Automatic detection of full ring galaxy candidates in SDSS

Lior Shamir,^{1*}

¹Kansas State University, Manhattan, KS 65506, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

A full ring is a form of galaxy morphology that is not associated with a specific stage on the Hubble sequence. Digital sky surveys can collect many millions of galaxy images, and therefore even rare forms of galaxies are expected to be present in relatively large numbers in image databases created by digital sky surveys. Sloan Digital Sky Survey (SDSS) data release (DR) 14 contains $\sim 2.6 \cdot 10^6$ objects with spectra identified as galaxies. The method described in this paper applied automatic detection to identify a set of 443 ring galaxy candidates, 104 of them were already included in the Buta + 17 catalogue of ring galaxies in SDSS, but the majority of the galaxies are not included in previous catalogues. Machine analysis cannot yet match the superior pattern recognition abilities of the human brain, and even a small false positive rate makes automatic analysis impractical when scanning through millions of galaxies. Reducing the false positive rate also increases the true negative rate, and therefore the catalogue of ring galaxy candidates is not exhaustive. However, due to its clear advantage in speed, it can provide a large collection of galaxies that can be used for follow-up observations of objects with ring morphology.

 Keywords:
 Catalogs — techniques: image processing — methods: data analysis — galaxies:

 peculiar
 dozen colliding ring galaxies from SDSS based on reports of

1 INTRODUCTION

The deployment of autonomous digital sky surveys has enabled the creation of very large databases of galaxy images, and therefore even very rare types of galaxies are assumed to be present in these databases. One of the less common types of galaxies is ring galaxies. Ring galaxies can be separated into several

different types (?) such as bar-driven or tidially-driven resonance

rings (?), collisional rings (?), polar rings (?????), "Hoag-type" rings

(???), and spiral galaxies with ringed bars (?).

Ring galaxies can be classified by their visual morphology into

three major sub-classes (?): Empty rings (RE), rings with offcentre nucleolus (RN), and rings with knots or condensations (RK). Another classification scheme for ring galaxies based on their visual appearance separates ring galaxies into "O-rings", which have a smooth ring structure and a nucleolus in its centre, and "P-type" rings, which have a knotty structure or a nucleolus that is

not in the centre of the ring (?).

Some ring galaxy catalogues were created using manual analysis of the galaxies in the past six decades. The catalogue of peculiar galaxies of ? includes two empty ring galaxies, and the ? catalogue includes 69 ring galaxies. ? prepared a catalogue of a

volunteers in the Galaxy Zoo on-line forum. The ? catalogue included 157 polar ring galaxy candidates, and several of these galaxies were confirmed as polar rings (?). ? released an atlas of

collisional rings. ? identified 16 polar ring galaxy candidates. ?

collected a set of Southern ring galaxies. ? and (?) used citizen science annotations and classifications to identify ring galaxy candidates by using the Galaxy Zoo 1 and Galaxy Zoo 2 databases, respectively. These catalogues are efficient in the sense that they have good detection accuracy due to the superior ability of the human brain to analyze galaxy morphology, but because they require very intensive labour, even when using a large number of volunteers it is difficult to perform an exhaustive analysis of the entire image databases collected by modern digital sky surveys. That bandwidth limitation will be magnified when more powerful

sky surveys such as LSST see first light. ? used computer analysis to release a catalogue of 186 automatically identified ring galaxy candidates in PanSTARRS.

As digital sky surveys become increasingly more powerful, it is clear that manual analysis of the images is not sufficient for comprehensive detection of ring galaxies among millions of galaxy images.Thatreinforcestheuseofautomationtodetectringgalaxies. The ability to identify galaxy morphology automatically can lead to much larger collections of ring galaxies, which can also be useful whenmorepowerfuldigitalskysurveyssuchasLSSTstarttocollect data.

