A catalog of automatically detected ring galaxy candidates in PanSTARRS

Ian Timmis and Lior Shamir

Lawrence Technological University, MI 48075

ABSTRACT

We developed and applied a computer analysis method to detect ring galaxy candidates in the first data release of PanSTARRS. The method works by applying a low-pass filter, followed by dynamic global thresholding to search for closed regions in the binary mask of each galaxy image. Applying the method to $\sim 3 \cdot 10^6$ PanSTARRS galaxy images produced a catalog of 186 ring galaxy candidates based on their visual appearance.

Subject headings: catalogs — techniques: image processing — methods: data analysis — galaxies: peculiar

1. Introduction

Ring galaxies are rare irregular galaxies that are not on the Hubble classifications scheme. Theys & Spiegel (1976) proposed a separate classification scheme for ring galaxies that includes three sub-classes based on their visual appearance: Empty ring galaxies (RE), Ring galaxies with off-center nucleolus (RN), and ring galaxies with knots or condensations (RK). They also identified that most, although not all, ring galaxies have a companion (Theys & Spiegel 1977). Few & Madore (1986) separated ring galaxies into two sub-classes: ‘’P-type” rings, which have a knotty structure or an off-center nucleolus, and “O-rings”, characterized by a smooth ring structure and a centered nucleolus.

Ring galaxies include polar rings (Whitmore et al. 1990; Macciò et al. 2005; Reshetnikov & Sotnikova 1997; Finkelman et al. 2012; Reshetnikov & Combes 2015), collisional rings (Appleton & Struck-Marcell 1996), and Hoag-type objects (Longo et al. 2012). The “Hoag’s Object” (Hoag 1950; Brosch 1985; Schweizer et al. 1987) was discovered in 1950, and its discovery was followed by the identification of other ring galaxies.

Catalogs of rings galaxies were created in the past by manual observation. The early Arp (1966) catalog of peculiar galaxies contains two galaxies with visual appearance of an empty ring. The catalog of southern peculiar galaxies (Arp & Madore 1988) includes 69 systems identified as rings. Whitmore et al. (1990) compiled a list of 157 polar ring galaxy candidates, and about half a dozen of these objects were confirmed as polar ring galaxies by kinematic follow-up observations (Finkelman et al. 2012). Madore et al. (2009) prepared an atlas of collisional ring galaxies. Garcia-Ribera et al. (2015) discovered 16 polar ring galaxy candidates. Buta (1995) created a catalog of Southern ring galaxies. Moiseev et al. (2011) used crowdsourcing and non-scientists volunteers to prepare a catalog of ring galaxy candidates through the Galaxy Zoo citizen science campaign.

While manual analysis performed by expert or citizen scientists has provided useful catalogs of ring galaxies, the rapidly increasing data acquisition power of digital sky surveys such as the Large Synoptic Survey Telescopes (LSST) can potentially allow the identification of a very large number of ring galaxies among a total of billions of astronomical objects. Due to the large size of these databases, effective identification of these objects requires automation, leading to the development of automatic methods of identifying peculiar objects in large databases of galaxy images (Shamir 2012; Shamir & Wallin 2014; Shamir 2016). Here we describe an automatic image analysis method that can identify ring galaxies, and apply the method to mine through $\sim 3{\cdot}10^6$ galaxies imaged by the Panoramic Survey Telescope and Rapid Response System (Hodapp et al. 2004; Flewelling et al. 2016; Chambers et al. 2016) to compile a catalog of ring galaxy candidates.

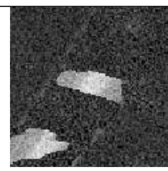
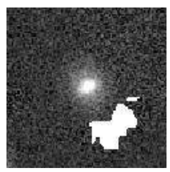
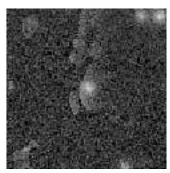
2. Methods

2.1. Data

The dataset was obtained from the Panoramic Survey Telescope and Rapid Response System (PanSTARRS) first data release (Hodapp et al. 2004; Flewelling et al. 2016; Chambers et al. 2016). The initial dataset includes 3,053,831 objects with r magnitude of less than 19. To avoid stars, the dataset included 2,394,452 objects identified as extended sources in all bands, and 659,379 additional objects that were not identified as extended objects in all bands, but their PSF magnitude subtracted by their Kron magnitude was larger than 0.05, and their r Petrosian radius was larger than 5.5". The images were then downloaded via the PanSTARRS *cutout* service as 120×120 JPG images, in a process similar to the image download done in (Kuminski & Shamir 2016). To avoid pressure on the PanSTARRS web server, one image was downloaded at a time, and therefore the processes required 62 days to complete.

