



Observatório  
Nacional

DISSERTAÇÃO DE MESTRADO

BUSCA POR PROTOAGLOMERADOS NO *DARK ENERGY SURVEY* COSMOS  
DEEP FIELD

GABRIEL VALIM CALÇADA

RIO DE JANEIRO

2024



Ministério da Ciência, Tecnologia, Inovações e Comunicações

Observatório Nacional

Programa de Pós-Graduação

Dissertação de Mestrado

BUSCA POR PROTOAGLOMERADOS NO *DARK ENERGY SURVEY* COSMOS  
DEEP FIELD

por

Gabriel Valim Calçada

Dissertação submetida ao Corpo Docente do Programa de Pós-graduação em Astronomia do Observatório Nacional, como parte dos requisitos necessários para a obtenção do Grau de Mestre em Astronomia.

Orientador: Prof. Dr. Ricardo Lourenço  
Correia Ogando

Rio de Janeiro, RJ – Brasil

Junho de 2024

V837

Valim Calçada, Gabriel

Busca por protoaglomerados no *Dark Energy Survey* Cosmos Deep Field [Rio de Janeiro] 2024.

xxiv, 89 p. 29,7 cm: graf. il. tab.

Dissertação (mestrado) - Observatório Nacional - Rio de Janeiro, 2024.

1. Galaxies: Protocluster. 2. Galaxies: High-redshift. 3. Galaxies: Clusters. I. Observatório Nacional. II. Título.

CDU 000.000.000

“BUSCA POR PROTOAGLOMERADOS NO *DARK ENERGY SURVEY* COSMOS  
DEEP FIELD”

GABRIEL VALIM CALÇADA

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ASTRONOMIA DO OBSERVATÓRIO NACIONAL COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM ASTRONOMIA.

Aprovada por:

---

Prof. Dr. Ricardo Lourenço Correia Ogando –  
Observatório Nacional - ON  
(Orientador)

---

Prof. Dr. Andre Luis Batista Ribeiro – Universidade  
Estadual de Santa Cruz - UESC

---

Prof. Dr. Jose Eduardo Telles – Observatório Nacional  
- ON

RIO DE JANEIRO, RJ – BRASIL

5 DE JUNHO DE 2024







# Agradecimentos

Gostaria de agradecer a Deus, em primeiro lugar. Minha fé em Deus sempre serviu de alicerce em cada um de meus passos e foi fundamental para esta conquista. Agradeço a minha noiva e aos meus pais, por todo apoio e suporte. Eu sou muito grato pela a consideração, preocupação e o suporte de vocês em cada etapa desse processo.

Meu agradecimento e gratidão ao meu orientador Ricardo Ogando por ter me proporcionado fazer essa pesquisa, por toda paciência, dedicação e orientação no desenvolvimento da pesquisa.

Minha gratidão ao professor Roderik Overzier que me recebeu como aluno de iniciação científica no Observatório Nacional, sendo o meu primeiro contato com a astronomia.

Agradeço a Christianne e a Giane que foram sempre amorosas, prestativas e sempre resolveram os nossos problemas com a maior agilidade possível.

Agradeço ao Observatório que me deu a oportunidade de trabalhar neste projeto, me fornecendo ainda uma mesa e um computador.

Agradeço a CAPES, que tornou a realização desta etapa possível ao me conceder uma bolsa de mestrado.



“Eu descobri em mim mesmo desejos que nada nesta terra podem satisfazer;  
a única explicação lógica é que eu fui feito para um outro mundo”.

---

- C. S. Lewis



BUSCA POR PROTOAGLOMERADOS NO *DARK ENERGY SURVEY* COSMOS  
DEEP FIELD

**RESUMO**

Aglomerados de galáxias são alicerces da estrutura em grande escala cuja abundância e respectivas massas fornecem informações sobre a organização do universo, além de ambientes especiais para a evolução de galáxias. Neste trabalho, foram utilizados dados do *Dark Energy Survey*, em particular Deep Fields, para buscar por candidatos a protoaglomerados, os progenitores de aglomerados de galáxias. Com um algoritmo para encontrar sobredensidades de galáxias, identificamos potenciais candidatos a protoaglomerados numa faixa de *redshift* de 0,5 a 2,0. Foram exploradas algumas propriedades dos candidatos e suas galáxias-membros. O desenvolvimento da metodologia e sua validação foi baseado em Narwal *et al.* (2020). Foram encontrados 176 candidatos e os resultados foram comparados a quatro catálogos da literatura na mesma região que a nossa. Como resultado de um match 3D entre os catálogos, foram identificados 150 candidatos inéditos que não apresentaram *match* com a literatura. Apenas 2 candidatos são comuns a todos os catálogos, evidenciando as incongruências entre eles. Realizamos uma análise visual dos 150 candidatos inéditos e encontramos 22 candidatos com múltiplos núcleos de sobredensidade, 30 candidatos apresentaram uma desproporção elevada entre as populações de galáxias azuis e vermelhas, identificadas com Gaussian Mixture Model, além de 12 candidatos nas bordas das células de detecção do mapa de sobredensidade. Foram eliminados 9 candidatos que apresentaram distorções no Diagrama Cor-Magnitude. A quantidade final de candidatos passou de 150 para 141. Os estudos de morfologia mostraram que há conexão com as cores, com galáxias dominadas por bojo sendo mais vermelhas e mais observadas a baixos *redshift*. Essa técnica pode ser aplicada a futuros levantamentos mais profundos e fornecer alvos para *follow-up* espectroscópico.



SEARCH FOR PROTOCLUSTERS IN DARK ENERGY SURVEY COSMOS DEEP  
FIELD

**ABSTRACT**

Galaxy clusters are the foundations of large-scale structure, whose abundance and respective masses provide information about the organization of the universe, as well as special environments for galaxy evolution. In this work, data from the Dark Energy Survey, particularly the Deep Fields, were used to search for protocluster candidates, the progenitors of galaxy clusters. Using an algorithm to find galaxy overdensities, we identified potential protocluster candidates in a redshift range from 0.5 to 2.0. Some properties of the candidates and their galaxies-member were explored. The development and validation of the methodology were based on Narwal *et al.* (2020). We found 176 candidates, and the results were compared to four literature catalogs in the same region as this work. As a result of a 3D match between the catalogs, 150 new candidates were identified that did not match any literature. Only 2 candidates appear in all catalogs, highlighting the inconsistencies between them. We conducted a visual analysis of the 150 new candidates and found 22 candidates with multiple density nuclei, 30 candidates showed a high disproportion between the populations of blue and red galaxies, identified with the Gaussian Mixture Model, and 12 candidates were at the edges of the detection cells of the density map. Nine candidates that showed distortions in the Color-Magnitude Diagram were eliminated. The final number of candidates went from 150 to 141. Morphological studies showed that there is a connection with colors, with bulge-dominated galaxies being redder and more observed at low redshifts. This technique can be applied to future deeper surveys providing targets for spectroscopic follow-up.



# Lista de Figuras

1.1	Representação do <i>redshift</i> fotométrico através de um espectro deslocado para o vermelho atravessando as bandas do sistema fotométrico ugriz em passos de 0,0, 0,4 e 0,8 de <i>redshift</i> . Fonte: Vanderplas <i>et al.</i> (2012) . . . . .	2
1.2	O primeiro <i>Deep Field</i> do James Webb. Aglomerado de galáxias SMACS 0723. Fonte: <a href="https://www.nasa.gov/image-feature/goddard/2022/nasa-s-webb-delivers-deepest-infrared-image-of-universe-yet">https://www.nasa.gov/image-feature/goddard/2022/nasa-s-webb-delivers-deepest-infrared-image-of-universe-yet</a> . . . . .	4
1.3	Imagem produzida pelo telescópio James Webb que apresenta um protoaglomerado a um <i>redshift</i> de 7,9. Fonte: NASA, ESA, CSA, T. Morishita (IPAC). Processamento de imagem: A. Pagan (STScI). . . . .	5
1.4	Distribuição de <i>redshifts</i> de protoclusters selecionados da literatura (círculos azuis cheios). Os dados para aglomerados abaixo de $z = 1,5$ foram retirados da compilação de aglomerados detectado em Bleem <i>et al.</i> (2015). Círculos vermelhos cheios são aglomerados em alto <i>redshift</i> , em $z > 1,6$ . A posição dos objetos ao longo do eixo polar não contém informações e é usado apenas para fins de visibilidade. <i>Redshifts</i> e idades (em Ganos) são indicados ao longo do eixo radial. Fonte: Overzier (2016). . . . .	6
1.5	Histograma de <i>redshift</i> de todos os dados do DES-SVA1 (azul) e das galáxias selecionadas (laranja). Fonte: Narwal <i>et al.</i> (2020) . . . . .	8
1.6	Curvas de nível do contraste de densidade suavizado exibindo as estruturas de sobredensidade no bin 1,04-1,24. Fonte: Figura 5 de Narwal <i>et al.</i> (2020). . . . .	9
1.7	Exemplo de uma tesselação de Voronoi tirada do artigo de Sochting <i>et al.</i> (2012). Esta seria uma fatia em $z \sim 2,7$ . . . . .	10
1.8	Distribuição do parâmetro fracdev para as amostras espirais e elípticas do <i>Sloan Digital Sky Survey</i> (SDSS). Fonte: Rodríguez e Padilla (2013). . . . .	14
2.1	Domo do telescópio Blanco no CTIO. Fonte: <a href="https://www.darkenergysurvey.org/the-des-project/instrument/">https://www.darkenergysurvey.org/the-des-project/instrument/</a> . . . . .	18
2.2	Representação dos CCDs DECam. Fonte: <a href="https://www.darkenergysurvey.org/the-des-project/instrument/">https://www.darkenergysurvey.org/the-des-project/instrument/</a> . . . . .	18
2.3	<i>Footprint</i> do DES em coordenadas equatoriais. Fonte Abbott <i>et al.</i> (2021). . . . .	19

2.4	Imagens coloridas (grz) em diferentes níveis de profundidade cobrindo metade de um dos chips no campo SN-X3. Fonte: Hartley <i>et al.</i> (2022). . . . .	19
2.5	Filtros da DECam: grizY. Fonte: Abbott <i>et al.</i> (2018) . . . . .	21
2.6	Filtros da VIRCам: YJHKs. Fonte: <a href="http://cosmos.ipac.caltech.edu">http://cosmos.ipac.caltech.edu</a> . . . . .	21
2.7	Exemplo da <i>query</i> do sistema do DES para obtenção do catálogo de galáxias. . . . .	22
2.8	Erros associados as magnitudes referentes as bandas griz extraídos do catálogo do DES. . . . .	25
2.9	Histograma de frequência de magnitude das bandas griz extraídos dos dados do DES. . . . .	26
2.10	RA e Dec do COSMOS extraído dos dados do DES. . . . .	26
2.11	Histograma de <i>counts</i> versus o <i>redshift</i> do campo do COSMOS. . . . .	27
2.12	Dados de BDF <i>fracdev</i> em função da magnitude na banda r das galáxias do catálogo G24. O painel da esquerda apresenta os dados completos e o painel da direita apresenta os dados após o corte em $BDF\ \text{fracdev} = 0,5_{-0,01}^{+0,005}$ e valores de $r(mag) > 24$ . As linhas vermelhas apresentam a zona do corte. . . . .	27
2.13	Histograma de <i>counts</i> versus o <i>redshift</i> dos <i>Deep Fields</i> . . . . .	28
2.14	Histograma de <i>counts</i> versus o <i>redshift</i> de Finoguenov <i>et al.</i> (2007). . . . .	29
2.15	Histograma de <i>counts</i> versus o <i>redshift</i> de Wen e Han (2011). . . . .	30
2.16	Dados de riqueza vs <i>redshift</i> de Wen e Han (2011). . . . .	30
2.17	Histograma de <i>counts</i> versus o <i>redshift</i> de Bellagamba <i>et al.</i> (2011). . . . .	31
2.18	Dados de riqueza vs <i>redshift</i> de Bellagamba <i>et al.</i> (2011). . . . .	31
2.19	Histograma de <i>counts</i> versus o <i>redshift</i> de Sochting <i>et al.</i> (2012). . . . .	32
2.20	Dados de riqueza vs <i>redshift</i> de Sochting <i>et al.</i> (2012). . . . .	32
3.1	Esquema representativo de busca de sobredensidades em diferentes <i>bins</i> de <i>redshift</i> . . . . .	34
3.2	Quantidade de galáxias em cada bin de <i>redshift</i> de um candidato. O ponto laranja representa o <i>bin</i> de <i>redshift</i> com maior quantidade de galáxias, sendo portanto escolhido como o <i>redshift</i> do candidato. . . . .	36
3.3	Exemplo de um CMD (painel da esquerda) e de um histograma número de galáxias distribuídas em função da cor (painel da direita) para o candidato de ID 42. . . . .	38
4.1	Histogramas de <i>redshift</i> dos dados do SVA1 gerados no trabalho de Narwal <i>et al.</i> (2020) (painel superior) e o reproduzido neste trabalho (painel inferior). . . . .	40
4.2	<i>Contour plot</i> do contraste de densidade exibindo as estruturas de sobredensidade no <i>bin</i> 1,01-1,21. O painel superior foi obtido no trabalho de Narwal <i>et al.</i> (2020) e o painel inferior foi obtido no presente trabalho, com a sobredensidade sendo destacada pelo círculo branco. . . . .	41

4.3	A direita, tem-se a porcentagem de <i>match</i> pelo tamanho do pixel referente ao histograma de sobredensidades. A esquerda, tem-se a relação entre número de <i>match</i> pelo tamanho do pixel. . . . .	42
4.4	Painéis variando o tamanho do pixel de 1 a 10, da esquerda para a direita contendo o primeiro <i>bin</i> de <i>redshift</i> . Primeira linha contém painéis de pixel 1, 2 e 3, segunda linha pixels 4, 5 e 6, terceira linha pixels 7, 8 e 9 e quarta linha pixel 10. . . . .	45
4.5	Fração de <i>matches</i> com o catálogo G24 em função do tamanho do pixel em Mpc, exceto para o pixel de tamanho igual a 2 Mpc por ter 100% de <i>match</i> por ser o catálogo G24. . . . .	46
4.6	<i>Match</i> entre o catálogo G24 e um catálogo gerado com pixel igual a 1 Mpc (1G24). . . . .	46
4.7	<i>Match</i> entre o catálogo G24 e um catálogo gerado com pixel igual a 2 Mpc (G24). . . . .	47
4.8	<i>Match</i> entre o catálogo G24 e um catálogo gerado com pixel igual a 3 Mpc (3G24). . . . .	47
4.9	<i>Contour plot</i> do contraste de densidade exibindo as estruturas de sobredensidade no primeiro (esquerda) e no último (direita) <i>bin</i> de <i>redshift</i> . . . .	48
4.10	<i>Match</i> 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de <i>clusters</i> de Sochting <i>et al.</i> (2012) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis. . . . .	49
4.11	<i>Match</i> 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de <i>clusters</i> de Finoguenov <i>et al.</i> (2007) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis. . . . .	50
4.12	<i>Match</i> 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de <i>clusters</i> de Wen e Han (2011) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis. . . . .	51
4.13	<i>Match</i> 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de <i>clusters</i> de Bellagamba <i>et al.</i> (2011) (pontos verdes) Os candidatos em comum estão representados pelos círculos azuis. . . . .	52
4.14	Valores de $dz$ ( $z_{W11} - z_{G24}$ ) versus <i>redshift</i> (Wen e Han, 2011). . . . .	53
4.15	Valores de $dz$ ( $z_{B11} - z_{G24}$ ) versus <i>redshift</i> (Bellagamba <i>et al.</i> , 2011). . . . .	53
4.16	Valores de $dz$ ( $z_{S12} - z_{G24}$ ) versus <i>redshifts</i> (Sochting <i>et al.</i> , 2012). . . . .	54
4.17	Valores de $dz$ ( $z_{F07} - z_{G24}$ ) versus <i>redshifts</i> (Finoguenov <i>et al.</i> , 2007). . . . .	54
4.18	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Wen e Han (2011) (pontos azuis). O painel da esquerda apresenta $dz$ vs $z$ dos candidatos com os aglomerados de Wen e Han (2011), a linha ponto-tracejada preta é a mediana de $dz$ , a linha tracejada preta é a média de $dz$ . . . . .	56

4.19	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Sochting <i>et al.</i> (2012) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os aglomerados de Sochting <i>et al.</i> (2012), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	56
4.20	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Bellagamba <i>et al.</i> (2011) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os aglomerados de Bellagamba <i>et al.</i> (2011), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	57
4.21	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Finoguenov <i>et al.</i> (2007) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os aglomerados de Finoguenov <i>et al.</i> (2007), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	57
4.22	Histograma da comparação entre os 500 catálogos <i>mock</i> gerados com o catálogo de Sochting <i>et al.</i> (2012) (S12). . . . .	58
4.23	Resultados da comparação entre o catálogo <i>mock</i> com o catálogo de Sochting <i>et al.</i> (2012) (S12). O círculo vermelho representa os <i>matches</i> e os pontos azuis são dados de S12 . . . . .	59
5.1	O painel da esquerda apresenta dados de r-i (mag) em função de i (mag) e o painel da direita a distribuição de r-i (mag) para o caso em que não foi aplicada a <i>k-correction</i> . Os dados são referentes ao candidato de ID 34 e <i>redshift</i> 0,85 do catálogo proposto neste trabalho. . . . .	62
5.2	Os painéis superior e inferior da esquerda apresentam dados de r-i (mag) e i-z (mag) em função de i (mag), respectivamente. Os painéis superior e inferior da direita apresentam o <i>countour plot</i> do candidato e o histograma de <i>redshift</i> das galáxias membros. Os dados são referentes ao candidato de ID 34 e <i>redshift</i> 0,85 do catálogo proposto neste trabalho. . . . .	62
5.3	Exemplos típicos de distribuição bimodal dos resultados obtidos neste trabalho para os IDs 23, 38, 39 e 42. . . . .	63
5.4	Relação de quantidade de galáxias vermelhas por azuis na faixa de <i>redshift</i> estudada. . . . .	64

5.5	O painel superior da esquerda apresenta dados de r-i (mag) em função de i (mag) e o painel superior da direita a distribuição de r-i (mag) para o caso em que foi aplicada a <i>k-correction</i> . Os dados são referentes ao candidato de ID 23 e <i>redshift</i> 0,71 do catálogo proposto neste trabalho. O painel inferior da esquerda apresenta dados de r-i (mag) em função de i (mag) e o painel inferior da direita a distribuição de r-i (mag) para o caso em que foi aplicada a <i>k-correction</i> . Os dados são referentes ao candidato de ID 147 e <i>redshift</i> 1,73 do catálogo proposto neste trabalho. . . . .	65
5.6	Os painéis superior e inferior apresentam o <i>contour plot</i> dos candidatos de ID 23 <i>redshift</i> 0,71 e 34 <i>redshift</i> 0,85, respectivamente, de <i>redshift</i> das galáxias membro. . . . .	66
5.7	Exemplo típico de um candidato que apresentou MN. Os painéis superior e inferior da esquerda apresentam dados de r-i (mag) e i-z (mag) em função de i (mag), respectivamente. Os painéis superior e inferior da direita apresentam o <i>countour plot</i> do candidato e o histograma de <i>redshift</i> das galáxias membros. Os dados são referentes ao candidato de ID 19 e <i>redshift</i> 0,75 do catálogo proposto neste trabalho. . . . .	68
5.8	Exemplo típico de um candidato que apresentou problema no CMD. O painel da esquerda apresenta dados de r-i (mag) em função de i (mag) e o painel da direita a distribuição de r-i (mag). Os dados são referentes ao candidato de ID 35 e <i>redshift</i> 0,97 do catálogo proposto neste trabalho. . .	68
5.9	Dados do parâmetro BDF <i>fracdev</i> em função do <i>redshift</i> para o catálogo de galáxias do COSMOS. . . . .	70
5.10	Dados do parâmetro BDF <i>fracdev</i> em função do <i>redshift</i> para o catálogo de galáxias de G24 (catálogo de candidatos encontrados neste trabalho). .	70
5.11	Divisão em duas populações nas cores azul e vermelha referente aos dados do parâmetro BDF <i>fracdev</i> em função do <i>redshift</i> para o catálogo de galáxias G24. As colunas representam respectivamente as cores g-r (painel superior), r-i (painel central) e i-z (painel inferior). . . . .	72
5.12	Histogramas do número de galáxias de G24 em função do BDF <i>fracdev</i> para a totalidade dos dados. Painel da esquerda apresenta a representação normal, e o da direita apresenta um zoom na faixa de 0,47 a 0,53. As linhas vermelhas verticais apresentam a faixa do possível corte. . . . .	73
5.13	Histogramas do número de galáxias de G24 em função do BDF <i>fracdev</i> com corte dos dados na região BDF <i>fracdev</i> = $0,5_{-0,01}^{+0,005}$ e valores de $r(mag) > 24$ . Painel da esquerda apresenta a representação normal, e o da direita apresenta um zoom na faixa de 0,47 a 0,53. As linhas vermelhas verticais apresentam a faixa do corte. . . . .	73

5.14	Dados do parâmetro BDF <code>fracdev</code> em função do <i>redshift</i> para o catálogo de galáxias de G24 após o corte na faixa de BDF <code>fracdev</code> = $0,5_{-0,01}^{+0,005}$ e valores de $r(mag) > 24$ . . . . .	74
5.15	BDF <code>fracdev</code> em função do <i>redshift</i> para o catálogo de galáxias de G24. O painel superior apresenta dados sem corte e o painel inferior apresenta os dados após o corte em BDF <code>fracdev</code> = $0,5_{-0,01}^{+0,005}$ e valores de $r(mag) > 24$ . . . . .	75
5.16	BDF <code>fracdev</code> em função do <i>redshift</i> para o catálogo de galáxias de G24 após o corte em BDF <code>fracdev</code> = $0,5_{-0,01}^{+0,005}$ e valores de $r(mag) > 24$ . Os candidatos do catálogo são representados em cores diferentes. . . . .	76
A.1	As imagens superior e inferior da coluna da esquerda apresentam dados de r-i (mag) e i-z (mag) em função de i (mag), respectivamente. As imagens superior e inferior da coluna da direita apresentam o <i>countour plot</i> do candidato e o histograma de <i>redshift</i> das galáxias membros, referentes ao candidato de ID 12 e <i>redshift</i> 0,68. . . . .	85
A.2	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Wen e Han (2011) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os <i>clusters</i> de Wen e Han (2011), a linha ponto-tracejada azul é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	86
A.3	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Sochting <i>et al.</i> (2012) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os <i>clusters</i> de Sochting <i>et al.</i> (2012), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	86
A.4	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Bellagamba <i>et al.</i> (2011) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os <i>clusters</i> de Bellagamba <i>et al.</i> (2011), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	87
A.5	O painel da direita apresenta os <i>matches</i> entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Finoguenov <i>et al.</i> (2007) (pontos azuis). O painel da esquerda apresenta dz vs z dos candidatos com os <i>clusters</i> de Finoguenov <i>et al.</i> (2007), a linha ponto-tracejada preta é a mediana de dz, a linha tracejada preta é a média de dz. . . . .	87

# Lista de Tabelas

2.1	Características dos Filtros Disponíveis na DECam (Flaugher <i>et al.</i> 2015) e na VIRCam (Emerson <i>et al.</i> 2006, Dalton <i>et al.</i> 2006) . . . . .	20
2.2	Dados de Campo, de posição, de área, de tempo de exposição, de profundidade dos campos, de número de fontes detectadas (após a aplicação de máscaras) e de número de chips excluídos devido à luz difusa ou outros defeitos de qualidade de imagem em grande escala. Tabela baseada no artigo de Hartley <i>et al.</i> (2022). . . . .	20
2.3	Descrição das variáveis presente na <i>query</i> que resultou no catálogo do presente trabalho. Fonte: Abbott <i>et al.</i> (2021) . . . . .	23
2.4	Resultados dos catálogos comparados, suas respectivas áreas e faixas de <i>redshift</i> . F07 (Finoguenov <i>et al.</i> 2007), W11 (Wen e Han, 2011), B11 (Bellagamba <i>et al.</i> 2011), S12 (Sochting <i>et al.</i> 2012). . . . .	28
4.1	Efeito do tamanho do pixel variando de 1 a 10 Mpc, no número de <i>match</i> de candidatos considerando o domínio de cada um dos trabalhos referenciados. F07 (Finoguenov <i>et al.</i> , 2007), W11 (Wen e Han, 2011), B11 (Bellagamba <i>et al.</i> , 2011) . . . . .	44
4.2	Resultado do <i>match</i> entre os catálogos da literatura e este trabalho. Os valores com ‘*’ referem-se a quantidade de <i>clusters</i> no domínio comparado. G24 refere-se ao presente trabalho. . . . .	49
5.1	Dados de ID, Característica e Situação de candidatos que apresentaram múltiplos núcleos (MN), distorções no Diagrama Cor-Magnitude (CMD), grande diferença entre populações de galáxias azuis e vermelhas (P) e candidatos que apresentam uma concentração de galáxias na extremidade do pixel (C). . . . .	67



# Sumário

<b>Lista de Figuras</b>	<b>xv</b>
<b>Lista de Tabelas</b>	<b>xxi</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Prólogo	1
1.2 Motivação e objetivo	3
1.3 Aglomerados de galáxias	3
1.4 Busca por aglomerados de galáxias	6
1.5 Catálogos de protoaglomerados	7
1.6 Morfologia de galáxias	11
<b>2 Dados</b>	<b>17</b>
2.1 <i>Dark Energy Survey</i>	17
2.2 Catálogos de protoaglomerados na literatura	28
<b>3 Metodologia</b>	<b>33</b>
3.1 Algoritmo de busca de candidatos	33
3.2 Busca de protoaglomerado no DES	35
3.3 <i>Match</i> 2D com dados da literatura	36
3.4 <i>Match</i> 3D com dados da literatura	37
3.5 Teste de <i>matches</i> com bootstrap	37
3.6 Estudo dos candidatos inéditos	38
<b>4 Resultados da validação do algoritmo</b>	<b>39</b>
4.1 Validação com Narwal et al. (2020)	39
4.2 Determinação do tamanho do pixel	42
4.3 Resultados com <i>match</i> 2D com a literatura	47
4.4 Resultados do <i>match</i> 3D com a literatura	55
4.5 Validação com catálogos <i>bootstrap</i>	57

<b>5</b>	<b>Resultados e análises dos candidatos</b>	<b>61</b>
5.1	Análise dos candidatos inéditos . . . . .	61
5.2	Morfologia das galáxias . . . . .	69
<b>6</b>	<b>Conclusões</b>	<b>77</b>
	<b>Referências Bibliográficas</b>	<b>79</b>
<b>A</b>	<b>Busca com passo 0,1, sem interseção</b>	<b>85</b>
<b>B</b>	<b>Dados de GMM e CMD dos candidatos inéditos de G24</b>	<b>89</b>

# Capítulo 1

## Introdução

### 1.1 Prólogo

A astronomia é o estudo do universo, sendo uma das ciências mais antigas da humanidade, com registros de observações astronômicas datando da civilização mesopotâmica. A Astronomia, e suas respectivas áreas como astrofísica e cosmologia, estudam uma ampla gama de fenômenos, desde a formação do universo até a evolução das estrelas e galáxias. Os astrônomos usam uma variedade de métodos para estudar o universo, começando por observações e usando simulações computacionais para entender os complexos fenômenos do universo. A história da astronomia está repleta de descobertas, como de planetas, estrelas e galáxias de diferentes tipos. Algumas descobertas são esperadas, outras nem tanto, como a expansão do universo.