2 GALAXY IMAGE ANALYSIS METHOD

The data source used in this study is the set of galaxies with spectra in SDSS DR14. SDSS DR14 contains a total of $\sim 4.8 \cdot 10^6$ IDs of objects with spectra, and $\sim 2.6 \cdot 10^6$ of these objects are labeled by SDSS pipeline as galaxies. The mean redshift of these galaxies is 0.38 (σ =0.24), and the mean g magnitude is \sim 20.56 (σ =2.09). The

image of each galaxy was obtained by using the cutout service of

SDSS as was done in (?). In summary, the images are downloaded as 120×120 JPG colour images. Since galaxies have different sizes, each galaxy was downloaded several times until 25% or less of the pixels on the edges of the image have gray value of less than 125. The initial scale was set to 0.25" per pixel, and it increased by 0.05" until 25% or less of the pixels on the edges are not bright, which

means that the galaxy fits inside the image (?).

The JPG images are used because they combine information from the different bands, providing a simple image format that contains information about the morphology of each object in a manner that is easier to process by machine vision. While the original FITS format allows to make accurate photometric measurements, that accuracy is not required for machine vision systems for the purpose of broad morphological analysis. Therefore, the simple JPG format provides an efficient mechanism

for both manual (?) and automatic (??) analysis.

Downloading that large dataset of galaxy images required \sim 16 days. The image analysis method is similar to the method

used in **?**. Each image was converted to a binary map such that all pixels above the threshold were set to 1, and the pixels below the threshold were set to 0. The initial threshold was set to 50, and increased by five until it reached 200.

For each threshold, the image is inverted, and a 4-connected labeling algorithm is applied to label all objects in the inverted image. If more than one object is detected, it means that the image contained background areas that are inside foreground objects, and therefore could be rings. Since a galaxy can contain many small ares inside the arms, if the size of the background area is less than 10% of the foreground galaxy the algorithm ignores that background area and does not consider it as a ring candidate. The algorithm is implemented as part of the Ganalyzer galaxy image analysis tool

(??).

3 RING GALAXY CANDIDATES

The method described in Section ?? and also explained in (?)

detected ring galaxy candidates, as listed in Table ??. The galaxies are provided with their catalogue number, right ascension and declination of each object.

The images of the galaxies are shown by Figures ??, ??, ??, ??,

and??, showing candidate resonance ringgalaxies, collisional rings, rings with an off-centre nucleus, rings with no obvious nucleus, and other rings, respectively.