Images that contain substantial noise or artifacts are difficult to analyze correctly, and can trigger false positives as will be explained in Section 2.2. Due to the large scale of the initial dataset, even a low rate of false detections can lead to an unmanageable resulting dataset. Because compression algorithms are more efficient when the signal is smooth, clean images of real galaxies tend to have a smaller compressed file size, and therefore artifacts and noisy images can be rejected by their compressed file size (Kuminski & Shamir 2016). Table 1 shows examples of galaxy images and their file sizes. Based on empirical observations, a threshold was set so that only images with file sizes of less than 5.5KB were analyzed, and larger files were rejected.

Table 1: Examples of clean galaxy images and artifacts or noisy images in PanSTARRS. The file size provides a simple mechanism to reject noisy images.

PanSTARRS object ID	File size (KB)	Image
102230806134866752	9.40	
103480451533225122	9.58	
103570759842751155	9.43	
100840464055080903	3.17	
104720155726185389	3.10	
104941422843081464	3.88	

2.2. Galaxy image analysis

Each image is smoothed by utilizing a median filter with window size of 5×5 to facilitate noise reduction, and converted to grayscale. The image is then converted into its binary mask using a dynamic threshold. The dynamic threshold starts with a minimum of 30, and is incremented iteratively until it reaches the gray level of 200. The conversion of the original ring galaxy into a binary map is displayed by Figure 1.

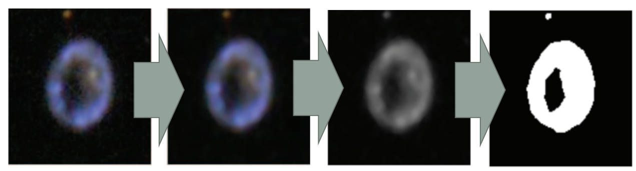


Fig. 1.— The stages in converting the original image into its binary map.

Fig. 2.— Dynamic thresholding binary maps. The Images show binary maps with thresholds of 20 (left), 44, 83, 99, and 115. The figure shows that when using a graylevel of 44 the ring is identified in the binary mask, while other graylevel thresholds show no ring.

For each threshold level the binary mask is computed, and a search for a ring inside the foreground is done using a Flood Fill algorithm (Asundi & Wensen 1998). Flood fill is an algorithm typically associated with the "bucket fill" tool in painting programs. Here we used a stack-based

4-connected version of the flood fill algorithm, which is a non-recursive process starting with an initial pixel and then analyzes the four pixels surrounding it. Each of these four pixels is flagged, and then the neighbors of each of them are also added. That continues until all pixels are flagged, or no neighbors with value of 0 remain. In that case it is determined that no path of pixels of value 0 to the edge exist, and therefore the image is suspected as a ring galaxy. However, if a pixel that is on the edge of the image is flagged, the algorithm stops and it is determined that no ring exists in that graylevel threshold.

The flood fill algorithm is applied for each pixel in the binary mask. If the flood fill algorithm finishes without reaching a pixel that is on the edge of the frame, the number of pixels in the closed area are counted, and divided by the number of foreground pixels. If the number of pixels in the closed area is less than 10% of the number of foreground pixels, it is assumed that the closed area is too small to be considered a ring galaxy. Figure 3 shows an example of closed areas in the binary mask that can be considered candidate rings (left), and small areas in the binary mask of the same image that are merely local grayscale variations (right).

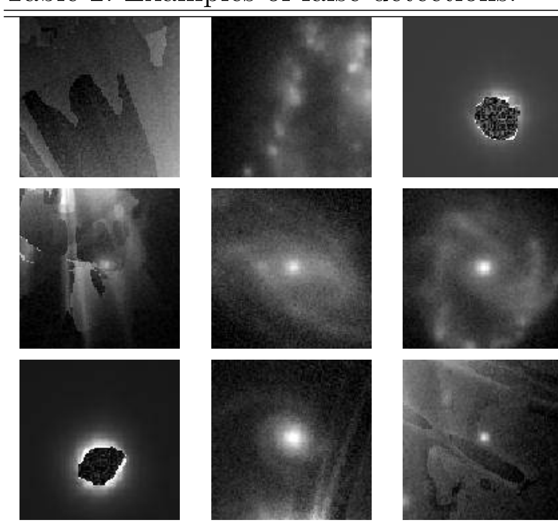
Fig. 3.— Comparison of the size of the closed area to the size of the foreground.