Em 1929, o astrofísico Edwin Hubble utilizou dados do telescópio de Mount Wilson de 2,5 m de diâmetro para identificar estrelas variáveis Cefeidas. Calculando as distâncias até elas, Hubble percebeu que a distância para a, então nebulosa, hoje galáxia de Andrômeda (M31) era de, aproximadamente, 1 milhão de anos luz, cujo valor era maior do que o diâmetro estimado da Via Láctea, significando que esse sistema estelar só poderia estar localizado além das fronteiras da Via Láctea. Em estudos posteriores, observou-se que galáxias mais distantes pareciam estar afastando-se do Sistema Solar, sendo esse efeito chamado de Lei de Hubble (Hubble, 1929), na qual foi possível estabelecer uma relação entre o *redshift* de um objeto e sua distância.

O *redshift* ocorre quando a luz emitida por um objeto que está afastando-se do observador é deslocada para o vermelho. Este deslocamento é causado pelo efeito Doppler, que, nesse caso, ocorre por consequência da expansão do universo, algo inesperado na época. O *redshift* tornou-se um parâmetro muito usado para determinar a distância de galáxias, pois dentre os difíceis degraus da escada de distância, é uma medida muito fácil de ser obtida - apesar das dificuldades observacionais de fazer-se grandes levantamentos espectroscópicos. Uma forma de calcular o *redshift* é com o *redshift* fotométrico, que

permite estimar a distância de galáxias e quasares com base nas variações de suas magnitudes observadas através de diferentes filtros fotométricos, sem a necessidade de medições espectroscópicas detalhadas.

A Figura 1.1 apresenta o espectro da estrela Vega em diferentes *redshifts*. Com um espectrógrafo, pode-se medir a posição das linhas de emissão com precisão. Na Figura 1.1, as 5 áreas cinzentas na parte inferior são as curvas de resposta dos filtros usados no sistema ugriz. Diferentemente da espectroscopia, no *redshift* fotométrico só consegue-se resolver grandes padrões, como quebras no contínuo, atravessando os filtros e por isso a precisão é bem menor.

Existem duas maneiras principais de se medir o photo-z. Uma, similar ao espectroscópico, pelo ajuste de espectros de comparação (Brammer *et al.*, 2008b). A outra, através de métodos de aprendizado de máquina, os quais em geral dependem de uma amostra de treinamento de *redshifts* espectroscópicos, e portanto nem sempre são adequadas para estudos a altos *redshifts* por conta das amostras espectroscópicas incompletas (Bonnert *et al.*, 2016).

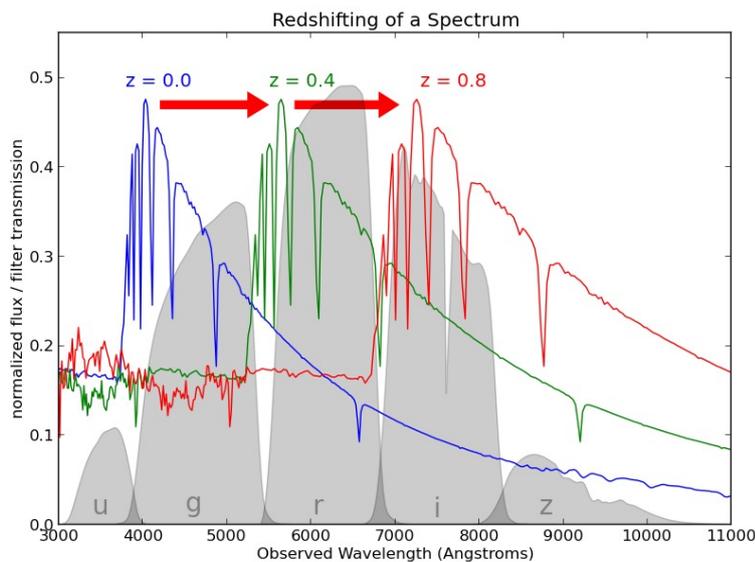


Figura 1.1. Representação do *redshift* fotométrico através de um espectro deslocado para o vermelho atravessando as bandas do sistema fotométrico ugriz em passos de 0,0, 0,4 e 0,8 de *redshift*. Fonte: Vanderplas *et al.* (2012)

De Hubble até os dias de hoje, muitos levantamentos foram realizados, mas o primeiro grande levantamento digital foi o *Sloan Digital Sky Survey* (SDSS, York *et al.* (2000)). Os dados do SDSS iniciaram uma nova fase na catalogação de aglomerados de galáxias, com até mais de cem mil aglomerados detectados (Hao *et al.*, 2010; Koester *et al.*, 2007; Miller *et al.*, 2005; Wen *et al.*, 2012).

Na sequência de grandes levantamentos digitais tem-se o *Dark Energy Survey* (DES, Dark Energy Survey Collaboration *et al.* (2016)), um conjunto de dados que vem sendo explorado para estudar a distribuição de matéria no universo e as propriedades da energia

escura. A detecção de objetos a altos *redshifts* foi pouco explorada pela colaboração (Sánchez *et al.*, 2022) - em parte, por não ser um projeto chave.

Entretanto, com os dados públicos do DES *Data Release 1* (Abbott *et al.*, 2018), a comunidade científica pôde começar a explorar essa frente. Por exemplo, Narwal *et al.* (2020) utilizaram dados do *Dark Energy Survey Science Verification 1* (DES-SVA1), uma área inicial menor, porém mais profunda que o DES em operação, e identificaram um candidato promissor a superaglomerado em  $z \sim 1,1$ .

Neste trabalho, foram utilizados dados do DES para encontrar candidatos a protoaglomerados. O DES é um levantamento em cinco bandas (grizY) no hemisfério Sul, que inclui duas estratégias de observação: *wide* e *deep* (Hartley *et al.*, 2022, Sánchez *et al.*, 2022) possuindo o *deep* uma magnitude limite *i-band* de 26,75 (AB,  $10\sigma$ , 2 arcsec). A estratégia *wide* possui 5.000 *graus*<sup>2</sup>, utilizou a Câmera de Energia Escura (DECam), localizada no telescópio Blanco de 4m, no Chile e possui 10 campos de supernova. Uma estratégia aplicada para detecção de supernovas consistiu em repetidas observações de determinadas áreas do céu, quase semanais, ao longo de vários anos de levantamento, sendo essas áreas denominadas de *Deep Fields* (Abbott *et al.*, 2019). Tais observações puderam ser combinadas, ou co-adicionadas, em observações mais profundas que possibilitaram os estudos de objetos mais distantes. No capítulo 2 são discutidos os dados do DES em mais detalhes.

## 1.2 Motivação e objetivo

Aglomerados de galáxias e seus progenitores são rastreadores dos picos de densidade da matéria do Universo e onde a formação de galáxias ocorre de maneira diferenciada em relação ao campo. É fundamental encontrar e entender as propriedades desses ambientes.

Com o aglomerado identificado e caracterizado, os astrônomos podem realizar estudos adicionais para investigar a dinâmica do aglomerado, sua evolução ao longo do tempo cósmico e seu papel na formação e evolução das estruturas em larga escala do universo.

Neste trabalho, fez-se uma busca e análise de candidatos a protoaglomerados de galáxias aplicando a técnica de sobredensidades de galáxias na faixa de *redshift* de  $0,5 \leq z \leq 2,0$  utilizando dados dos *Deep Fields* do *Dark Energy Survey Year 3*, mais especificamente do campo *Cosmic Evolution Survey* (COSMOS, Scoville *et al.* (2007)).

## 1.3 Aglomerados de galáxias

Aglomerados de galáxias são sistemas únicos que oferecem uma visão privilegiada do universo, sendo fundamentais para a compreensão da estrutura em grande escala. Sua abundância e massa fornecem informações valiosas sobre a distribuição da matéria no universo e são ambientes propícios para a formação e evolução das galáxias (Cautun *et al.*,

2014, Overzier, 2016). Eles são compostos de uma variedade de componentes, incluindo galáxias, matéria escura, luz estelar difusa e um meio *intracluster* quente.

A matéria escura é um componente misterioso do universo que não interage com ondas eletromagnéticas, mas influencia gravitacionalmente a matéria ao seu redor. Os aglomerados de galáxias são os locais onde a matéria escura é mais abundante, representando cerca de 80% da massa total do aglomerado.

As galáxias são os componentes mais visíveis dos aglomerados de galáxias e representam cerca de 5% da massa total do aglomerado.

Nas últimas décadas, mais de 100 mil aglomerados de galáxias, ou grupos, foram identificados a partir de levantamentos ópticos ou infravermelhos, sendo que a maioria desses aglomerados identificados estão em  $z < 0,5$ . Aglomerados de alto *redshifts*, com  $z > 1$ , são mais difíceis de observar por sua distância.

Como exemplo de um campo profundo, a Figura 1.2 apresenta uma imagem do primeiro *Deep Field* do James Webb, apresentando o aglomerado de galáxias SMACS 0723, cuja identificação é clara, mas objetos mais distantes misturam-se no fundo de múltiplas fontes ao longo da linha de visada.



Figura 1.2. O primeiro *Deep Field* do James Webb. Aglomerado de galáxias SMACS 0723. Fonte: <https://www.nasa.gov/image-feature/goddard/2022/nasa-s-webb-delivers-deepest-infrared-image-of-universe-yet>

Protoaglomerados são estruturas precursoras de aglomerados de galáxias. A Figura 1.3 apresenta uma imagem de um protoaglomerado obtida com o telescópio espacial James Webb. O estudo dos protoaglomerados é importante porque eles podem ajudar a entender

como os aglomerados de galáxias formam-se.

No entanto, não há uma definição única de protoaglomerado. Uma definição simples para protoaglomerado é que se trata de uma estrutura que irá, eventualmente, colapsar em um aglomerado de galáxias. Pode-se observar, na Figura 1.4 um exemplo da quantidade de aglomerados e protoaglomerados catalogados em função do *redshift*, como compilado por Overzier (2016). Os protoaglomerados não são tão numerosos como os aglomerados. O motivo principal para essa distribuição tornar-se mais rarefeita com *redshift* é a dificuldade de detecção. Quanto mais longe, mais difícil torna-se a detecção, seja para aglomerado ou protoaglomerado.

Os *surveys* têm aprofundado em *redshifts* cada vez mais altos, resultando em campos profundos disponíveis para estudos e buscas por aglomerados e protoaglomerados mais distantes.

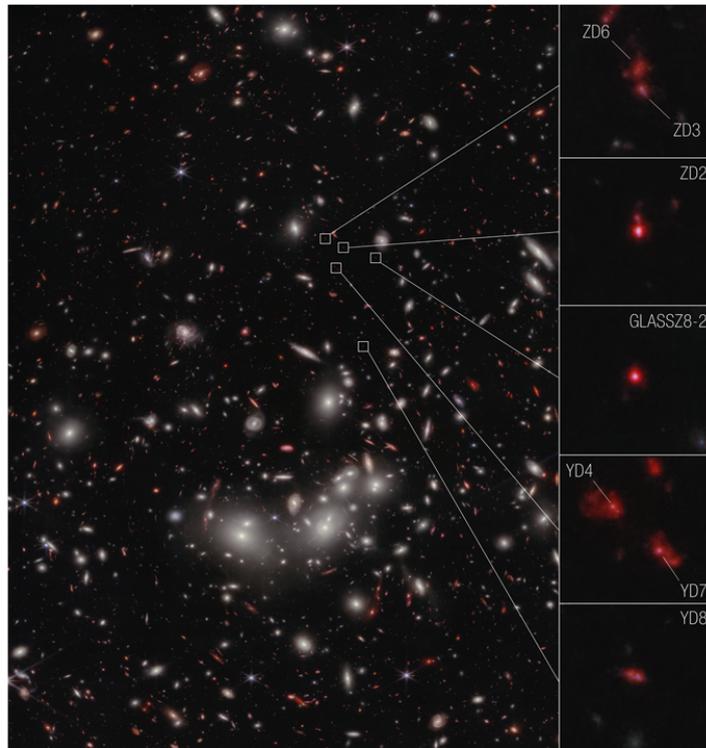


Figura 1.3. Imagem produzida pelo telescópio James Webb que apresenta um protoaglomerado a um *redshift* de 7,9. Fonte: NASA, ESA, CSA, T. Morishita (IPAC). Processamento de imagem: A. Pagan (STScI).

A fronteira das buscas de protoaglomerados atualmente está na faixa de  $1,5 < z < 2$ , sendo  $z \sim 2$  uma época importante para aglomerados de galáxias massivos, pois é quando eles estão formando-se e crescendo rapidamente (Overzier, 2016). É comum encontrar na literatura artigos em que os autores, em uma primeira etapa, apresentam alguns candidatos a protoaglomerados (Brinch *et al.* 2023), ou até mesmo candidatos únicos (Lammers *et al.* 2022, Toshikawa *et al.* 2024, Brinch *et al.* 2024), para posteriormente confirmarem usando espectroscopia (Jiang *et al.* 2018, Diener *et al.* 2015, Capak *et al.* 2011).

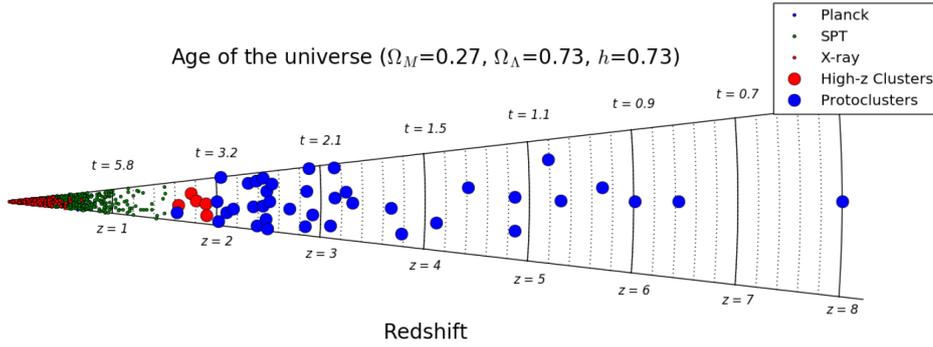


Figura 1.4. Distribuição de *redshifts* de protoclusters selecionados da literatura (círculos azuis cheios). Os dados para aglomerados abaixo de  $z = 1,5$  foram retirados da compilação de aglomerados detectado em Bleem *et al.* (2015). Círculos vermelhos cheios são aglomerados em alto *redshift*, em  $z > 1,6$ . A posição dos objetos ao longo do eixo polar não contém informações e é usado apenas para fins de visibilidade. *Redshifts* e idades (em Ganos) são indicados ao longo do eixo radial. Fonte: Overzier (2016).

Para entender a evolução de galáxias em protoaglomerados Muldrew *et al.* (2018) utilizaram a simulação *Millennium* com base no modelo cosmológico  $\Lambda$ CDM. Para a simulação *Millennium* foram necessários modelos auxiliares que descrevem a formação de galáxias. Com este objetivo, os autores utilizaram um modelo semianalítico que consistiu em um método que relaciona e conecta as propriedades observadas das populações de galáxias e sua evolução com o *redshift*. Os autores reproduziram o comportamento da formação estelar tanto em protoaglomerados quanto no campo. A taxa de formação estelar atingiu seu ápice nos protoaglomerados antes do que no campo e diminui à medida que se afasta dos protoaglomerados. Os resultados mostraram que 80% da massa estelar dos protoaglomerados formou-se em  $z = 1,4$ , enquanto que, no mesmo período, foi formada 45% da massa estelar no campo, achado parecido com o de Chiang *et al.* (2017).

## 1.4 Busca por aglomerados de galáxias

Os aglomerados - e protoaglomerados - de galáxias servem de rastreadores dos picos de densidade da matéria na composição do Universo, sendo ambientes especiais para estudar a evolução de galáxias (Muldrew *et al.*, 2018).

Existem vários métodos para se encontrar aglomerados de galáxias. Todos se resumem em encontrar sobredensidades em “fatias” de universo, que podem ser feitas através do filtro de cores, como a sequência vermelha (Gladders e Yee, 2005), ou através do *redshift* fotométrico (Wen *et al.*, 2012). Seja como for, o objetivo de se fatiar o universo em bins de *redshift* é evitar efeitos de projeção ao longo de uma linha de visada, que poderia integrar concentrações de galáxias ao longo de diferentes distâncias.

O Método de sequência vermelha do aglomerado (*Red Sequence Cluster Method*, Glad-

ders e Yee, 2005, Bellagamba *et al.*, 2011) é uma técnica empregada na detecção de aglomerados de galáxias com base nas características das suas populações estelares. Este método explora a tendência de galáxias em aglomerados de terem cores vermelhas, delimitando uma sequência vermelha no diagrama cor-magnitude, onde, em geral, a Galáxia Mais Brilhante (BCG, em inglês) do aglomerado ocupa a ponta mais brilhante dessa sequência.

A aplicação do método depende dos dados fotométricos presentes em uma determinada região do céu na qual pretende-se executar a busca. A partir dos dados fotométricos, é construído um diagrama que mostra a relação entre a cor e a magnitude das galáxias na região observada. A tendência para que as galáxias tenham cores mais vermelhas dentro do aglomerado é devido à predominância de estrelas velhas e evoluídas nessas galáxias. Em seguida, as galáxias são selecionadas como candidatas a pertencerem ao aglomerado. Esta seleção é baseada na proximidade das galáxias com a sequência vermelha. Uma vez identificadas as galáxias que fazem parte da sequência vermelha, o aglomerado pode ser confirmado e caracterizado por meio de observações adicionais em diferentes comprimentos de onda ou com espectroscopia. Estas análises subsequentes permitem determinar com mais precisão propriedades como a distância do aglomerado, sua massa, quem são as galáxias-membro, entre outras características.

A detecção de aglomerados por meio de sobredensidade de galáxias é uma técnica usada para identificar regiões do céu que contêm um grande número de objetos em comparação com sua vizinhança. Neste caso, escolhe-se em um catálogo, como por exemplo o DES, para um levantamento das galáxias em uma determinada região do céu. Utilizando técnicas estatísticas, como algoritmos de identificação de sobredensidades, procura-se regiões onde a densidade de galáxias é significativamente maior do que a média esperada para o universo, regiões identificadas como *clusters* de galáxias. A partir da análise dos dados, os *clusters* de galáxias podem ser caracterizados em termos de sua massa, tamanho, e forma. Como exemplo de aplicação desta técnica tem-se o trabalho de Narwal *et al.* (2020) que determinou um candidato a superaglomerado de galáxias em  $s \sim 1, 1$ .

## 1.5 Catálogos de protoaglomerados

Na sequência, são apresentados os trabalhos realizados por Narwal *et al.* (2020), para o DES SVA1, e Finoguenov *et al.* (2007), Wen e Han (2011), Bellagamba *et al.* (2011), e Sochting *et al.* (2012) para a região de COSMOS, um dos campos mais observados da astronomia. Estes catálogos foram utilizados como fonte de comparação com o presente trabalho.

Narwal *et al.* (2020) aplicaram a técnica de sobredensidade de galáxias, que serviu como referência no desenvolvimento deste trabalho. Narwal *et al.* (2020) utilizaram dados do *Dark Energy Survey Science Verification 1* (DES-SVA1), umas das primeiras do DES

a ser observada, ficar pública e possuía magnitude limite  $i = 23.0$ . Para comparação, a estratégia *wide* do DES para o ano 1 possui magnitude limite  $i = 22.5$  (Drlica-Wagner *et al.*, 2018) e o *deep* possui  $i = 26.75$  (Hartley *et al.*, 2022). Essa escolha deu-se pelo fato de ser um campo mais profundo que o DES típico e por abranger uma grande área do céu, de  $\sim 140 \text{ deg}^2$ . Durante o desenvolvimento da estratégia do levantamento DES concluiu-se que seria melhor uma área de verificação com essas características. Os autores usaram o *redshift* fotométrico medido com o método de aprendizado de máquina *Skynet* Bonnett *et al.* (2016). Esses *redshifts* tem erros de cerca de 0,1. Logo, os autores dividiram o conjunto de dados em fatias (*bins*), com sobreposições, de *redshift* de 0,20, em intervalos de 0,03. Dessa forma, um *bin* vai de 0,8 até 1,0 e o seguinte vai de 0,83 até 1,03, e assim por diante. O fato de objetos poderem ser detectados em fatias consecutivas suaviza o sinal das aglomerações ao longo dos *bins* e ajuda a reduzir os efeitos dos erros relativamente grandes do photo-z (Sánchez *et al.*, 2017). A faixa analisada foi entre 0,2 e 1,24 como visto na Figura 1.5.

