

Universidade Federal do Rio Grande do Sul
Instituto de Física

Caracterização de galáxias de baixo brilho
superficial identificadas através de busca por
similaridade

Isadora da Silva Espíndola

Porto Alegre

2023

Universidade Federal do Rio Grande do Sul
Instituto de Física

**Caracterização de galáxias de baixo brilho superficial
identificadas através de busca por similaridade**

Isadora da Silva Espíndola

Trabalho de Conclusão de Curso apresentado
à Universidade Federal do Rio Grande do Sul,
como requisito necessário para obtenção do grau
de Bacharel em Física.

Orientadora: Prof^ª. Dra. Cristina Furlanetto

Porto Alegre

2023

Agradecimentos

Gostaria de agradecer ao meu companheiro Christian Garcia por todo apoio e por escolher caminhar ao meu lado. Aos meus familiares e amigos próximos que acompanharam e se fizeram presentes ao longo dessa minha longa jornada acadêmica.

Agradecer aos professores que em diferentes espaços e momentos da minha trajetória acadêmica me mostram que a educação é revolucionária e transformadora. Em especial, aos professores da UFPEL Dr. Álvaro Ayala Filho e Dr. Valdemar das Neves Vieira por cada ensinamento científico e pessoal que me proporcionaram – sempre serão meus exemplos de educadores.

A todos os colegas e amigos que a física me presenteou ao longo desses anos. Não poderia deixar de agradecer ao "pessoal do café da física": Neto, Amanda, Marco, Gustavo, Tainá e Naamã, por todas as risadas, conversas, brincadeiras e abraços neste último semestre. Com vocês o semestre com certeza foi muito melhor. Também aos amigos do Laboratório de Astronomia da UFRGS, em especial à Ana Antonini, que principalmente nessas últimas semanas foi super acolhedora. Pessoal, muito obrigada pelo carinho, parceria e por falas como: "Estamos estudando há muito tempo (ou estressados e cansados), vamos dar uma voltinha no campus?", ou ainda, "Está tudo bem? Vamos sentar um pouco lá fora por alguns minutos?". Esses momentos com certeza foram muito importantes durante esta difícil reta final e é um privilégio estar me formando com vocês.

A minha orientadora, Dr^a. Cristina Furlanetto, pelos ensinamentos e companheirismo, nesses últimos seis meses. Poucos estudantes tem a oportunidade de terem ao seu lado orientadores que acrescentem não somente a sua pesquisa científica, mas também nas suas perspectivas de futura mulher pesquisadora.

Aos membros da banca Prof^a. Dr^a. Ana Chies Santos e o doutorando Rodrigo Flores de Freitas por aceitarem prontamente à contribuírem com meu trabalho. E por fim, ao Programa Institucional de Bolsas de Iniciação Científica BIC Meninas na Ciência da UFRGS pelo apoio financeiro.

Resumo

As galáxias de baixo brilho superficial (LSBGs, do inglês *Low Surface Brightness Galaxies*) são galáxias que apresentam brilho superficial central típico de $\mu_0 \geq 23.0$ mag arcsec², constituindo uma classe de objetos abundante e importante para o entendimento do Universo. Entretanto, detectar e caracterizar esses objetos é extremamente desafiador devido ao seu baixo brilho superficial. Neste estudo, foram caracterizadas 100 galáxias de baixo brilho superficial identificadas através da busca de similaridade nos dados do terceiro ano do Dark Energy Survey, sendo 41 delas novas candidatas a LSBGs e as outras 59 LSBGs já confirmadas na literatura. A caracterização foi realizada utilizando a técnica de fotometria superficial por meio do uso dos softwares SExtractor e GALFIT, com adaptações feitas no código público Galfitpywrap, e executado para todas as galáxias da amostra. Conclui-se que o método automatizado de caracterização foi validado, embora precise ser aperfeiçoado. Além disso, a maioria das novas candidatas encontradas pela busca por similaridade teve sua natureza de LSBG confirmada, o que demonstra o potencial dos métodos implementados.

Palavras-chave: GALÁXIAS DE BAIXO BRILHO SUPERFICIAL, FOTOMETRIA, BUSCA POR SIMILARIDADE

Abstract

Low Surface Brightness Galaxies (LSBGs) are galaxies that have a typical central surface brightness of $\mu_0 \geq 23.0$ mag arcsec², constituting an abundant and important class of objects for understanding the Universe. However, detecting and characterizing these objects is extremely challenging due to their low surface brightness. In this study, 100 low surface brightness galaxies were characterized by identifying them through similarity searches in the third year data of the Dark Energy Survey. Of these, 41 were new LSBG candidates and the other 59 were previously confirmed LSBGs in the literature. Characterization was performed using the technique of surface photometry by using the software SExtractor and GALFIT, with adaptations made to the public code Galfitpywrap, and executed for all galaxies in the sample. It is concluded that the automated characterization method was validated, although it needs to be improved. In addition, the majority of the new candidates found by similarity search had their LSBG nature confirmed, demonstrating the potential of the implemented methods.

Keywords: LOW SURFACE BRIGHTNESS GALAXIES, PHOTOMETRY, SIMILARITY SEARCH.

Sumário

1	Introdução	1
1.1	Galáxias de baixo brilho superficial	3
1.2	Detecção de LSBGs	6
1.2.1	Método de busca por similaridade	7
1.3	Fotometria superficial	10
1.4	Objetivos e estrutura do trabalho	11
2	Metodologia	13
2.1	Dados	13
2.2	Caracterização de LSBGs	14
2.2.1	SOURCE EXTRACTOR	15
2.2.2	GALFIT	16
2.2.3	Galfitpywrap	16
3	Resultados e Discussão	21
3.1	Validação do método	21
3.2	Análise das novas candidatas a LSBGs	23
4	Sumário e Conclusões	27
4.1	Perspectivas	28
	Referências	29

Lista de Figuras

Figura 1.1 – Perfil de Sérsic para diferentes índices n . (Peng et al., 2002)	2
Figura 1.2 – Exemplo de UDG: DF44. Créditos: Teymoor Saifollahi and NASA/HST (HST Proposal 14643, PI: van Dokkum)	4
Figura 1.3 – Vizinhos da LSBG chave ID=157441790, onde d é a distância à chave e i é índice do vizinho em ordem de distância. Em vermelho estão os objetos não presentes no catálogo de LSBGs de Tanoglidis et al. (2021a) e em verde, temos os objetos presentes neste catálogo. Figura retirada do trabalho de Tidball (2022).	9
Figura 1.4 – Em azul, histograma da distância dos 25000 vizinhos mais próximos do objeto de ID=157441790. Em laranja, o histograma da distância dos vizinhos mais próximos da mesma chave que estão presentes no catálogo de Tanoglidis et al. (2021a). Imagem retirada de (Tidball, 2022).	9
Figura 2.1 – Os painéis representam as imagens original (esquerda), modelo (centro), residual (direita) de saída após a execução do código Galfitpywrap adaptado para três imagens diferentes cujos modelos foram bem ajustados.	18
Figura 2.2 – Os painéis representam as imagens original (esquerda), modelo (centro), residual (direita) de saída após a execução do código Galfitpywrap adaptado para três imagens diferentes cujos modelos não foram bem ajustados.	19
Figura 3.1 – Gráfico $g_T \times g$, onde g_T são as medidas de magnitude de Tanoglidis et al. (2021a) e g as medidas realizadas neste trabalho. A reta identidade é utilizada para visualizarmos quão distintas são as medidas em ambos os trabalhos.	21
Figura 3.2 – Gráfico $R_{e,T} \times R_e$, onde $R_{e,T}$ são as medidas de raio efetivo de Tanoglidis et al. (2021a) e R_e as medidas realizadas neste trabalho.	22
Figura 3.3 – Gráfico $\epsilon_T \times \epsilon$, onde ϵ_T são as medidas de elipticidade de Tanoglidis et al. (2021a) e ϵ as medidas realizadas neste trabalho.	23
Figura 3.4 – Gráfico $n_T \times n$, onde n_T são as medidas de índice de Sérsic de Tanoglidis et al. (2021a) e n as medidas realizadas neste trabalho.	24
Figura 3.5 – Gráfico representando raio efetivo R_e em função da magnitude na banda g para a amostra de 59 LGBGs confirmadas, em azul, e 41 LSBGs candidatas, em vermelho, dentre elas o subconjunto de 9 LSBGs com resíduo significativo marcadas com um círculo preto.	25

Figura 3.6 – Nesta figura tanto a magnitude na banda g quanto o raio efetivo estão em função do índice de Sérsic. Note que em (a) o eixo da magnitude cresce da direita para a esquerda. Em ambos os gráficos temos em azul as LGBGs confirmadas, as LSBGs candidatas em vermelho e as LSBGs com resíduo significativo marcadas com um círculo preto. 26

1 Introdução

A vastidão do universo abriga inúmeras galáxias, diversas em suas propriedades, mas cada uma delas com um conjunto único de características. Uma galáxia é uma enorme coleção de estrelas, gás, poeira e matéria escura que são mantidos juntos pela gravidade. As estrelas em uma galáxia variam em tamanho, cor e temperatura, e são tipicamente organizadas em aglomerados e grupos que formam padrões distintos. A distribuição das estrelas dentro de uma galáxia está associada à sua forma, com galáxias espirais apresentando braços espirais e galáxias elípticas com uma forma mais esferoidal.

As galáxias são sistemas variados com uma ampla gama de morfologia e propriedades. Por exemplo, os tamanhos característicos de seus componentes luminosos variam de 0.1 kpc a dezenas de kiloparsecs, enquanto as luminosidades ópticas e massas estelares estão na faixa de 10^3 a 10^{12} em unidades solares (Cimatti et al., 2020).

Com relação aos seus componentes materiais, podem incluir estrelas, gás neutro e ionizado, poeira, nuvens moleculares, entre outros. Os processos físicos associados a esses elementos desempenham papéis importantes na evolução de galáxias e ajudam a moldar a estrutura e comportamento geral das galáxias.

Além disso, as galáxias podem existir em uma variedade de ambientes, habitando de regiões isoladas (também chamadas de campo) até densos aglomerados de galáxias. Aglomerados são conjuntos de galáxias que são mantidos unidos pela gravidade e podem conter de algumas centenas a milhares de galáxias. Dentro desses aglomerados, as galáxias podem interagir umas com as outras, podendo haver troca de gás e estrelas ou até mesmo fundirem-se e formar galáxias maiores. As interações entre galáxias podem ocorrer intra-aglomerado Gunn e Gott (1972), ou ainda, por interações gravitacionais com companheiras próximas Merritt (1983). Entretanto, mesmo quando não houver troca, a estrutura da galáxia pode se alterar em interação com galáxias vizinhas devido a efeitos de maré. Essas interações podem ter um impacto profundo na evolução das galáxias, moldando sua morfologia, composição e até mesmo seu destino geral (Boselli et al., 2014; Muzzin et al., 2012; Peng et al., 2012).

Um importante fator para caracterizar as galáxias é seu brilho superficial, sendo ele uma medida da intensidade da luz por unidade de área, que é uma maneira de caracterizar galáxias através da sua distribuição de brilho. A distribuição de brilho superficial revela importantes informações sobre a estrutura interna da galáxia. Essas distribuições de brilho podem ser descritas através do perfil Sérsic, uma função analítica que descreve como a intensidade I de uma galáxia varia com a distância R do seu centro. O perfil de Sérsic é

dado por,

$$I(R) = I_e \exp \left\{ -b_n \left[\left(\frac{R}{R_e} \right)^{1/n} - 1 \right] \right\}, \quad (1.1)$$

onde I_e é a intensidade no raio efetivo R_e , a constante b_n é definida em termos do parâmetro n , chamado índice de Sérsic, que controla o grau de curvatura do perfil. O valor de b_n ¹ pode ser calculado ou analiticamente de forma aproximada ou numericamente (Graham e Driver, 2005; Ciotti e Bertin, 1999). Outro parâmetro muito importante deste perfil é o raio efetivo da galáxia ou raio de meia-luz, R_e , definido como o raio no qual metade da luz do modelo está contida (assumindo que a galáxia tem simetria esférica intrínseca ou uma simetria circular quando vista no plano do céu).