193	148.591	51.243	194	152.806	53.516		195	141.639	48.011	196	118.186	25.786
197	164.499	51.017	198	151.799	48.755		199	150,764	45.597	200	115.717	22.112
201	127 439	32 611	202	138.063	39 126		203	178 757	48 786	204	126 713	4 4197
201	127.435	52.011	202	138.005	55.120		203	176.757	48.780	204	120.713	4.4157
205	149.480	55.911	206	172.044	60.538		207	176.594	6.8419	208	152.865	6.6629
209	170.926	7.5019	210	170.435	9.0033		211	172.277	8.9885	212	195.901	8.9922
213	197.627	9.0227	214	170.275	9.6956		215	143.277	8.1112	216	184.366	67.558
217	183.089	68.120	218	183,552	68.354		219	157,983	50.684	220	158.091	50.698
	100.000	50.120	210	100.002	54.000		222	130.030	50.001	220	135.012	50.050
221	198.592	53.077	222	189.000	54.220		223	1/0.9/3	53.848	224	175.943	54.442
225	180.213	54.591	226	185.335	54.761		227	242.073	38.176	228	244.923	36.088
229	245.153	36.365	230	228.312	48.495		231	214.146	55.481	232	204.799	57.900
222	246 222	20.702	224	225 190	47.967		225	240.001	44 509	226	242 520	43.459
233	240.322	36.792	254	235.189	47.807		255	240.961	44.508	230	242.529	43.438
237	49.9495	-0.221	238	150.004	8.5004		239	152.192	9.1769	240	174.041	10.055
241	159.028	45.131	242	142.741	8.9121		243	133.555	6.9731	244	144.494	10.299
245	124 360	7 6804	246	1/13 128	11 715		247	170 282	15 122	2/18	128 260	27 860
245	124.305	7.0804	240	145.128	11.715		24/	170.282	15.155	240	120.205	27.800
249	143.547	33.935	250	226.157	40.372		251	157.685	40.057	252	231.672	46.582
253	242.420	39.406	254	251.975	32.123		255	222.494	52.609	256	211.155	54.793
257	213 933	54 043	258	211 272	43 273		259	195 093	47 445	260	218 954	42 539
201	212.221	44.000	200	205 402	46.974		200	162 410	42.012	200	105 450	44.440
201	215.2/1	44.008	202	205.495	40.874		205	103.419	42.012	204	105.450	44.449
265	190.911	44.094	266	173.248	43.993		267	178.491	44.535	268	209.072	12.177
269	215.581	47.935	270	204.184	51.544		271	168.958	41.409	272	190.235	42.905
272	172 122	12 806	274	166 / 85	6 3174		275	188 7/8	6 6070	276	175 650	7 0617
275	1/3.125	42.800	274	100.485	0.3174		275	100.740	0.0979	270	175.050	7.0017
277	186.338	42.850	278	240.874	24.456		279	180.837	39.608	280	238.010	26.315
281	245.992	21.820	282	193.027	39.818		283	229.797	7.4837	284	200.312	8.5059
285	227 996	6 0962	286	213 297	8 4694		287	217 921	8 0661	288	229 609	5 2192
200	227.550	5.0404	200	223.257	6.1651		201	402 572	54 472	200	223.005	44.044
289	230.929	5.0401	290	227.691	0.3001		291	193.573	51.172	292	233.115	41.811
293	238.011	39.112	294	237.939	38.960		295	232.652	42.717	296	201.442	40.103
297	236.668	28.128	298	214.405	38,164		299	230.211	33,388	300	240.475	31.892
201	242.052	20.220	200	254.446	20 5 40		200	250.222	26.474	204	240.446	20.005
301	242.052	30.739	302	251.416	20.549		303	250.602	26.474	304	240.146	28.965
305	241.110	7.6260	306	243.344	7.0388		307	225.839	11.308	308	248.723	23.211
309	211 579	36 833	310	223 355	33 283		311	231 612	28 338	312	232 668	27 989
212	246.065	0.0001	214	224 520	0.628		215	244 775	0.200	210	228.020	0.0270
212	340.005	0.0001	514	324.529	-0.038		312	344.775	-0.296	310	556.050	0.0370
317	115.757	45.120	318	124.647	54.488		319	331.985	0.3701	320	131.475	59.715
321	20.2657	-0.300	322	29.2562	-0.278		323	23.9027	0.0149	324	49.2956	0.1095
325	180 /11/	14 055	326	203 026	12 220		227	170 574	15 287	378	105 005	10 248
325	100.414	14.055	320	203.520	15.550		327	175.574	13.287	320	133.303	40.240
329	189.013	39.046	330	136.555	26.672		331	124.107	20.652	332	117.587	17.169
333	221.580	31.938	334	229.761	29.016		335	144.877	33.526	336	156.226	35.127
337	195.074	34.944	338	204.919	33.689		339	142.756	26.819	340	197.703	34.078
3/11	108 162	34.065	3/12	203 709	33 300		3/13	209 014	32 716	344	161 604	33 780
341	150.102	54.005	342	203.705	33.305		545	205.014	52.710	544	101.004	33.785
	107 100	20.012		200.405	20.000			205 072			202.450	24.000
345	197.488	30.913	346	200.106	30.602		347	205.873	31.003	348	203.169	31.986
349	202.401	32.400	350	196.213	31.725		351	159.327	30.371	352	223.132	25.337
353	229.458	24.139	354	234,594	22.445		355	233.594	23,501	356	225.305	21.006
555	2251150	2200	551	2011001	22.115		555	200.001	20.001	550	220.000	21.000
257	224 024	10 240	250	222.400	22.270		250	240.220	16 206	200	204 692	26.220
357	234.024	18.348	358	222.488	22.278		359	240.338	16.306	360	204.683	26.328
361	211.438	25.392	362	228.788	21.329		363	233.159	19.884	364	227.019	22.308
365	236.443	21.568	366	239.405	20.755		367	236.073	16.952	368	226.009	21.072
					1	l						1
360	220 020	10 915	270	246 651	14 004	l	271	222 107	22 000	272	202 172	22 020
309	229.920	19.015	570	240.031	14.090		5/1	222.107	22.990	572	200.172	23.029
				1	1	l						
				407		l						ac
373	242.006	54.611	374	137.512	22.851	l	375	144.277	25.502	376	145.949	26.374
				1	1	l						
377	162.641	27.772	378	124.292	15.915		379	166.420	29.146	380	127.104	18.132
381	131.427	19.725	382	134.827	17.588		383	120.268	11.429	384	177.835	26.471
	-											-
				1	1	l						
385	120 352	11 916	386	139 225	19 302	l	387	172 1/17	26 381	388	167 699	26 375
505	120.332	11.510	560	133.233	19.302	l	507	1/2.14/	20.301	000	107.050	20.373
				1	1	l						
202	453.000	24 722	205	170.007	27.540	l	201	476.016	20.055	202	202.011	35.047
388	153.964	24.728	390	170.867	27.510	l	391	176.846	28.055	392	203.811	25.044
					1	l						1
					1	l						1
393	162.443	22.669	394	164.961	24.057	l	395	120.475	9.6258	396	126.934	12.233
					1	l						
					1	l						1
397	197.893	21.554	398	161.396	20.692	l	399	198.673	21.793	400	192.947	21.670
						l						
				1	1	l						
401	233 208	15 027	402	216 520	19 5/0	l	402	139 772	16 857	404	186 972	19 429
401	200.200	10.007	+02	210.320	10.040	l	-+03	133.123	10.007	-+04	100.372	13.430
						1				•		
405	176 106	20 125	106	204 050	19 714		407	225 227	12 007	109	236 542	12 002
405	176.196	20.125	406	204.959	18.714		407	235.337	12.987	408	236.543	12.982
405	176.196	20.125	406	204.959	18.714		407	235.337	12.987	408	236.543	12.982