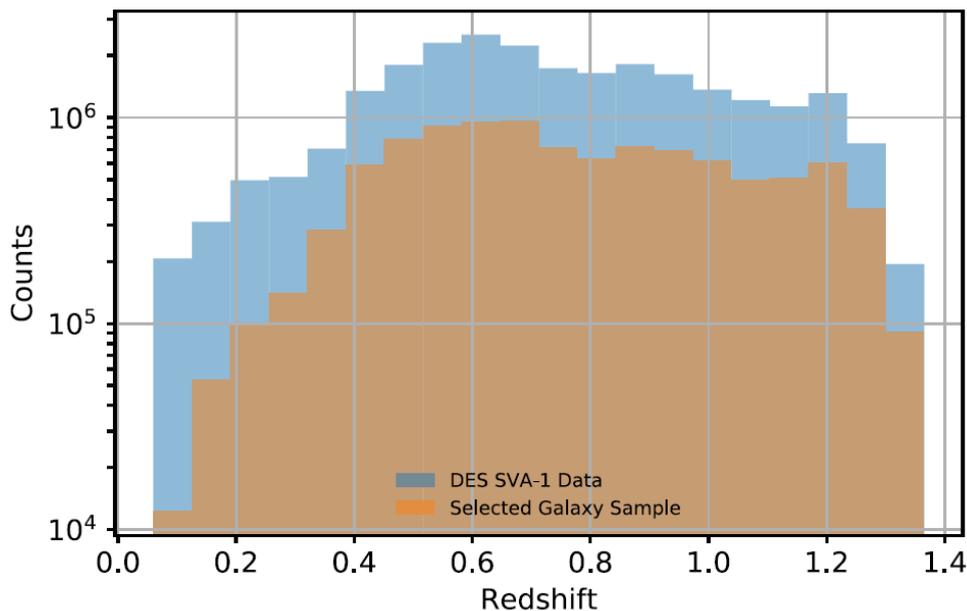


Figura 1.5. Histograma de *redshift* de todos os dados do DES-SVA1 (azul) e das galáxias selecionadas (laranja). Fonte: Narwal *et al.* (2020)

As faixas de *redshift* foram divididas em pixels<sup>1</sup> de 5 Mpc. Os autores perceberam que aumentando o tamanho dos *pixels* causavam uma redução da sobredensidade próximo das bordas do aglomerado, e tamanhos menores de *pixel* produziam um efeito desprezível. A Figura 1.6 apresenta as curvas de nível do contraste de densidade.

<sup>1</sup>aqui chamamos de pixel o elemento da matriz produzida pelo histograma 2D no Python Matplotlib

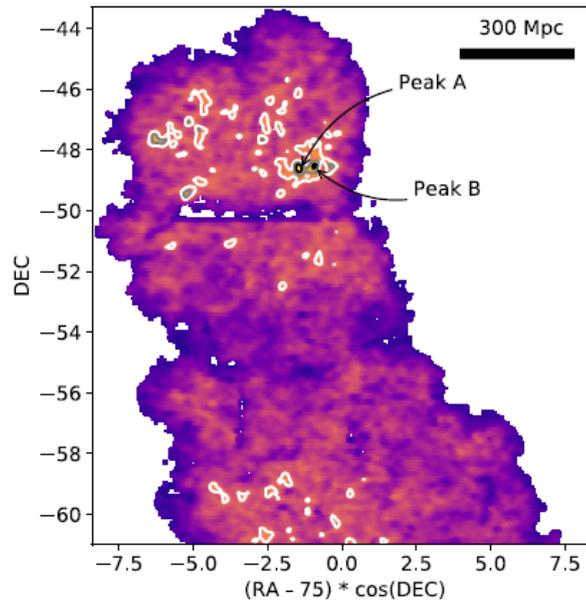


Figura 1.6. Curvas de nível do contraste de densidade suavizado exibindo as estruturas de sobredensidade no bin 1,04-1,24. Fonte: Figura 5 de Narwal *et al.* (2020).

A seguir, fazemos uma breve descrição dos métodos utilizados pelos autores cujos trabalhos serviram para comparação e validação da técnica aplicada neste trabalho na região do COSMOS.

No estudo conduzido por Finoguenov *et al.* (2007) foi realizada uma busca por aglomerados de galáxias nos primeiros 36 apontamentos do XMM-Newton no campo COSMOS. Os candidatos a aglomerados foram identificados por meio de um filtro *wavelet* na emissão estendida de raios-x. A verificação desses candidatos foi realizada com base em uma análise de concentração de galáxias em fatias de *redshift* de espessura de 0,1 – 0,2, utilizando o catálogo fotométrico multibanda do campo COSMOS da área do XMM-Newton e sendo a faixa de *redshift* igual a  $0,08 < z < 1,25$  e  $i_{AB} < 25$ . O estudo identificou 72 aglomerados.

No estudo conduzido por Bellagamba *et al.* (2011) foi apresentado o *optimal matched filter* para a detecção de aglomerados de galáxias em levantamentos no óptico. Este filtro otimiza a relação sinal-ruído, aproveitando o conhecimento sobre as propriedades físicas intrínsecas das galáxias e as características do ruído nos dados. O método concentra-se nas magnitudes das galáxias, suas posições e *redshifts* fotométricos. O método foi aplicado ao campo COSMOS e os resultados foram comparados com detecções de aglomerados obtidas por diferentes métodos e catálogos da literatura, sendo Finoguenov *et al.* (2007) uma das literaturas comparadas e que também foi utilizada no presente trabalho.

No estudo conduzido por Wen e Han (2011) sobre aglomerados de galáxias em alto *redshift*, os autores conseguiram identificar 631 aglomerados com o Telescópio Canada–France–Hawaii (CFHT), 202 no campo profundo (*Deep Field*) do CFHT, 737 no campo do Spitzer *Wide-area InfraRed Extragalactic Survey* (SWIRE) e 187 no campo

do COSMOS com *redshift* variando na faixa de  $0,1 \leq z \leq 1,6$ . O autores escreveram um algoritmo de detecção de aglomerados para identificar aglomerados de alto *redshift*. O algoritmo utiliza dados de *redshift* fotométrico e uma combinação de cortes de cor e informação de sequência vermelha para distinguir galáxias membros dos aglomerados de galáxias do campo. Com essa amostra de aglomerados de alto *redshift*, os autores estudaram a evolução das cores das Galáxias mais Brilhantes dos Aglomerados, em inglês, *Brightest Cluster Galaxies* (BCGs).

O estudo conduzido por Sochting *et al.* (2012) no campo do COSMOS resultou na identificação de uma amostra de grupos e aglomerados de galáxias com *redshifts* intermediários a alto. Foi utilizado um algoritmo que detectava picos de densidade em uma distribuição gerada por tesselação de Voronoi (vide Figura 1.7) em *bins* sobrepostos de *redshift* fotométrico sendo uma fatia de *redshift*  $\delta_z = 0,02 * (1 + z)$ . O catálogo resultante contém 1.780 estruturas, com *redshifts* variando de 0,2 a 3,0.

A distribuição de *redshifts* das estruturas detectadas mostra variações significativas, com picos proeminentes que sugerem a presença de estruturas de grande escala no campo estudado pelos autores. Foi possível detectar candidatos a superaglomerados nos *redshifts*  $z = 0,35, 0,72, 0,94, 1,12, 1,27, 1,45, 2,0$  e  $2,52$ , porém, os autores não disponibilizaram mais informações com a promessa de continuar a pesquisa, impossibilitando uma análise dos mesmos. Até o presente momento, não houve publicações relacionadas a essa parte do trabalho. Em  $z = 2,9$ , uma aglomeração compacta de grupos e aglomerados de galáxias foi identificada, sugerindo a presença de outra estrutura semelhante a um superaglomerado, que é a candidata ao maior *redshift* até agora. Foram identificados seis novos superaglomerados dentre essas detecções.

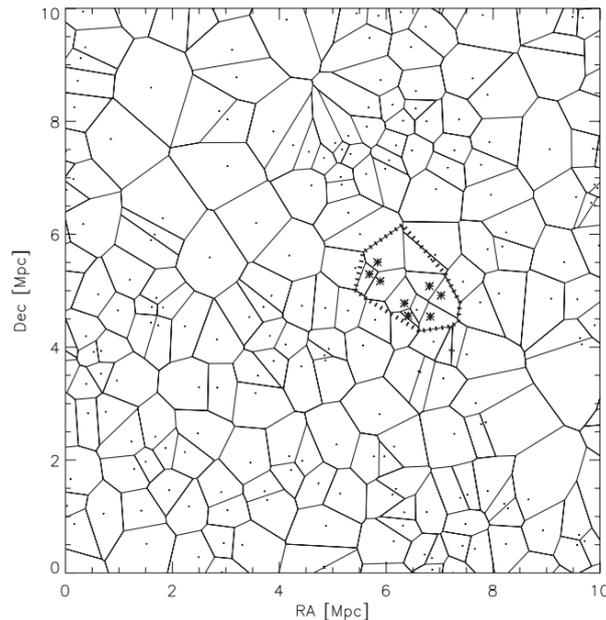


Figura 1.7. Exemplo de uma tesselação de Voronoi tirada do artigo de Sochting *et al.* (2012). Esta seria uma fatia em  $z \sim 2,7$ .

## K-correction

Para análise de propriedades fotométricas em diferentes *redshifts* é necessário corrigir os fluxos luminosos para as mudanças nos comprimentos de onda dos filtros (Oke e Sandage, 1968). A *k-correction* é importante para interpretação das observações fotométricas de galáxias, especialmente quando considera-se a relação entre magnitude e *redshift* (Poggianti 1997). A compreensão detalhada da *k-correction* torna-se essencial para estabelecer conexões significativas entre galáxias locais e distantes, devido às incertezas associadas aos dados à medida que se alcança *redshifts* mais altos e magnitudes mais fracas. A *k-correction* é definida como o termo corretivo aplicado às magnitudes observadas em diversas bandas espectrais devido ao *redshift* cosmológico, desempenha um papel vital na correção e na análise das observações astronômicas. No entanto, é importante ressaltar que a *k-correction* não leva em conta os efeitos da evolução galáctica, exigindo, em alguns casos, a aplicação de correções evolutivas adicionais.

Abordagens comuns para calcular essas correções são baseadas no ajuste de modelos das distribuições de energia espectral (no inglês, *spectral energy distributions*, SED) observadas e, portanto, requerem fotometria multibanda (Chilingarian *et al.* 2010). Chilingarian *et al.* (2010) demonstrou que, em casos de filtros ópticos e infravermelhos (NIR) as *k-corrections* podem ser aproximadas com polinômios bidimensionais de baixa ordem de apenas dois parâmetros: *redshift* e uma cor observada. Com essa abordagem, os autores apresentaram funções de ajuste polinomial para *k-corrections* nos filtros do *Sloan Digital Sky Survey* (SDSS) *ugriz*, no *United Kingdom Infrared Telescope* (UKIRT) *Wide Field Camera* YJHK, *Johnson-Cousins* UBVRcIc e no *Two Micron All Sky Survey* JHKs para galáxias com *redshift*  $z < 0,5$ , com base em valores obtidos por meio do ajuste combinado óptico-NIR SEDs de um conjunto de 105 galáxias construídas a partir do SDSS *Data Release 7* (DR7) e do *UKIRT Infrared Deep Sky Survey* DR5 fotometria usando o *Virtual Observatory*.

Poggianti (1997) propôs correções K (*k-corrections*) e evolutivas para os tipos de galáxias da classificação de Hubble E, Sa e Sc para os filtros U, B, V, R, I, J, H, K do sistema fotométrico e os filtros gri do sistema modificado até o *redshift*  $z = 3$ . Os autores estudaram a dependência dessas correções em relação à escala de tempo da formação estelar em galáxias elípticas. As correções foram calculadas com base em um modelo que reproduz o espectro integrado da galáxia na faixa de 1.000 a 25.000 Å.

## 1.6 Morfologia de galáxias

Determinar a morfologia de galáxias é importante por fornecer pistas sobre a sua origem e evolução. No entanto, a determinação destas propriedades requer classificação confiável no processo de produção de grandes catálogos para posterior estudo estatístico.

Além disso, o entendimento da morfologia das galáxias em *redshift* baixo é crucial para qualquer comparação significativa com galáxias em elevados *redshifts*. A maioria dos dados que se tem a respeito da morfologia de galáxias são obtidas de forma visual por astrônomos profissionais e amadores (ciência cidadã), tendo sido classificadas e catalogadas inúmeras galáxias (Bamford *et al.*, 2009).

Mas as técnicas de aprendizado de máquina requerem treinamento, feito por humanos. Estudos de classificações morfológicas visuais são apresentados por Bamford *et al.* (2009) em análises das relações entre a morfologia, cor, ambiente e massas estelares. Segundo Bamford *et al.* (2009), a cor e morfologia são sensíveis a massa estelar e suas frações diferem do ambiente em que as galáxias são formadas. Os autores afirmaram que embora a cor seja sensível ao ambiente de desenvolvimento, a morfologia apresenta tendências ambientais mais fracas e concluíram que uma pequena parte das relações entre morfologia e cor podem ser devido à variação da massa estelar. Em todos os ambientes e independentemente de sua morfologia, em sua maioria as galáxias apresentam cor vermelha devido as grandes massas estelares e as galáxias de baixa massa estelar são prioritariamente azuis em ambientes de baixa densidade e vermelhas em ambientes de alta densidade. Embora galáxias com morfologia do tipo *early-type* (ETG) apresentem maiores frações de vermelhas, esse fator é menos influente que o efeito do ambiente em relação à dependência da fração vermelha com a massa estelar. A relação entre cor-densidade é principalmente devido a variações nas frações de cor dentro de uma morfologia fixa. Esse fato ocorre na fração de galáxias espirais que exibem cores vermelhas, principalmente em massas estelares baixas. Os autores concluíram que há uma dependência ambiental da cor que vai além da morfologia. O efeito do ambiente na cor das galáxias de azul para vermelho pode ocorrer em escalas de tempo mais curtas do que a transição de morfologias espirais para tipos iniciais.

Os aglomerados de galáxias podem ser utilizados para avaliar como o ambiente afeta as galáxias uma vez que estes aglomerados apresentam distribuição de densidade com regiões de alta e baixa concentração. Ribeiro *et al.* (2023) estudaram o efeito do ambiente do aglomerado no crescimento tardio de galáxias do tipo ETG. Os autores mostraram que o tamanho e crescimento destes sistemas é de longo prazo e pode durar até a presente data. É de fundamental importância determinar o centro do sistema do aglomerado que corresponda ao mínimo do poço de potencial gravitacional. A partir da caracterização do centro, pode-se definir os perfis de densidade de massa, definir raios característicos do aglomerado e a morfologia das galáxias com relação ao centro.

A classificação de galáxias e os estudos de morfologia têm se tornado um grande desafio devido à complexidade e ao vasto volume de dados obtidos pelo avanço tecnológico (Lahav *et al.*, 1995). Nesse contexto, é cada vez mais útil a aplicação de ferramentas baseadas em algoritmos matemáticos para auxiliar nessas análises. Um exemplo desse tipo de trabalho é a pesquisa desenvolvida por Bom *et al.* (2024), na qual foram aplicadas técnicas de *Deep*

*Learning* (DL) para a análise de imagens. Os autores obtiveram um catálogo de galáxias e suas respectivas classificações morfológicas para áreas ainda inexploradas. Os resultados obtidos por Bom *et al.* (2024) apresentados nos diagramas de cor versus magnitude ( $g - r$ ) mostraram que as ETG tendem a ter coloração avermelhada, enquanto as *late-type galaxies* (LTG) são mais azuis. Os resultados foram obtidos com altos valores de pureza e completeza, tendo o método de DL obtido 98,5% de pureza. Com relação as análises de imagens, o estudo utilizou um modelo para prever a qualidade das imagens, categorizando objetos que não se enquadram em ETG ou LTG. Os resultados permitiram concluir que a maioria dos objetos não classificados como galáxias de tipo LTG ou ETG tiveram uma baixa probabilidade de serem imagens confiáveis. Neste caso, os autores concluíram que a seleção de imagens confiáveis leva a uma diminuição nos erros associados a classificação das galáxias. Outro achado interessante foi que entre as imagens não confiáveis pode haver objetos extraordinários como Galáxias Antenas. O estudo propôs um novo Catálogo de Valor Agregado (VAC) de morfologias de galáxias na área coberta pelo S-PLUS DR3.

Uma maneira de caracterizar a morfologia das galáxias é por meio do estudo da fração do fluxo de luz do bojo pelo fluxo de luz do disco da galáxia. O parâmetro que relaciona esses fluxos de luz é o *fracdev*, que é obtido por meio de observações fotométricas. Este parâmetro tem como função definir o tipo de perfil de brilho de galáxias, sendo que *fracdev* igual a um para um perfil de de Vaucouleurs puro (Galáxias Elípticas); e *fracdev* igual a zero para um perfil exponencial puro (Galáxias Espirais) (Vincent e Ryden, 2005, Rodríguez e Padilla, 2013). A Figura 1.8 apresenta a relação entre o parâmetro *fracdev* com a morfologia das galáxias do *Sloan Digital Sky Survey* (SDSS). A linha contínua azul representa as galáxias espirais e pode-se perceber que o pico destas galáxias está presente em *fracdev* = 0 diminuindo consideravelmente em direção a um. A linha tracejada verde representa as galáxias elípticas e possui um pico de concentração de galáxias próximo a *fracdev* = 1 diminuindo consideravelmente em direção a zero.

Uma generalização da lei de de Vaucouleurs foi proposta por Sérsic (1963), denominada perfil Sérsic que é uma função matemática que descreve como a intensidade do brilho de uma galáxia varia com a distância do seu centro. Existem diversas abordagens para realizar esse ajuste, 1D, para o perfil radial, e 2D, para a imagem inteira. No trabalho de de Souza *et al.* (2004) foi desenvolvido um software aplicado na classificação morfológica de Galáxias denominado BUDDA (Bulge/Disk Decomposition Analysis). BUDDA é um código dedicado a realizar uma decomposição bojo/disco bidimensional diretamente a partir das imagens de galáxias. O componente bojo é ajustado com um perfil generalizado de Sérsic, enquanto os discos possuem um perfil exponencial. Barras e outras subestruturas, como lentes, anéis, barras internas e discos internos são estudadas com as imagens residuais obtidas através da subtração dos modelos de bojo e disco das imagens originais.

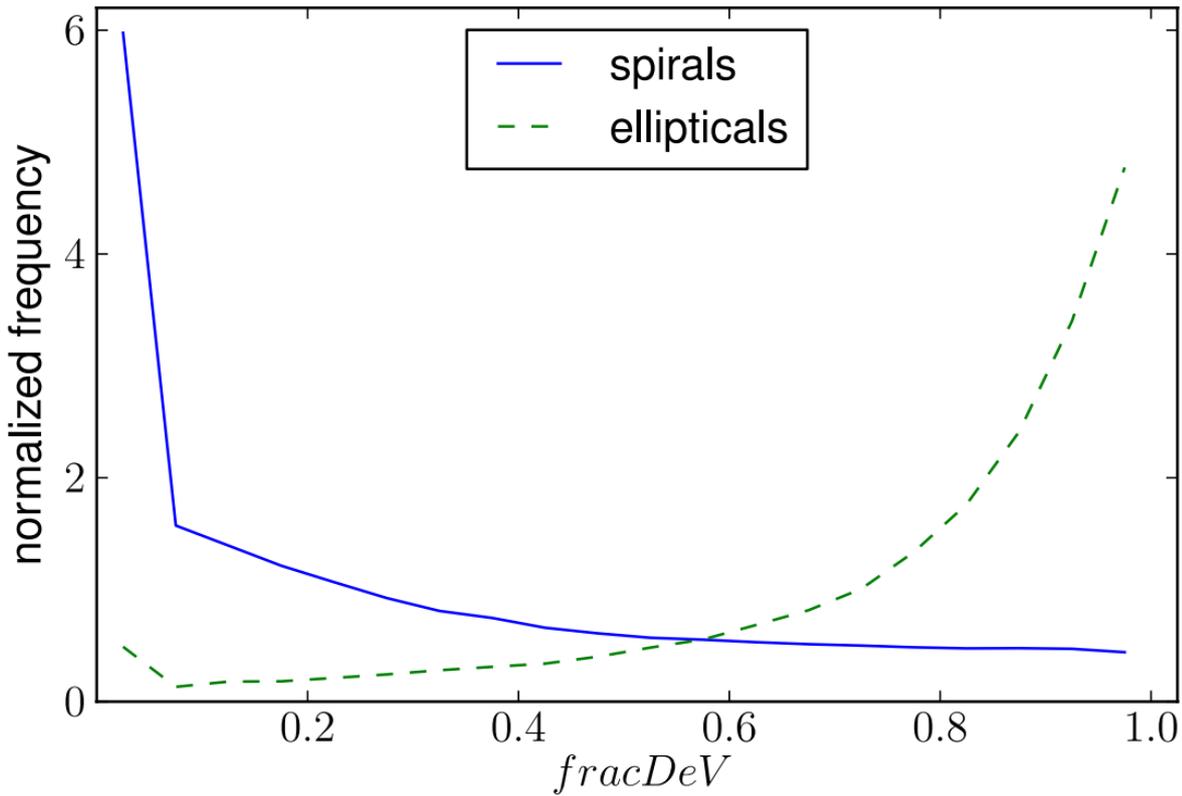


Figura 1.8. Distribuição do parâmetro  $fracdev$  para as amostras espirais e elípticas do *Sloan Digital Sky Survey* (SDSS). Fonte: Rodríguez e Padilla (2013).

Nos capítulos seguintes são apresentados os dados utilizados neste trabalho, a metodologia aplicada, a validação da metodologia e dos resultados, uma apresentação e análise dos resultados obtidos e as conclusões. No capítulo 2, são apresentados os dados relativos à região denominada de COSMOS proveniente do *Deep Field* do DES, destacando os detalhes dos dados em que foi feita a busca por candidatos a protoaglomerados de galáxias. No capítulo 3, é apresentada a metodologia abordada para a detecção de sobredensidade de galáxias destacando detalhes a respeito de algoritmos escritos e técnicas utilizadas na análise. Em seguida, no capítulo 4, foi apresentada a validação do algoritmo proposto baseado na técnica de busca por sobredensidades utilizado para identificar esses aglomerados, garantindo assim a precisão e confiabilidade dos resultados obtidos. Os resultados foram comparados com a literatura, resultando o catálogo de candidatos a protoaglomerados denominado por nós de G24. Posteriormente, no capítulo 5, discutiu-se os principais resultados obtidos a partir do catálogo G24 proposto, fornecendo informações sobre a distribuição e propriedades desses candidatos. Ainda no capítulo de resultados, foram apresentadas as propriedades dos aglomerados e um estudo sobre a morfologia das

galáxias presentes no catálogo G24. Ao final da dissertação, no capítulo 6 de conclusões, é apresentado o sumário dos resultados obtidos em cada etapa da pesquisa. Discutem-se as implicações desses resultados para o entendimento atual da formação e evolução de aglomerados de galáxias, bem como suas contribuições para áreas correlatas da astronomia.



# Capítulo 2

## Dados

### 2.1 *Dark Energy Survey*

O *Dark Energy Survey* (DES) é um projeto de pesquisa astronômico que busca entender a natureza da energia escura. Os dados do projeto foram obtidos com o Telescópio Blanco (Figura 2.1), localizado no Observatório Interamericano Cerro Tololo (CTIO), em La Serena, Chile. O telescópio possui 4m de diâmetro e em seu foco primário está a Câmera de Energia Escura (DECam, Flaugher *et al.* 2015) que é composta por um mosaico de CCDs (Figura: 2.2) com um total de 570 megapixels.

O conjunto de telescópio, tempo de exposição e a DECam permite uma observação profunda sendo os limites de magnitude para as imagens na banda *i* ou *i-band* de 23,3 mag para o *Wide*, o levantamento principal (Abbott *et al.* 2021), enquanto para os *Deep Fields* (DF) é de 26,75 mag (Hartley *et al.*, 2022).



Figura 2.1. Domo do telescópio Blanco no CTIO.

Fonte: <https://www.darkenergysurvey.org/the-des-project/instrument/>.

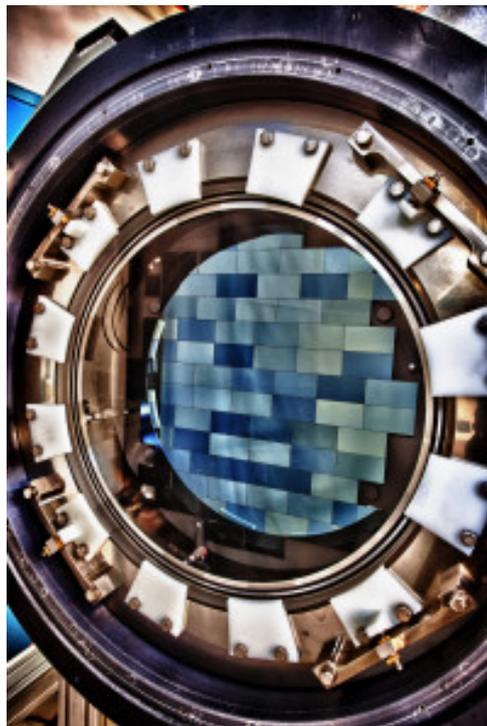


Figura 2.2. Representação dos CCDs DECAM.

Fonte: <https://www.darkenergysurvey.org/the-des-project/instrument/>.