A Figura 1.1 mostra o perfil de Sérsic para diferentes valores de n . É possível ver que quanto menor o valor de n maior é a inclinação da curva na região central da galáxia (sua distribuição de luz é mais homogênea).

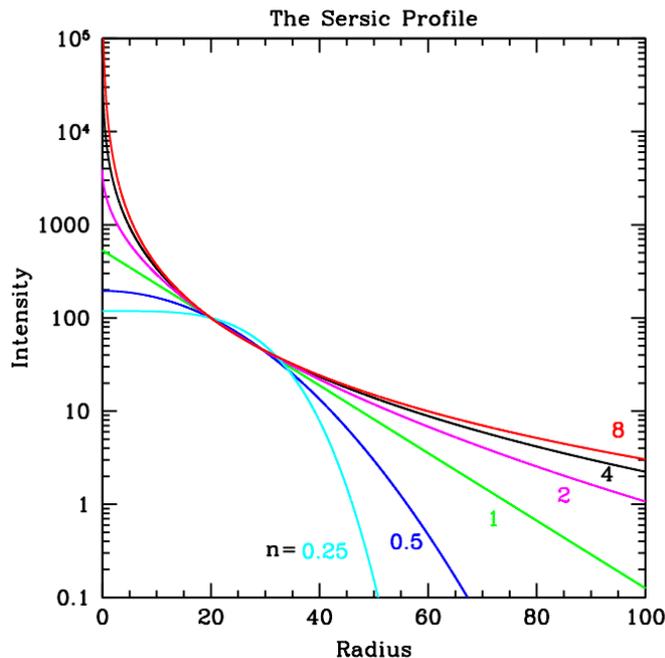


Figura 1.1 – Perfil de Sérsic para diferentes índices n . (Peng et al., 2002)

Em geral é necessário mais de um perfil para ajustar o perfil de brilho de galáxias. As componentes mais esferoidais (galáxias elípticas, bojo de espirais) podem ser descritas com perfil de Sérsic com n mais alto e a componente de disco com $n \sim 1$. Entretanto, há uma classe de galáxias que possui, dentre outras características, um perfil radial com distribuição de luz pouco concentrada na região central: Galáxias de baixo brilho superficial (LSBGs, do inglês *Low Surface Brightness Galaxies*).

¹ Capaccioli (1989) desenvolveu uma das primeira aproximações para o valor de b_n , $b_n = 1.9992n - 0.3271$ para $0.5 < n < 10$.

Acredita-se que estes objetos compõem uma relevante porção da matéria do Universo local (Driver, 1999; O’Neil e Bothun, 2000). As LSBGs são, em sua maioria, galáxias anãs com uma baixa densidade de estrelas e por isso, são objetos tênues e difusos, sendo assim, muito difícil de serem observadas. Esse é inclusive um dos motivos de apenas mais recentemente, dados os avanços observacionais, termos encontrado um maior número delas. A primeira LSBG foi detectada somente no final dos anos de 1980 (Bothun et al., 1987) e com o avanço tecnológico espera-se que cada vez mais essas galáxias difusas sejam encontradas.

No entanto, apesar do enorme avanço científico-tecnológico que atingimos nos dias atuais, estudar galáxias de baixo brilho superficial ainda é desafiador, pois os telescópios de forma geral têm como foco objetos astronômicos brilhantes. Além disso, muitos levantamentos fotométricos possuem grandes áreas, mas pouca profundidade, o que dificulta a detecção desse tipo de objeto. Por outro lado, alguns levantamentos são de campos profundos, porém de área limitada, em geral com foco em ambientes densos, como os de aglomerados de galáxias. Esses levantamentos acabam priorizando o estudo de LSBGs em aglomerados de galáxias e pouco se sabe das características desses objetos em ambientes mais isolados.

De encontro a isso, as novas gerações de levantamentos fotométricos como o Legacy Survey of Space and Time (LSST)¹, por exemplo, será capaz de encontrar objetos fracos devido à maior profundidade, que será atingida ao longo de 10 anos de observação.

Dessa forma, torna-se necessário o desenvolvimento de novas abordagens técnicas que sejam eficientes na análise desta quantidade gigantesca de dados e que otimizem o processo de encontrar objetos astronômicos de interesse. Com intuito de colaborar com a criação de novas técnicas, neste trabalho trazemos a caracterização de galáxias de baixo brilho superficial detectadas em um trabalho anterior por Tidball (2022) via técnica de busca por similaridade.

Apresentaremos mais detalhes a respeito destas abordagens técnicas e sobre nosso objeto de estudo nas seções que seguem.

1.1 Galáxias de baixo brilho superficial

Galáxias de baixo brilho superficial são convencionalmente definidas como galáxias com brilho superficial central mais fracos do que o céu noturno (Bothun et al., 1997). Na literatura atualmente encontra-se que estes objetos astronômicos possuem brilho superficial central típico na banda B em torno de $\mu_0(B) \geq 23.0$ mag arcsec².

O primeiro pressuposto teórico da sua existência foi em meados dos anos 1970 por Disney (1976), porém a primeira observação dessa classe de objetos somente foi possível uma década depois com a descoberta de Malin 1 (maior galáxia espiral já encontrada,

¹ <https://www.lsst.org/>

(Bothun et al., 1987)). Embora sejam objetos de difícil observação, estima-se que elas podem ser responsáveis por cerca de 15% da massa dinâmica no Universo atual (Driver, 1999). Mais recentemente trabalhos estimam que de 30% a 60% das galáxias do Universo Local sejam LSBGs (O’Neil e Bothun, 2000).

As LSBGs têm um conjunto único de características que as diferenciam de outros tipos de galáxias. Uma de suas características definidoras é o baixo brilho superficial, o que significa que emitem menos luz por unidade de área do que outras galáxias, pois possuem uma baixa densidade de estrelas. Isso as torna difíceis de detectar e estudar.

Segundo Prole et al. (2019), as LSBGs quanto às suas propriedades de massa estelar e metalicidade, são similares a galáxias anãs, embora hajam LSBGs de tamanhos físicos e massas do halo diferentes destas. Tipicamente, as LSBGs possuem massa estelares entre 10^7 a $10^9 M_{\odot}$ e raios efetivos da ordem de 1 kpc. Recentemente van Dokkum et al. (2015) encontraram LSBGs excepcionalmente grandes ($r_e \geq 1.5$ kpc) no aglomerado Coma e as denominaram “galáxias ultra-difusas” (UDGs, do inglês *Ultra-Diffuse Galaxies*, Figura 1.2) desde então, a comunidade tem voltado sua atenção novamente a estes objetos, surgindo assim diversas propostas para explicar suas origens.

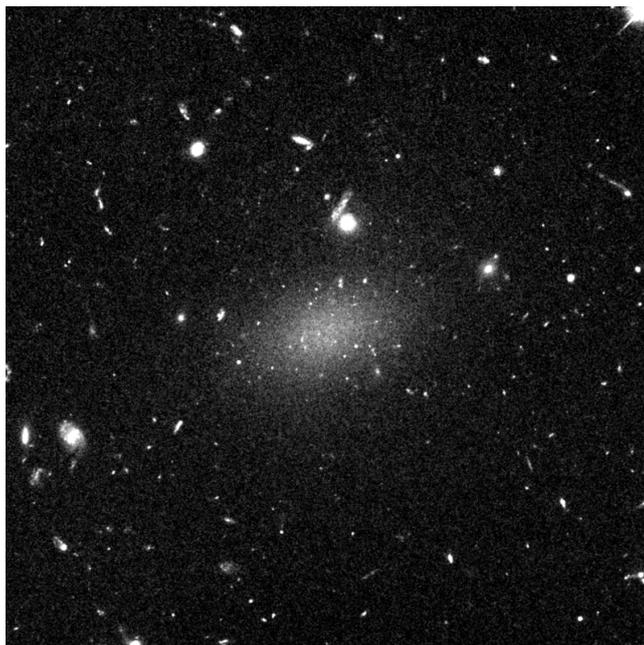


Figura 1.2 – Exemplo de UDG: DF44. Créditos: Teymoor Saifollahi and NASA/HST (HST Proposal 14643, PI: van Dokkum)

As galáxias de baixo brilho superficial são um grupo diversificado de galáxias que abrangem uma ampla gama de tamanhos físicos, cores e ambientes. Elas podem ser encontrados em uma variedade de locais em todo o universo, desde pequenos satélites ultra-tênuas da Via Láctea (McConnachie, 2012) até enormes aglomerados de galáxias como Perseus (Wittmann et al., 2017), Coma (van Dokkum et al., 2015) e Fornax (Venholá

et al., 2017).

Com relação à variação de propriedades com ambiente, as encontradas em aglomerados de galáxias são mais avermelhadas e esféricas, com deficiência de gás e, portanto, compostas por estrelas velhas. Por outro lado, as LSBGs encontradas em isolamento são mais azuis, irregulares e possuem alta massa de hidrogênio neutro, importante componente para formação de novas estrelas (Prole et al., 2019; Lee et al., 2020).

Apesar de seu baixo brilho superficial, as LSBGs são objetos de estudo fascinantes porque fornecem informações sobre a formação e evolução das galáxias. Por exemplo, o fato de poderem ser encontrados em uma variedade de ambientes e formas, sugere que podem ter se formado por diferentes mecanismos. Compreender como essas galáxias se formaram e evoluíram pode lançar luz sobre os processos maiores que moldam o universo como um todo.

Vários cenários de formação de LSBGs e UDGs foram propostos na literatura, entre os quais destacamos três: o primeiro propõe a ideia de “galáxias que falharam”, em que as UDGs que não conseguiram gerar uma quantidade típica de estrelas devido a processos relacionados ao ambiente. Outro cenário seriam galáxias anãs estendidas que de certa forma “inflaram” por processos internos. E por último a proposta de que os progenitores da UDGs e LSBGs foram estendidos devido a efeitos de maré causados por galáxias massivas vizinhas. Embora hajam diversos cenários para a explicação de sua origem, Prole et al. (2019) sugerem que múltiplas origens podem contribuir para a formação e evolução desses objetos.

Outro ponto em debate é a respeito da matéria escura que essas galáxias possuem, principalmente após alguns trabalhos recentes que indicavam a detecção de uma UDG sem matéria escura (van Dokkum et al., 2018). Assim, estudar parâmetros estruturais (brilho, raio, concentração de luz, etc) das galáxias é um importante passo para uma descrição quantitativa das LSBGs.

O estudo do Universo de baixa luminosidade permanece relativamente inexplorado devido às limitações dos telescópios ópticos em observar objetos com esta característica. No entanto, as LSBGs oferecem uma ferramenta vital para melhorar nossa compreensão da extremidade mais fraca da função da luminosidade das galáxias. Uma vantagem significativa de estudar LSBGs é sua abundância. Estudos mostraram que as LSBGs são a maioria das galáxias no Universo (McGaugh e de Blok, 1998). Como tal, oferecem uma excelente oportunidade para investigar os processos de formação e evolução galáctica, em particular, no regime de menores massas. Além disso, proporcionam também uma oportunidade para investigar os processos de formação estelar em ambientes de baixa densidade estelar no interior da galáxia.

O desafio de detectar LSBGs motiva o desenvolvimento de novos métodos de detecção, mais eficientes, que possam aumentar significativamente a amostra de objetos dessa classe de forma que possamos entendê-los mais profundamente. Entendendo-os mais

profundamente, estaremos conhecendo melhor o universo das galáxias nos regimes de baixa luminosidade e baixa massa.

1.2 Detecção de LSBGs

As LSBGs são tipicamente mais fracas e têm menor brilho superficial do que as galáxias típicas, tornando-as difíceis de detectar e estudar. Frente a isto, na literatura nos últimos anos tem-se buscado diversas maneiras e métodos inovadores para realizar a detecção destes objetos, embora ainda hajam limitações com respeito aos telescópios já citadas anteriormente.

Dentre os métodos voltados à detecção podemos citar a inspeção visual, método mais comum e naturalmente utilizados pelos astrônomos. No caso das LSBGs, a sua natureza difusa as torna mais difíceis de detectar nas imagens do que as galáxias mais brilhantes, dessa forma, a inspeção visual de imagens profundas às vezes pode revelar estas candidatas em potencial. Um fator importante é a profundidade das imagens, uma maneira de detectar LSBGs é por meio de imagens profundas com grandes telescópios, que podem ser obtidas pelo empilhamento de várias imagens do mesmo campo.