413176.27819.9661414207.76719.4351415224.10414.5421416235.30916.629417213.32314.342418208.30816.165419210.31416.058420235.33312.337421167.19416.712422157.99816.320423177.65617.823424180.43417.898425163.70116.631426166.61717.345427189.96616.422428183.20818.269429345.5026.2461430119.33211.20643526.8014-10.2143643611.937433151.75831.983434130.78311.08543526.8014-10.21436190.801-2.00344122.84757.139434257.93464.11243926.101264.836400162.799.265144122.84757.13942212.79518.56743416.31221.614.4316.3124.4316.715	409	238.880	12.895	ĺ	410	152.186	16.807	l	411	177.394	18.705	412	169.589	19.543
417213.32314.3421418208.30816.1651419210.31416.0581420235.33312.337421167.19416.712422157.99816.320423177.65617.823424180.43417.898425163.70116.631426166.61717.345427189.96616.422428183.20818.269429345.5026.2461430119.33211.20643126.8014-10.21432128.53611.987433151.75813.983434130.78311.085435176.9321.8262436190.801-2.003434190.3831.5136438257.93464.112439261.01264.836440162.7399.2651441222.84757.139422122.79518.56744316.31232.7621111	413	176.278	19.966		414	207.767	19.435		415	224.104	14.542	416	235.309	16.629
421167.19416.712422157.99816.320423177.65617.823424180.43417.898425163.70116.631426166.61717.345427189.96616.422428183.20818.269429345.5026.2461430119.33211.20643126.8014-10.21432128.53611.987433151.75813.983434130.78311.085435176.9321.8262436190.801-2.003437190.3831.5136438257.93464.112439261.01264.836440162.7399.2651441222.84757.139422122.79518.567443116.31232.7621111	417	213.323	14.342		418	208.308	16.165		419	210.314	16.058	420	235.333	12.337
425 163.701 16.631 426 166.617 17.345 427 189.966 16.422 428 183.208 18.269 429 345.502 6.2461 430 119.332 11.206 431 26.8014 -10.21 432 128.536 11.987 433 151.758 13.983 434 130.783 11.085 435 176.932 1.8262 436 190.801 -2.003 437 190.383 1.5136 438 257.934 64.112 439 261.012 64.836 440 162.739 9.2651 441 222.847 57.139 442 122.795 18.567 443 16.312 32.762 1	421	167.194	16.712		422	157.998	16.320		423	177.656	17.823	424	180.434	17.898
429 345.502 6.2461 430 119.332 11.206 431 26.8014 -10.21 432 128.536 11.987 433 151.758 13.983 434 130.783 11.085 435 176.932 1.8262 436 190.801 -2.003 437 190.383 1.5136 438 257.934 64.112 439 261.012 64.836 440 162.739 9.2651 441 222.847 57.139 442 122.795 18.567 443 116.312 32.762 1	425	163.701	16.631		426	166.617	17.345		427	189.966	16.422	428	183.208	18.269
433 151.758 13.983 434 130.783 11.085 435 176.932 1.8262 436 190.801 -2.003 437 190.383 1.5136 438 257.934 64.112 439 261.012 64.836 440 162.739 9.2651 441 222.847 57.139 442 122.795 18.567 443 116.312 32.762 1 <t< td=""><td>429</td><td>345.502</td><td>6.2461</td><td></td><td>430</td><td>119.332</td><td>11.206</td><td></td><td>431</td><td>26.8014</td><td>-10.21</td><td>432</td><td>128.536</td><td>11.987</td></t<>	429	345.502	6.2461		430	119.332	11.206		431	26.8014	-10.21	432	128.536	11.987
437 190.383 1.5136 438 257.934 64.112 439 261.012 64.836 440 162.739 9.2651 441 222.847 57.139 442 122.795 18.567 443 116.312 32.762 1	433	151.758	13.983		434	130.783	11.085		435	176.932	1.8262	436	190.801	-2.003
441 222.847 57.139 442 122.795 18.567 443 116.312 32.762	437	190.383	1.5136		438	257.934	64.112		439	261.012	64.836	440	162.739	9.2651
	441	222.847	57.139		442	122.795	18.567		443	116.312	32.762			
Table 1: Ring galaxy candidates identified automatically														