O DES abrange uma área de 5000 graus quadrados, sendo o levantamento realizado em cinco bandas, grizY. A Figura 2.3 mostra uma representação esquemática do *footprint* do DES. Os *Deep Fields* do DES totalizam 30 graus quadrados de área, sendo eles 10 campos

de Supernova, observados com mais frequência, mais o campo do *Cosmic Evolution Survey* (COSMOS), que é amplamente explorado na literatura (Finoguenov *et al.*, 2007, Wen e Han, 2011, Bellagamba *et al.*, 2011, Sochting *et al.*, 2012). Essas regiões foram observadas mais vezes com o intuito de encontrar supernovas, permitindo a detecção de fontes que não seriam detectadas por exposições simples. As varias exposições combinadas (processo de coadição) formam os *Deep Fields*. A Figura 2.4 apresenta um exemplo de imagens coadicionadas com três níveis de profundidade que foram produzidas por Hartley *et al.* (2022), sendo elas SE\_TRUTH (*Single-epoch truth*), COADD\_TRUTH (*Coadd truth*) e DEEPEST (*Deepest*), da esquerda para a direita, respectivamente. O tipo SE\_TRUTH representa imagens com tempo de exposição efetiva de 10 vezes uma exposição nominal do campo *Wide* do DES, o termo COADD\_TRUTH refere-se a imagens que foram geradas com tempo de exposição efetiva com 10 vezes o tempo de exposição médio integrado por pixel que o campo *Wide* do DES e a DEEPEST apresenta imagens em que todas as exposições consideradas úteis foram incluídas na imagem coadicionada. É possível notar como as fontes vão aparecendo no fundo da imagem a medida que a profundidade aumenta.

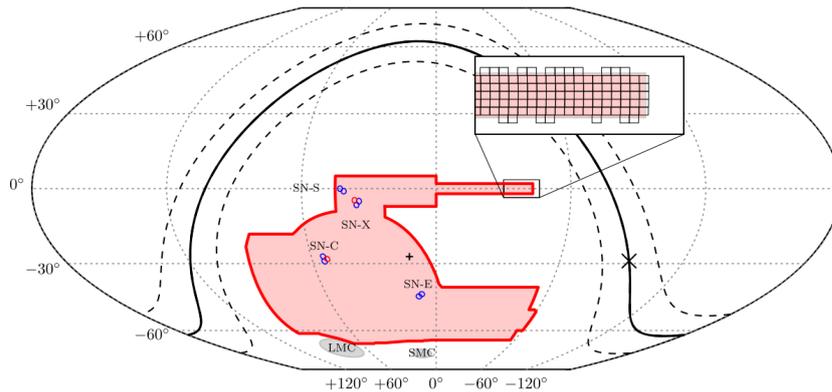


Figura 2.3. *Footprint* do DES em coordenadas equatoriais. Fonte Abbott *et al.* (2021).

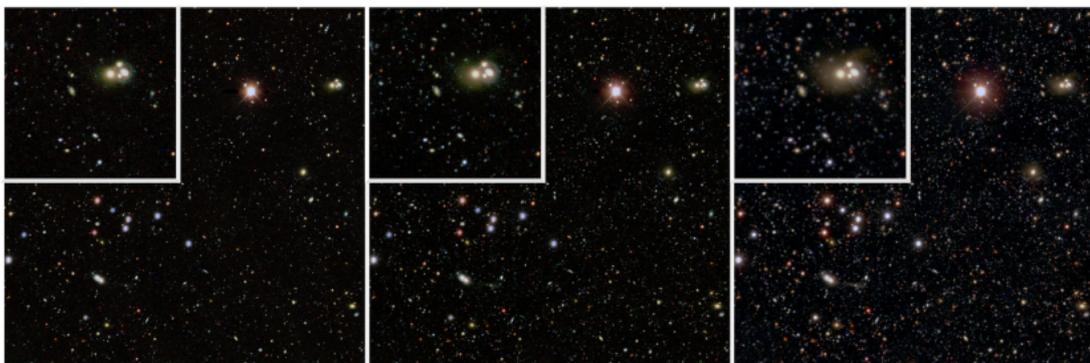


Figura 2.4. Imagens coloridas (grz) em diferentes níveis de profundidade cobrindo metade de um dos chips no campo SN-X3. Fonte: Hartley *et al.* (2022).

A correspondência com a VISTA VIRC*am* (Emerson *et al.* 2006, Dalton *et al.* 2006)

foi interessante por expandir a quantidade de bandas para o infravermelho. Após a correspondência com as imagens da VISTA VIRCam, o catálogo passou a ter informações fotométricas em oito faixas espectrais, ugrizJHK, para mais de 1,6 milhão de fontes. A Tabela 2.1 mostra as características gerais dos filtros da DECam e da VIRCam, e as Figuras 2.5 e 2.6 mostram as passagens de banda entregues dos filtros da DECam (g, r, i, z, Y) e da VIRCam (Y, J, H, Ks), respectivamente. A correspondência entre os catálogos foi diminuiu a área analisada, tendo em vista que a VISTA VIRCam é menos profunda que os DF do DES, o catálogo final dos *Deep Fields* contou com uma redução de 30 graus quadrados para 5,88. Mais informações, como posição, área e profundidade dos *Deep Fields* podem ser vista na Tabela 2.2.

Filtro	$\Lambda(nm)$	FWHM (nm)
DECam g	398 - 548	150
DECam r	568 - 716	148
DECam i	710 - 857	147
DECam z	850 - 1002	152
DECam Y	953 - 1065	112
VIRCam J	1253,47	180
VIRCam H	1645,34	300
VIRCam Ks	2153,99	300

Tabela 2.1: Características dos Filtros Disponíveis na DECam (Flaugher *et al.* 2015) e na VIRCam (Emerson *et al.* 2006, Dalton *et al.* 2006)

Campo	RA (J2000)	Dec (J2000)	Área	$i_{lim}$	Num. Fontes	Nº Chips Excl.
SN-C3	52.6484	-28.1000	1,70	25,06	462.739	13
SN-X3	36.4500	-4.6000	1,52	25,04	430.555	16
SN-E2	9.5000	-43.9980	1,42	25,06	386.946	17
COSMOS	150.1166	+2.2058	1,24	25,54	403147	3
Total	-	-	5,88	-	1.683.387	49

Tabela 2.2: Dados de Campo, de posição, de área, de tempo de exposição, de profundidade dos campos, de número de fontes detectadas (após a aplicação de máscaras) e de número de chips excluídos devido à luz difusa ou outros defeitos de qualidade de imagem em grande escala. Tabela baseada no artigo de Hartley *et al.* (2022).

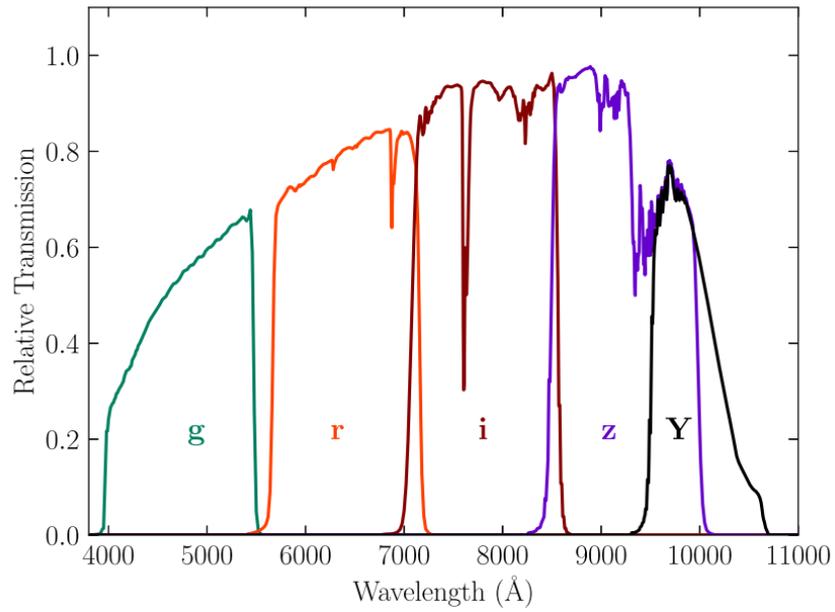


Figura 2.5. Filtros da DECam: grizY. Fonte: Abbott *et al.* (2018)

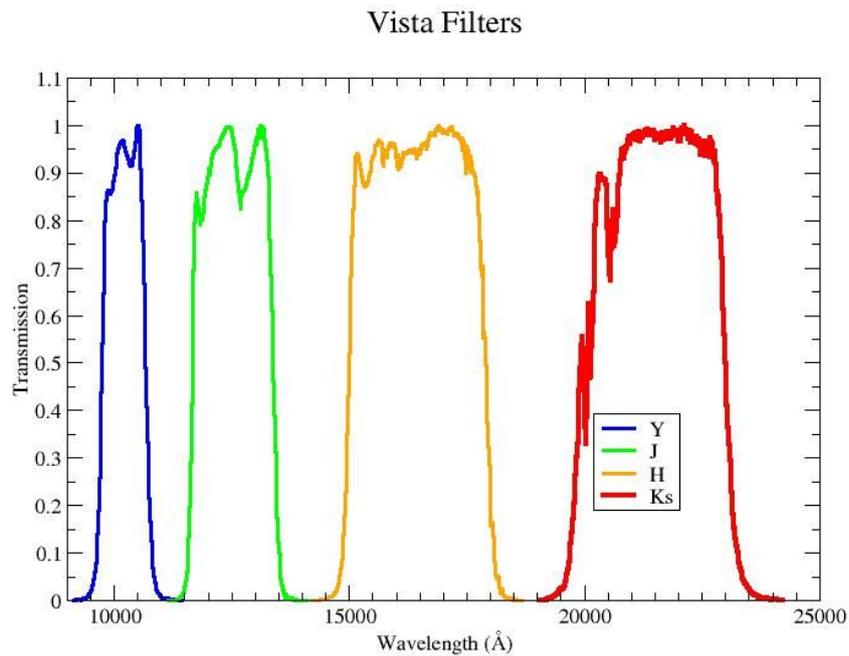


Figura 2.6. Filtros da VIRCam: YJHKs. Fonte: <http://cosmos.ipac.caltech.edu>

Os dados utilizados no presente trabalho são provenientes da seguinte *query* (Figura 2.7) do sistema do DES. Uma descrição de cada variável usada na *query* está presente na Tabela 2.3.

```
SELECT id, ra, dec, tilename, BDF_MAG_DERED_CALIB_U as mag_u,  
BDF_MAG_DERED_CALIB_G as mag_g, BDF_MAG_DERED_CALIB_R as mag_r,  
BDF_MAG_DERED_CALIB_I as mag_i, BDF_MAG_DERED_CALIB_Z as mag_z,  
BDF_MAG_DERED_CALIB_J as mag_J, BDF_MAG_DERED_CALIB_H as mag_H,  
BDF_MAG_DERED_CALIB_KS as mag_Ks, BDF_MAG_ERR_DERED_U as magerr_u,  
BDF_MAG_ERR_DERED_G as magerr_g, BDF_MAG_ERR_DERED_R as magerr_r,  
BDF_MAG_ERR_DERED_I as magerr_i, BDF_MAG_ERR_DERED_Z as magerr_z,  
BDF_MAG_ERR_DERED_J as magerr_J, BDF_MAG_ERR_DERED_H as magerr_H,  
BDF_MAG_ERR_DERED_KS as magerr_Ks, BDF_FRACDEV, BDF_T, BDF_T_RATIO  
FROM Y3_DEEP_FIELDS_PHOTOM  
WHERE flags=0 and flags_nir=0 and mask_flags=0 and mask_flags_nir=0;  
> y3_deep_fields.fits
```

Figura 2.7. Exemplo da *query* do sistema do DES para obtenção do catálogo de galáxias.

Variável	Descrição	# colunas
id	Número de identificação <i>Deep Fields Y3</i>	1
ra	Ascensão Reta, J2000	1
dec	Declinação do objeto, J2000 (graus)	1
tilename	Nome do campo e número do <i>chip</i>	1
<i>BDF_MAG_DERED_CALIB_{U, G, R, I, Z, J, H, KS}</i>	Magnitude do modelo <i>Bulge+Disk</i> com correção de calibração campo-a-campo e correção de avermelhamento Galáctico aplicada	8
<i>BDF_MAG_ERR_DERED_{U, G, R, I, Z, J, H, KS}</i>	Incerteza na magnitude do modelo <i>Bulge+Disk</i> com correção de avermelhamento Galáctico aplicada	8
<i>BDF_FRACDEV</i>	Fração de luz no componente de de Vaucouleurs do modelo <i>Bulge+Disk</i>	1

Tabela 2.3: Descrição das variáveis presente na *query* que resultou no catálogo do presente trabalho. Fonte: Abbott *et al.* (2021)

O catálogo proveniente da seleção presente na Figura 2.7 apresentou as seguintes informações: magnitudes limite para  $g$  de 25,3 mag, para  $r$  de 24,54 mag, para  $i$  de 24,16 mag e para  $z$  de 24,0 mag. Foram feitos cortes nas magnitudes selecionando as galáxias em  $g < 30$ ,  $r < 30$ ,  $i < 26$  e  $z < 30$  e também foram feitos cortes nos erros relativos às magnitudes onde foram selecionadas galáxias com erro menor que 0,5 presente na Figura 2.8. As linhas vermelhas presentes na Figura 2.8 apresentam a magnitude limite do campo do COSMOS no catálogo de DF do DES. Com esses cortes removemos fontes espúrias e usamos a banda  $i$  como referência de limite para a magnitude. A distribuição das magnitudes, a distribuição de RA e Dec e a distribuição de *redshift* do campo COSMOS podem ser observadas na Figura 2.9, 2.10 e 2.11, respectivamente. Note que vamos um pouco além em profundidade na banda  $i$  para permitir uma maior completeza ao tentar encontrar candidatos a aglomerações em altos *redshifts*, como explorado por Sánchez *et al.* (2022). Os vazios presentes na Figura 2.10 devem-se a máscaras que foram aplicadas em busca de eliminar estrelas brilhantes ou defeitos na imagem.

Um outro parâmetro importante para nossa análise é o *fracdev* que tem aplicações em muitos estudos como apresentado anteriormente. Entretanto, este parâmetro tem algumas limitações como: problemas de resolução angular dos dados observacionais, nem todas as galáxias são bem representadas por modelos de decomposição bojo-disco (Kimm e Yi, 2007). No DES, devido ao fato do *fracdev* ser mal ajustado em objetos com magnitude aparente fraca, não se tem informações suficientes para obter um *fracdev* preciso. Esses objetos se concentram em *fracdev*= 0,5, que é o *prior* do modelo utilizado, por não ter um ajuste adequado. Devido a essas limitações, foi realizado o seguinte corte na variável BDF  $\text{fracdev}_{-0,01}^{+0,005}$  em  $r(\text{mag}) > 24$ , como apresentado na Figura 2.12, na hora de analisar esse parâmetro.

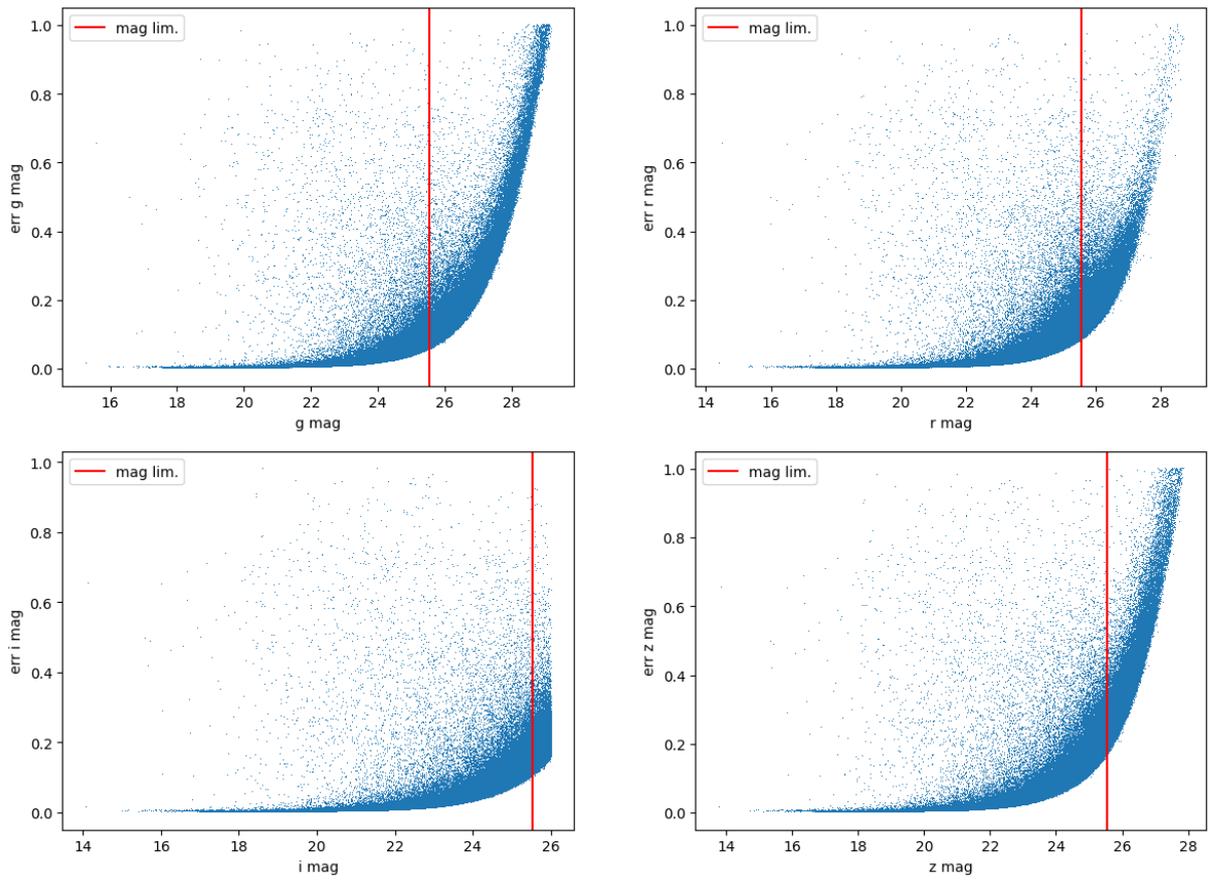


Figura 2.8. Erros associados as magnitudes referentes as bandas griz extraídos do catálogo do DES.

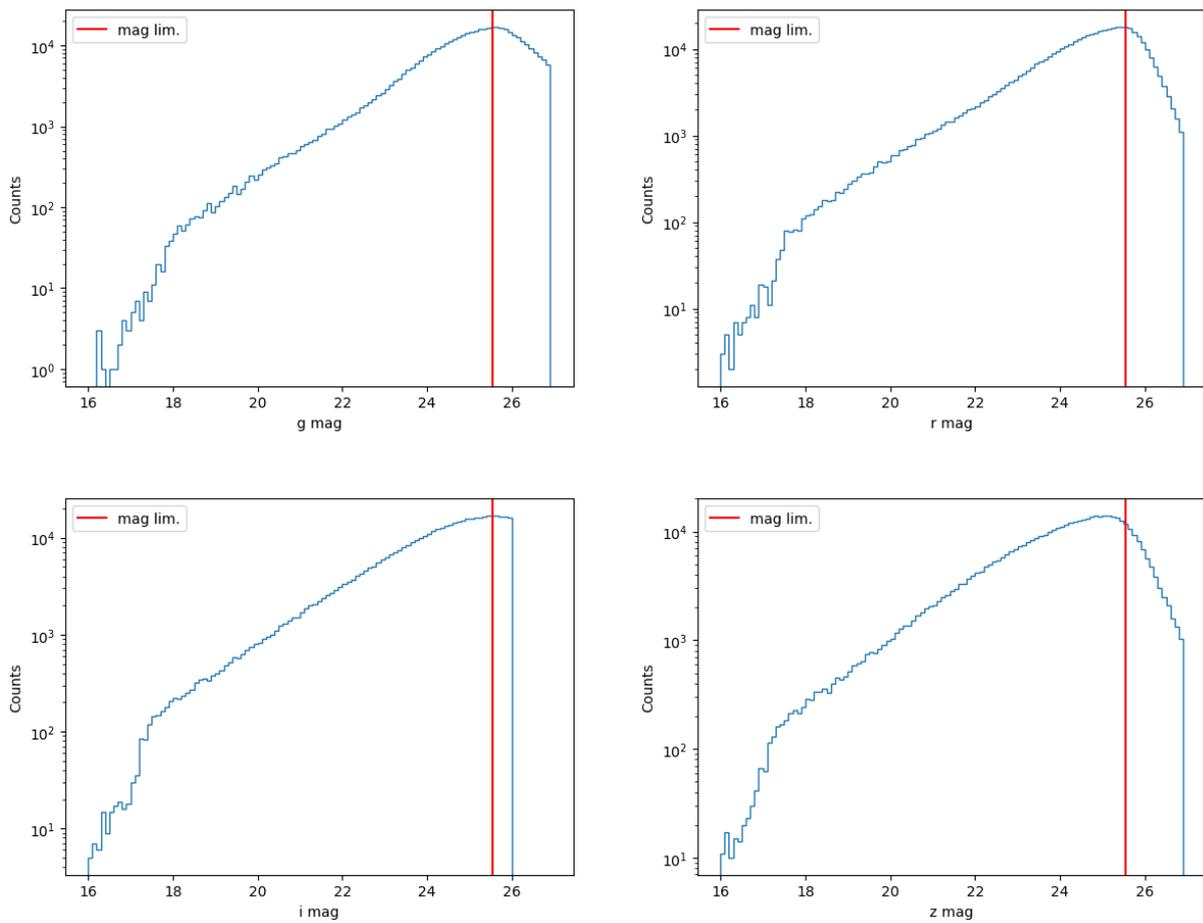


Figura 2.9. Histograma de frequência de magnitude das bandas griz extraídos dos dados do DES.

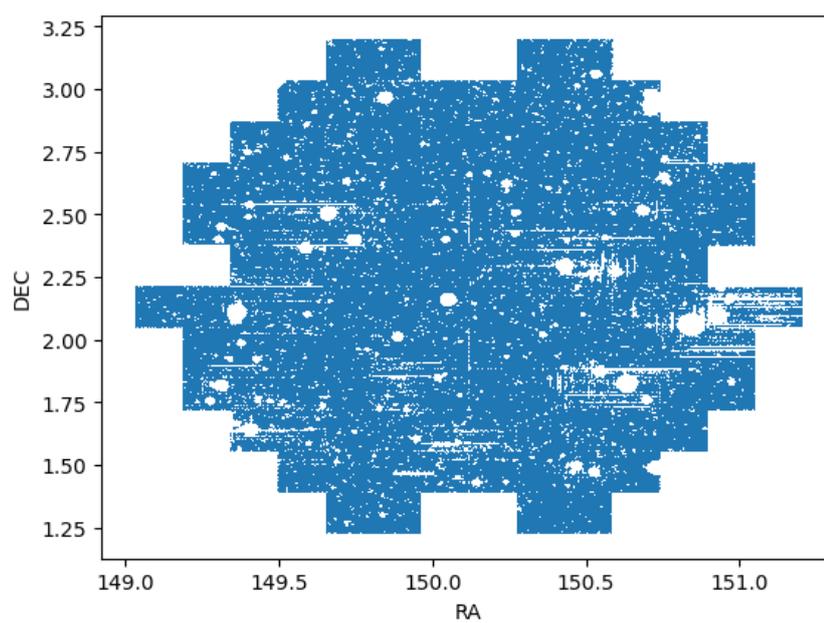


Figura 2.10. RA e Dec do COSMOS extraído dos dados do DES.