Na literatura há uma diversidade de abordagens utilizadas para a detecção desses objetos. Por exemplo, Zaritsky et al. (2019) com dados do Legacy Surveys², um levantamento de imagens de três bandas cobrindo 14.000 graus quadrados do céu extragaláctico, subtraíram objetos mais brilhantes das imagens e aplicaram filtros para realçar as características difusas destes objetos e encontraram 275 candidatas a UDGs no aglomerado de Coma Já Prole et al. (2019) utilizaram uma ferramenta de detecção de objetos astronômicos em imagens chamada MTOBJECTS (Max-Tree Object; Teeninga et al. (2015)) e aplicaram vários critérios de seleção baseados em propriedades fotométricas e estruturais (como brilho superficial e raio efetivo), obtendo um total de 212 candidatas a UDGs de campo contidas no Kilo-Degree Survey (KiDS)³.

Técnicas paramétricas para a detecção de objetos como estas, embora tendam a ser mais precisas e por este motivo mais confiáveis, necessitam de um tempo de processamento computacional significativamente grande e uma minuciosa análise feita de forma manual. Além disso, elas se tornam demasiadamente custosas à medida que a quantidade de dados aumenta.

Além de técnicas dessa natureza, mais recentemente o uso de aprendizado de máquina (ML, do inglês *machine learning*) tornou-se uma ferramenta cada vez mais popular na Astronomia para analisar grandes quantidades de dados e extrair informações significativas. Na detecção de objetos em imagens não seria diferente e as redes neurais convolucionais (CNN, do inglês *Convolutional Neural Networks*), por exemplo, constituem o estado da

² <https://www.legacysurvey.org/>

³ <https://kids.strw.leidenuniv.nl/>

arte de ML para esta tarefa.

Para identificar e classificar LSBGs em imagens astronômicas de grandes áreas do céu, Tanoglidis et al. (2021a) utilizaram técnicas de *Support Vector Machines* para criar uma amostra de candidatas à LSBGs. Eles treinaram um modelo em tabelas de dados (contendo propriedades como magnitude, fluxo, etc) para fazer de forma automatizada um corte de ~ 420000 objetos para ~ 44000 que foram classificadas pelo modelo como galáxias de baixo brilho superficial. Depois realizaram inspeção visual, ajustaram um perfil de Sérsic e selecionaram galáxias que possuíam parâmetros compatíveis com os típicos de LSBGs, chegando ao final com uma amostra de ~ 23790 LSBGs. Posteriormente, Tanoglidis et al. (2021b) utilizaram esta amostra para treinar uma rede neural convolucional que separa LSBGs de artefatos de baixo brilho superficial em imagens.

Embora seja uma potente ferramenta, estas técnicas de ML requerem grandes amostras de dados rotulados para treinamento, que ainda não estão disponíveis, uma vez que justamente estamos buscando por novas candidatas. Além é claro, do seu custo computacional e sua interpretabilidade limitada, pois as CNNs costumam ser chamadas de “caixas pretas”, já que pode ser difícil entender como a rede chegou à sua saída, tornando este um complicador.

Neste sentido, buscar por métodos automatizados de detecção mais simples e sem a necessidade de grandes quantidades de dados é promissor para a Astronomia. Na seção seguinte, explicamos brevemente a respeito do método de similaridade e os resultados alcançados por Tidball (2022), nos quais este trabalho se baseia.

1.2.1 Método de busca por similaridade

O problema de busca por similaridade, também conhecido como busca por k-vizinhos mais próximos (kNN, do inglês *K-nearest neighbors*), consiste em encontrar um item que seja mais próximo (mais similar) ao item de referência, também chamado de dado chave (Wang et al., 2014). Este método é usualmente utilizado por ferramentas de busca para encontrar textos, imagens e outros objetos específicos similares à chave (Bhatia e Vandana, 2010). Ele possui assim, uma alta aplicabilidade como, por exemplo, a identificação de fraudes, sendo uma técnica muito eficiente e eficaz para estes fins.

Quando estamos pensando em métodos computacionais, uma das estruturas mais básicas para armazenar dados computacionalmente é o vetor. O algoritmo de kNN é baseado na verificação das distâncias entre cada vetor de um banco de dados e obtém como saída quais desses itens são similares (mais próximos) do item que foi escolhido como chave. Sua simplicidade é sua principal vantagem, no entanto, o requisito de memória e a complexidade de computação também são fatores relevantes para sua execução. Uma alternativa a esse problema é a utilização do kNN aproximado ao invés da busca absoluta.

Em um trabalho recente Tidball (2022) utilizaram um tipo de kNN aproximado chamado *hashing sensível à localidade* (LSH, do inglês *Locality-sensitive hashing*) para

encontrar LSBGs por similaridade em catálogos de levantamentos astronômicos públicos. O LSH é um método para encontrar de forma aproximada os vizinhos mais próximos em espaços de alta dimensão. Ele é baseado na ideia de que pontos semelhantes no espaço de alta dimensão provavelmente estarão próximos um do outro em um subespaço de dimensão inferior. A principal ideia por trás dele é usar uma função *hash* para mapear os dados em *buckets*. Dados semelhantes serão mapeados para o mesmo *bucket* com alta probabilidade. Isso permite consultas rápidas ao vizinho mais próximo de uma chave, pois só precisamos procurar entre os dados que pertencem ao mesmo *bucket* que a chave (Leskovec et al., 2020).

Os algoritmos kNN aproximados como o LSH fornecem um bom compromisso entre precisão e eficiência computacional, pois podem não encontrar os k vizinhos mais próximos exatos de um novo ponto de dados, mas podem fornecer uma boa aproximação com custo computacional significativamente reduzido. Eles se tornam assim extremamente úteis em casos como o nosso, no qual não temos uma grande quantidade de dados de treinamento para utilizar em alguma rede neural, por exemplo.

Em seu trabalho Tidball (2022) abordou o problema de encontrar LSBGs utilizando a busca de vizinhos aproximada com o LSH. A partir dos dados do catálogo DES Y3 Gold coadd (v2.2) do levantamento fotométrico Dark Energy Survey contendo ~ 400 milhões de objetos, selecionou ~ 11 milhões de objetos que possuíam característica relativamente próximas das características típicas das LSBGs, em termos de tamanho, magnitude e cor. Aplicando o LSH nesse catálogo com os objetos previamente selecionados, ele encontrou objetos visualmente muito similares a chaves retiradas do catálogo de LSBGs de Tanoglidis et al. (2021a). Um exemplo da aplicação deste método encontra-se na Figura 1.3. Entre os objetos similares encontrados, muitos já faziam parte do catálogo de Tanoglidis et al. (2021a), porém outros eram novos candidatos a LSBGs, demonstrando a generalidade do método.

A Figura 1.4 mostra o número de objetos similares em função da distância (no espaço de parâmetros onde o LSH foi aplicado) até o dado usado como chave, onde foram considerados os 25000 vizinhos mais próximos. Notamos que o método de similaridade implementado por Tidball (2022) além de recuperar várias LSBGs já catalogadas por Tanoglidis et al. (2021a) encontra uma grande quantidade de objetos similares que são novas candidatas a LSBGs. Com somente uma galáxia e baixo brilho superficial (chave utilizada), foram encontradas 12260 galáxias da amostra de Tanoglidis et al. (2021a) entre os 25000 vizinhos mais próximos da chave. O restante dos objetos, objetos são visualmente muito similares a LSBGs e não estão presentes no catálogo, como no exemplo da figura 1.3.

Como será descrito na seção 1.4, o escopo deste trabalho é caracterizar essas novas candidatas a LSBGs, encontradas pelo método de busca por similaridade.

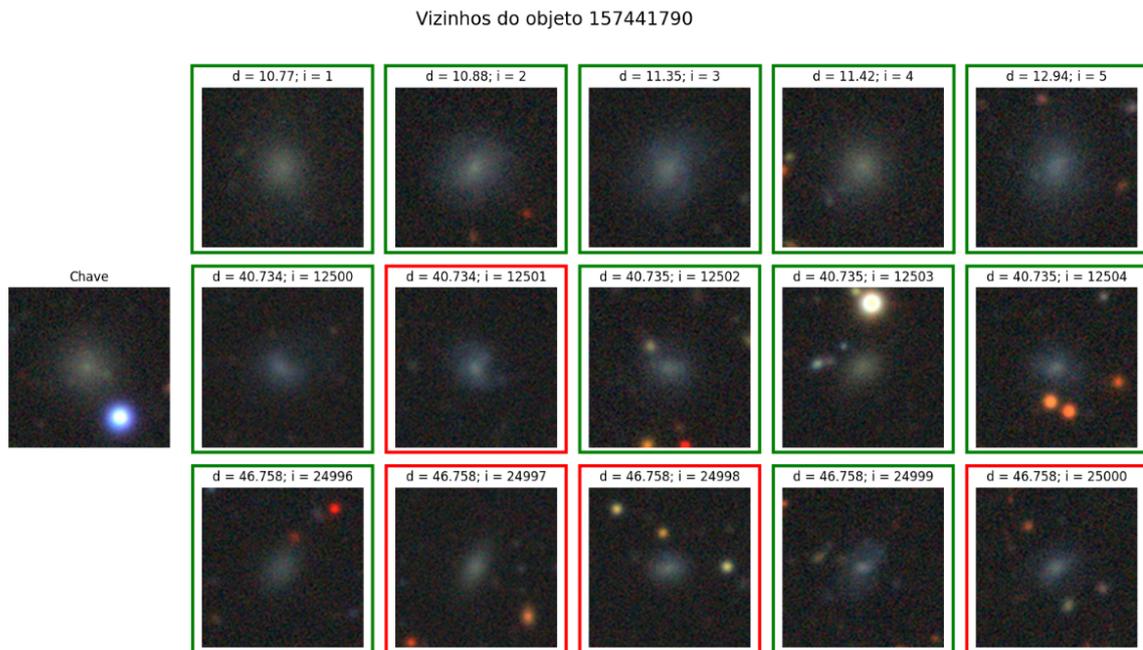


Figura 1.3 – Vizinhos da LSBG chave ID=157441790, onde d é a distância à chave e i é índice do vizinho em ordem de distância. Em vermelho estão os objetos não presentes no catálogo de LSBGs de Tanoglidis et al. (2021a) e em verde, temos os objetos presentes neste catálogo. Figura retirada do trabalho de Tidball (2022).

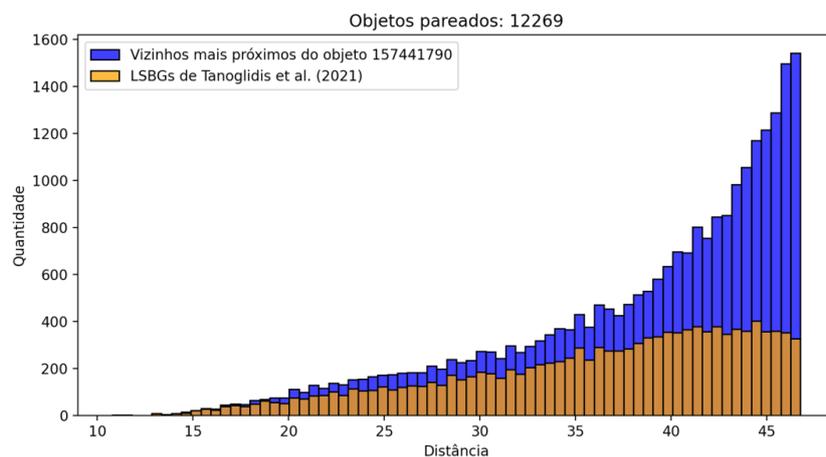


Figura 1.4 – Em azul, histograma da distância dos 25000 vizinhos mais próximos do objeto de ID=157441790. Em laranja, o histograma da distância dos vizinhos mais próximos da mesma chave que estão presentes no catálogo de Tanoglidis et al. (2021a). Imagem retirada de (Tidball, 2022).