3.1 Comparison of the ring galaxy candidates to previous catalogues

The galaxies in the catalogue were compared to the catalogue of

275 polar ring galaxy candidates in SDSS (?). That catalogue showed 23 galaxies with a full ring that fit the morphology of the

target galaxies shown in Table ??. The catalogue IDs of these galaxies are 7, 239, 240, 241, 243, 244, 245, 246, 249, 253, 254, 255, 256, 259, 260, 261, 263, 265, 267, 268, 270, 272, 274.

Comparison to the galaxies in Table ?? shows that none of these

galaxies were also included in Table ??. Therefore, Table ?? is clearly not a complete set of all SDSS galaxies with a full rung morphology, and many relevant galaxies with a full ring still exist in the SDSS database.

Comparing to the ring galaxies identified by ?, one (CGCG 222-022) of the 12 ring galaxies is included in this catalogue. The fact that just one galaxy is included in the catalogue shows that many more ring galaxies still exist in the SDSS database.

The list of automatically identified galaxies was also compared to the ring galaxies that were identified in SDSS by using

citizen science (?). The ? catalogue contains 3,962 galaxies that volunteers identified manually by visually inspecting the images through an on-line web-based platform. From the 443 galaxies

identified automatically, 104 are included in the ? catalogue. The

careful manual inspection process used in (?) is clearly more accurate than any existing computer algorithm. However, the manual classification and annotation requires substantial labour, and therefore less than $3\cdot 10^5$ galaxies were examined. The method described in this paper is automatic, and was applied to a much larger dataset of ~ $2.6\cdot 10^6$ galaxies, and therefore includes

very many galaxies that were not examined by ?.

It can be expected that many of the objects listed in Table ?? have been identified previously and are part of existing catalogues. Table ?? shows the ring galaxy candidates that were also identified in previous studies.