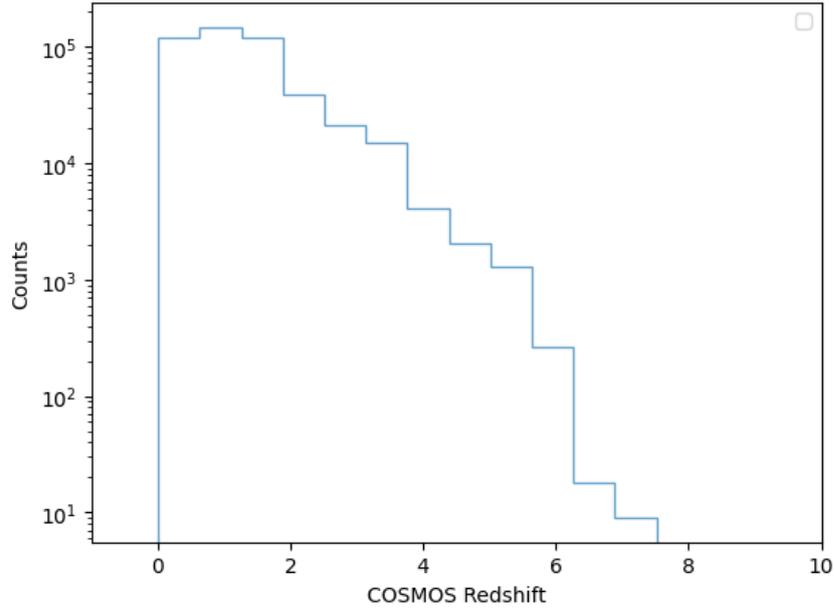


Figura 2.11. Histograma de *counts* versus o *redshift* do campo do COSMOS.

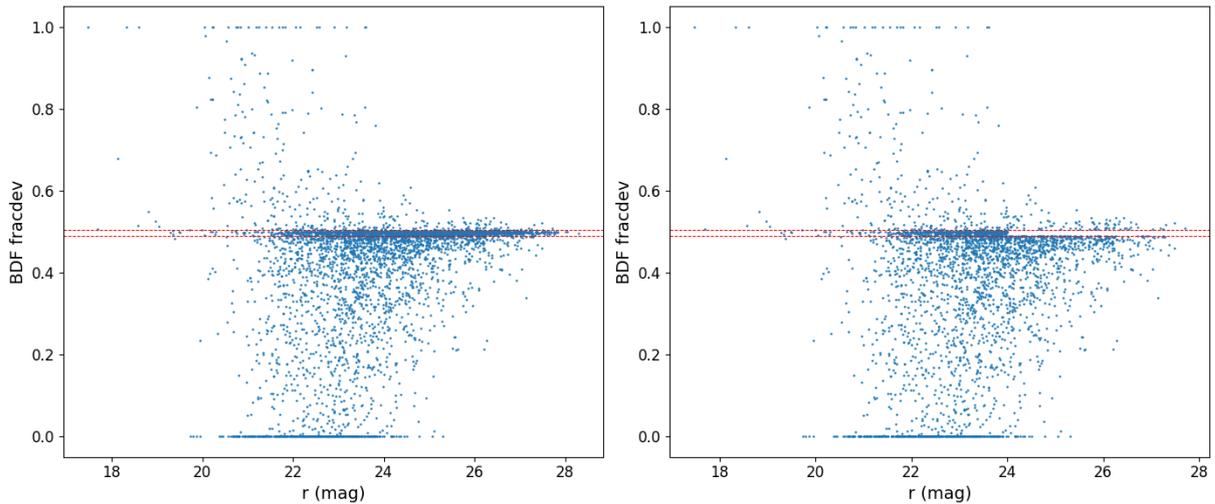


Figura 2.12. Dados de BDF *fracdev* em função da magnitude na banda r das galáxias do catálogo G24. O painel da esquerda apresenta os dados completos e o painel da direita apresenta os dados após o corte em  $BDF\ fracdev = 0,5^{+0,005}_{-0,01}$  e valores de  $r(mag) > 24$ . As linhas vermelhas apresentam a zona do corte.

A variável correspondente ao photo-z utilizada no presente trabalho foi o  $Z\_PEAK$  que é dada pelo *redshift* médio do pico mais provável do *redshift* (Hartley *et al.* 2022). A representação da distribuição de *redshift* utilizada está presente na Figura 2.13. Note que a partir de  $z$  maior que 2 a distribuição cai bastante prejudicando a busca por estruturas, por isso restringimos nossa análise a uma faixa restrita de *redshift*. Estes *redshifts* estão presentes no DES *Data Management* (DESDM) como catálogo Y3\_DEEP\_FIELDS\_PHOTOZ (EAZY) de Hartley *et al.* (2022) e foram calculados usando o software de ajuste de *template* EAZY (Brammer *et al.*, 2008a). Como as informações de photo-z são provenientes

de um outro catálogo, pois o catálogo de fotometria não possui informações a respeito do photo-z, foi necessário fazer um *match* entre eles.

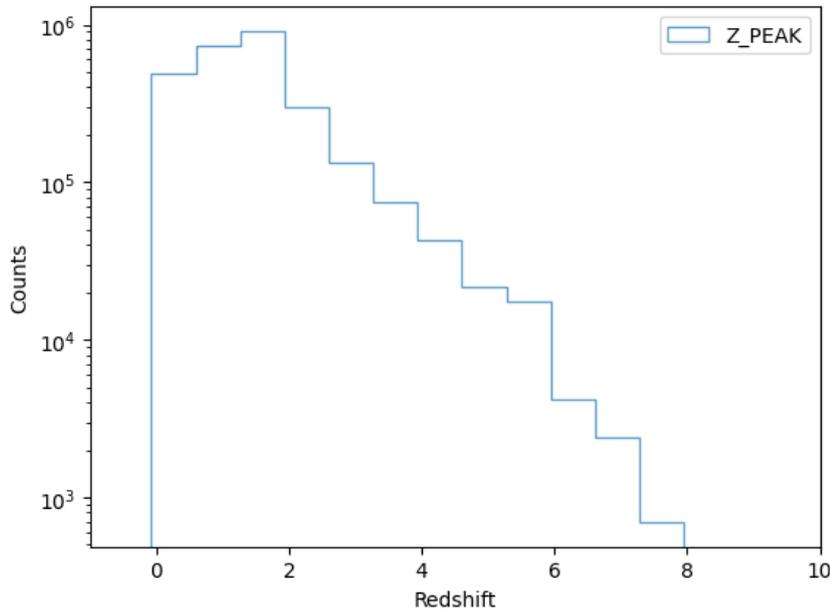


Figura 2.13. Histograma de *counts* versus o *redshift* dos *Deep Fields*.

## 2.2 Catálogos de protoaglomerados na literatura

O presente trabalho não é o primeiro a explorar aglomerações de galáxias em alto *redshift*, em particular no campo COSMOS. Vários autores produziram catálogos de protoaglomerados com diferentes propriedades que mencionamos na seção 1.5. Neste contexto, foram selecionados catálogos pela proximidade com a presente análise, possibilitando a comparação. Os catálogos utilizados para comparar com os resultados do presente trabalho foram obtidos por Finoguenov *et al.* (2007), Bellagamba *et al.* (2011), Wen e Han (2011) e Sochting *et al.* (2012). A Tabela 2.4 apresenta dados dos catálogos disponíveis na literatura que serviram de referência nas análises deste trabalho.

Estudo	Nº de Cand/Clusters	Área	Faixa de z
F07	72	2,1 <i>deg</i> <sup>2</sup>	0,08-1,25
W11	187	2 <i>deg</i> <sup>2</sup>	0,18-1,61
B11	119	2 <i>deg</i> <sup>2</sup>	0,12-0,8
S12	1780	2 <i>deg</i> <sup>2</sup>	0,20-2,98

Tabela 2.4: Resultados dos catálogos comparados, suas respectivas áreas e faixas de *redshift*. F07 (Finoguenov *et al.* 2007), W11 (Wen e Han, 2011), B11 (Bellagamba *et al.* 2011), S12 (Sochting *et al.* 2012).

Finoguenov *et al.* (2007) (F07) utilizaram os dados dos primeiros 36 apontamentos do XMM-Newton no COSMOS onde obteve um número de aglomerados igual a 72. O tamanho da área do COSMOS que eles estudaram foi de 2,1 *deg*<sup>2</sup> e a faixa de *redshift*

é de 0,08-1,25, a magnitude limite foi de  $i_{AB} < 25$ . O gráfico contendo a distribuição de *redshifts* encontra-se na Figura 2.14, porém, a relação entre riqueza e *redshift* não foi indicada pelos autores.

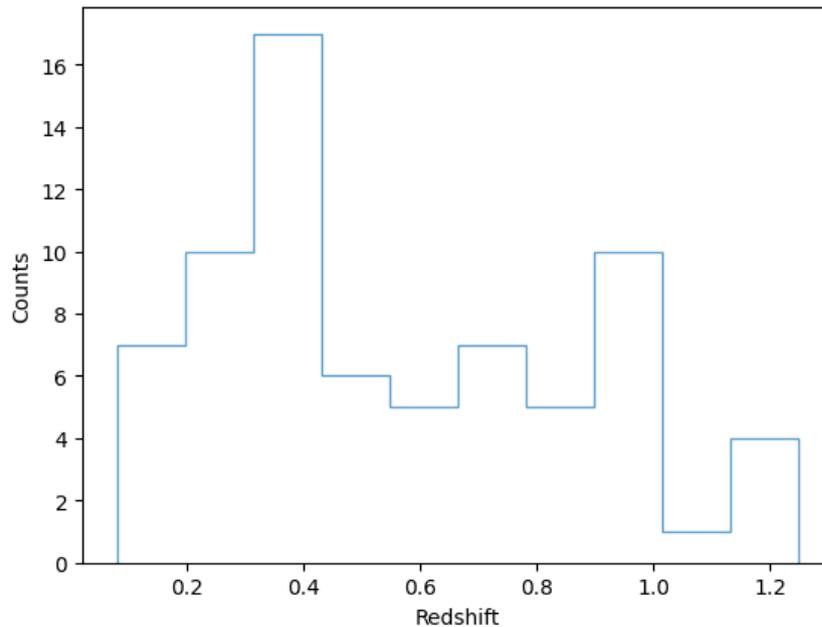


Figura 2.14. Histograma de *counts* versus o *redshift* de Finoguenov *et al.* (2007).

Wen e Han (2011) (W11) utilizaram os dados de Canada-France-Hawaii Telescope (CFHT) *wide field*, COSMOS e do *Spitzer Wide-area InfraRed Extragalactic Survey (SWIRE) field*, tendo identificado 631, 187 e 737 aglomerados, respectivamente. O foco da comparação com esse trabalho é a área do COSMOS, logo, conseguiu um número de aglomerados igual a 187. O tamanho da área estudada foi de  $2 \text{ deg}^2$  e a faixa de *redshift* é de 0,18-1,61, a magnitude limite do estudo foi de  $i < 25$  nesse artigo. Os autores identificaram 127 novos aglomerados (dentre os 187 ditos previamente). Desses, 43 foram encontrados em altos *redshift* 38 desses são novas descobertas. O gráfico contendo a distribuição de *redshift* encontra-se na Figura 2.15 e a relação entre riqueza e *redshift* na Figura 2.16

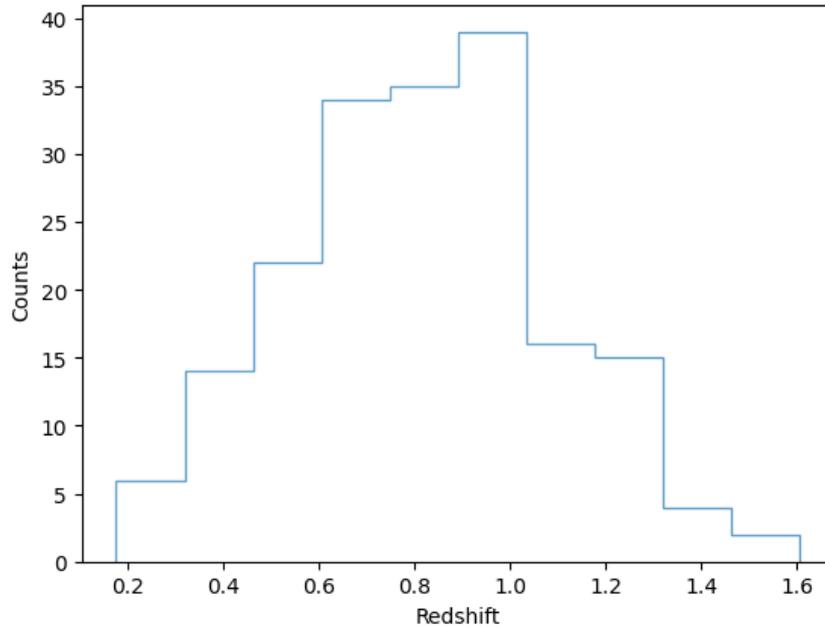


Figura 2.15. Histograma de *counts* versus o *redshift* de Wen e Han (2011).

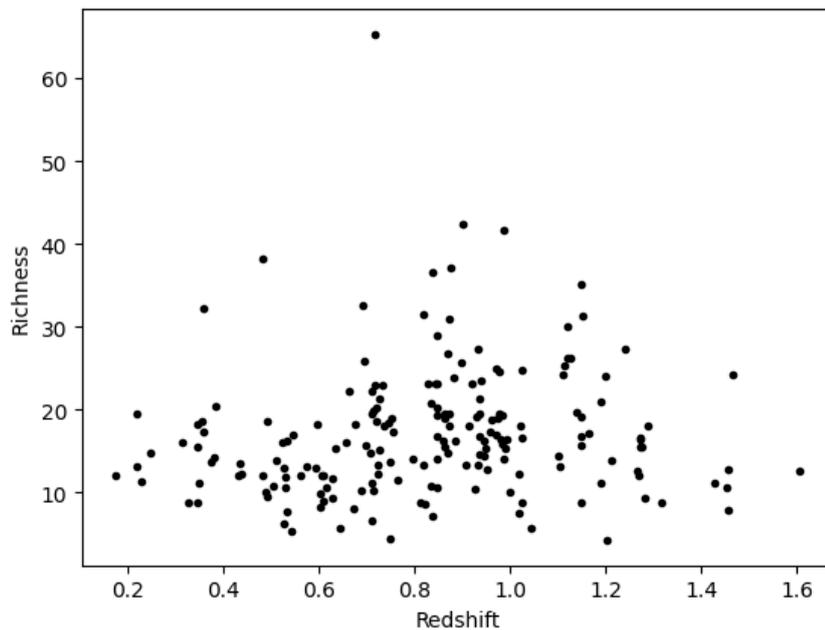


Figura 2.16. Dados de riqueza vs *redshift* de Wen e Han (2011).

Bellagamba *et al.* (2011) (B11) utilizaram os dados do COSMOS, tendo obtido um número de candidatos igual a 119. O tamanho da área do COSMOS que eles estudaram foi de  $2 \text{ deg}^2$  e a faixa de *redshift* é de 0,12-0,8, que é um pouco menor que F07. O gráfico contendo a distribuição de *redshift* encontra-se na Figura 2.17 e a relação entre riqueza e *redshift* na Figura 2.18

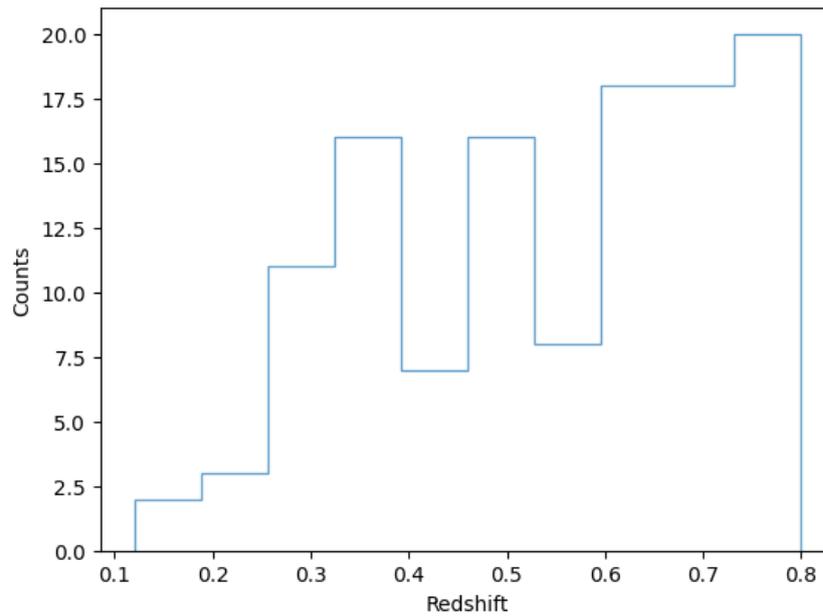


Figura 2.17. Histograma de *counts* versus o *redshift* de Bellagamba *et al.* (2011).

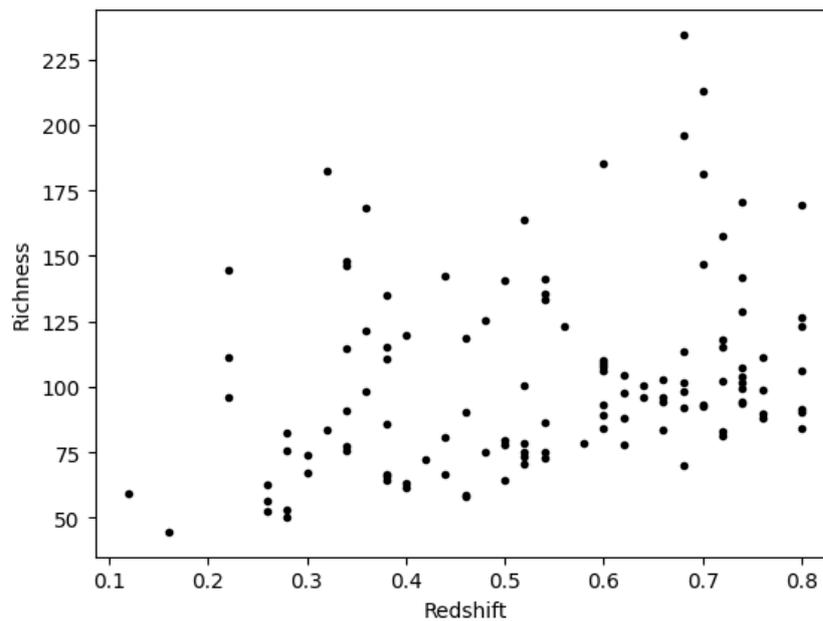


Figura 2.18. Dados de riqueza vs *redshift* de Bellagamba *et al.* (2011).

Sochting *et al.* (2012) utilizaram os dados do COSMOS, tendo neste catálogo 607.617 objetos e detectado 1780 aglomerados. O tamanho da área do COSMOS que eles estudaram foi de  $2 \text{ deg}^2$  e a faixa de *redshift* é de 0,2-3,0, a magnitude limite foi de  $i_{AB} < 21,5 + 3,0 \times z_{ph}$ . O gráfico contendo a distribuição de *redshift* encontra-se na Figura 2.19 e a relação entre riqueza e *redshift* na Figura 2.20

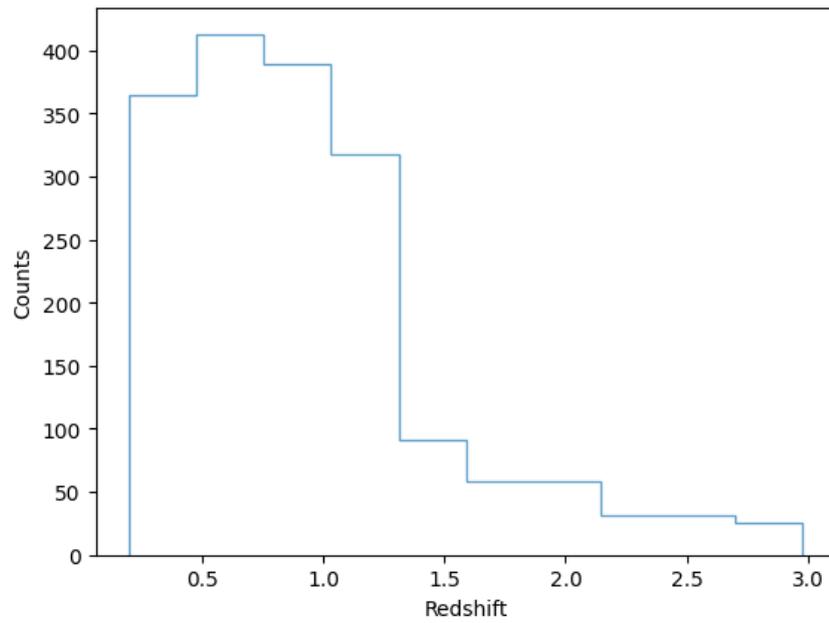


Figura 2.19. Histograma de *counts* versus o *redshift* de Sochting *et al.* (2012).

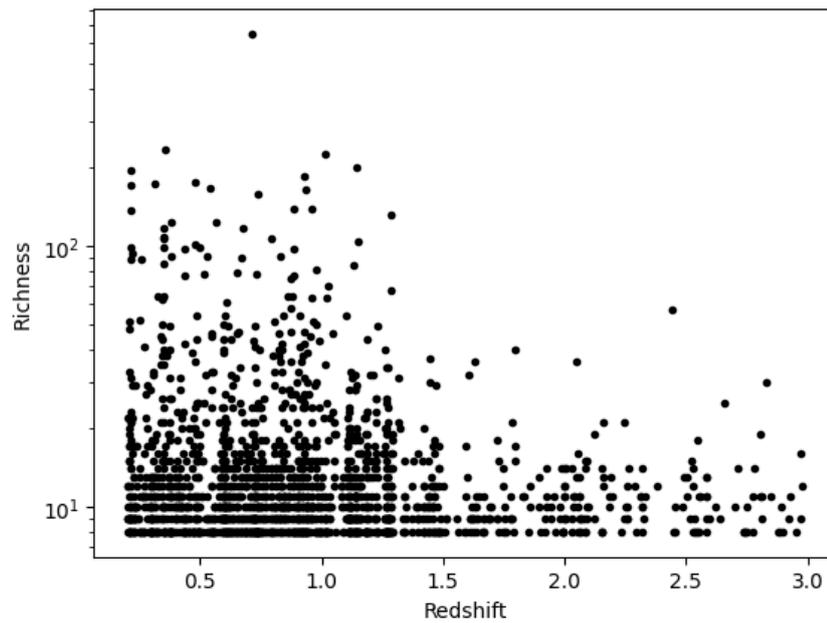


Figura 2.20. Dados de riqueza vs *redshift* de Sochting *et al.* (2012).

Esses catálogos serão utilizados como base de comparação entre eles e nosso catálogo de candidatos na seção 3.4.

# Capítulo 3

## Metodologia

Este capítulo apresenta a metodologia aplicada na determinação de candidatos a protoaglomerados de galáxias, sua validação e caracterização. As seções 3.1 e 3.2, respectivamente, apresentam a técnica e a metodologia de busca por protoaglomerados de galáxias baseada na localização de regiões de alta densidade. Nas seções 3.3 e 3.4, respectivamente, são apresentadas a metodologia de estudo de *match* 2D e 3D com a literatura. Na seção 3.5, é apresentada a técnica de validação dos resultados baseada no *match* com um catálogo *Bootstrap*. Na seção 3.6, é apresentada a estratégia de estudo dos candidatos inéditos e estudos de morfologia.

Inicialmente, foram avaliadas técnicas para a localização de protoaglomerados de galáxias na faixa de  $[0,5-2,0]$ . A técnica adotada foi a de busca por sobredensidade de galáxias. Para a escrita do algoritmo de busca, utilizou-se a metodologia aplicada por (Narwal *et al.*, 2020). Os autores utilizaram a técnica de sobredensidades para localizar picos de concentração de galáxias em fatias de *redshift* e identificaram um candidato em  $z \sim 1,1$ . Uma vez escolhida a técnica de sobredensidades, foram escritos algoritmos em linguagem Python e validados pelos resultados de Narwal *et al.* (2020), tendo sido identificado o mesmo candidato a superaglomerado encontrado pelos autores. Posteriormente, foi feita uma busca por candidatos em uma região previamente definida. Os candidatos encontrados pelo algoritmo de busca foram comparados com catálogos da literatura e um catálogo *mock*. Finalmente, o catálogo de candidatos foi estudado com o objetivo de validá-lo.