1.3 Fotometria superficial

O brilho superficial μ é o fluxo por unidade de área emitido pelo objeto. Na Terra, o medimos como fluxo por unidade de ângulo sólido que chega ao observador. Já a distribuição de brilho superficial mostra como varia o fluxo por unidade de área ao longo da galáxia, e os perfis radiais de brilho superficial como ocorre a variação ao longo do raio da galáxia (Souza e Oliveira, 2017).

Desde de quando utilizavam-se chapas fotográficas para fazer fotometria de galáxias, um dos principais métodos para estudar a distribuição de luz das galáxias é o modelamento utilizando funções analíticas – uma técnica conhecida como ajuste paramétrico (Peng et al., 2002). Esse método busca descrever a forma e o perfil de brilho superficial das galáxias permitindo a obtenção de informações importantes sobre sua estrutura e dinâmica.

A caracterização paramétrica de galáxias de baixo brilho superficial é uma tarefa desafiadora devido às suas formas irregulares, morfologias complexas e brilho fraco. Embora o perfil Sérsic seja comumente usado para modelar a distribuição radial de luz de galáxias, isso pode ser um desafio para LSBGs, pois sua natureza difusa pode dificultar o modelamento e a estimativa do raio efetivo e outras propriedades necessárias para tal finalidade. Em particular, determinar o raio efetivo da LSBG é crucial para estimar propriedades físicas, como fluxo, que pode ser significativamente afetado pela escolha do raio.

No caso em que o ajuste paramétrico utiliza o perfil de Sérsic, o perfil de brilho superficial pode ser obtido multiplicando-se o logaritmo negativo da (1.1) por 2.5, resultando em

$$\mu(R) = \mu_e + \frac{2.5b_n}{\log(10)} \left[\left(\frac{R}{R_e} \right)^{1/n} - 1 \right]. \quad (1.2)$$

De acordo com Graham e Driver (2005), a partir do valor de μ_e , o brilho superficial no raio efetivo R_e , e conhecendo o valor índice n , podemos calcular μ_0 e $\bar{\mu}_e$ brilhos superficiais central e médio dentro do raio efetivo, respectivamente. No centro do perfil tem-se, pela equação (1.2),

$$\mu(R=0) \equiv \mu_0 = \mu_e - \frac{2.5b_n}{(\ln 10)} \quad (1.3)$$

$$\mu_0 = \mu_e - 1.822, \quad n = 1. \quad (1.4)$$

Já o brilho superficial médio dentro do raio efetivo de acordo com Graham e Driver (2005) o $\bar{\mu}_e$ é dado pela equação,

$$\bar{\mu}_e = \mu_e - 2.5 \log[f(n)], \quad (1.5)$$

$$= \mu_e - 0.699, \quad n = 1. \quad (1.6)$$

onde $f(n)$ é dado por,

$$f(n) = \frac{2ne^{b_n}}{b_n^{2n}} \int_0^{b_n} e^{-x} x^{2n-1} dx \quad (1.7)$$

$$= \frac{ne^{b_n}}{b_n^{2n}} \Gamma(2n), \quad (1.8)$$

onde Γ é a função Gama completa, e a magnitude dentro de um raio efetivo e a magnitude total são dadas por,

$$m(< R_e) = \bar{\mu}_e - 2.5 \log(\pi R_e^2) \quad (1.9)$$

$$m_{tot} = \bar{\mu}_e - 2.5 \log(2\pi R_e^2). \quad (1.10)$$

1.4 Objetivos e estrutura do trabalho

Esse trabalho se propõe a caracterizar uma amostra de novas candidatas a LSBGs que foram detectadas através do método de busca por similaridade (ver seção 1.2.1). A caracterização das candidatas identificadas se dará através de medidas de parâmetros estruturais e fotométricos. Posteriormente, utilizaremos os intervalos de valores encontrados para LSBGs do trabalho de Tanoglidis et al. (2021a) para classificarmos a nossa amostra de novas candidatas. Com isso, poderemos confirmar quais das novas candidatas identificadas pelo método de busca por similaridade são de fato LSBGs, além de criar um método automatizado que poderá ser aplicado a grandes amostras de candidatas a LSBGs.

Em virtude de precisarmos restringir a pesquisa para que fosse possível realizá-la dentro do prazo do TCC, aplicaremos o método para uma pequena amostra de objetos identificados pelo busca por similaridade. Futuramente aplicaremos para amostras maiores, validando a proposta de identificação de novas LSBGs de Tidball (2022).

O trabalho se estruturará da seguinte forma: No capítulo 2 apresentamos os dados e a metodologia utilizada, incluindo a descrição dos softwares usados para realizar a caracterização das LSBGs. Posteriormente, na no capítulo 3, apresentaremos as medidas realizadas ao longo do trabalho juntamente com as análises e apresentação dos resultados obtidos. Por fim, no capítulo 4, retomamos as discussões trabalhadas ao longo do texto trazendo o fechamento e perspectivas futuras.

2 Metodologia

Neste capítulo, primeiro apresentaremos os dados utilizados para caracterizar as LSBGs. Descreveremos a amostra de candidatas selecionadas pelo método de busca por similaridade, bem como a amostra de LSBGs utilizada para validar nossa amostra de candidatas. Além disso, forneceremos detalhes sobre os parâmetros, programas e métodos utilizados na caracterização das LSBGs.

2.1 Dados

Neste trabalho utilizamos dados do Dark Energy Survey (DES)¹, que é um levantamento fotométrico de grande área que visa estudar a natureza da energia escura. Os dados deste levantamento foram obtidos com uma câmera digital de 570 megapixels chamada Dark Energy Camera (DECam), montada no telescópio Blanco de 4 metros no Observatório Interamericano Cerro Tololo, no Chile. O DES começou a operar em 2013 e concluiu suas observações em 2019, cobrindo uma área de cerca de um oitavo do céu e seus dados já processados estão disponíveis publicamente.

Utilizamos os dados de dois trabalhos mencionados anteriormente: o catálogo de LSBGs publicado por Tanoglidis et al. (2021a)² e uma amostra de galáxias similares a LSBGs obtida a partir de uma LSBG conhecida (chave) por Tidball (2022).

Visto que o objetivo principal do trabalho é criar uma ferramenta de caracterização para confirmação da natureza de LSBGs das galáxias selecionadas pelo método de busca por similaridade e considerando o tempo limitado para a realização do trabalho, optamos por usar uma subamostra de 100 candidatas a LSBGs de Tidball (2022), que correspondem aos primeiros 100 vizinhos de uma chave escolhida aleatoriamente. Esta subamostra contém 59 galáxias em comum com o catálogo de Tanoglidis et al. (2021a) e 41 novas candidatas a LSBGs.

Ambos os trabalhos de Tanoglidis et al. (2021a) e Tidball (2022) utilizaram dados dos primeiros três anos de observação do DES, com base no catálogo DES Y3 Gold. Embora os catálogos apresentem propriedades medidas a partir de imagens nas três bandas fotométricas (g, r, i), neste trabalho escolhemos utilizar apenas imagens na banda g, já que apresentam melhor qualidade e, neste primeiro momento, não tínhamos interesse em estudar a cor dessas galáxias.

A partir do identificador dos objetos da nossa amostra de 100 galáxias selecionadas pelo método de busca por similaridade, utilizamos a plataforma científica NOIRLab Astro

¹ <https://des.ncsa.illinois.edu/>

² <https://des.ncsa.illinois.edu/releases/other/y3-lsbg>

Data Lab³ para obter as respectivas imagens da banda g em arquivo FITS. Obtivemos uma imagem de 1.2 x 1.2 arcmin em torno da candidata a LSBG. Esse tamanho de imagem foi escolhido de forma a conter as LSBGs de maior tamanho e também ter uma área suficientemente grande em torno da galáxia para medir adequadamente o céu.

Cada imagem FITS possui suas respectivas informações na extensão cabeçalho, como tempo de exposição, coordenadas de um ponto de referência, data de observação, informações sobre o CCD entre outras. Como os dados foram obtidos com o mesmo instrumento, alguns parâmetros têm o mesmo valor para todas as imagens, a saber: magnitude de ponto zero $m_{zpt} = 30$ mag e *pixel scale* igual a 0.27 arcsec/pixel.

É importante destacar aqui que a largura à meia altura (FWHM, do inglês *full width at half maximum*) da função de espalhamento pontual (PSF, do inglês *Point spread function*)⁴ pode ser obtida a partir da análise de estrelas não saturadas na região no entorno da galáxias a ser analisada. Uma das maneiras de fazer isto é utilizando a ferramenta IMEXAMINE do IRAF⁵ e o programa SAOImageDS9⁶, que ajusta e permite a visualização do perfil radial de brilho das estrelas escolhidas. Dado que este processo é mais difícil de ser automatizado e considerando a limitação de tempo para a realização deste trabalho, utilizamos um único valor de FWHM ($FWHM_m = 1.14$ arcsec) da PSF gaussiana para todos os objetos igual ao valor médio da FWHM para a banda g do DES (Sevilla-Noarbe et al., 2021).

2.2 Caracterização de LSBGs

Um dos primeiros passos para trabalhar com análises fotométricas de LSBGs é a remoção de objetos vizinhos (de primeiro plano e de fundo) que podem interferir na análise da distribuição de brilho superficial, causando contaminação na luz da candidata de interesse. Muitas vezes haverá estrelas ou galáxias distantes ou até mesmo defeitos da imagem no entorno ou sobreposto à LSBG. Dessa forma, é preciso mascarar esses objetos de alto brilho superficial na imagem e também objetos que podem estar muito próximos da candidata de interesse. Isso pode ser feito com o software SExtractor (descrito na seção 2.2.1).

Para uma boa análise de galáxias de baixo brilho superficial, é crucial realizar um ajuste cuidadoso do fundo de céu. Isso se deve ao fato de que o brilho superficial dessas galáxias é comparável ao brilho do céu noturno, o que torna difícil separar a luz da galáxia da luz de fundo. Além disso, é comum utilizar um ajuste de perfil de Sérsic (equação

³ <https://datalab.noirlab.edu/>

⁴ A PSF é uma representação matemática de como a luz de uma fonte pontual é distribuída sobre uma imagem.

⁵ IRAF é o Image Reduction and Analysis Facility, um sistema de software de propósito geral para a redução e análise de dados astronômicos. Disponível em <https://iraf.net/>

⁶ Ferramenta de exibição e visualização de imagens em formato FITS para dados astronômicos. Disponível em <https://ds9.si.edu>.

1.1) para modelar a distribuição radial de luz dessas galáxias. Para realizar esse ajuste, pode-se utilizar softwares como o GALFIT (descrito na seção 2.2.2), que é capaz de modelar a distribuição bidimensional de luz de uma galáxia em termos de uma função analítica. Dessa forma, o ajuste do fundo de céu e o ajuste de perfil de Sérsic são etapas cruciais para uma análise precisa de galáxias de baixo brilho superficial.

Por se tratar de um problema relativamente comum na caracterização de objetos astronômicos, antes de desenvolver um novo método automatizado para enfrentar esse problema, fizemos uma busca na literatura e encontramos um software que se adequava muito bem aos nossos propósitos, chamado Galfitpywrap⁷, que concatena o uso do SExtractor e do GALFIT. A seguir descreveremos em detalhes os softwares utilizados. Na seção 2.2.3 descreveremos o Galfitpywrap e como usamos ele para realizar medidas fotométricas da nossa amostra de candidatas a LSBGs, bem como todos os procedimentos adotados para a obtenção dos resultados apresentados no capítulo 3.

2.2.1 SOURCE EXTRACTOR

O SOURCE EXTRACTOR (SEXTRACTOR) é um programa que detecta e realiza medidas de fontes em imagens astronômicas, gerando um catálogo com as suas posições, magnitudes e outros parâmetros de interesse (Bertin e Arnouts, 1996).

Entre as principais características do SEXTRACTOR está a sua habilidade em processar imagens muito grandes rapidamente e com pouca intervenção humana. Dessa maneira, ele torna-se uma importante ferramenta na análise de dados de sondagens de grandes áreas do céu, sendo comumente utilizado para fazer fotometria de forma automatizada.