Processing of a small 120×120 galaxy image using a single core of an Intel Xeon E5-1650 requires ∼2.1 seconds to complete.

2.3. False detections

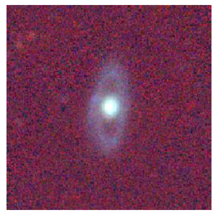
When mining through a very large number of galaxies, even a small rate of false detections can lead to an unmanageable database. Of over three million images that were tested, the algorithm detected 2490 galaxies in which manual inspection showed no ring. These galaxies included artifacts, saturated objects, and regular galaxies. Table 2 shows examples of false detections of the algorithm. As can be seen in the example images, overlapping arms or stars nearby a spiral galaxy can lead to false detections. Saturated objects can also be mistakenly identified as rings. However, these objects are fairly rare, and the false detection rate is less than 0.1% of the initial set of galaxies.

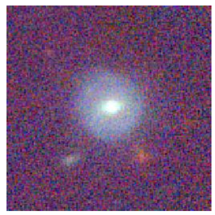
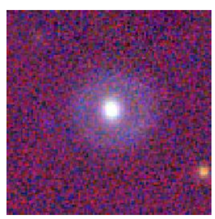
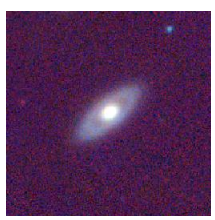
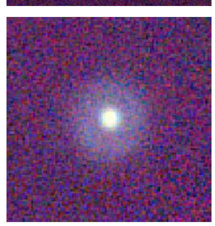
Table 2: Examples of false detections.

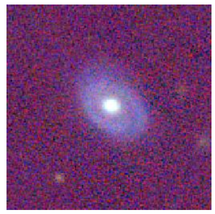
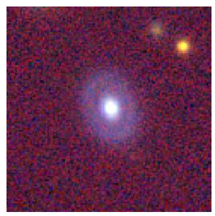


3. Ring galaxy candidates

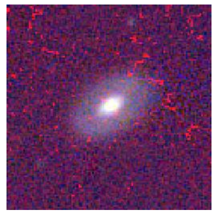
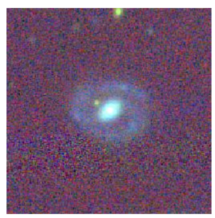
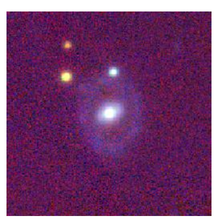
The ring galaxy candidates that were detected with their right ascension and declination coordinates are shown in Table 3. The table contains candidate ring galaxies with centered nucleolus, off-centered nucleolus, ring galaxies with knotty structure, interacting systems, and galaxies that have rings and arms.

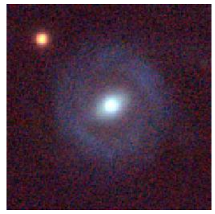
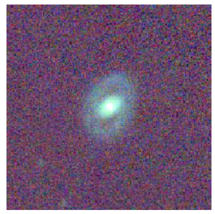
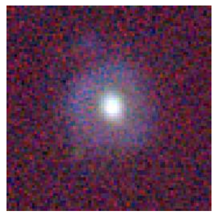
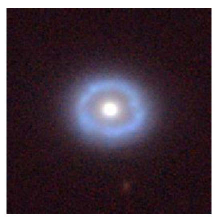
No.	Image			
1				
2				
3				
4				
5				
6				
7				

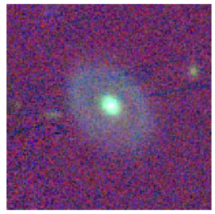
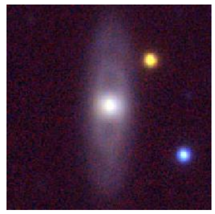
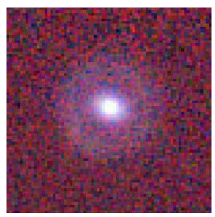
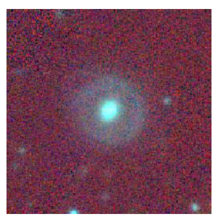
8				
9				
10				
11				
12				
13				
14				