3.2 Distribution and photometry of the ring galaxy candidates

As mentioned in Section $\ref{eq:section}$, the galaxies in the catalogue

described in Section **??** are galaxies detected among the subset of SDSS DR14 galaxies that have spectra. The galaxies included in the

? catalogue are also galaxies with nuclear spectra. Because the galaxies are galaxies with spectra, their distribution in the sky is not uniform, but a distribution that corresponds to the spectroscopy survey of SDSS DR14. Therefore, the majority of the

ring galaxy candidates are in the RA range of 120° -240°. Figure **??**

shows the distribution of the galaxies in Table ?? combined with

the galaxies of the ? catalogue by their redshift. The figure shows the number of galaxies, as well as their frequency among the galaxies with spectra in DR14 in the same redshift range. The Petrosian radius of all galaxies is larger than 5.5", which is large

enough to allow the identification of the galaxy morphology (?).

The graph shows that the frequency of the ring galaxy candidates in the catalogue starts to decline when the redshift is higher than 0.08. That can be explained by the less detailed

morphology of the imaged galaxies when the redshift gets higher, which does not allow clear identification of morphological details such as the presence of a full ring. The low frequency in the 0-0.02 range can be explained by a higher number of objects misidentified as galaxies by the SDSS pipeline, but are in fact not extra-galactic objects. In any case, in the redshift range of 0-0.02 the number of detected ring galaxy candidates is very small, and

does not allow meaningful statistical analysis. Figure ?? shows the

distribution of the ring galaxy candidates in Table **??**. As the graph shows, the distribution is not substantially different from the

distribution in the ? catalogue.

Figure ?? shows the colour differences between the ring

galaxy candidates including the galaxies of the ? catalogue, and all other DR14 galaxies with spectra and Petrosian radius larger than 5.5". The graph shows that the u-g, r-i, and i-z declined with the increase in redshift for the galaxies identified as ring galaxies. That decline is in opposite trend to the other galaxies with spectra and Petrosian

radiuslargerthan5.5". Also, the colour of the ringgalaxy candidates changed in a more moderate manner with the redshift compared to the general galaxy population in SDSS DR14. That can be explained by the more morphologically homogeneous population in the set ring galaxy candidates, compared to the population of galaxies in SDSS. It should be noted that the majority of ring galaxy candidates are selected from the Galaxy Zoo 2 dataset, which are not a random selection of galaxies.

Figure ?? shows the colour differences between the ring

galaxies in Table ?? and the other SDSS DR14 galaxies with spectra, and Petrosian radius larger than 5.5". The graphs show no significant differences between the colour of the ring galaxy candidates and the colour of other galaxies, with the exception of the u-g colour. The u-g of the ring galaxy candidates is lower in all redshift ranges compared to the u-g colour of the other galaxies in SDSS DR14 that have Petrosian radius larger than 5.5". The mean u-g of the ring galaxy candidates is 1.482±0.018, while the mean u-g of all other DR14 galaxies with Petrosian radius larger than 5.5" is 1.572±0.0007, and therefore the difference is statistically significant (P < 0.001). The difference in the blue colour can be explained by the fact that rings in star-forming galaxies have a larger visible contract, and therefore can be detected more easily in distant galaxies compared to the redder rings in the same redshift ranges. That can therefore increase the number of blue galaxies among ring galaxies compared to the general galaxy population.

The graphs also show substantial difference in all colours for galaxies in the redshift range of 0-0.02. That can be explained by stars identified by error as galaxies in the SDSS photometric pipeline. However, due to the small number of ring galaxies in that range no meaningful statistical analysis of the difference is

possible. It should be noted that the galaxies in the ? catalogue are bright and large objects selected by Galaxy Zoo 2, and are much

larger than the objects in Table ??. The mean Petrosian radius (r

band) of the galaxies in the (?) is \sim 19.28", while it is \sim 9.67" for the galaxies in Table ??.

4 LIMITATIONS OF THE METHOD

Themethodusedinthisstudyaimsatanalyzingverylargedatabases of galaxy images that might be too big to analyze manually, even

when using crowdsourcing. That can only be done by automation. However, due to the very large databases of galaxy images, even a small false positive rate can lead to a very large number of false positive instances that becomes very difficult to handle manually. For instance, in the database used in this paper an algorithm with detection accuracy of 99% (which is normally considered extremely high in machine vision standards), would generate a dataset of ~ 2.6 \cdot 10⁵ false positives. Therefore, a practical application of the method requires to minimize the false positive rates. Since machine vision clearly does not meet the accuracy level of the human brain, achieving a low false positive rates require the sacrifice of some of the true positives.