O programa realiza as detecções dos objetos a partir da determinação do fundo do céu da imagem e se um dado *pixel* pertence ao objeto ou ao céu, conforme os parâmetros limites definidos previamente pelo usuário no arquivo de configuração. Dessa forma, ele distingue o céu e objetos distintos, calcula propriedades dessas fontes, e posteriormente ele escreve um catálogo com estas informações em um arquivo de saída.

Cada vez que é executado, o SEXTRACTOR procura um arquivo de configuração no qual o usuário estabelece, por exemplo, a área mínima de detecção e os parâmetros de detecção, entre outros parâmetros que o programa irá usar para detectar, separar e modelar as fontes presentes na imagem FITS fornecida.

Dada uma imagem FITS de entrada, o SEXTRACTOR identifica grupos de *pixels* vizinhos cujos valores de intensidade são superiores a um certo valor limite considerado para detecção. Cada conjunto desses *pixels* vizinhos é dividido em uma ou mais fontes individuais, em um processo de *deblending*. Após a etapa de identificação das fontes e *deblending*, são disponibilizados os valores calculados, como magnitude, fluxo, coordenadas centrais, dentre outros, em um arquivo de saída. Também como saída é produzida

⁷ <https://github.com/Grillard/GalfitPyWrap>

uma imagem contendo o mapeamento dos *pixels* associados a cada objeto (imagem de segmentação).

Para LSBGs, é importante que o usuário tenha cuidado com esses parâmetros de configuração do SExtractor, uma vez que o brilho superficial das LSBGs é muito próximo daquele do fundo do céu.

Utilizamos o SExtractor para detectar e classificar todos os objetos contidos em cada imagem. Desta forma, com ele em conjunto com o SAOImageDS9, podemos identificar a LSBG da imagem e calibrar a melhor combinação de valores para os parâmetros responsáveis por distinguir os *pixels* dos objetos e do fundo do céu, e também entre objetos distintos.

2.2.2 GALFIT

O GALFIT (Galaxy Fitting, Peng et al. (2002)) é um programa, escrito na linguagem C, de ajuste bidimensional da distribuição de luz das galáxias através da sua decomposição em funções analíticas. Cada função tem um conjunto de parâmetros livres que podem ser ajustados para reproduzir o perfil de brilho superficial de uma imagem. Vale ressaltar que mais de uma função pode ser utilizada para representar diferentes componentes da galáxia, por exemplo, uma para o bojo e outra para o disco.

Inicialmente o GALFIT define um modelo inicial utilizando chutes iniciais dos parâmetros, que são definidos pelo usuário no seu arquivo de configuração. Este modelo é então convoluído com a PSF. O ajuste do GALFIT se dá através da minimização do χ^2 residual entre a imagem observada e o modelo convoluído, ajustando todos os parâmetros livres simultaneamente utilizando um algoritmo chamado Levenberg-Marquardt. O χ^2 é por definição:

$$\chi^2 = \sum_{x=1}^{nx} \sum_{y=1}^{ny} \frac{(F_{x,y} - M_{x,y})^2}{\sigma_{x,y}^2}, \quad (2.1)$$

onde $F_{x,y}$ é o fluxo do pixel (x, y) , $M_{x,y}$ é o fluxo dado pelo modelo ajustado para o mesmo *pixel*, nx e ny são, respectivamente, as dimensões x e y da imagem, $\sigma_{x,y}^2$ é a variância nos *pixels* x e y e, portanto, $\sigma_{x,y}^{-2}$ é o peso dado a cada *pixel*.

Uma vez encontrados os parâmetros do modelo que fornece o melhor ajuste à imagem original, são geradas as imagens de modelo e resíduo e também um arquivo de saída com os valores das medidas realizadas, tais como: magnitude, raio efetivo, índice Sérsic, valor médio do céu, etc.

2.2.3 Galfitpywrap

A caracterização das galáxias da nossa amostra foi feita com o código Galfitpywrap⁸. Esse é um código desenvolvido em linguagem *Python* que concatena a execução do

⁸ <https://github.com/Grillard/GalfitPyWrap>

SEXTRACTOR (seção 2.2.1) e GALFIT (seção 2.2.2) para identificar, separar e modelar objetos de interesse em imagens astronômicas.

Os arquivos de configuração e parâmetros do SEXTRACTOR precisam ser definidos pelo usuário, o único requisito é que ele crie uma imagem de segmentação. Além disso, o usuário também deve definir os modelos a serem ajustados no arquivo de configuração do GALFIT.

O Galfitpywrap usa o SEXTRACTOR para identificar as fontes contidas na imagem e criar a máscara de todos os objetos, exceto o central, que é a candidata a LSBG. A máscara é salva em um arquivo FITS para ser usada posteriormente pelo GALFIT para excluir os *pixels* de objetos brilhantes e objetos no entorno da candidata a LSBG no momento do seu modelamento.

Os parâmetros medidos pelo SEXTRACTOR para a candidata a LSBG (como magnitude, raio, elipticidade, ângulo de posição) e para o fundo do céu são passados para o arquivo de configuração do GALFIT como chutes iniciais dos parâmetros do modelo a ser ajustado. Foi considerado somente um perfil de Sérsic como modelo para descrever a distribuição de brilho das candidatas a LSBG. Como o SEXTRACTOR não estima o valor do índice de Sérsic n , usamos o chute inicial desse parâmetro igual a 1, pois esse é o valor típico de n para LSBGs.

Por fim, o Galfitpywrap executa o GALFIT e gera os seus arquivos de saída. A Figura 2.1 ilustra o ajuste realizado pelo Galfitpywrap para três LSBG da nossa amostra.

Na etapa seguinte, inspecionamos visualmente todas as imagens resultantes e categorizamos em dois grupos de acordo com a quantidade de fluxo residual na subtração do modelo nos dados. Para os casos em que o ajuste não foi satisfatório (baseado no valor do χ^2 e na análise visual da imagem residual), modificamos os parâmetros relacionados à criação da máscara ou alteramos os chutes iniciais de forma manual e executamos novamente o Galfitpywrap. Os parâmetro relacionados à máscara foram alterados pois estes são responsáveis pela segmentação dos *pixels* e por identificar quais *pixels* são pertencentes a objetos distintos contido na mesma imagem.

No entanto, mesmo assim verificamos que para 9 imagens (9% da nossa amostra) ainda encontramos fluxos residuais significativos (>20% do fluxo original). Uma possível razão para não ter ocorrido o ajuste do modelo corretamente se dá pelo fato de utilizarmos um único Sérsic, pois principalmente as galáxias com algum indício de estrutura ou irregulares não foram bem modeladas, como nos exemplos da Figura 2.2.

Esses objetos com fluxos residuais significativos foram considerados separadamente em alguns casos nas análises do capítulo 3.

Acreditamos que uma possível maneira de melhorar esse problema seria a inclusão de um outro perfil de Sérsic ou incluir uma fonte nuclear pontual em alguns casos, o que está além do escopo deste trabalho.

Embora espera-se que a utilização de um modelo de PSF específico para cada um dos

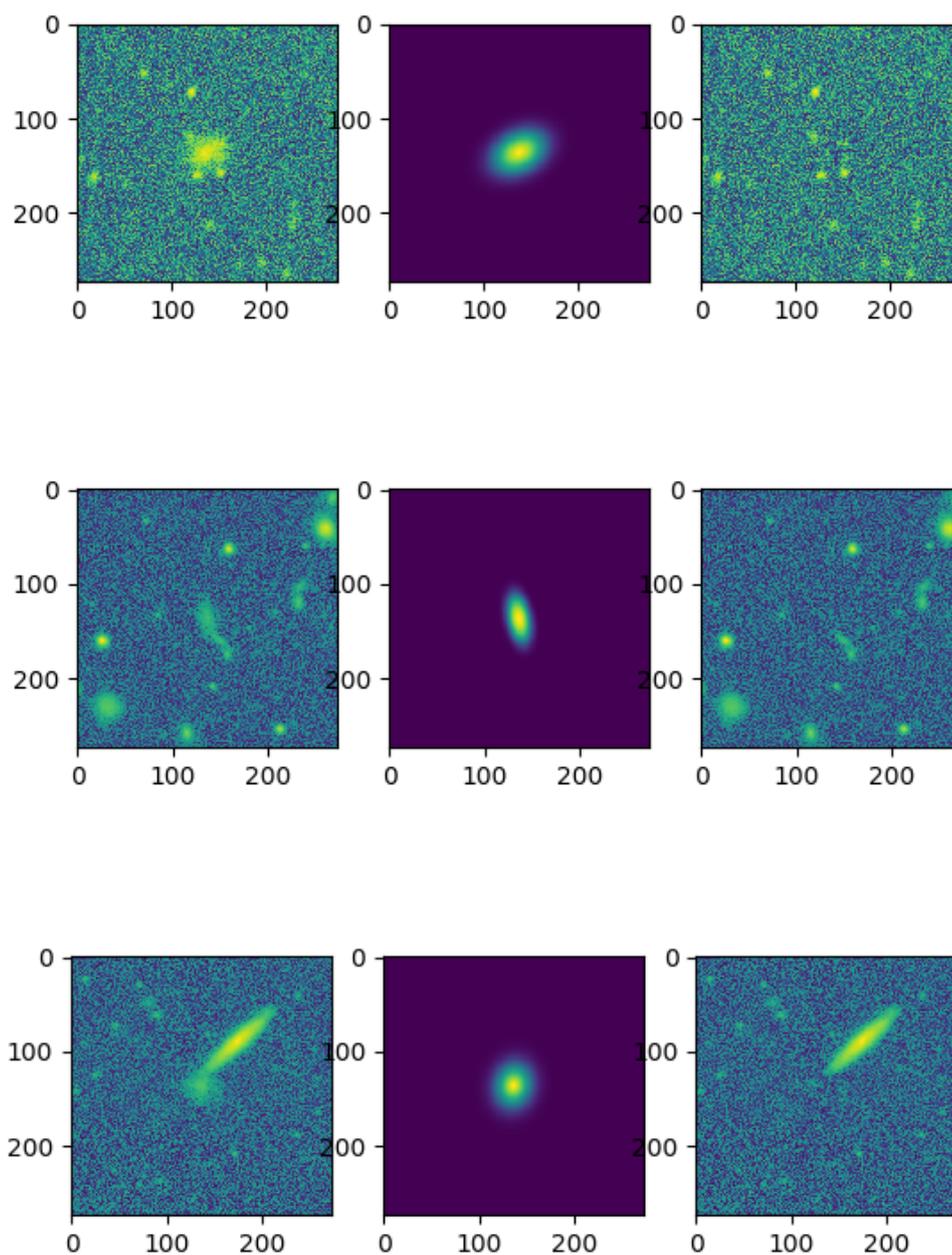


Figura 2.1 – Os painéis representam as imagens original (esquerda), modelo (centro), residual (direita) de saída após a execução do código Galfitpywrap adaptado para três imagens diferentes cujos modelos foram bem ajustados.

objetos obtenha um melhor ajuste, fica claro que para uma grande parte da nossa amostra o ajuste de um único perfil Sérsic foi muito satisfatória como podemos ver na Figura 2.1.