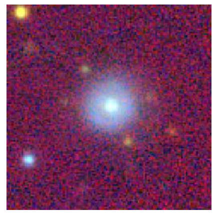
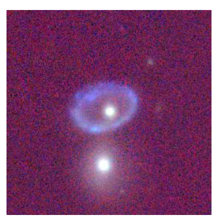
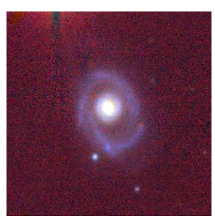
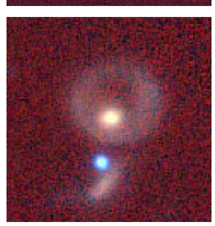
15				
16				
17				
18				
19				
20				
21				

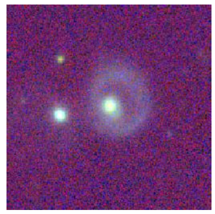
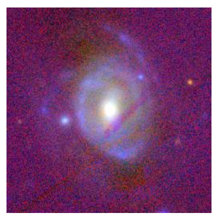
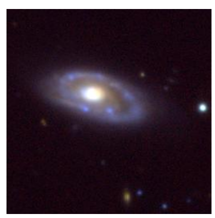
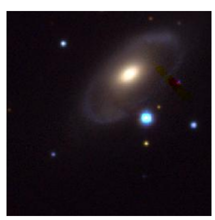
22				
23				
24				
25				
26				
27				
28				

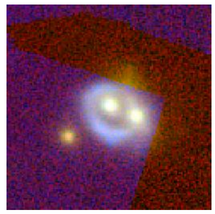
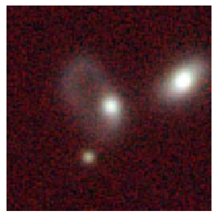
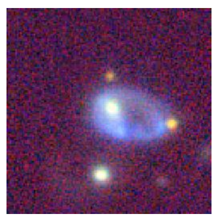
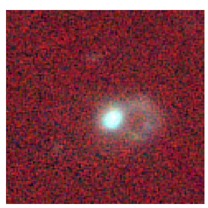
29				
30				
31				
32				
33				
34				
35				

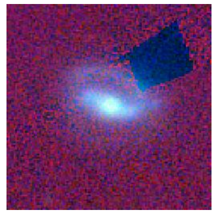
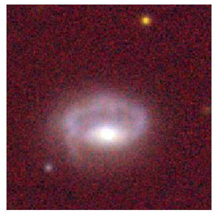
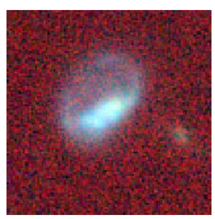
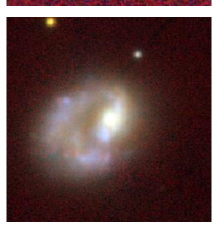
36				
37				
38				
39				
40				
41				
42				

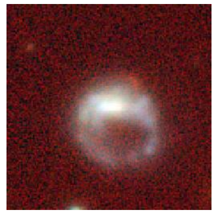
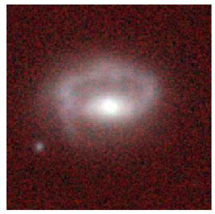
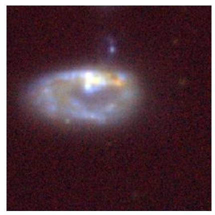
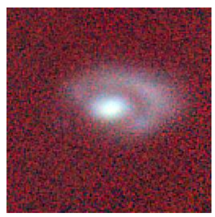
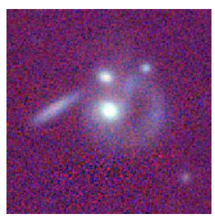
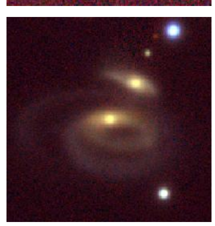
43				
44				
45				
46				
47				
48				
49				