### 3.1 Algoritmo de busca de candidatos

Foi aplicado o método de sobredensidades para encontrar os candidatos a aglomerados escalonando o tamanho físico dos pixels conforme a distância em Mpc. O algoritmo foi construído na seguinte ordem:

- 1) São lidas informações de um catálogo de galáxias, no caso deste trabalho, o catálogo

foi gerado pelo *pipeline* do DES;

- 2) Dividiu-se as galáxias em *bins* de *redshift* com tamanho igual a 0,1 e passo de 0,03,
- 3) Fez-se uma busca por sobredensidades de galáxias da seguinte forma. Divide-se a área de busca em pixels de tamanho igual a 2 Mpc, e checa-se a quantidade de galáxias em cada pixel, utilizando a função `np.histogram2d()`. Determina-se o *threshold* igual a 95% do valor do pico máximo para dizer se determinado pixel possui ou não uma sobredensidade. Tomou-se como *threshold* 95% do valor do pixel de maior densidade de cada *bin*, ou seja, o pixel que possuir um valor  $\geq$  a 95% do valor do pixel mais denso é considerado uma sobredensidade. A Figura 3.1 apresenta um esquema representativo de busca por sobredensidade em *bins* predefinidos de aplicação da função de busca `np.histogram2d()`;
- 4) Após definir os picos de sobredensidade, fez-se uma busca pelas galáxias membro de cada candidato utilizando a função `search_around_sky()`. Esta função determina as galáxias que estão a uma determinada distância de um dado ponto;

O algoritmo foi desenvolvido em linguagem Python utilizando bibliotecas do `astropy` e do `numpy` (Astropy Collaboration *et al.*, 2022, Harris *et al.*, 2020).

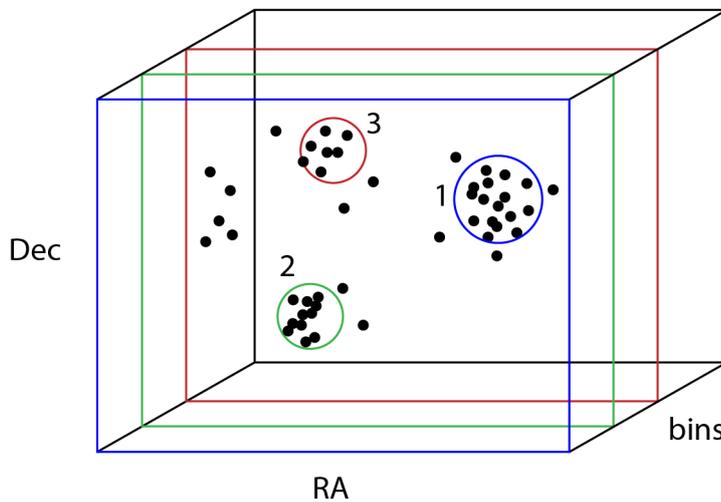


Figura 3.1. Esquema representativo de busca de sobredensidades em diferentes *bins* de *redshift*.

Embora os protoaglomerados tenham uma escala de tamanho físico característica e, por exemplo, Narwal *et al.* (2020) utilize um tamanho de 5 Mpc, realizou-se uma verificação do tamanho ideal para a busca. Com esse fim, foram testados diferentes tamanhos de pixels e fez-se uma comparação do catálogo resultante com aqueles fornecidos pela

literatura. Para definir o tamanho dos pixels mencionados anteriormente, foi realizada uma análise variando o tamanho dos mesmos em Mpc. O estudo foi feito para pixels de 1 Mpc até 10 Mpc de largura. Os resultados dessas análises são apresentados na seção 4.2.

Em seguida, cada *bin* de *redshift* foi analisado buscando por sobredensidades de galáxias, ou seja, por possíveis candidatos a protoaglomerados. Tendo em vista o *bin* em que os autores identificaram a sobredensidade e, portanto, o candidato a protoaglomerado, foi feita uma busca na mesma área do DES-SVA1.

## 3.2 Busca de protoaglomerado no DES

Foram estudados diversos métodos para identificar sobredensidades de galáxias, incluindo FOF (friends-of-friends) (Farrens *et al.*, 2011), ou de diferentes populações, como por exemplo *Lyman Break Galaxies* (Toshikawa *et al.* 2016). Para o desenvolvimento deste trabalho foi escolhido o método de busca por sobredensidades como o utilizado por Narwal *et al.* (2020). Esta escolha deu-se pela simplicidade de programação do método de sobredensidade.

Destá forma, validada a metodologia apresentada por Narwal *et al.* (2020), foram realizadas buscas por sobredensidade nos *Deep Fields* (DF) do DES, com ênfase no campo do COSMOS. Escolheu-se este campo, visto que ele possui numerosos estudos na literatura (Sochting *et al.* (2012), Dong *et al.* (2023), Lazar *et al.* (2023), Reshetnikov *et al.* (2023)) sendo o mais contíguo dentre os DFs, tendo em vista a Tabela 2.2.

O campo do COSMOS possui  $\sim 1,24deg^2$ , a profundidade na banda *i* chega a 25,54 mag e possui 403.147 fontes. Os dados foram divididos em *bins* de *redshift*, sendo a faixa analisada  $0,5 \leq z \leq 2$ , estes com uma largura de 0,1. Tal intervalo de *redshift* foi escolhido baseado na distribuição de *redshift* presente na Figura 2.13, que apresenta o histograma de *counts* versus o *redshift* dos *Deep Fields*.

Inicialmente, foi feita uma abordagem com passo de 0,1 para a separação em *bins* de *redshift* sem considerar a interseção entre eles a fim de compreender a aplicação da técnica, ou seja, utilizando-se de um caso mais simples para fins de prototipagem. Posteriormente, foram utilizados *bins* de *redshift* de 0,1, mas com passos de 0,03 de interseção entre *bins*. A interseção entre *bins* resulta em detecções que atravessam múltiplos *bins* contíguos, mas que produz uma suavidade maior na distribuição ao longo do photo-z, melhorando nossa capacidade de definir em que *z* está localizado o candidato.

Logo, para uma detecção, foram feitas buscas ao longo de *bins* contíguos com o objetivo de encontrar o *bin* onde há a maior quantidade de galáxias e selecionar o photo-z dos candidatos. A Figura 3.2 apresenta um resultado típico onde o ponto de maior quantidade de galáxias dentre os *bins* é selecionado para compor o catálogo de candidatos.

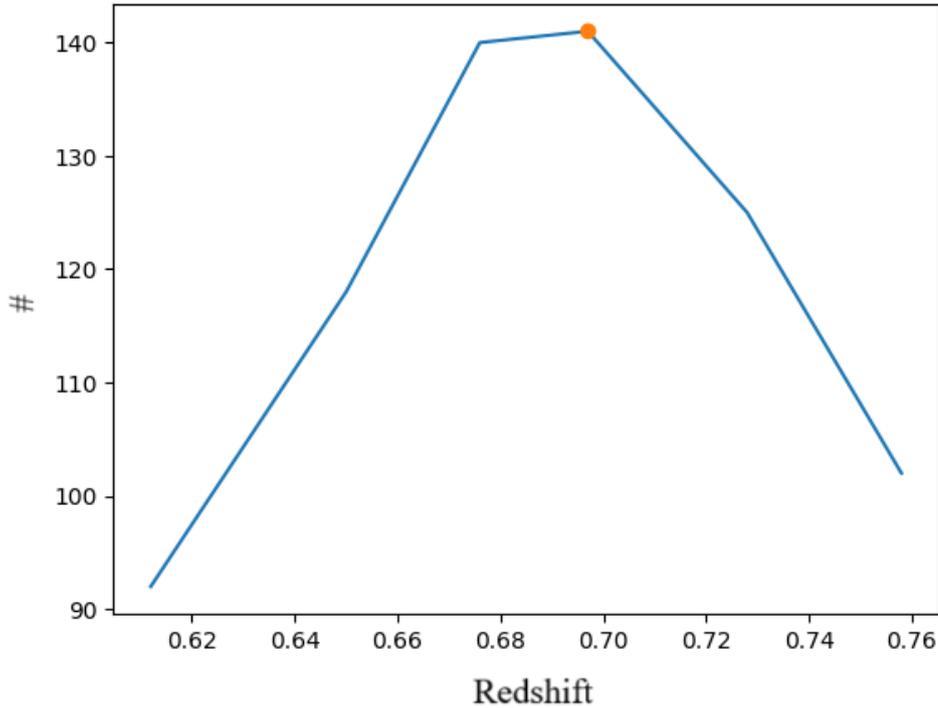


Figura 3.2. Quantidade de galáxias em cada bin de *redshift* de um candidato. O ponto laranja representa o *bin* de *redshift* com maior quantidade de galáxias, sendo portanto escolhido como o *redshift* do candidato.

Terminada a busca por candidatos, montou-se um arquivo com os dados (v.g. RA, Dec,  $z$ ) dos candidatos, além de um arquivo para cada candidato com suas galáxias-membro e suas propriedades (v.g. RA, Dec, magnitudes, respectivos erros, e fração de bojo/disco). Com o catálogo principal de candidatos foi feita uma comparação com catálogos da literatura (Finoguenov *et al.*, 2007 (F07), Wen e Han, 2011 (W11), Bellagamba *et al.*, 2011 (B11) e Sochting *et al.*, 2012 (S12)).

### 3.3 *Match* 2D com dados da literatura

Inicialmente, foi feito um estudo do *match* 2D para entender as limitações do photo- $z$ . Neste contexto, foram realizadas comparações diretas entre os catálogos da literatura e entre a literatura e este trabalho. Por exemplo, aglomerados que pareciam estar próximos angularmente e deram *match* 2D, deixaram de dar *match* quando foi adicionado o *redshift* no *match*, como discutido na próxima seção. Uma análise baseada apenas nas coordenadas de RA e Dec pode ajudar a entender as limitações do photo- $z$  no estudo por meio, por exemplo, do resíduo de photo- $z$  (*redshift* da literatura menos *redshift* deste trabalho) em função do *redshift*. Os resultados dessas análises são apresentados na seção 4.3.

### 3.4 *Match* 3D com dados da literatura

O estudo do *match* 3D, considerando o *redshift*, foi realizado para considerar a distância angular e na linha de visada entre os candidatos a protoaglomerados.

Esta etapa levou a uma redução na quantidade de correspondências identificadas entre os catálogos, tanto da literatura quanto neste trabalho. Esta redução era esperada pois ao incorporar a variável  $z$  na análise, a comparação tornou-se mais refinada, resultando em um corte mais criterioso na quantidade de correspondências identificadas.

Um dos parâmetros livres do nosso método de busca por protoaglomerados é o tamanho do pixel onde é calculada a sobredensidade. No estudo, foram avaliados diferentes tamanhos de pixels de modo a verificar qual produzia o maior número de *matches* com a literatura, chegando-se ao valor de 2Mpc para o tamanho do pixel.

O catálogo de Sochting *et al.* (2012) apresenta um grande número de detecções na área, o que, após a imposição de restrições na faixa de  $z$  analisada, permitiu uma análise mais adequada do *match* entre os conjuntos de dados.

É importante ressaltar que Sochting *et al.* (2012) fizeram buscas considerando aglomerações contendo um número reduzido de galáxias. Outra informação importante que os autores forneceram foi que para  $z > 1,3$ , as estruturas com luminosidades mais baixas não eram detectadas devido ao limite de profundidade dos dados que eles utilizaram. Desta forma, objetivando comparar os resultados dos autores com este trabalho, foi feito um corte nos dados a partir de  $z > 1,3$  e na quantidade de galáxias por aglomerado. Na escala de riqueza deles, cortamos aqueles aglomerados a partir de 16 de modo a evitar o *match* espúrio com os numerosos pequenos aglomerados.

Neste trabalho, na comparação com catálogos da literatura, as correspondências entre S12 e os catálogos de F07, W11 e B11, foi aplicada a restrição de  $z > 1,3$  em S12 e isto não afetou, significativamente, a quantidade de *match* entre eles. Os autores também encontraram 6 superaglomerados, porém não deram informações para que fosse possível analisá-los. Os resultados dessas análises são apresentados na seção 4.4.

### 3.5 Teste de *matches* com bootstrap

Como o catálogo de Sochting *et al.* (2012) tem 1485 de aglomerados, após os cortes mencionados acima, bem maior que os demais da literatura, avaliou-se a possibilidade de ocorrência de *matches* aleatórios entre os catálogos analisados. Para confirmar que os *matches* obtidos não eram aleatórios, foi utilizada a técnica de *Bootstrap* (Efron, 1979) para randomizar os dados do catálogo do presente trabalho e procurar por correspondências com Sochting *et al.* (2012). Não tentamos reproduzir o grau de correlação dos dados originais nesse processo, simplesmente randomizando suas coordenadas.

Os resultados dessas análises são apresentados na seção 4.5.

### 3.6 Estudo dos candidatos inéditos

Foram considerados candidatos inéditos aqueles que não apresentaram *match* com a literatura e que passaram por uma seleção por meio da análise das populações de galáxias a partir da separação em grupos de galáxias vermelhas e azuis como o resultado típico apresentado na Figura 3.3.

Desta forma, foi aplicada a rotina *Gaussian Mixture Model* (GMM) (Pedregosa *et al.*, 2011) para dividir a população de galáxias desses candidatos inéditos entre galáxias vermelhas e azuis. O GMM é um modelo estatístico que supõe que os dados são gerados por uma combinação de múltiplas distribuições gaussianas, cujos parâmetros são desconhecidos. No caso escolhemos duas distribuições representando as populações vermelhas e azuis.

Para confirmar a eficiência da rotina *K-Correct* foi feita uma comparação dos dados não corrigidos com os corrigidos, colocando lado a lado os resultados. Foi feita uma análise visual dos dados afim de averiguar se os Diagrama Cor-Magnitude (CMDs) estariam bem definidos. Os candidatos que não apresentaram no CMD uma estrutura bem definida, com populações azuis ou vermelhas misturadas, foram excluídos do catálogo final.

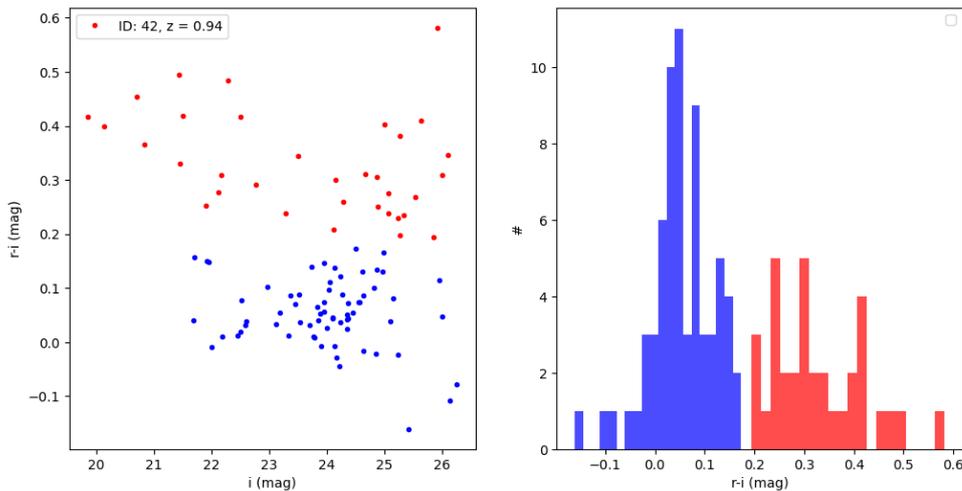


Figura 3.3. Exemplo de um CMD (painel da esquerda) e de um histograma número de galáxias distribuídas em função da cor (painel da direita) para o candidato de ID 42.

Os resultados dessas análises são apresentados na seção 5.1

Em uma etapa seguinte, foi feito um estudo das características morfológicas do catálogo de candidatos por meio da variável BDF `fracdev` em função do *redshift* e o comportamento das populações azuis e vermelhas de galáxias de acordo com as cores g-r, r-i e i-z.

# Capítulo 4

## Resultados da validação do algoritmo

Este capítulo apresenta os resultados de validação do algoritmo proposto para a busca de sobredensidades. A técnica de sobredensidade faz uma busca por regiões densas, necessitando de poucos parâmetros para ser aplicada. Para a realização das buscas de candidatos a protoaglomerados, foi feito um algoritmo em linguagem Python.

A validação do algoritmo foi feita a partir da reprodução de um trabalho da literatura (Narwal *et al.*, 2020). No trabalho, que foi adotado para validar os resultados, os autores também utilizaram a técnica de sobredensidades para localizar picos de concentração de galáxias em fatias de *redshift* e identificaram um candidato a superaglomerado em  $z \sim 1,1$ . Nos testes de validação do código, foi identificada uma sobredensidade no mesmo local no qual os autores encontraram o candidato a superaglomerado.

Na sequência, fez-se um estudo do tamanho ótimo do pixel variando-o de 1 a 10Mpc. Tendo determinado o tamanho ótimo do pixel, fez-se uma busca por candidatos na região do COSMOS do DES, na faixa de *redshift* [0,5 a 2,0]. Os candidatos encontrados pelo algoritmo de busca foram comparados com catálogos da literatura. Na etapa final, o catálogo de candidatos foi analisado visando validar os resultados obtidos.

### 4.1 Validação com Narwal et al. (2020)

No estudo preliminar conduzido para testar a aplicação da técnica, foi possível reproduzir o resultado obtido por Narwal *et al.* (2020). Para a busca do candidato a superaglomerado, dividiu-se os *redshifts* das galáxias em *bins* e fez-se histograma 2D de cada *bin*. Considerando o *bin* no qual os autores identificaram o candidato a superaglomerado, foi feita uma busca na mesma região.

Narwal *et al.* (2020) selecionaram somente objetos com maior certeza de serem galáxias com pureza de, aproximadamente, 90%. Foram removidas galáxias cujos erros eram maiores que 1 mag em todas as bandas griz. Assim, como no artigo de Narwal *et al.* (2020), a seleção dos dados foi apresentada na forma de um histograma dos *redshifts* do DES-SVA1, ver Figura 4.1, cuja seleção foi feita aproximando o melhor possível dos

dados utilizados pelos autores. Os painéis apresentados na Figura 4.1 representam os histogramas de *redshift* dos dados do SVA1 gerado por Narwal *et al.* (2020) e o reproduzido neste trabalho.

Após a seleção dos dados do catálogo, conduziu-se uma busca por sobredensidades. Como resultado, obteve-se o *contour plot* do contraste de densidade exibindo as estruturas de sobredensidade no *bin* [1,01-1,21] apresentado no painel inferior da Figura 4.2. Pode-se observar na Figura 4.2 (painel inferior) que neste trabalho obteve-se a sobredensidade no intervalo de Dec de -50 a -48 e RA de -2,5 a 0,0, coincidindo com a mesma região reportada pelos autores, ver painel superior da Figura 4.2. Desta forma, considerou-se o algoritmo validado por ter reproduzido o resultado obtido por Narwal *et al.* (2020). Tendo validado o algoritmo a partir do trabalho de Narwal *et al.* (2020), foi realizada a busca por sobredensidade nos DF do DES.

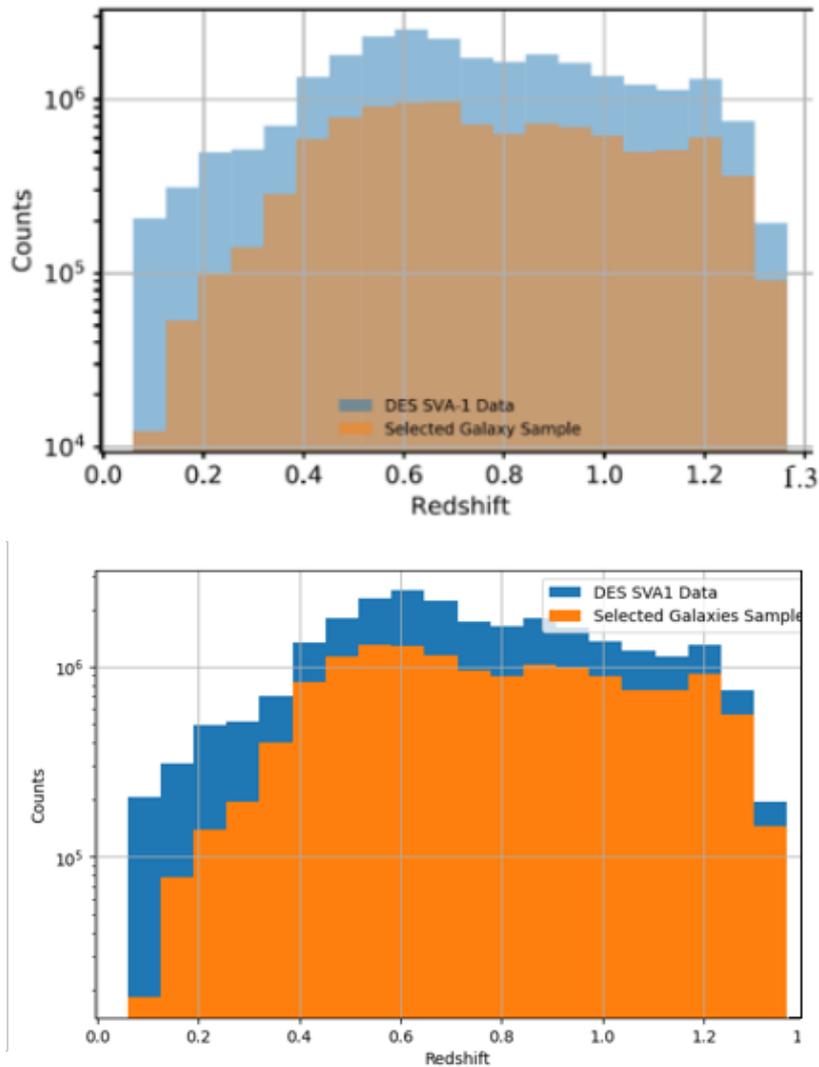


Figura 4.1. Histogramas de *redshift* dos dados do SVA1 gerados no trabalho de Narwal *et al.* (2020) (painel superior) e o reproduzido neste trabalho (painel inferior).

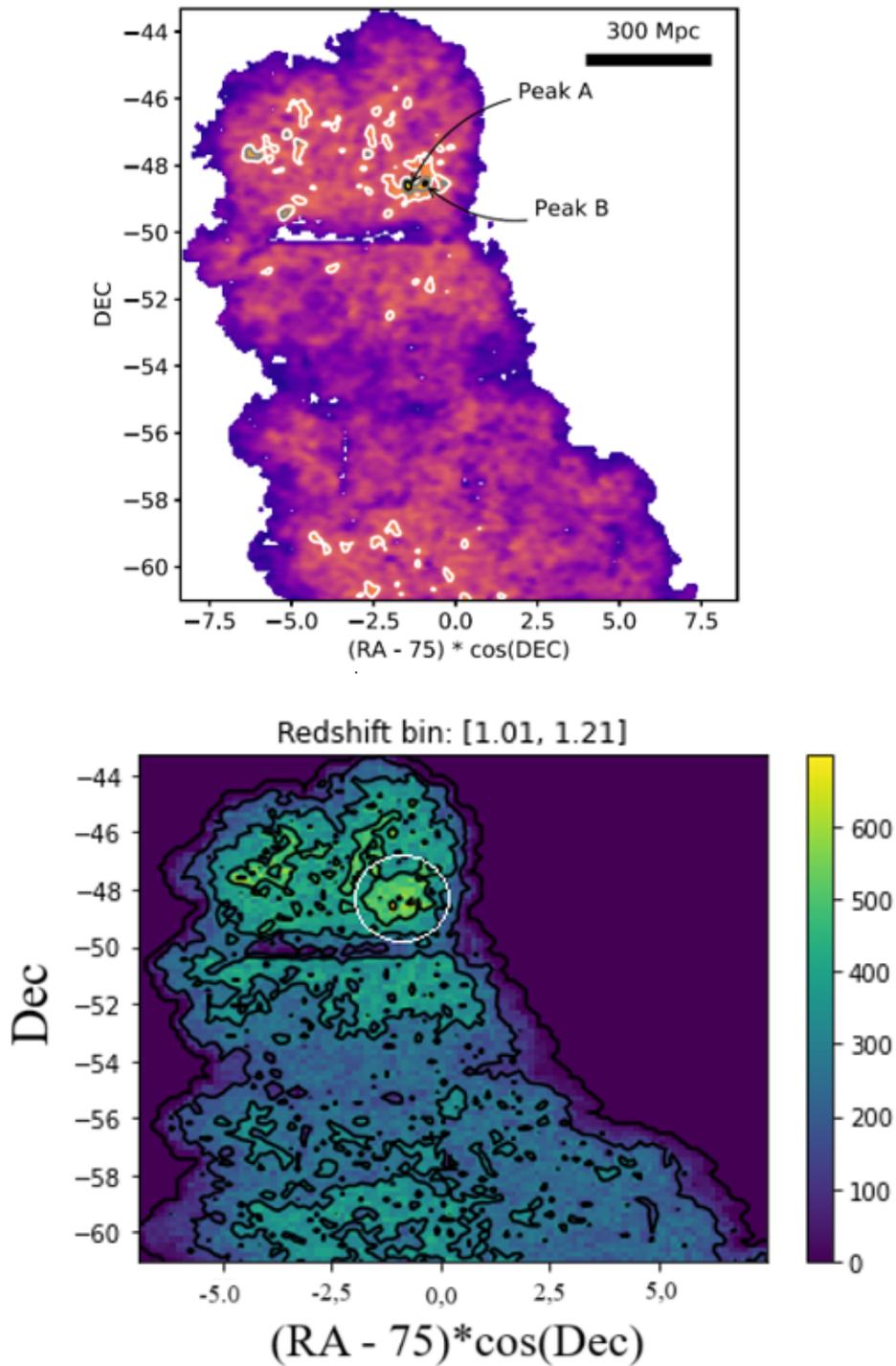


Figura 4.2. *Contour plot* do contraste de densidade exibindo as estruturas de sobredensidade no *bin* 1,01-1,21. O painel superior foi obtido no trabalho de Narwal *et al.* (2020) e o painel inferior foi obtido no presente trabalho, com a sobredensidade sendo destacada pelo círculo branco.