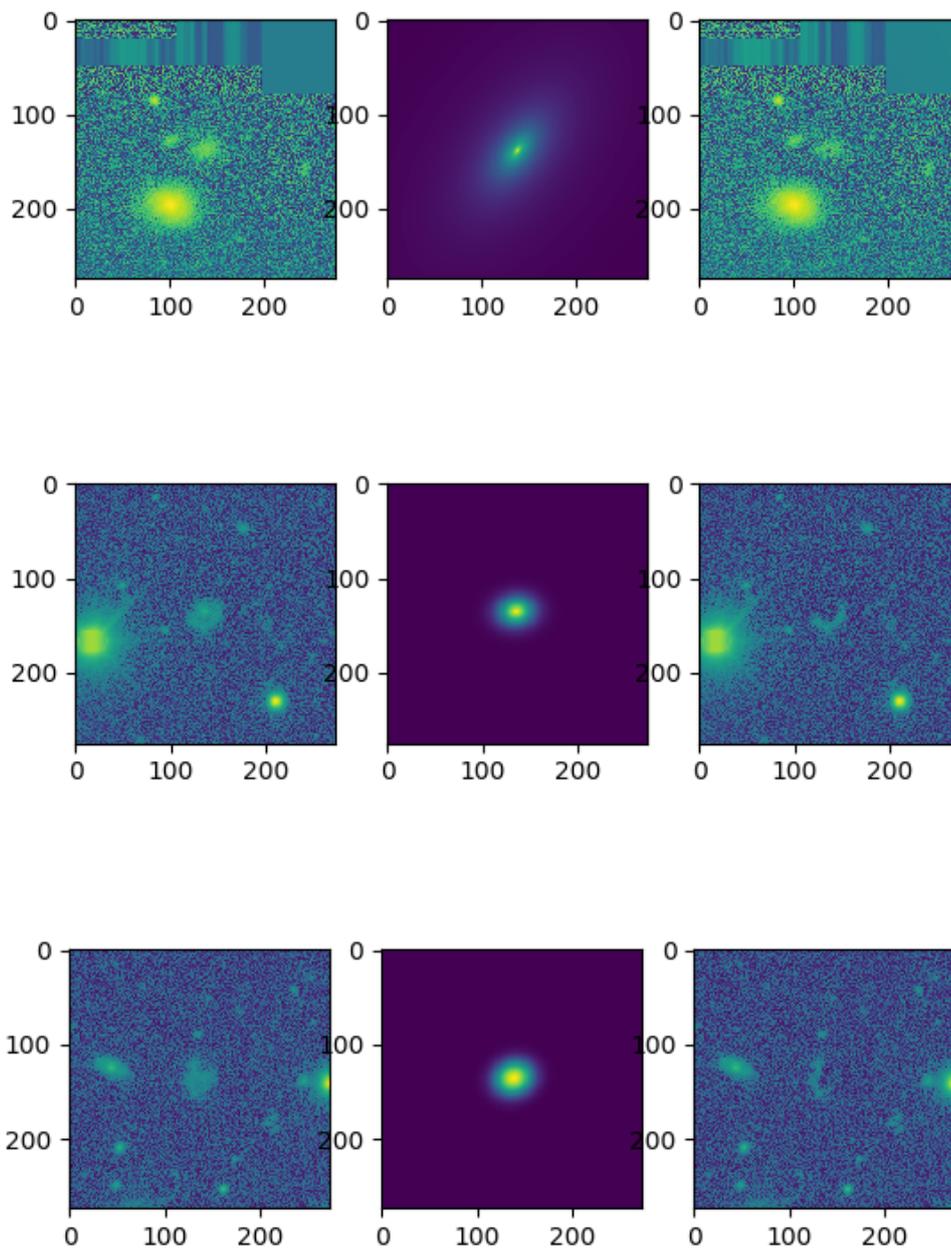


Figura 2.2 – Os painéis representam as imagens original (esquerda), modelo (centro), residual (direita) de saída após a execução do código Galfitpywrap adaptado para três imagens diferentes cujos modelos não foram bem ajustados.

3 Resultados e Discussão

A discussão a seguir se aprofunda nas análises realizadas em nossa amostra de dados. Apresentaremos uma validação dos métodos aplicados através da recuperação de LSBGs confirmadas contidas no catálogo de Tanoglidis et al. (2021a) e análise das candidatas a LSBGs.

3.1 Validação do método

Como validação da nossa metodologia, dada a amostra de 100 galáxias detectadas pelo método de similaridade (seção 1.2.1) por Tidball (2022), identificamos qual a intersecção de objetos entre este catálogo e o disponibilizado por Tanoglidis et al. (2021a), resultando em 59 LSBGs confirmadas no catálogo de Tanoglidis et al. (2021a) e 41 novas candidatas a LSBGs. Feito isso, foi possível produzir comparações das medidas realizadas por Tanoglidis et al. (2021a) e as nossas. Embora nossa amostra seja pequena, estatisticamente falando, as tendências não devem mudar de um método de medida para outro.

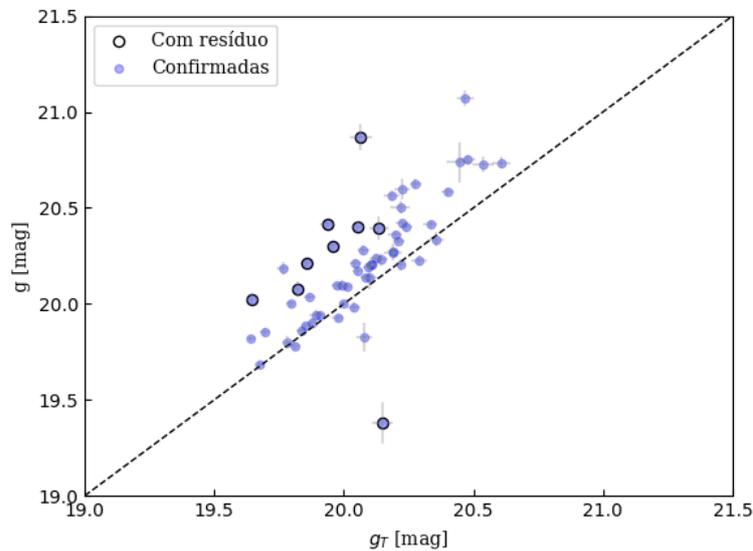


Figura 3.1 – Gráfico $g_T \times g$, onde g_T são as medidas de magnitude de Tanoglidis et al. (2021a) e g as medidas realizadas neste trabalho. A reta identidade é utilizada para visualizarmos quão distintas são as medidas em ambos os trabalhos.

Iniciamos comparando as medidas de magnitude na banda g , como mostra a Figura 3.1. Nela usamos a reta identidade como guia para sabermos o quão próximos estão os valores das medidas de ambos os trabalhos. Nota-se que as nossas medidas possuem valores maiores de magnitude, as LSBGs medidas pelo nosso procedimento são mais fracas.

Calculamos que a diferença média entre as magnitudes é de $\langle g - g_T \rangle = 0.15$ e o desvio padrão dos pontos em relação a essa média é $\sigma_g = 0.21$.

Pode haver várias razões para essa diferença. Vale a pena notar que a máscara que usamos pode ser um fator contribuinte. Podemos estar mascarando muito os objetos no entorno das LSBGs ou o Tanoglidis et al. (2021a) mascarou menos que o necessário e com isso as medidas deles podem sofrer contaminação de outros objetos, diminuindo sua magnitude e consequentemente descrevendo objetos mais brilhantes. Além disso, dada a dependência da magnitude medida com o raio efetivo R_e , como mostra a equação 1.10, se medimos R_e diferentes, as medidas de magnitude serão também diferentes.

Ao encontro disso, na Figura 3.2 as medidas de raio efetivo possuem uma maior concentração ligeiramente abaixo da reta identidade. Exceto pelos pontos com os maiores erros associados, os raios efetivos que medimos são menores que encontrados em Tanoglidis et al. (2021a). Calculamos que a diferença média entre os raios efetivos é de $\langle R_e - R_{e,T} \rangle = 0.05$ arcsec e o desvio padrão é $\sigma_{R_e} = 1.15$ arcsec. Vejamos que o desvio padrão associado a nossa medida é bem significativo, porém isso se dá devido a contribuição dos pontos com maior erro associado que dispersam a distribuição.

Essa diferença entre os valores de raios efetivos medidos pelo dois trabalhos pode ser causadas pelo uso de uma PSF gaussiana com a mesma FWHM para todas as LSBGs da nossa amostra, enquanto Tanoglidis et al. (2021a) usou uma PSF diferente para cada imagem. Isso pode explicar a diferença observada nas medidas de magnitude apresentadas anteriormente.

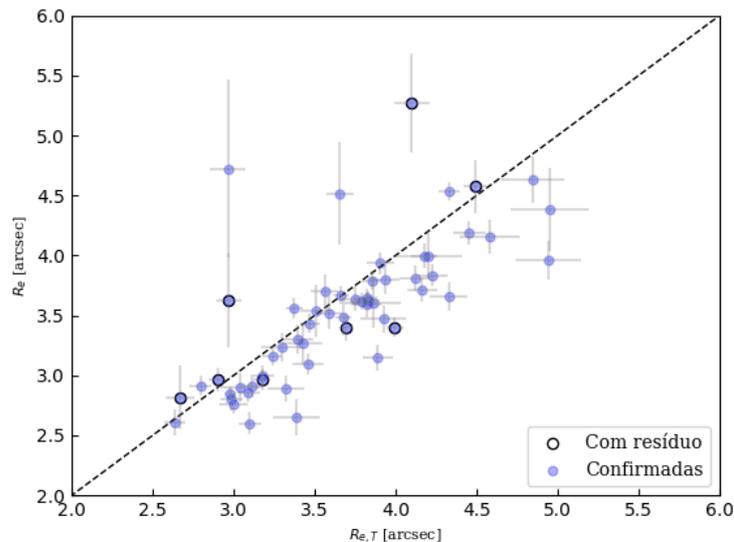


Figura 3.2 – Gráfico $R_{e,T} \times R_e$, onde $R_{e,T}$ são as medidas de raio efetivo de Tanoglidis et al. (2021a) e R_e as medidas realizadas neste trabalho.

Analisamos também a elipticidade ϵ das galáxias, conforme mostra a Figura 3.3. A

maioria das medidas encontra-se ligeiramente abaixo da reta identidade, ou seja, estamos medindo objetos com a elipticidades menores. Como $\epsilon = 1 - b/a$, nossa razão b/a (onde a é semieixo maior da elipse e b o semieixo menor) é maior, que leva a conclusão que para nós as LSBGs tendem a ser mais esferoidais. Obtemos que a diferença média entre as medidas de elipticidade é de $\langle \epsilon - \epsilon_T \rangle = -0.02$ e o desvio padrão é $\sigma_\epsilon = 0.09$. Essas diferenças entre os valores medidos pelo dois trabalhos também podem ser causadas pelo uso de uma PSF gaussiana com a mesma FWHM para todas as LSBGs da nossa amostra, enquanto Tanoglidis et al. (2021a) usou uma PSF diferente para cada imagem.

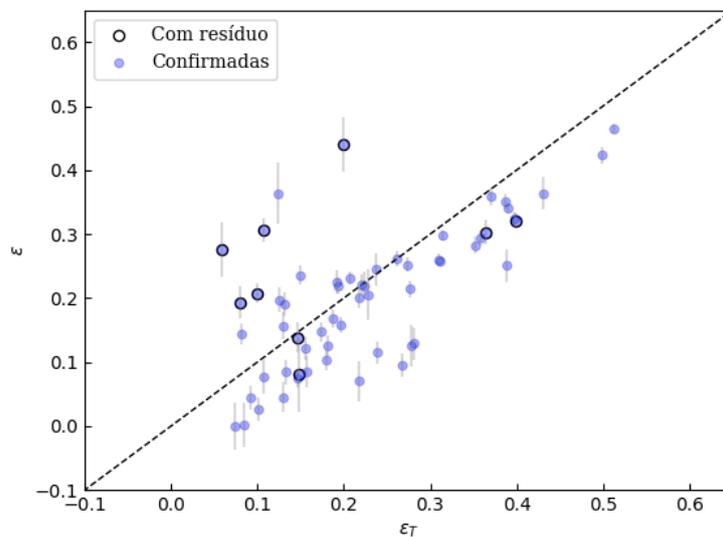


Figura 3.3 – Gráfico $\epsilon_T \times \epsilon$, onde ϵ_T são as medidas de elipticidade de Tanoglidis et al. (2021a) e ϵ as medidas realizadas neste trabalho.

Também podemos analisar de forma indireta a tendência obtida para a distribuição de brilho do modelo das galáxias através dos valores medidos em cada trabalho para o índice de Sérsic, como apresentado na Figura 3.4. A diferença média entre os valores de índice de Sérsic é de $\langle n - n_T \rangle = -0.05$ e o desvio padrão é $\sigma_n = 0.54$. Como podemos ver, medimos valores de n ligeiramente menores, ou seja, a distribuição de luz do modelo é mais suave. Isso pode também explicar em parte a diferença encontrada nas medidas de magnitudes.