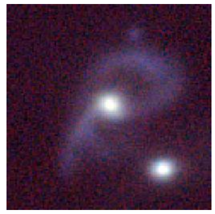
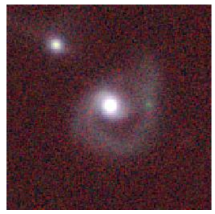
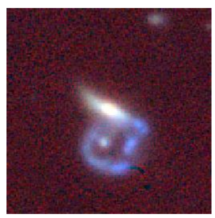
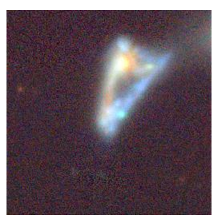
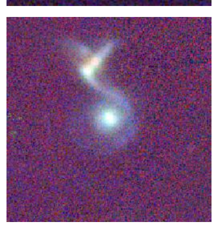
50				
51				
52				
53				
54				
55				
56				

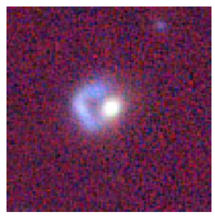
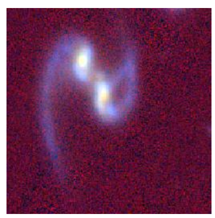
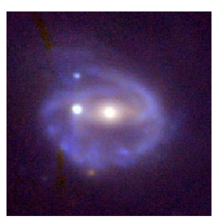
57				
58				
59				
60				
61				
62				
63				

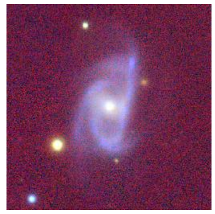
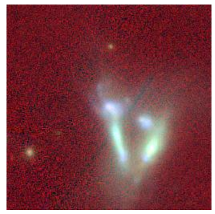
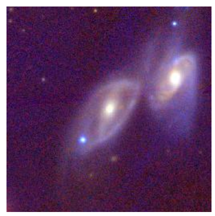
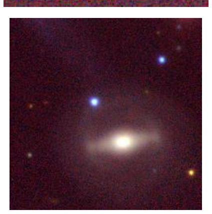
64				
65				
66				
67				
68				
69				
70				

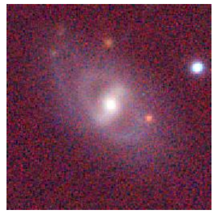
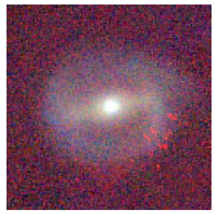
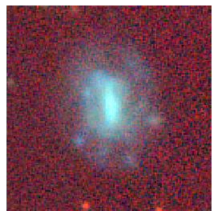
71				
72				
73				
74				
75				
76				
77				

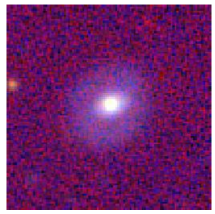
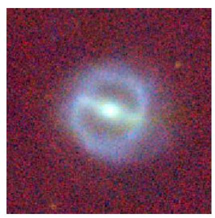
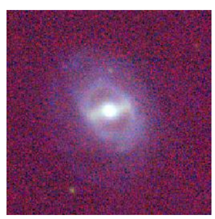
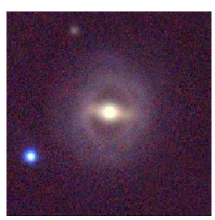
78				
79				
80				
81				
82				
83				
84				

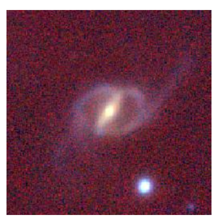
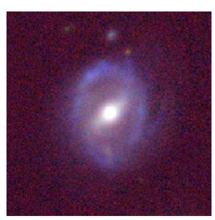
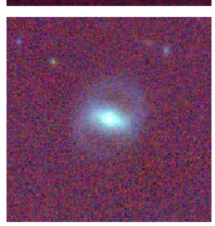
85				
86				
87				
88				
89				
90				
91				

