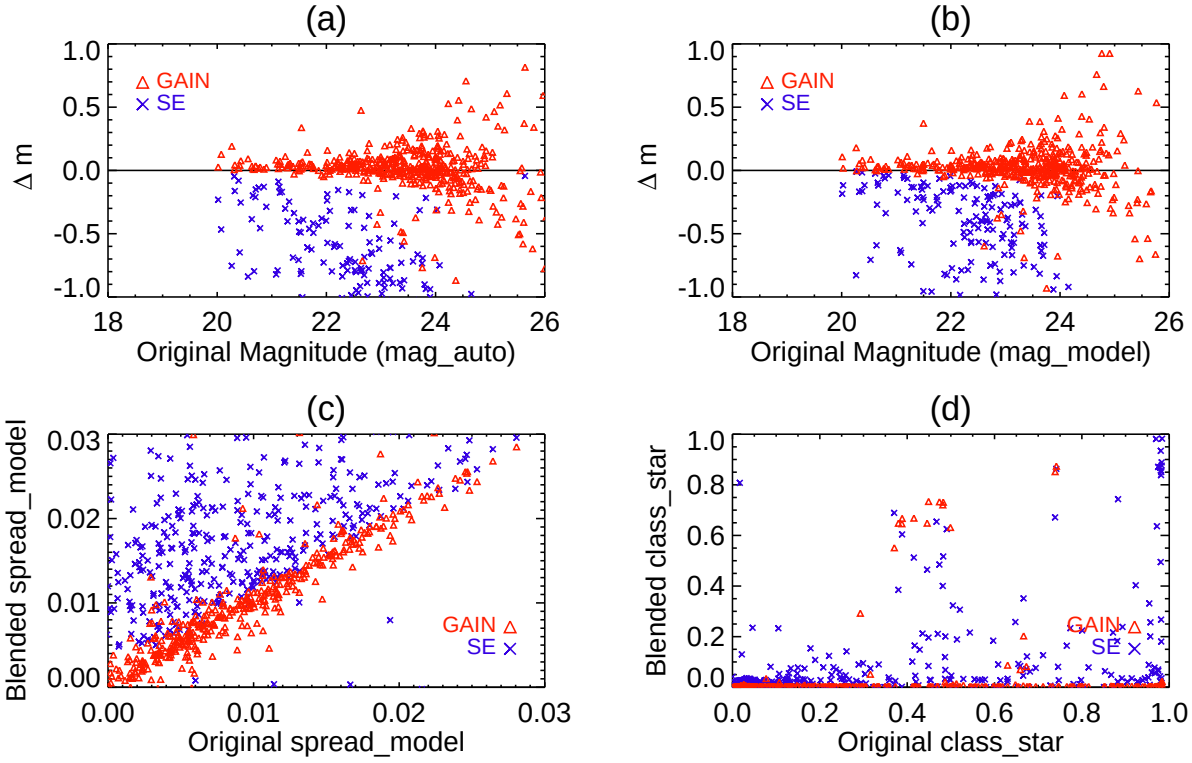

Fig. 6.— Comparison of photometry measurements and star/galaxy separation quantities with SExtractor using the "global background" setting. (a)(b) Offsets between truth table magnitudes and magnitude measurements from altered images for "artificially" blended objects. (c)(d) Comparison of star/galaxy separation quantities for "artificially" blended objects.