3.2 Análise das novas candidatas a LSBGs

Utilizamos nosso script adaptado Galfitpywrap (seção 2.2.3) para modelar todos os objetos da nossa amostra selecionados pelo método de busca por similaridade. A partir dele foi possível estimar as seguintes propriedades físicas de nosso interesse: magnitude na banda g, raio efetivo, índice de Sérsic e a elipticidade. Também calculamos o brilho

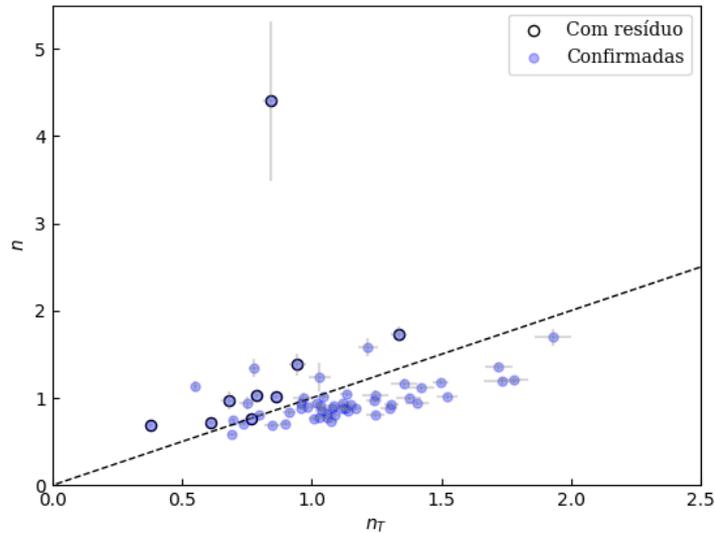


Figura 3.4 – Gráfico $n_T \times n$, onde n_T são as medidas de índice de Sérsic de Tanoglidis et al. (2021a) e n as medidas realizadas neste trabalho.

superficial médio dentro do raio efetivo, $\bar{\mu}_e$ de acordo com as equações descritas na seção 2.2.3. Após separamo-as em dois grupos para análise: LSBGs confirmadas (59 objetos contidos também no catálogo de Tanoglidis et al. (2021a)) e LSBGs candidatas (41 novos objetos), dentre essas, o subconjunto de imagens que não foi possível obter um bom modelo ($>20\%$ do fluxo original nos dados).

Na Figura 3.5 apresentamos os raios efetivos em função da magnitude na banda g para ambas as amostras e também os histogramas das distribuições de ambas as quantidades. Podemos notar que a distribuição das magnitudes das candidatas está deslocada para a esquerda, assumindo valores de magnitude na banda g maiores, ou seja, indicando que estas são mais fracas que as confirmadas. Também podemos ver que as medidas seguem a tendência de quanto maior o raio efetivo R_e menor o valor da magnitude, ou seja, mais brilhantes as LSBGs. Além disso, mesmo as imagens que apresentam resíduo significativo seguem a mesma tendência dos outros pontos. Isso é esperado dado que magnitude total e R_e estão correlacionados e que as quantidades da figura tendem a estar correlacionadas independentemente da qualidade do modelo.

É importante notar que aqui estamos comparando as medidas de ambas as subamostras feitas com o mesmo método, então os vieses nos parâmetros apresentados na seção anterior se manifestam da mesma forma para ambas as amostras. Em resumo, a partir das nossas medidas, temos que as novas candidatas tendem a ser menores e mais fracas que as LSBGs confirmadas. Isso pode indicar que o método de busca por similaridade está também detectando objetos que passaram despercebidos por Tanoglidis et al. (2021a), devido a algum corte em magnitude ou brilho superficial que os autores fizeram naquele trabalho.

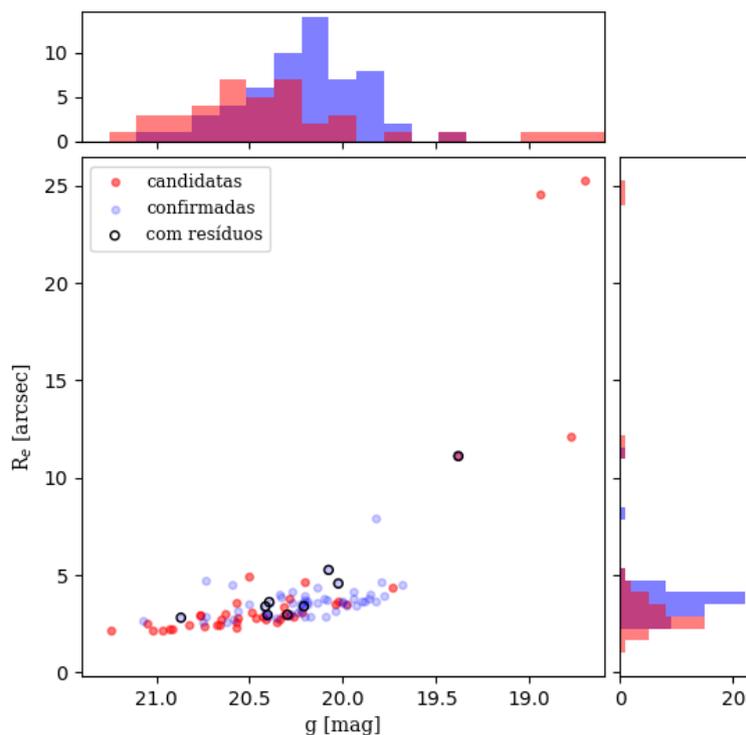
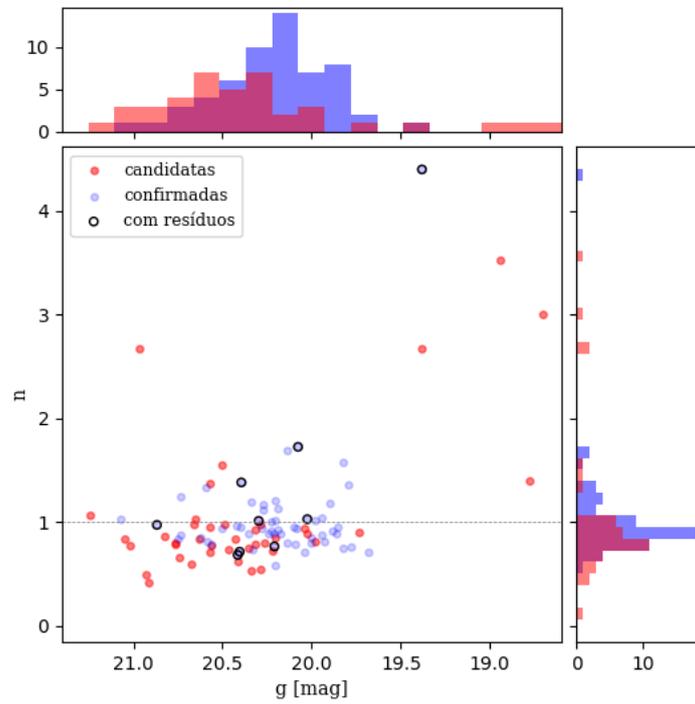


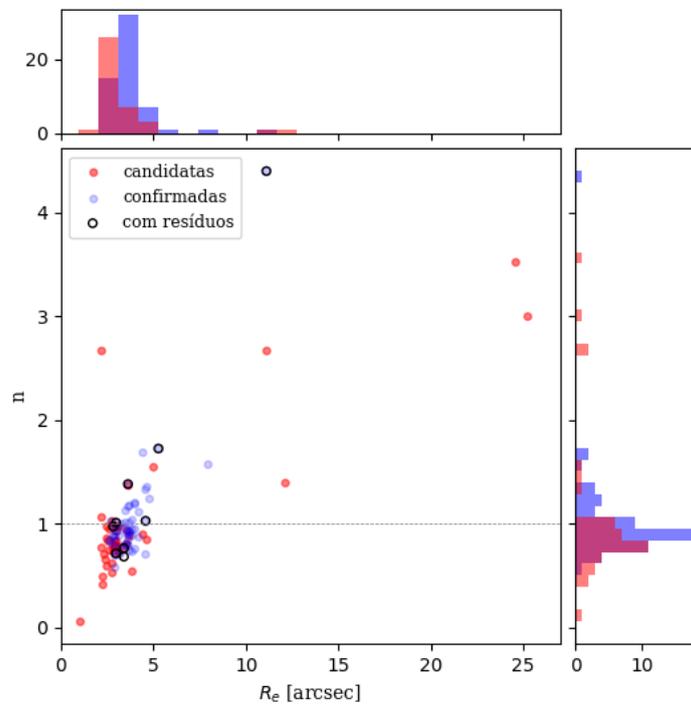
Figura 3.5 – Gráfico representando raio efetivo R_e em função da magnitude na banda g para a amostra de 59 LGBGs confirmadas, em azul, e 41 LSBGs candidatas, em vermelho, dentre elas o subconjunto de 9 LSBGs com resíduo significativo marcadas com um círculo preto.

Outro parâmetro calculado para caracterização da amostra, além da magnitude na banda g e o raio efetivo, é seu índice de Sérsic. Podemos notar na Figura 3.6 que ambas as subamostras estão com as medidas em torno de valores de $n = 1$, como esperado para LSBGs, já que possuem um perfil de brilho menos concentrado como representado pelas curvas com $n \leq 1$ figura 1.1. Notamos que as novas candidatas tendem a possuir n levemente menores do que as LSBGs confirmadas e esse resultado independe da magnitude (painel a) ou do raio efetivo (painel b) dos objetos. Há algumas novas candidatas que possuem n muito elevado ($n > 2.5$), o que indica que estas provavelmente não são LSBGs.

No geral, para a amostra de 100 galáxias utilizada, nossa análise aqui valida o método de busca de similaridade implementado por Tidball (2022). Ela também confirma que nosso método é capaz de recuperar as LSBGs detectadas e caracterizadas por Tanoglidis et al. (2021a). Ressaltamos que ainda é preciso realizar a validação para todo o domínio de chaves utilizadas em Tidball (2022). Entretanto, considerando que os 100 objetos analisados neste trabalho são uma boa representação dos vizinhos mais próximos das chaves de Tidball (2022), esperamos que a distribuição das propriedades desta amostra seja muito semelhante à distribuição das propriedades de toda a amostra de Tidball (2022).



(a)



(b)

Figura 3.6 – Nesta figura tanto a magnitude na banda g quanto o raio efetivo estão em função do índice de Sérsic. Note que em (a) o eixo da magnitude cresce da direita para a esquerda. Em ambos os gráficos temos em azul as LGBGs confirmadas, as LSBGs candidatas em vermelho e as LSBGs com resíduo significativo marcadas com um círculo preto.

4 Sumário e Conclusões

As LSBGs são um grupo diversificado e fascinante de objetos que abrangem uma ampla gama de tamanhos físicos, cores e ambientes. Diferentemente da maioria das galáxias estudadas até hoje, onde a sua região central é a parte mais brilhante de toda a galáxia, as LSBGs são muito difusas e têm seus brilhos superficiais comparáveis aos do céu noturno.

Elas são uma classe de objetos muito abundante e importante para o entendimento do Universo, uma vez que constituem uma porção significativa da matéria no Universo Local. Entretanto, devido a seu baixo brilho superficial, detectar e caracterizar estes objetos é muito desafiador. Além disso, a crescente quantidade de dados produzidos pelos atuais e futuros levantamentos fotométricos de grandes áreas do céu requerem o desenvolvimento de técnicas cada vez mais eficientes de detecção e processamento dos dados.

A partir da proposta de Tidball (2022) de detectar LSBGs através do método de busca por similaridade, foram encontrados objetos muito similares às chaves (LSBGs) presentes no catálogo de Tanoglidis et al. (2021a). Alguns destes objetos já faziam parte deste catálogo, enquanto outros eram novas candidatas.

Neste trabalho caracterizamos propriedades físicas de 100 galáxias de baixo brilho superficial identificadas por Tidball (2022) e presentes nos dados do terceiro ano do Dark Energy Survey. Destas, 41 eram novas candidatas a LSBGs e 59 LSBGs confirmadas, ou seja, já presentes no catálogo de Tanoglidis et al. (2021a). A caracterização se deu utilizando a técnica de fotometria superficial através do uso dos softwares `SEXTRACTOR` e `GALFIT` por intermédio das adaptações feitas no código público `Galfitpywrap` e sua execução para todas as galáxias da nossa amostra.