## 4.2 Determinação do tamanho do pixel

Nesta etapa fatiou-se a faixa analisada de *redshift* em intervalos de 0,03, de forma que houvesse interseção entre os *bins*. Por exemplo, *bins* consecutivos seriam: [0,5-0,6], [0,53-0,63], [0,56-0,66] e assim sucessivamente. Desta forma, um objeto poderia aparecer em *bins* diferentes, sendo essa técnica utilizada para compensar o erro associado ao *photo-z* (Sánchez *et al.*, 2017). A transição entre fatias de *redshift* da detecção de candidatos foi rastreada de maneira mais suave. Um erro maior que o *redshift* espectroscópico é apresentado pelo *photo-z*, então galáxias de um mesmo candidato podem ser distribuídas ao longo de *bins*, que, fisicamente, seriam impossíveis de serem habitados se as distâncias fossem calculadas diretamente do *photo-z*.

No processo de busca pelos candidatos, devido a interseção entre os *bins*, encontrou-se multiplicidade de objetos. Para resolver a possível multiplicidade do mesmo objeto presente em *bins* diferentes e contíguos, foi escrito um programa em Python para a seleção dos candidatos encontrados em diferentes *bins*. Desta forma, procurou-se o *match* ao longo dos *bins* contíguos de um candidato, traçou-se um perfil de contagem de membros e foi selecionado como *redshift* o *bin* com maior contagem como visto na Figura 3.2.

Fez-se um estudo do efeito do tamanho do pixel na busca por sobredensidades, tal que a quantidade de *match* com a literatura fosse a maior possível, gerando a Figura 4.3. A partir da análise da Figura 4.3, escolheu-se o tamanho do *pixel* como 2 Mpc, a partir da qual, observou-se uma queda na correspondência entre os catálogos à medida que o tamanho do pixel aumentava. É natural que as aglomerações tenham uma escala física de tamanho variável e maior que a típica de 1 Mpc para aglomerados no universo local. Além disso, cada catálogo da literatura possui uma função de seleção diferente, de modo que usá-los como tabela-verdade não necessariamente corresponde à verdade. Não obstante, é um exercício interessante para entender-se como o algoritmo funciona.

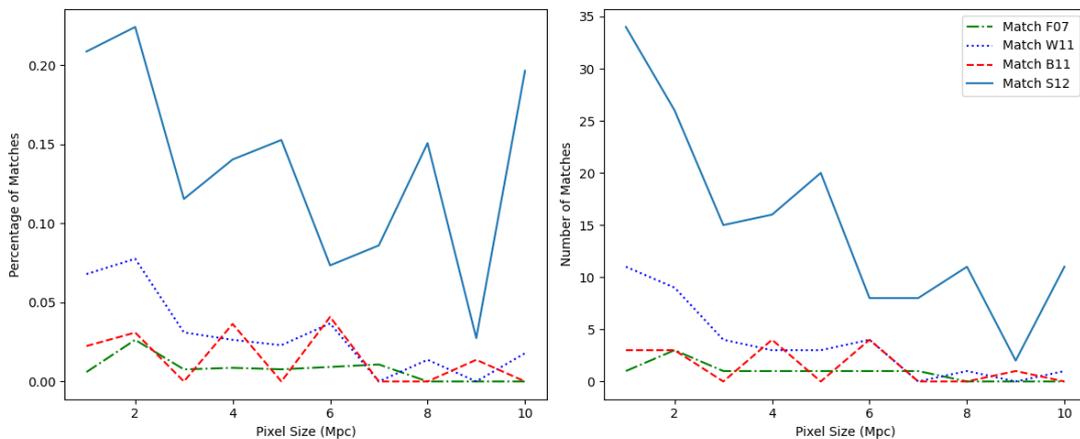


Figura 4.3. A direita, tem-se a porcentagem de *match* pelo tamanho do pixel referente ao histograma de sobredensidades. A esquerda, tem-se a relação entre número de *match* pelo tamanho do pixel.

Observou-se que na comparação do número de candidatos e de *matches* em função do número de *pixels*, uma maior quantidade de candidatos na mesma área de estudo dos levantamentos não significa maior taxa de *match*. Todas as informações a respeito do comportamento da quantidade de *match* com respeito ao tamanho do pixel são apresentadas na Tabela 4.1. A Tabela 4.1 apresenta para cada valor de pixel, o número de *match* alcançado considerando a quantidade de candidatos encontrados na área de estudo e a sua, respectiva, porcentagem para cada um dos trabalhos analisados. O catálogo gerado pela busca de sobredensidades com pixel de tamanho igual a 2 Mpc foi denominado de G24 sendo composto de um total de 176 candidatos encontrados na região do COSMOS na faixa de *redshift* de 0,5 - 2,0.

Pixel (Mpc)	F07			W11			B11			S12		
	# Match	# Cand. <sup>a</sup>	(%)	# Match	# Cand. <sup>a</sup>	(%)	# Match	# Cand. <sup>a</sup>	(%)	# Match	# Cand. <sup>a</sup>	(%)
1	1	168	1.0	11	162	7.0	3	134	2.0	34	163	21.0
2	3	114	3.0	9	116	8.0	3	97	3.0	26	116	22.4
3	1	131	1.0	4	129	3.0	0	116	0.0	15	130	12.0
4	1	116	1.0	3	114	3.0	4	110	4.0	16	114	14.0
5	1	130	1.0	3	131	2.0	0	94	0.0	20	131	15.0
6	1	109	1.0	4	109	4.0	4	98	4.0	8	109	7.0
7	1	93	1.0	0	93	0.0	0	84	0.0	8	93	9.0
8	0	73	0.0	1	73	1.0	0	64	0.0	11	73	15.0
9	0	73	0.0	0	73	0.0	1	73	1.0	2	73	3.0
10	0	56	0.0	1	56	2.0	0	56	0.0	11	56	20.0

Tabela 4.1: Efeito do tamanho do pixel variando de 1 a 10 Mpc, no número de *match* de candidatos considerando o domínio de cada um dos trabalhos referenciados. F07 (Finoguenov *et al.*, 2007), W11 (Wen e Han, 2011), B11 (Bellagamba *et al.*, 2011)

A Figura 4.4 apresenta painéis variando o tamanho do pixel de 1 a 10. Pode-se observar o tamanho do pixel em função da área estudada. A partir da análise visual da figura, percebe-se que a partir do valor de 4 Mpc há uma perda de definição do domínio estudado e a importância de ter áreas grandes para estudar a estrutura em grande escala no universo local.

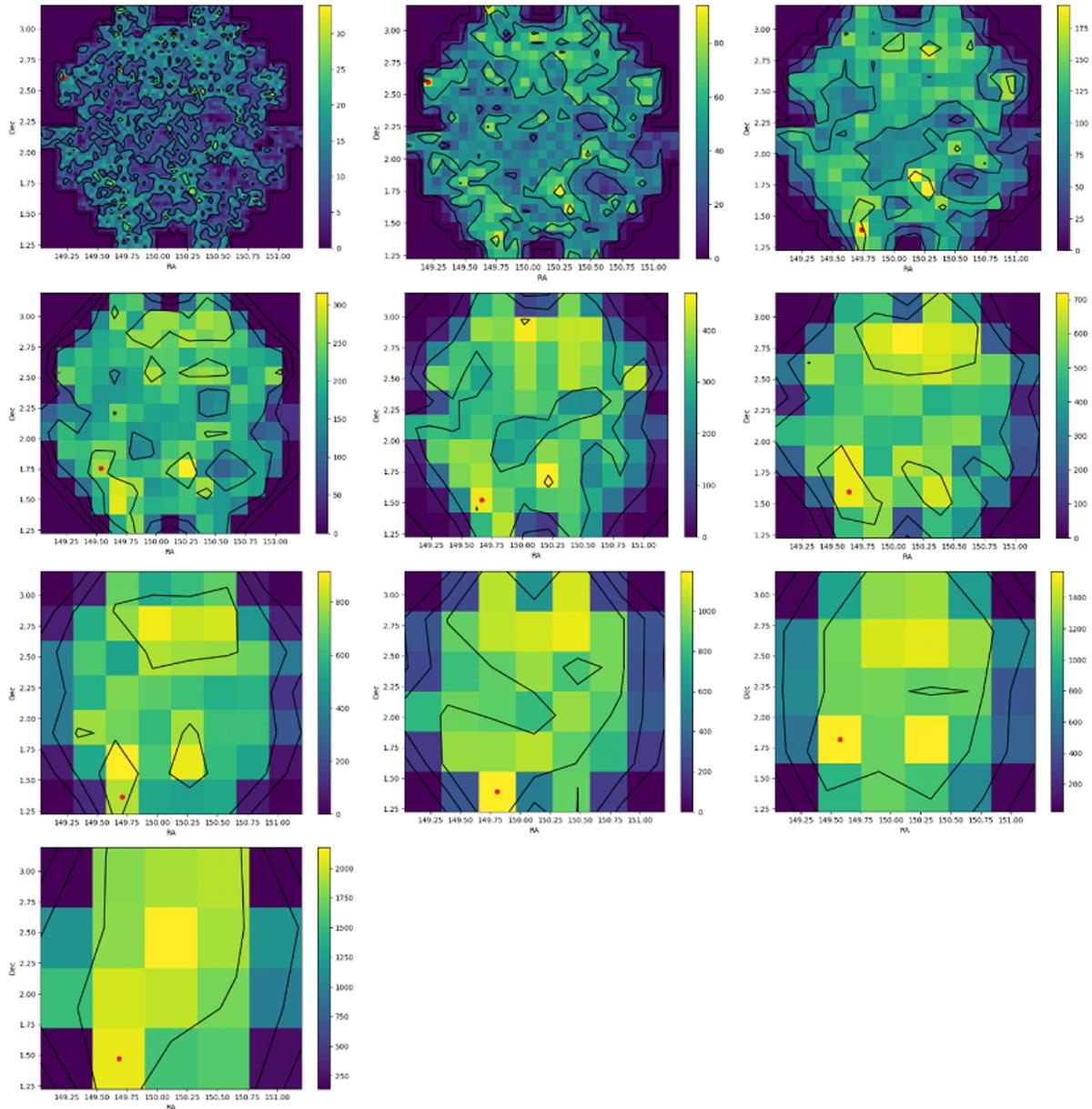


Figura 4.4. Painéis variando o tamanho do pixel de 1 a 10, da esquerda para a direita contendo o primeiro *bin* de *redshift*. Primeira linha contém painéis de pixel 1, 2 e 3, segunda linha pixels 4, 5 e 6, terceira linha pixels 7, 8 e 9 e quarta linha pixel 10.

Na sequência, fez-se uma análise de auto-consistência e estudou-se o efeito do tamanho do pixel com relação aos nossos próprios candidatos de G24. A Figura 4.5 apresenta a porcentagem de *matches* (número de *match* com o catálogo G24 dividido pela quantidade de detecções do catálogo comparado) com o catálogo G24 em função do tamanho do pixel

em Mpc.

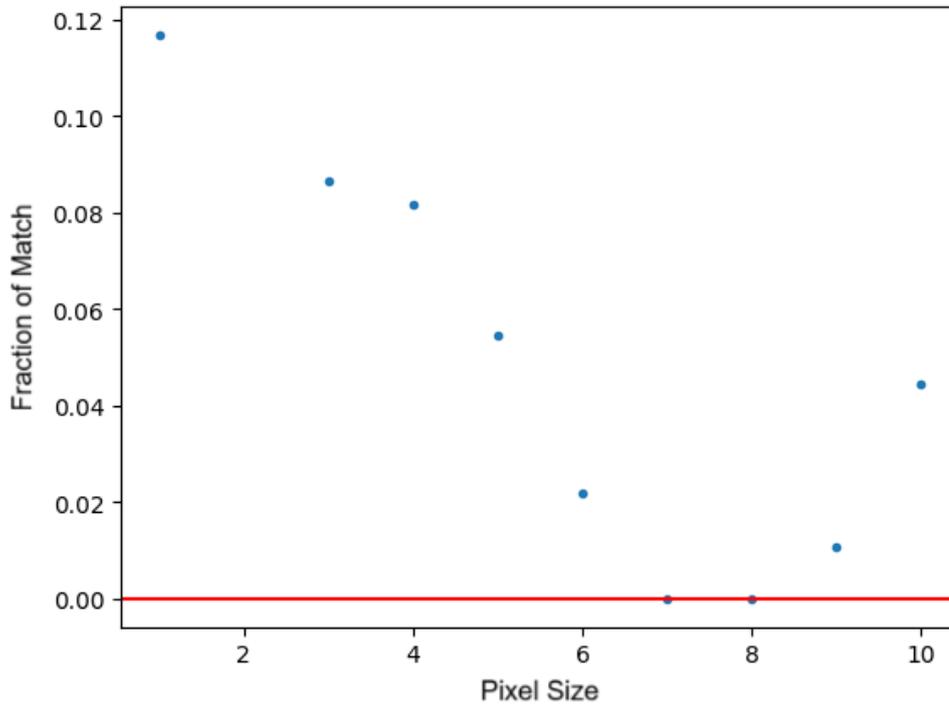


Figura 4.5. Fração de *matches* com o catálogo G24 em função do tamanho do pixel em Mpc, exceto para o pixel de tamanho igual a 2 Mpc por ter 100% de *match* por ser o catálogo G24.

As Figuras 4.6, 4.7 e 4.8 apresentam os *matches* entre o catálogo G24 e os catálogos gerados com tamanho de pixel 1 Mpc (1G24), 2 Mpc (G24) e 3 Mpc (3G24), respectivamente. Em um desenvolvimento orientado a testes é importante que o G24 alcance correspondência total com ele mesmo.

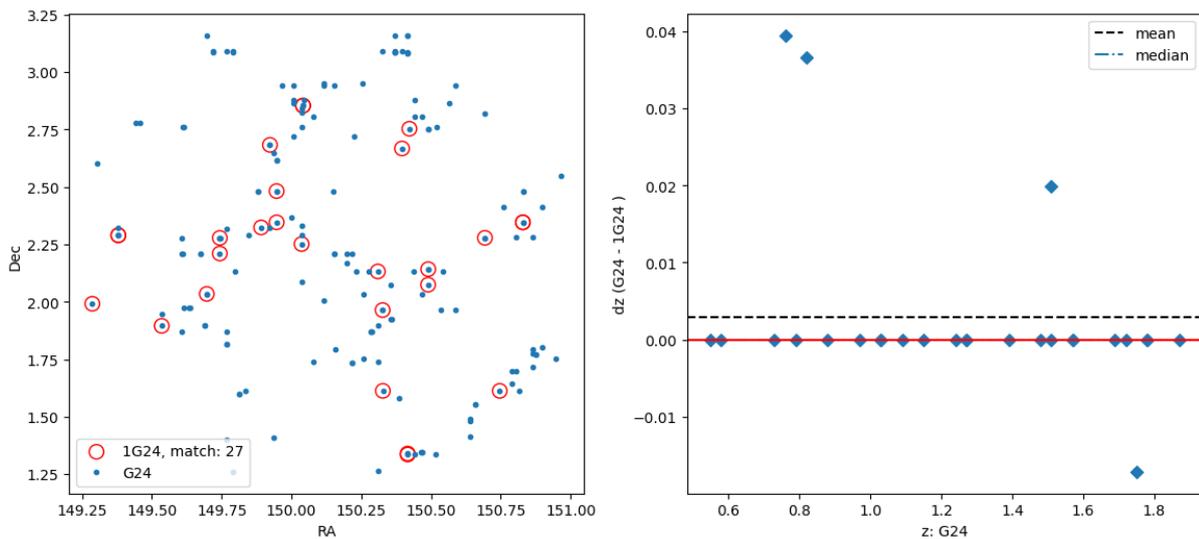


Figura 4.6. *Match* entre o catálogo G24 e um catálogo gerado com pixel igual a 1 Mpc (1G24).

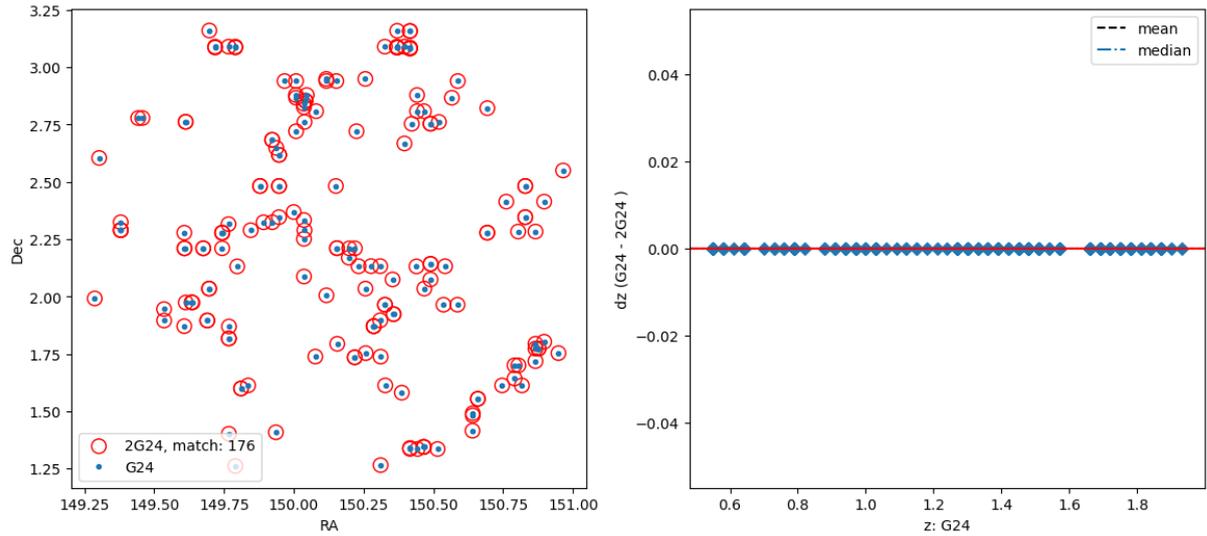


Figura 4.7. *Match* entre o catálogo G24 e um catálogo gerado com pixel igual a 2 Mpc (G24).

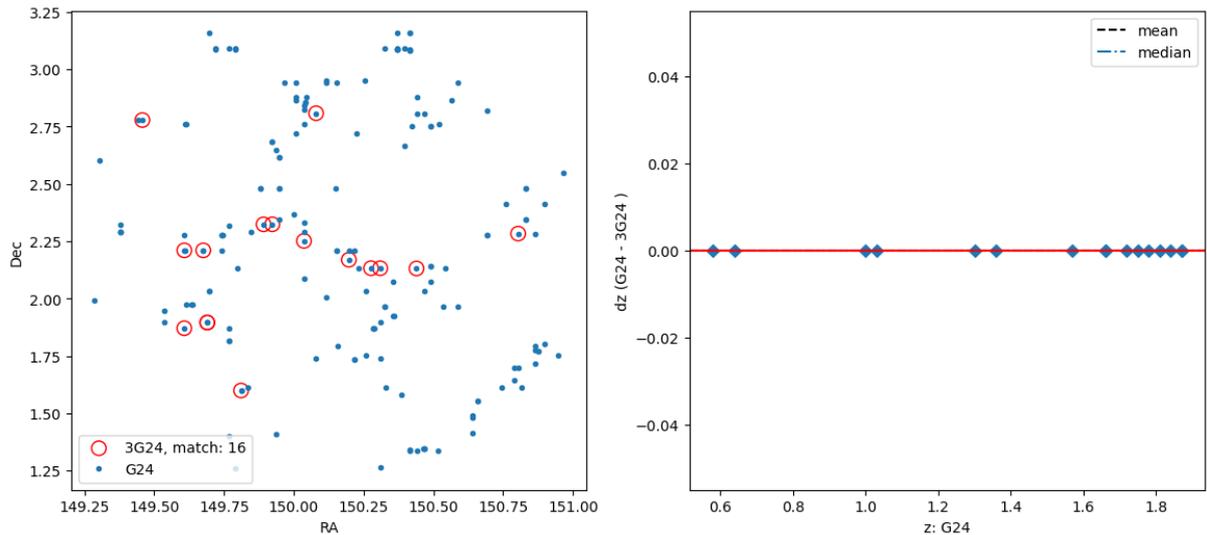


Figura 4.8. *Match* entre o catálogo G24 e um catálogo gerado com pixel igual a 3 Mpc (3G24).

### 4.3 Resultados com *match* 2D com a literatura

A validação dos resultados obtidos neste estudo foi feita comparando com quatro catálogos conhecidos da literatura (Finoguenov *et al.* 2007, Wen e Han 2011, Bellagamba *et al.* 2011, Sochting *et al.* 2012).

A Figura 4.9 apresenta dados típicos de *contour plot* de densidade que foram produzidos no presente trabalho. O painel da esquerda representa o *contour plot* de densidade do primeiro *bin* (0,5-0,6) e o painel da direita representa o *contour plot* de densidade do último *bin* (1,9 - 2,0).

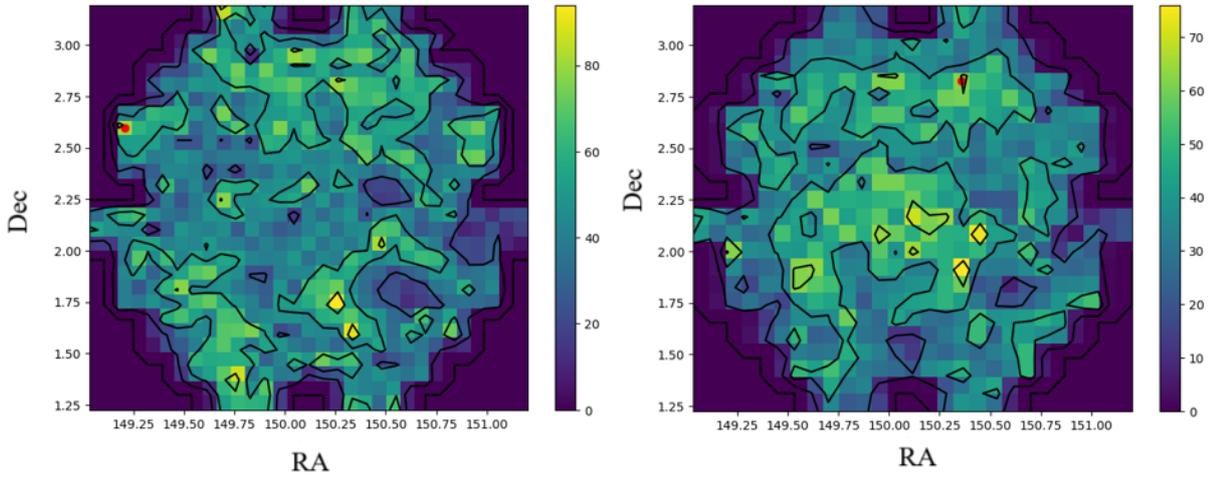


Figura 4.9. *Contour plot* do contraste de densidade exibindo as estruturas de sobredensidade no primeiro (esquerda) e no último (direita) *bin* de *redshift*.

Como dito anteriormente, a busca por sobredensidades de galáxias gerou um catálogo que foi denominado de G24. Nesta etapa do trabalho, o G24 foi comparado com outros catálogos da literatura. Esses catálogos incluem os trabalhos de Finoguenov *et al.* (2007) (F07), Wen e Han (2011) (W11), Bellagamba *et al.* (2011) (B11) e Sochting *et al.* (2012) (S12). Como apresentado na Tabela 2.4, cada catálogo possui uma área e uma faixa de *redshift*. As áreas e faixas de *redshift* de comparação foram restringidas aos domínios de cada estudo, por exemplo, na comparação do G24 com F07, a área foi restringida a  $1,24 \text{ graus}^2$  (área de G24), essa restrição da área foi feita a partir da comparação dos valores mínimos e máximos de RA e Dec de cada aglomerado comparado. E na restrição da faixa de *redshift* limitou-se de acordo com os valores máximos e mínimos de *redshift*, no caso da comparação do G24 com F07 a faixa analisada foi de  $[0,5-1,25]$  ( $z = 0,5$  referente a G24 e  $z = 1,25$  referente a F07). Os códigos foram escritos em Python visando fazer o *match* 2D entre cada catálogo.