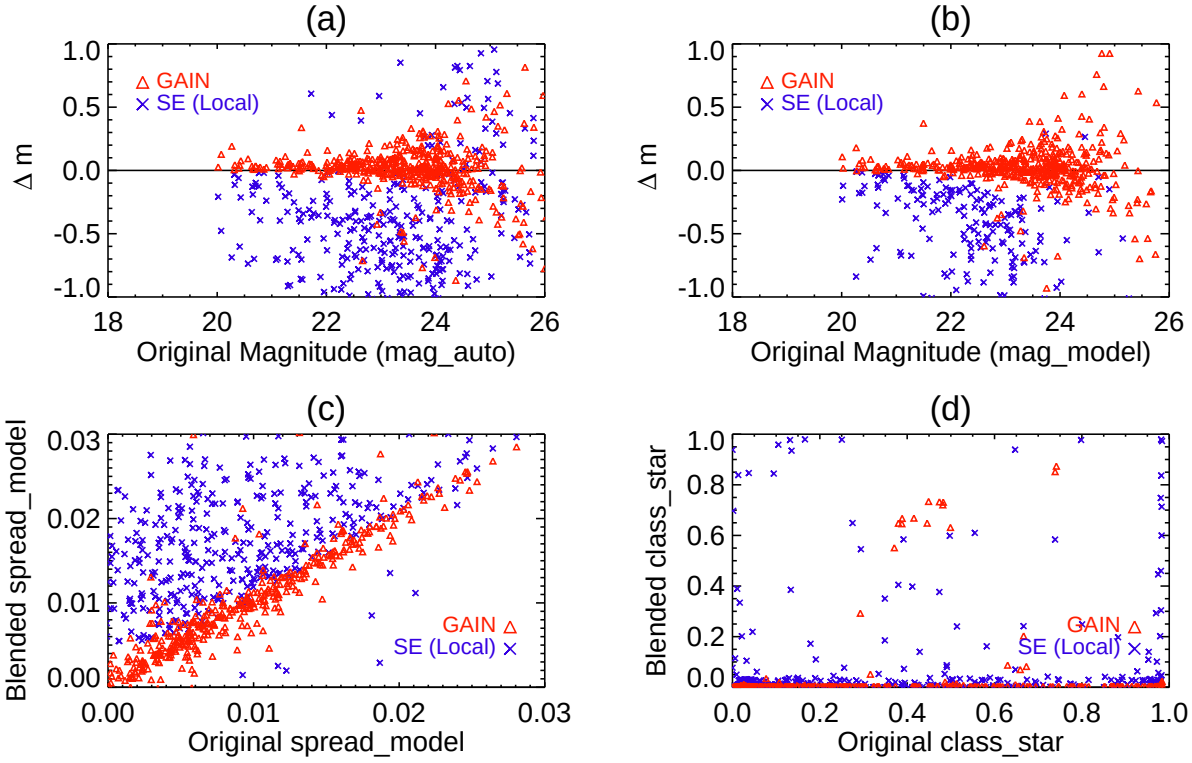

Fig. 7.— Comparison of photometry measurements and star/galaxy separation quantities with SExtractor using the "local background" setting. (a)(b) Offsets between truth table magnitudes and magnitude measurements from altered images for "artificially" blended objects. (c)(d) Comparison of star/galaxy separation quantities for "artificially" blended objects.

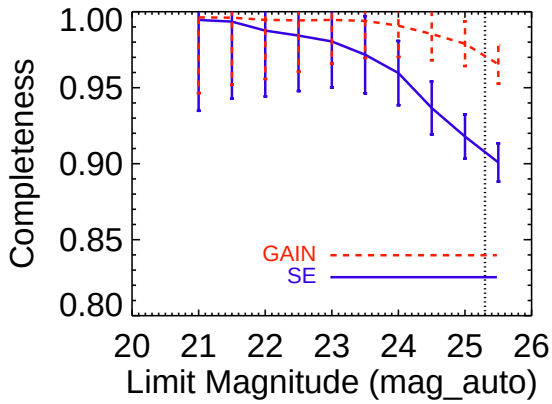

Fig. 8.— Completeness of the C1 catalog (blue solid line, from SExtractor) and completeness of the combination (red dashed line) of C1 and C2 (C2 from GAIN) as computed in Section 4.2.1. The non-negligible incompleteness of SExtractor cataloging can be improved by GAIN all the way to 25.5 mag. The vertical dotted line marks the 10 sigma limiting magnitude of the image that the test is performed on. The errors in this plot are estimated through assuming poisson distribution for the samples.

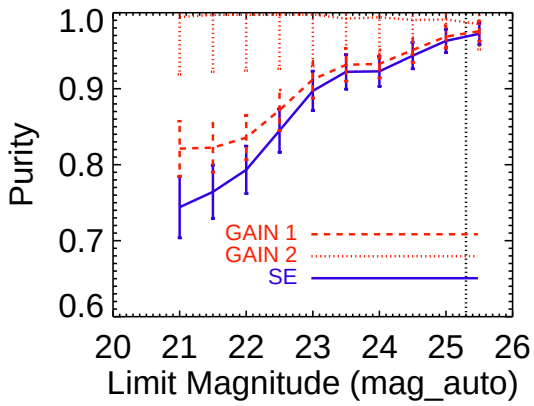

Fig. 9.— Purity of C1 (blue Solid line) by SExtractor and C2 (red dotted line) by GAIN and purity of the combination of C1 and C2 (red dashed line) as computed in Section 4.2.2. The purity of the combination of C1 and C2 is dragged down by the purity of C1. The vertical dotted line marks the 10 sigma limiting magnitude of the image that the test is performed on. The errors in this plot are estimated through assuming poisson distribution for the samples.

and purity presented in this section should not be taken as estimations for real astronomical images, as we are imposing exaggerated deblending difficulty on the testing images.