Figure 1. SDSS candidates of resonance ring galaxies. Figure 2. SDSS candidates of collisional ring galaxies.

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	

Figure 3. SDSS candidates of ring galaxies with off-centre nucleus.



Figure 4. SDSS candidates of ring galaxies with no obvious nucleus inside the ring.



Figure 5. Other ring galaxy candidates in SDSS not included in Figures ??, ??, and ??.

As mentioned in Section **??**, pixels below the threshold level of 50 were considered not sufficiently bright and were ignored. The JPEG threshold of 50 is in some cases high, and can lead to the

exclusion of many ring galaxies such as the Hoag object (?), which has a clear but relatively dim ring compared to some other ring galaxies. However, lowering the threshold leads to a high number

of false positives. For instance, Figure ?? shows examples of objects that are not ring galaxies, but the algorithm would have flagged them as rings if a lower graylevel threshold would have been applied. Such objects are very common in the SDSS dataset, and many of them are flagged as galaxies by the SDSS photometric analysis pipeline. Since the method is designed to work with very large databases without the use of manual analysis, it sacrifices the detection of true positives, as even a small rate of false positives leads to an unmanageable output that requires a substantial step of manual analysis.

As discussed above, avoiding false positives is an important requirement, since due to the very large size of the database even a small false positive rate can make the method unusable. As a result, the detection method also has a high true negative rate, and many ring galaxies might not be detected by the method. To test the behavior of the methods and characterize the ring galaxies

that it might fail to detect, ring galaxies from (?) that were not

detected by the method were examined. Figure ?? shows the first

galaxies from the (?) Galaxy Zoo sample that were not detected by the method. The figure also shows the binary transformation of each image with different threshold levels.

For galaxies 2, 3, 5, and 6, a small background area surrounded



Figure 6. The number and frequency of the ring galaxy candidates in Table

?? combined with the galaxies of the Buta (2017) catalogue. The line shows the number of galaxies in each redshift range, and the bars show the frequency in the entire galaxy population.





by foreground pixels can be seen. However, these areas are smaller than 10% of the foreground, and therefore these galaxies are not flagged as ring candidates. In galaxies 1, 4, 6, and 8 part of the ring can be seen in the binary transform, but in none of the threshold levels the ring is complete in the sense that the background is completely surrounded by foreground pixels. For instance, in galaxy 4 the ring opens in the top right part of the galaxy. That happens because the ring is dimmer in that part, and the pixels in that part of the ring do not pass the threshold of the rest of the ring. In ring 6, the lower left part of the ring is dimmer than the rest of the ring, and therefore the ring cannot be detected by the method. In galaxy 3 the luminosity of the area inside the ring is not consistent, and therefore the ring is connected to the nucleus of the galaxy in the binary mask of the image. The same can also be seen in galaxy 5 and galaxy 2. In galaxy 7, the ring is made of a slightly bluer colour, but the pixel intensity of the ring is not higher than the intensity of the pixels between the ring and the nucleus. Since the method first converts the pixels to grayscale, rings that are visible because they have different colour than the rest of the galaxy will not be detected.

These examples show that the method is mostly dependent on the consistency of the luminosity of the ring, as well as the part of the galaxy inside the ring. Rings that their luminosity varies might not be detected by the method because parts of the ring might not pass the luminosity threshold, leaving the ring in the binary mask open. The same is also for variation inside the ring. If the luminosity inside the ring varies, some parts inside the ring might pass the luminosity threshold and prevent the detection of the ring. Therefore, the method will not always detect ring galaxies that the luminosity of the ring or the parts inside it is not consistent across the different areas.

5 CONCLUSION

While ring galaxies are relatively rare, it can be assumed that the number of ring galaxies within a certain set of galaxies increases with the size of the dataset. The application of automatic identification can therefore allow the detection of such galaxies in very large databases, and is not limited by the availability of human resources that can scan the database manually. When much larger databases such as then Large Synoptic Survey Telescope (LSST) are collected, automatic detection will be able to identify many more ring galaxies.