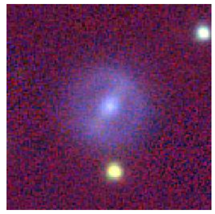
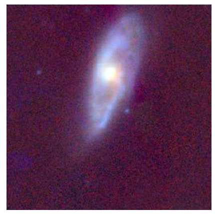
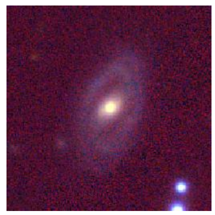
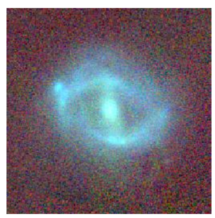
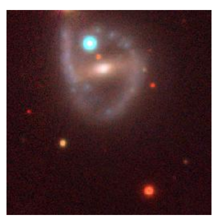
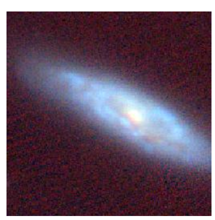
92	
93	
94	
95	
96	
97	
98	

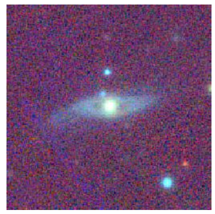
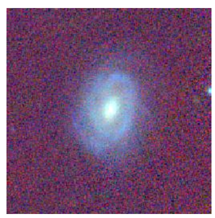
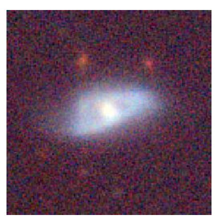
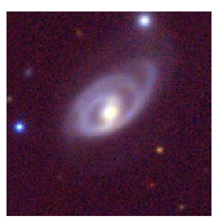
99				
100				
101				
102				
103				
104				
105				

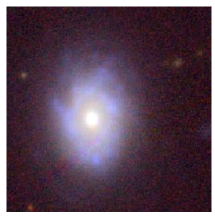
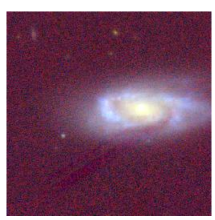
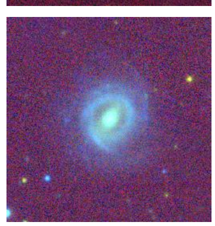
106				
107				
108				
109				
110				
111				
112				

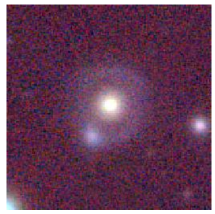
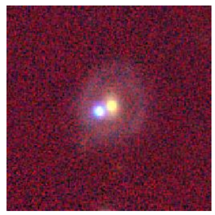
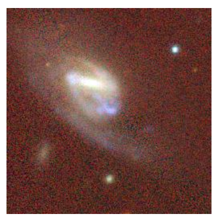
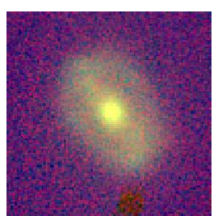
113				
114				
115				
116				
117				
118				
119				

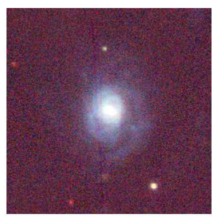
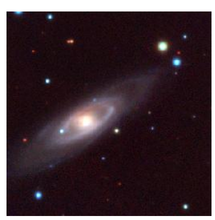
120				
121				
122				
123				
124				
125				
126				

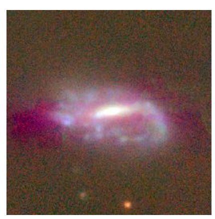
127					
128					
129					
130					
131					
132					
133					

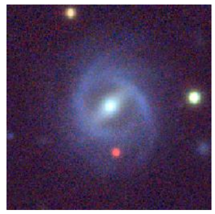
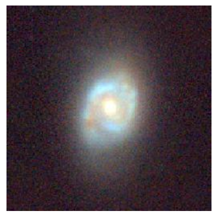
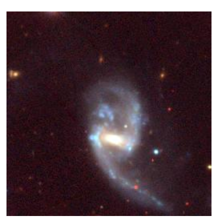
134				
135				
136				
137				
138				
139				
140				

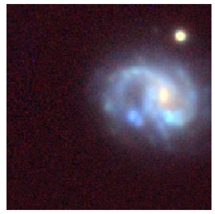
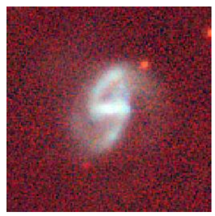
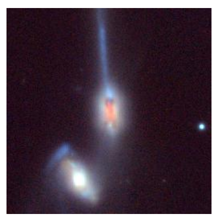
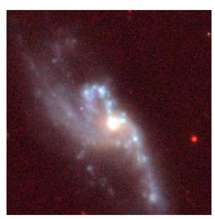
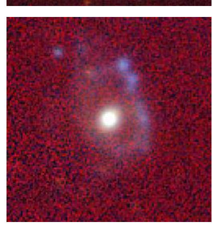
141				
142				
143				
144				
145				
146				
147				