4.2.1. Completeness

For the completeness test, we first run SExtractor on the unaltered coadd image, and call the resulting catalog Truth Table 1 (TT1). The GAIN run on this original, unaltered coadd produces a catalog of blended sources not detected in TT1, and we call this list of additional sources Truth Table 2 (TT2). The combination of TT1 and TT2 is then used as the total truth table (TT) in our completeness and purity tests. We then insert simulated BCGs, run SExtractor on the image to produce Catalog 1 (C1), and run the GAIN package to search for blended sources and extract another catalog called Catalog 2 (C2). We then take all objects brighter than some magnitude from TT and match them to the C1 and C2 catalogs derived from the image containing simulated BCGs.

The matching is done in descending order of brightness for objects in the Truth Tables. For one object in the TTs, we search for the object that is nearest in C1 or C2. If the nearest neighbor from C1 or C2 is separated less than 5 pixels to the TTs object, we claim this object as *matched*. Once a C1/C2 object is matched, it is removed from the list available for matching. The matching procedure is done at different magnitude limit for the TTs sample. To ensure that the completeness evaluation is least subject to photometry measurement scatters, When selecting TTs sample above certain magnitude, we ease the magnitude limit on C1 or C2 by 1 mag. We evaluate completeness by computing the ratio between the number of *matched* objects in a TTs sample and the *total* number of objects in the TTs sample. We compute this quantity for SExtractor by matching TT1 to C1 and for deblending improved catalogs using our package by matching TTs to the combination of C1 and C2. The result is shown in Figure 8.

After the image is altered by simulated BCGs, a small but non-negligible fraction of sources are missed from SExtractor data production, especially at the faint end. After improving deblending with our package, the fraction of missed sources can be lowered to a satisfying level out to magnitude 25.5.

4.2.2. Purity upon Deblending

The purity is tricky to test as there is no clear definition for "real" astronomical objects in our test. Here we focus on evaluating the number of spurious detections introduced by the deblending procedure rather than categorizing objects as "real" or not.

To test the purity of the catalogs, we match C1 or C2 from altered coadd images to Truth Tables from un-altered coadd images. The procedure is similar to the completeness test in Section 4.2.1, except that we match Catalogs to the Truth Tables instead of match Truth Tables to Catalogs. When selecting C1 or C2 sample above certain magnitude limit, we also ease the magnitude limit on TTs by 1mag. We match C1 to T1, C2 to the combination of T1 and T2 and also the combination of C1 and C2 to the combination of T1 and T2. To evaluate the number of spurious detections produced during deblending, we calculate purity as the ratio between the number of *matched* objects in C1 or C2 sample and the *total* number of objects in the sample. The result is shown in Figure 9. In this plot, GAIN 1 is the computed purity for the combination of C1 and C2, while GAIN2 is the computed purity of C2 alone. The purity of SExtractor (SE) is evaluated as the purity of C1 alone.

In Figure 9, purity of C1 (SE) lowers toward the bright end, and is less than 80% at magnitude 21. Purity of the combination of C1 and C2 (GAIN 1) is also un-satisfying dragged by spurious detections in C1. For C2 (GAIN 2), the sources contained in C2 are consistent with sources contained in T1 and T2 to 99%. Figure 9 indicates that while performing the deblending procedure, SExtractor is likely to introduce spurious detections but GAIN is not.

In Figure 9, part of the explanation to C1's unsatisfying purity is that blended objects have deteriorated photometry and astrometry measurement (see discussion in 4.1), that the matching procedure fails for these sources. This tension can be relaxed if using GAIN to remeasure photometry for objects that have "blended" flag from SExtractor.

5. Summary

Deep astronomical images face deblending challenges, especially in the crowded cores of galaxy

clusters. Current deblending algorithms are not optimized to handle this problem. To take full advantage of the opportunity offered by new surveys like the DES, we need better methods for extracting accurate galaxy lists in cluster cores. In this paper, we describe a relatively simple approach to sorting out blended sources in these crowded regions. The design of this GAIN package includes two innovative features.

1. This package makes use of the Laplacian of an intensity image for blended source detection. In deblending procedures, one of the biggest challenges occurs when the intensity contrast between blended sources is too low to trigger detection. This problem can often be solved by making use of image intensity gradient. Future improvements in optical telescope data production software may wish to make use of the information in image intensity gradient to help deblending.