Identificamos em nossa amostra total que 9% dos objetos não obtiveram um bom ajuste somente com um perfil de Sérsic, resultando em resíduos que podem ser maiores do que 20% do fluxo original.

Usamos uma subamostra de LSBGs confirmadas para validar as nossas medidas e encontramos resultados similares aos de Tanoglidis et al. (2021a), embora com alguns vieses que podem ser explicados pelo uso de um PSF com FWHM inadequado. Esses ocorreram principalmente nas medidas de raio efetivo R_e , que afeta outras medidas também, como magnitude. Nossos resultados, para as já confirmadas, indicam galáxias de baixo brilho superficial menores e mais fracas o que pode estar relacionado ao fato de usarmos a mesma PSF para a análise de todas as galáxias.

Em geral, encontramos que as novas LSBGs são menores, mais fracas e mais difusas (menor n) do que as LSBGs confirmadas, mas que para a maioria das novas candidatas os parâmetros fotométricos e estruturais medidos estão no mesmo intervalo daqueles das LSBGs confirmadas por Tanoglidis et al. (2021a). Somente $\sim 5\%$ das novas LSBGs possuem parâmetros que não estão de acordo com os parâmetros das LSBGs de Tanoglidis et al.

(2021a). Isso indica uma capacidade do método de busca por similaridade desenvolvido por Tidball (2022) de encontrar objetos menos brilhantes do que o método empregado por Tanoglidis et al. (2021a). A caracterização dessas galáxias menos brilhantes que não foram encontradas por Tanoglidis et al. (2021a) é muito importante, pois explora um regime novo e relevante para entendermos a formação e evolução das galáxias.

Concluimos então que nosso método automatizado de caracterização foi validado, embora deve ser aperfeiçoado e que a grande maioria das novas candidatas encontradas pelo método de busca por similaridade teve sua natureza de LSBG confirmada, o que demonstra o potencial do método implementado por Tidball (2022) para encontrar novas LSBGs.

4.1 Perspectivas

Como perspectivas para trabalhos futuros, buscaremos expandir a amostra de candidatas à LSBGs e utilizar métodos estatísticos para melhorar e aumentar a acurácia dos parâmetros estimados. Construiremos PSFs individuais para cada galáxia para melhorar o modelo de distribuição de luz das imagens, e também realizaremos o ajuste multibanda com intuito de investigar as cores das galáxias da amostra.

Referências

- Bertin, E. e Arnouts, S. (1996). SExtractor: Software for source extraction. *Astronomy and Astrophysics*, 117:393–404.
- Bhatia, N. e Vandana (2010). Survey of Nearest Neighbor Techniques. (*IJCSIS*) *International Journal of Computer Science and Information Security*, página arXiv:1007.0085.
- Boselli, A., Voyer, E., Boissier, S., Cucciati, O., Consolandi, G., Cortese, L., Fumagalli, M., Gavazzi, G., Heinis, S., Roehly, Y., e Toloba, E. (2014). The GALEX Ultraviolet Virgo Cluster Survey (GUViCS). IV. The role of the cluster environment on galaxy evolution. *Astronomy and Astrophysics*, 570:A69.
- Bothun, G., Impey, C., e McGaugh, S. (1997). Low-Surface-Brightness Galaxies: Hidden Galaxies Revealed. *Publications of the ASP*, 109:745–758.
- Bothun, G. D., Impey, C. D., e Malin, D. F. (1987). The discovery of a highly flattened, optically dominated galaxy at a large redshift. *The Astrophysical Journal*, 313:596.
- Capaccioli, M. (1989). Photometry of early-type galaxies and the $r^{1/4}$ law. In Corwin, Harold G., J. e Bottinelli, L., editores, *World of Galaxies (Le Monde des Galaxies)*, páginas 208–227.
- Cimatti, A., Fraternali, F., e Nipoti, C. (2020). *Introduction to galaxy formation and evolution: From primordial gas to present-day galaxies*. Cambridge University Press.
- Ciotti, L. e Bertin, G. (1999). Analytical properties of the $R^{1/m}$ law. *Astronomy and Astrophysics*, 352:447–451.
- Disney, M. J. (1976). Visibility of galaxies. *Nature*, 263(5578):573–575.
- Driver, S. P. (1999). The Contribution of Normal, Dim, and Dwarf Galaxies to the Local Luminosity Density. *Astrophysical Journal*, 526(2):L69–L72.
- Graham, A. W. e Driver, S. P. (2005). A Concise Reference to (Projected) Sérsic $R^{1/n}$ Quantities, Including Concentration, Profile Slopes, Petrosian Indices, and Kron Magnitudes. *Publications of the Astron. Soc. of Australia*, 22(2):118–127.
- Gunn, J. E. e Gott, J. Richard, I. (1972). On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *Astrophysical Journal*, 176:1.
- Lee, J. H., Kang, J., Lee, M. G., e Jang, I. S. (2020). The Nature of Ultra-diffuse Galaxies in Distant Massive Galaxy Clusters: A370 in the Hubble Frontier Fields. *Astrophysical Journal*, 894(1):75.

- Leskovec, J., Rajaraman, A., e Ullman, J. D. (2020). *Mining of Massive Datasets*. Cambridge University Press, 3 edição.
- McConnachie, A. W. (2012). The Observed Properties of Dwarf Galaxies in and around the Local Group. *Astronomical Journal*, 144(1):4.
- McGaugh, S. S. e de Blok, W. J. G. (1998). Testing the Dark Matter Hypothesis with Low Surface Brightness Galaxies and Other Evidence. *Astrophysical Journal*, 499(1):41–65.
- Merritt, D. (1983). Relaxation and tidal stripping in rich clusters of galaxies. I. Evolution of the mass distribution. *Astrophysical Journal*, 264:24–48.
- Muzzin, A., Wilson, G., Yee, H. K. C., Gilbank, D., Hoekstra, H., Demarco, R., Balogh, M., van Dokkum, P., Franx, M., Ellingson, E., Hicks, A., Nantais, J., Noble, A., Lacy, M., Lidman, C., Rettura, A., Surace, J., e Webb, T. (2012). The Gemini Cluster Astrophysics Spectroscopic Survey (GCLASS): The Role of Environment and Self-regulation in Galaxy Evolution at $z \sim 1$. *Astrophysical Journal*, 746(2):188.
- O’Neil, K. e Bothun, G. (2000). The space density of galaxies through $b(0) = 25.0$ magnitudes per inverse arcsecond squared. *The Astrophysical Journal*, 529(2):811.
- Peng, C. Y., Ho, L. C., Impey, C. D., e Rix, H.-W. (2002). Detailed Structural Decomposition of Galaxy Images. *Astronomical Journal*, 124(1):266–293.
- Peng, Y.-j., Lilly, S. J., Renzini, A., e Carollo, M. (2012). Mass and Environment as Drivers of Galaxy Evolution. II. The Quenching of Satellite Galaxies as the Origin of Environmental Effects. *Astrophysical Journal*, 757(1):4.
- Prole, D. J., van der Burg, R. F. J., Hilker, M., e Davies, J. I. (2019). Observational properties of ultra-diffuse galaxies in low-density environments: field UDGs are predominantly blue and star forming. *Monthly Notices of the RAS*, 488(2):2143–2157.
- Sevilla-Noarbe, I., Bechtol, K., Kind, M., Carnero Rosell, A., Becker, M., Drlica-Wagner, A., Gruendl, R., Rykoff, E., Sheldon, E., Yanny, B., Alarcon, A., Allam, S., Amon, A., Benoit-Lévy, A., Bernstein, G., Bertin, E., Burke, D., Carretero, J., Choi, A., e Wilkinson, R. (2021). Dark energy survey year 3 results: Photometric data set for cosmology. *The Astrophysical Journal Supplement Series*, 254:24.
- Souza, O. F. K. d. e Oliveira, S. M. d. F. (2017). *Astronomia e Astrofísica*. Livraria da Física.
- Tanoglidis, D., Čiprijanović, A., e Drlica-Wagner, A. (2021a). DeepShadows: Separating low surface brightness galaxies from artifacts using deep learning. *Astronomy and Computing*, 35:100469.

- Tanoglidis, D., Drlica-Wagner, A., Wei, K., Li, T. S., Sánchez, J., Zhang, Y., Peter, A. H. G., Feldmeier-Krause, A., Prat, J., Casey, K., Palmese, A., Sánchez, C., DeRose, J., Conselice, C., Gagnon, L., Abbott, T. M. C., Aguena, M., Allam, S., Avila, S., Bechtol, K., Bertin, E., Bhargava, S., Brooks, D., Burke, D. L., Rosell, A. C., Kind, M. C., Carretero, J., Chang, C., Costanzi, M., da Costa, L. N., De Vicente, J., Desai, S., Diehl, H. T., Doel, P., Eifler, T. F., Everett, S., Evrard, A. E., Flaughner, B., Frieman, J., García-Bellido, J., Gerdes, D. W., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hartley, W. G., Hollowood, D. L., Huterer, D., James, D. J., Krause, E., Kuehn, K., Kuropatkin, N., Maia, M. A. G., March, M., Marshall, J. L., Menanteau, F., Miquel, R., Ogando, R. L. C., Paz-Chinchón, F., Romer, A. K., Roodman, A., Sanchez, E., Scarpine, V., Serrano, S., Sevilla-Noarbe, I., Smith, M., Suchyta, E., Tarle, G., Thomas, D., Tucker, D. L., Walker, A. R., e DES Collaboration (2021b). Shadows in the Dark: Low-surface-brightness Galaxies Discovered in the Dark Energy Survey. *Astrophysical Journal*, 252(2):18.
- Teeninga, P., Moschini, U., Trager, S. C., e Wilkinson, M. H. F. (2015). Improved detection of faint extended astronomical objects through statistical attribute filtering. In Benediktsson, J. A., Chanussot, J., Najman, L., e Talbot, H., editores, *Mathematical Morphology and Its Applications to Signal and Image Processing*, páginas 157–168, Cham. Springer International Publishing.
- Tidball, M. (2022). Busca de galáxias de baixo brilho superficial por similaridade em grandes catálogos astronômicos.
- van Dokkum, P., Danieli, S., Cohen, Y., Merritt, A., Romanowsky, A. J., Abraham, R., Brodie, J., Conroy, C., Lokhorst, D., Mowla, L., O’Sullivan, E., e Zhang, J. (2018). A galaxy lacking dark matter. *Nature*, 555(7698):629–632.
- van Dokkum, P. G., Abraham, R., Merritt, A., Zhang, J., Geha, M., e Conroy, C. (2015). Forty-seven Milky Way-sized, Extremely Diffuse Galaxies in the Coma Cluster. *Astrophysical Journal*, 798(2):L45.
- Venhola, A., Peletier, R., Laurikainen, E., Salo, H., Lisker, T., Iodice, E., Capaccioli, M., Verdois Kleijn, G., Valentijn, E., Mieske, S., Hilker, M., Wittmann, C., van de Ven, G., Grado, A., Spavone, M., Cantiello, M., Napolitano, N., Paolillo, M., e Falcón-Barroso, J. (2017). The Fornax Deep Survey with VST. III. Low surface brightness dwarfs and ultra diffuse galaxies in the center of the Fornax cluster. *Astronomy and Astrophysics*, 608:A142.
- Wang, J., Shen, H. T., Song, J., e Ji, J. (2014). Hashing for Similarity Search: A Survey. *arXiv e-prints*, página arXiv:1408.2927.

-
- Wittmann, C., Lisker, T., Ambachew Tilahun, L., Grebel, E. K., Conselice, C. J., Penny, S., Janz, J., Gallagher, John S., I., Kotulla, R., e McCormac, J. (2017). A population of faint low surface brightness galaxies in the Perseus cluster core. *Monthly Notices of the Royal Astronomical Society*, 470(2):1512–1525.
- Zaritsky, D., Donnerstein, R., Dey, A., Kadowaki, J., Zhang, H., Karunakaran, A., Martínez-Delgado, D., Rahman, M., e Spekkens, K. (2019). Systematically Measuring Ultra-diffuse Galaxies (SMUDGes). I. Survey Description and First Results in the Coma Galaxy Cluster and Environs. *Astrophysical Journal*, 240(1):1.