The collection of ringgalaxy candidates described in this paper is clearly not exhaustive, as evident by the differences between the

galaxies in this catalogue and the galaxies in the catalogues of ?, ?,

or ?. Automatic analysis is still not as accurate as the human eye and brain, especially for the non-trivial problem of galaxy image analysis. However, automatic analysis has the clear advantage of analyzing data much faster than any human or group of humans. The purpose of the approach described in this paper is to analyze very large databases of galaxies, under the assumption that even a small true positive rate can lead to large catalogues of ring galaxies.

Due to its higher sensitivity, the ? catalogue of manually classified ring galaxies in SDSS already contains 104 of the galaxies identified in this study. But because the computer analysis method can scan much more galaxies with no cost of human labour, the vast majority of the galaxies identified in this study are not included in previous catalogues. It should be mentioned that the set of galaxies with spectra used as the initial database is not a completely random subset of SDSS galaxies, but selected by a

certain algorithm (?).

While manual analysis of galaxy morphology has provided good collections of ring galaxies, the labour-intensive efforts required to compile such catalogues reduce the total number of galaxies that can be analyzed. As digital sky surveys are becoming increasinglymorepowerful, it is clear that automation will be required to analyze these databases and turn them into data products that enable scientific discoveries.

6 ACKNOWLEDGMENTS

This research was supported by NSF grants AST-1903823 and IIS1546079. I would like to thank the anonymous reviewer for the insightful comments that helped to improve the paper. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max



Lior Shamir









Figure 9. Colour differences between ring galaxy candidates of Table ?? and all galaxies with Petrosian radius larger than 5.5".

SDSS ring galaxy candidates

[h]

#	Identifier	Ring
		Reference
4	MCG+07-19-002	
16	VII Zw 466	?
23	MCG+05-23-004	
37	IIHz4	?
39	MCG+01-27-015	PGC 31038
60	IC 4074	
64	IC 1706	
04	10 1700	
78	NGC 7613	
82	NGC 4031	?
85	NGC 3754	
96	IC 1698	PGC 5261
101	MCG+08-15-041	?
110	IC 1010	
130	NGC 2740	?
140	NGC 5636	PGC 51785
141	IC 1007	PGC 51465
145	MCG+11-16-015	?
148	NGC 5876	?
153	NGC 6184	
15/	LIGC 10420	DCC 58285
155	UGC 10430	100 30303
100	00010010	
181	MCG+10-16-093	
183	MCG+09-19-213	?
189	UGC 6109	
100	MCG+08-18-057	PGC 2012/
203	MCG+08-22-038	10025124
200		
211	IC 699	?
216	MCG+11-15-044	
222		
222	NGC 4566	
225	MCG+09-20-062	
228	UGC 10342	?
240	IC 2941	?
250	UGC 9691	
265		2
205 202	10 38/4	r 2
202 202	10 2044 (GCG 222-022	: ?
292	NGC 5947	•
308	MCG+04-39-016	
316	UGC 12068	PGC 69089
319	MCG+00-56-009	
222		
323 276	IVICG+00-05-013	FUL 3928
520	.0.501	
328	IC 4135	
337	MCG+06-29-011	
2/11		
341	10100+00-29-059	
349	UGC 8484	PGC 47369
359	UGC 10134	
360	MCG+05-32-048	

14



374	IC 2441C		
395 405	MCG+02-21-005 UGC 6719	?	
413	MCG+03-30-094		
424	MCG+03-31-015		
435 440	UGC 6769 NGC 3429	? ?	

Figure10. Example of galaxies that would have been flagged as ringgalaxies below a graylevel threshold of 50.

Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean ScientistGroup,theChineseAcademyofSciences(LAMOST),Los Alamos National Laboratory, the Max Planck Institute for Astronomy (MPIA), the Max Planck Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Table 2. Galaxies that are part of previous catalogues

Lior Shamir



Figure 11. Ring galaxies from Struck (2010) catalogue that were not detected by the method, and the binary transformation with different thresholds.