2. This package uses an interpolation technique to separate blended light from multiple sources. This is an improvement comparing to two popular approaches: simply assigning pixels to blended sources which is inaccurate but computationally efficient, and simultaneously fitting profiles of multiple sources which is accurate but computationally inefficient. Our method provides a nice balance between accuracy and efficiency.

We have tested this package on DES coadd images. Our tests shows that it can increase the reliability of photometry for blended objects. It can also increase the completeness of blended source detection, while introducing only a modest number of spurious detections. Upon application to DES data, GAIN has been used to improve cluster galaxy detection for red sequence photometric redshift measurement and accurate modeling of cluster central galaxy light profile. It is also possible to apply GAIN to HST and SDSS images, although the various threshold values need to be tuned.

The authors are pleased to acknowledge support of Department of Energy grant DE-FG02-95ER40899, and also U.S. Department of Energy contract to SLAC no. DE-AC02-76SF00515. We thank members of the Dark Energy Survey Data Management project, members of the Dark Energy Survey Cluster Working Group, members of the Dark Energy Survey Galaxy Evolution Working Group, and the University of Michigan Dark Energy Survey team for encouragement and suggestions that greatly facilitated progress on this work. We thank Gary Bernstein, Ofer Lahav, Gus Evrard, Nacho Sevilla, Tom Diehl, and Alistair Walker for comments on the paper draft. We use DES Science Verification in this paper. We are grateful for the extraordinary contributions of our CTIO colleagues and the DES Camera, Commissioning and Science Verification teams in achieving the excellent instrument and telescope conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the DES Data Management organization.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the as-

sociated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

This paper has gone through internal review by the DES collaboration.

REFERENCES

Adelman-McCarthy, J. K. et al. 2006, ApJS, 162, 38

Beard, S. M., MacGillivray, H. T., and Thanisch, P. F. 1990, MNRAS, 247, 311

Bertalmio, M., Bertozzi, A., and Sapiro, G. 2001, in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 1, p. I

Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. 2000, in Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00 (New York, NY, USA, ACM Press/Addison-Wesley Publishing Co.), p. 417

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, and A. H. Rots, p. 435

Bertin, E. and Arnouts, S. 1996, A&AS, 117, 393

Criminisi, A., Perez, P., and Toyama, K. 2004, Image Processing, IEEE Transactions on, 13, 1200

Desai, S. et al. 2012, ApJ, 757, 83

Diehl, H. et al. 2014, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, p. 91490V

Diolaiti, E., Bendinelli, O., Bonaccini, D., Close, L., Currie, D., and Parmeggiani, G. 2000, A&AS, 147, 335

Federici, L., Fusi Pecci, F., Zavaroni, C., Buonanno, R., Corsi, C. E., Quintiliani, A., and Bingham, E. A. E. 1983, Ap&SS, 90, 405

Flaugher, B. L. et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Jarvis, J. F. and Tyson, J. A. 1981, AJ, 86, 476

Kron, R. G. 1980, ApJS, 43, 305

Lindeberg, T. 1993, Scale-space theory in computer vision (Springer)

Lupton, R. 2001, AJ, submitted

Maddox, S. J., Efstathiou, G., Sutherland, W. J., and Loveday, J. 1990, MNRAS, 243, 692

Mandelbaum, R. et al. 2005, MNRAS, 361, 1287

Meyer, F. 1992, in Image Processing and its Applications, 1992., International Conference on, p. 303

Mohr, J. J. et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Peng, C. Y., Ho, L. C., Impey, C. D., and Rix, H.-W. 2002, AJ, 124, 266

Peng, C. Y., Ho, L. C., Impey, C. D., and Rix, H.-W. 2010, AJ, 139, 2097

Petrosian, V. 1976, ApJ, 209, L1

Savage, R. S. and Oliver, S. 2007, ApJ, 661, 1339

Sethian, J. A. 1996, Proceedings of the National Academy of Sciences, 93, 1591

Simard, L. et al. 2002, The Astrophysical Journal Supplement Series, 142, 1

Stetson, P. B. 1987, PASP, 99, 191

Telea, A. 2004, Journal of Graphics Tools, 9, 23

Yee, H. K. C. 1991, PASP, 103, 396

York, D. G. et al. 2000, AJ, 120, 1579

This 2-column preprint was prepared with the AAS LaTeX macros v5.2.