148				
149				
150				
151				
152				
153				
154				

155				
156				
157				
158				
159				
160				
161				

162				
163				
164				
165				
166				
167				
168				

169					
170					
171					
172					
173					
174					
175					

176				
177				
178				
179				
180				
181				
182				

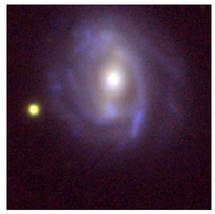
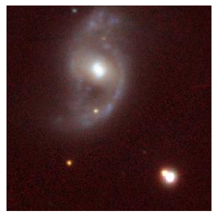
183				
184				
185				
186				

Table 3:: Ring galaxy candidates identified automatically

4. Conclusion

Autonomous sky surveys have enabled the acquisition of very large databases of image and other data, substantially increasing the discovery power of ground and space-based telescopes. To utilize the discovery power and turn these data into scientific discoveries, it is required to apply computational methods that can mine these very large databases. Since a substantial part of these data are in the form of images, full analysis of the data requires image analysis methodology. Here we use a simple and fast automatic image analysis method and apply it to the PanSTARRS first data release to detect ring galaxy candidates. Despite the simple nature of the image analysis method, it can find ring galaxies that are highly difficult to find without using automation, and it is sufficiently fast to be applied to much larger databases such as LSST.

5. Acknowledgments

This study was supported by NSF grant IIS-1546079.

The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

REFERENCES

Appleton, P., & Struck-Marcell, C. 1996, Fundamentals of Cosmic Physics, 16, 18

Arp, H. 1966, ApJS, 14, 1

Arp, H. C., & Madore, B. F. 1988, Contemporary Physics, 29, 99

Asundi, A., & Wensen, Z. 1998, Applied Optics, 37, 5416

Brosch, N. 1985, A&A, 153, 199

Buta, R. 1995, ApJS, 96, 39

Chambers, K. C., Magnier, E., Metcalfe, N., et al. 2016, arXiv preprint arXiv:1612.05560

Few, J. M., & Madore, B. F. 1986, MNRAS, 222, 673

Finkelman, I., SJ, J. G. F., & Brosch, N. 2012, MNRAS, 422, 2386

Flewelling, H., Magnier, E., Chambers, K., et al. 2016, arXiv preprint arXiv:1612.05243

Garcia-Ribera, E., Pérez-Montero, E., García-Benito, R., & Vílchez, J. 2015, in Highlights of Spanish Astrophysics VIII, Vol. 1, 372–372

Hoag, A. A. 1950, AJ, 55, 170

Hodapp, K., Kaiser, N., Aussel, H., et al. 2004, AN, 325, 636

Kuminski, E., & Shamir, L. 2016, ApJS, 223, 20

Longo, G., Capaccioli, M., & Busarello, G. 2012, Morphological and Physical Classification of Galaxies: Proceedings of the Fifth International Workshop of the Osservatorio Astronomico di Capodimonte, Vol. 178 (Springer Science & Business Media)

Macciò, A. V., Moore, B., & Stadel, J. 2005, ApJL, 636, L25

Madore, B. F., Nelson, E., & Petrillo, K. 2009, ApJS, 181, 572

Moiseev, A. V., Smirnova, K. I., Smirnova, A. A., & Reshetnikov, V. P. 2011, MNRAS, 418, 244

Reshetnikov, V., & Combes, F. 2015, MNRAS, 447, 2287

Reshetnikov, V., & Sotnikova, N. 1997, arXiv preprint astro-ph/9704047

Schweizer, F., Ford Jr, W. K., Jedrzejewski, R., & Giovanelli, R. 1987, ApJ, 320, 454

Shamir, L. 2012, Journal of Computational Science, 3, 181

—. 2016, PASP, 129, 024003

Shamir, L., & Wallin, J. 2014, MNRAS, 443, 3528

Theys, J., & Spiegel, E. 1976, ApJ, 208, 650

Theys, J. C., & Spiegel, E. 1977, ApJ, 212, 616

Whitmore, B. C., Lucas, R. A., McElroy, D. B., et al. 1990, AJ, 100, 1489

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