A Tabela 4.2 apresenta uma comparação entre as quantidades de *match* para cada um dos catálogos. Nesta tabela, são apresentadas as quantidades de aglomerados de cada catálogo, e o *match* entre eles, sendo que o valor entre parênteses refere-se a quantidade de aglomerados no domínio comparado.

As Figuras 4.10, 4.11, 4.12 e 4.13 apresentam, respectivamente, os resultados do *match* 2D entre este trabalho e os catálogos produzidos por Sochting *et al.* (2012), Finoguenov *et al.* (2007), Wen e Han (2011) e Bellagamba *et al.* (2011).

Estudo	Qtd. de <i>Clusters</i>	F07	W11	B11	S12
F07	72	-	26 (70)*	22 (63)*	150 (70)*
W11	187	26 (184)*	-	47 (144)*	413 (184)*
B11	119	22 (63)*	47	-	272
S12	1780	150 (1760)*	413 (1776)*	272 (1449)*	-
G24	105	13 (100)*	36 (99)*	14 (87)*	337 (100)*

Tabela 4.2: Resultado do *match* entre os catálogos da literatura e este trabalho. Os valores com ‘\*’ referem-se a quantidade de *clusters* no domínio comparado. G24 refere-se ao presente trabalho.

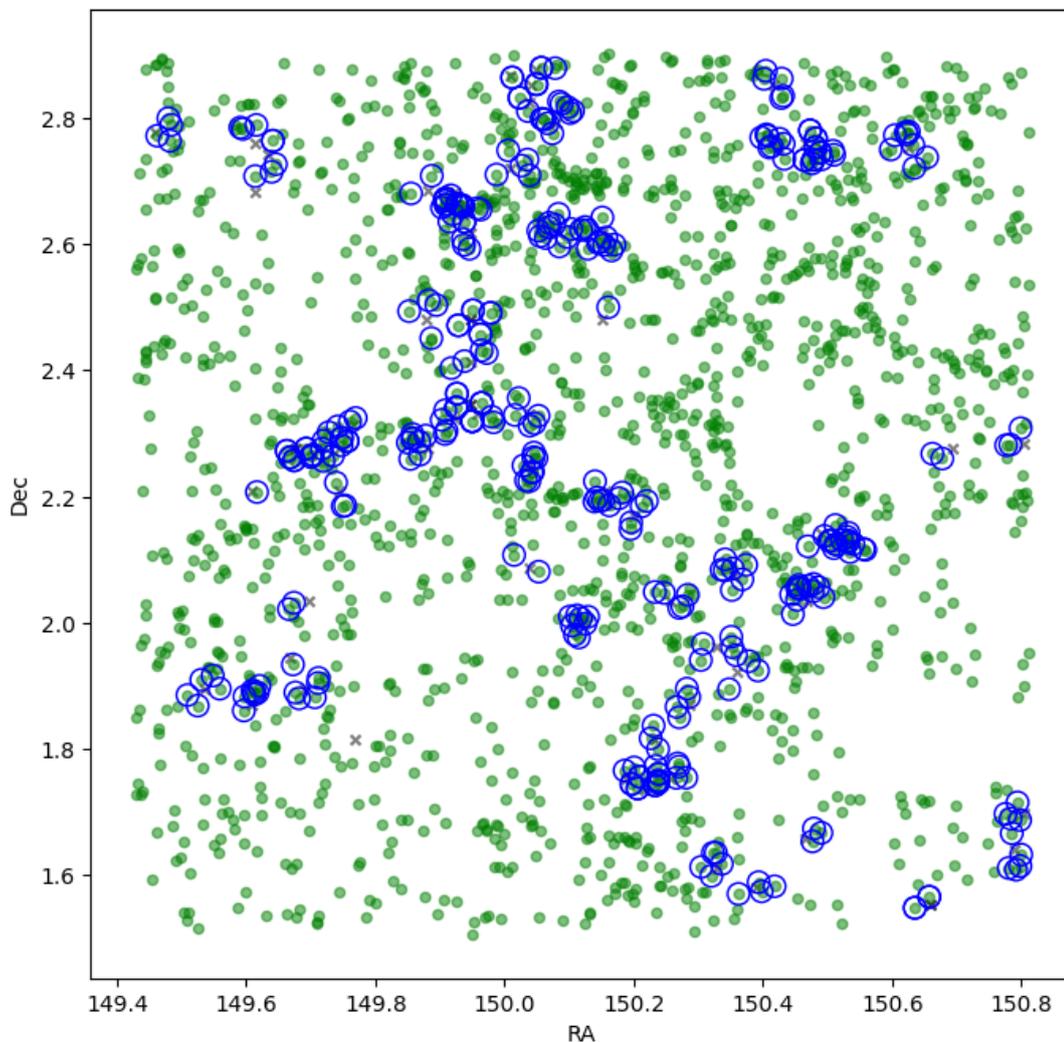


Figura 4.10. *Match* 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de *clusters* de Sochting *et al.* (2012) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis.

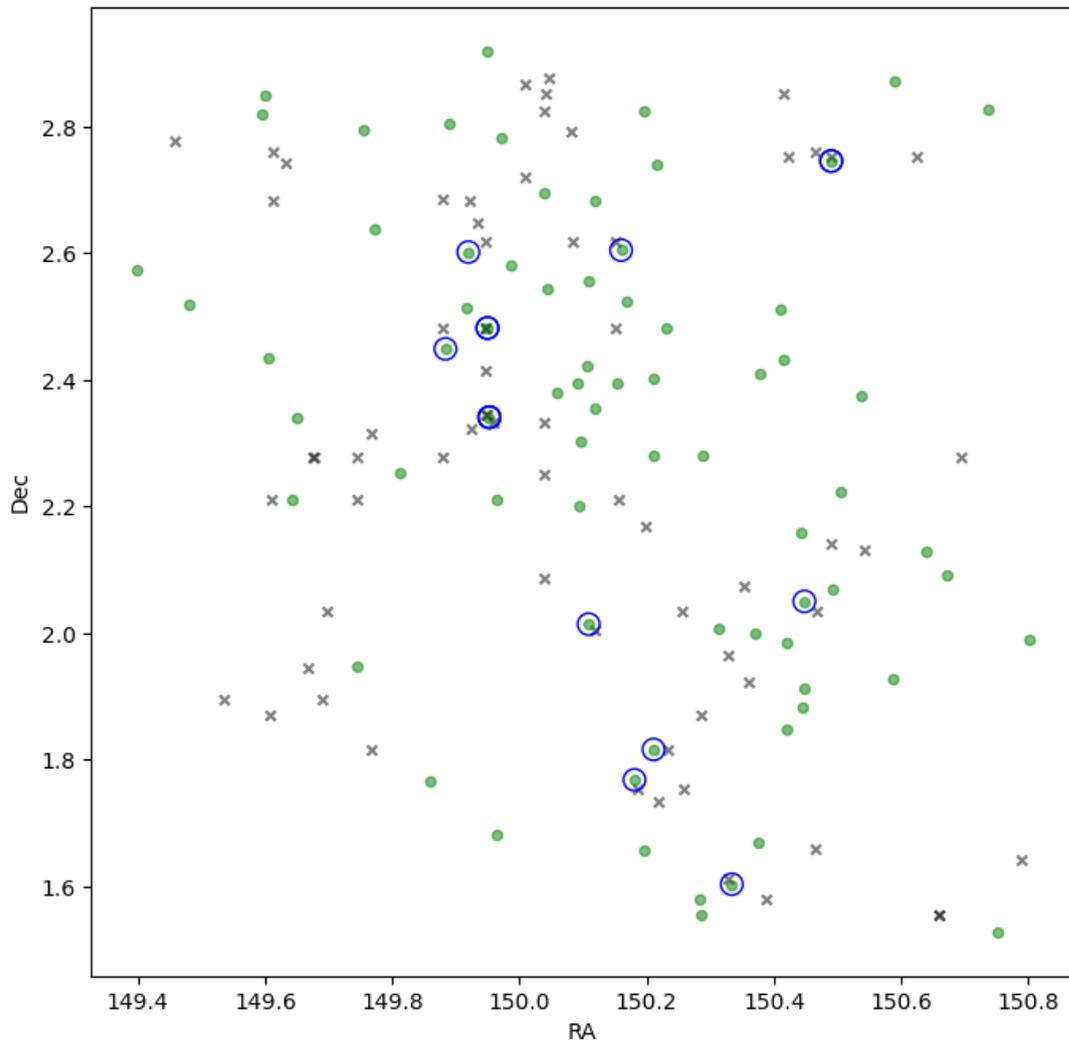


Figura 4.11. *Match 2D* entre os candidatos deste trabalho (cruz cinza) com o catálogo de *clusters* de Finoguenov *et al.* (2007) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis.

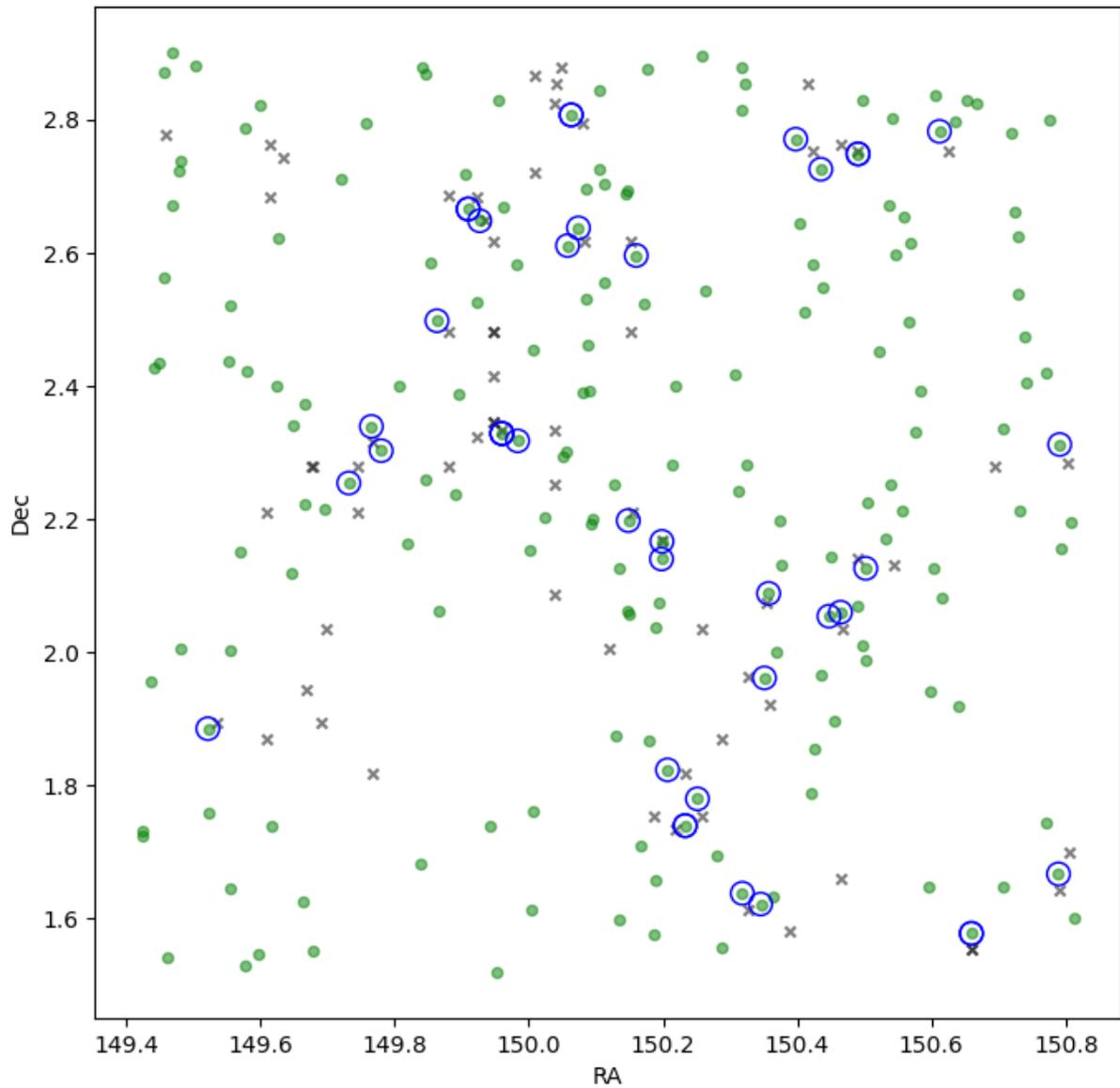


Figura 4.12. *Match 2D* entre os candidatos deste trabalho (cruz cinza) com o catálogo de *clusters* de Wen e Han (2011) (pontos verdes). Os candidatos em comum estão representados pelos círculos azuis.

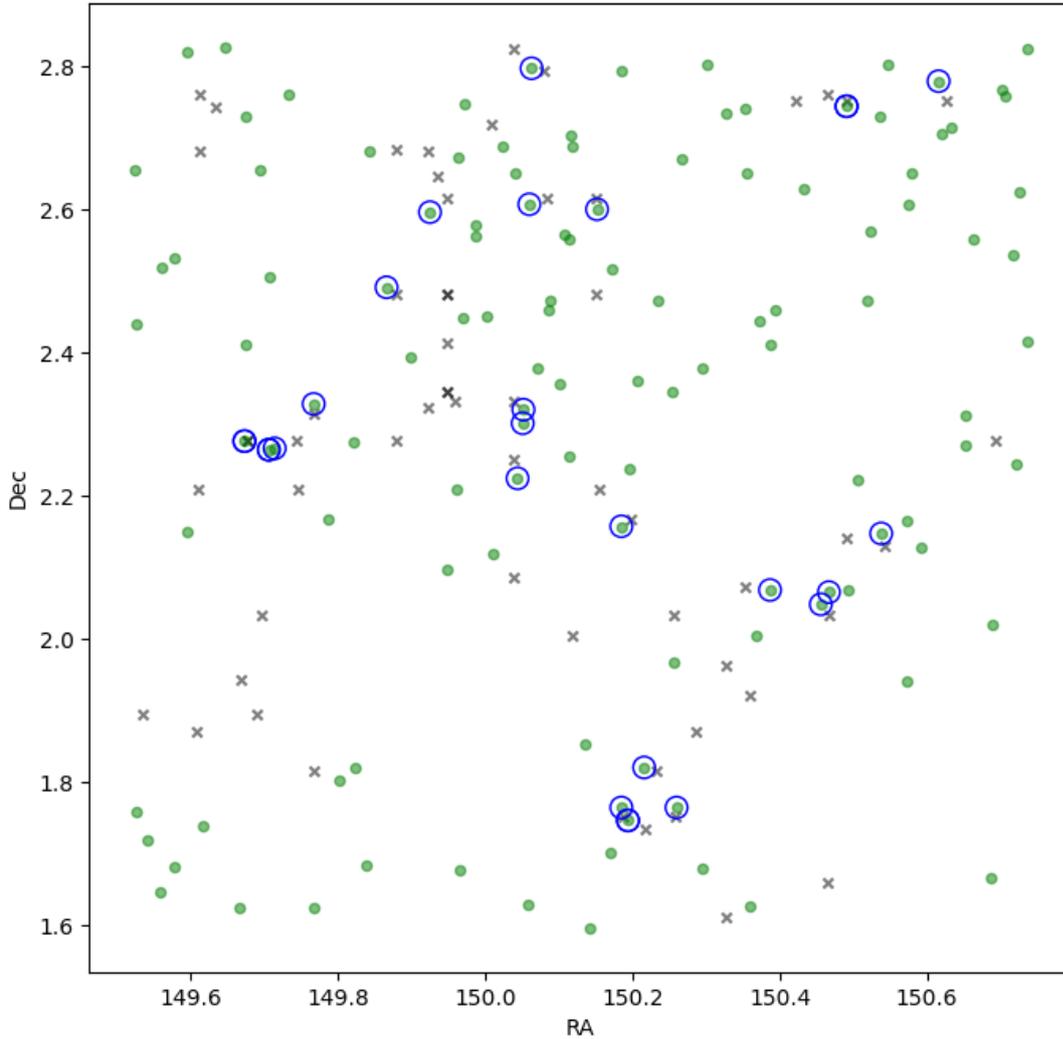


Figura 4.13. *Match* 2D entre os candidatos deste trabalho (cruz cinza) com o catálogo de *clusters* de Bellagamba *et al.* (2011) (pontos verdes) Os candidatos em comum estão representados pelos círculos azuis.

Da forma que foi feito o *match*, o pareamento entre os catálogos levou a falsos positivos, com multiplicidade de *match* para todos os catálogos comparados. As galáxias em diferentes distâncias podem superpor-se na mesma linha de visada, criando uma ilusão para o aspecto 2D da análise, levando a uma superestimativa do número de correspondências entre catálogos. Esse efeito já era esperado para este estudo 2D devido ao efeito de projeção.

Logo, para avaliar o efeito destas superposições, foram calculados os resíduos do *redshift* entre catálogos ( $dz$ ) em função do *redshift* dos candidatos. Nesta análise,  $dz$  significa o *redshift* do candidato deste trabalho subtraído do *redshift* do seu respectivo *match* com um dado catalogo da literatura. Como resultados, as Figuras 4.14, 4.15, 4.16 e 4.17 referem-se, respectivamente, aos resultados do *match* com os dados de Wen e Han (2011), Bellagamba *et al.* (2011), Sochting *et al.* (2012) e Finoguenov *et al.* (2007). Nestas figuras, os resíduos  $dz$  indicam que os objetos estão em *redshifts* bem diferentes apesar de

apresentarem RA e Dec próximos.

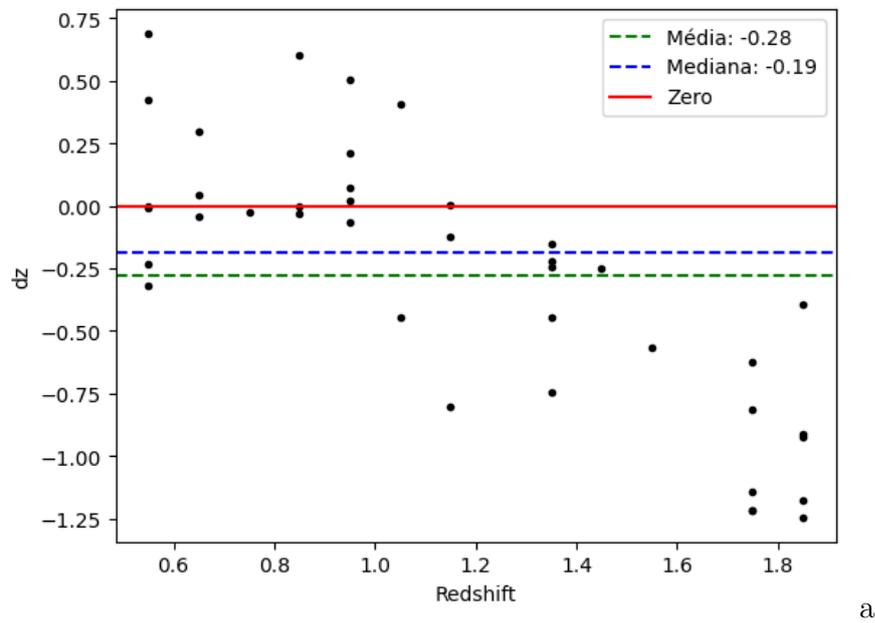


Figura 4.14. Valores de  $dz$  ( $z_{W11} - z_{G24}$ ) versus *redshift* (Wen e Han, 2011).

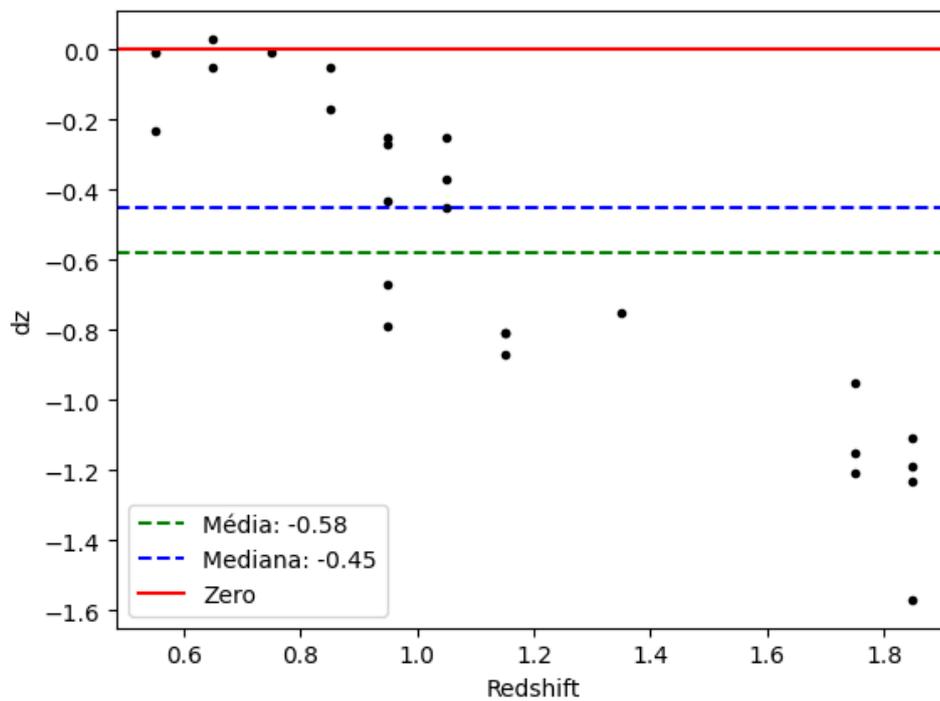


Figura 4.15. Valores de  $dz$  ( $z_{B11} - z_{G24}$ ) versus *redshift* (Bellagamba *et al.*, 2011).

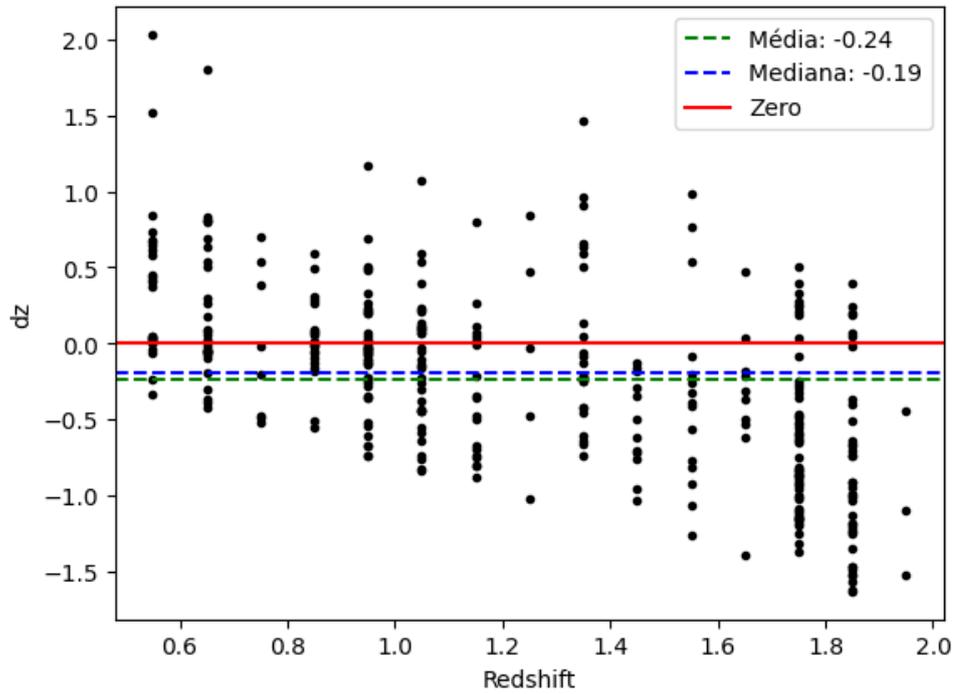


Figura 4.16. Valores de  $dz(z_{S12} - z_{G24})$  versus *redshifts* (Sochting *et al.*, 2012).

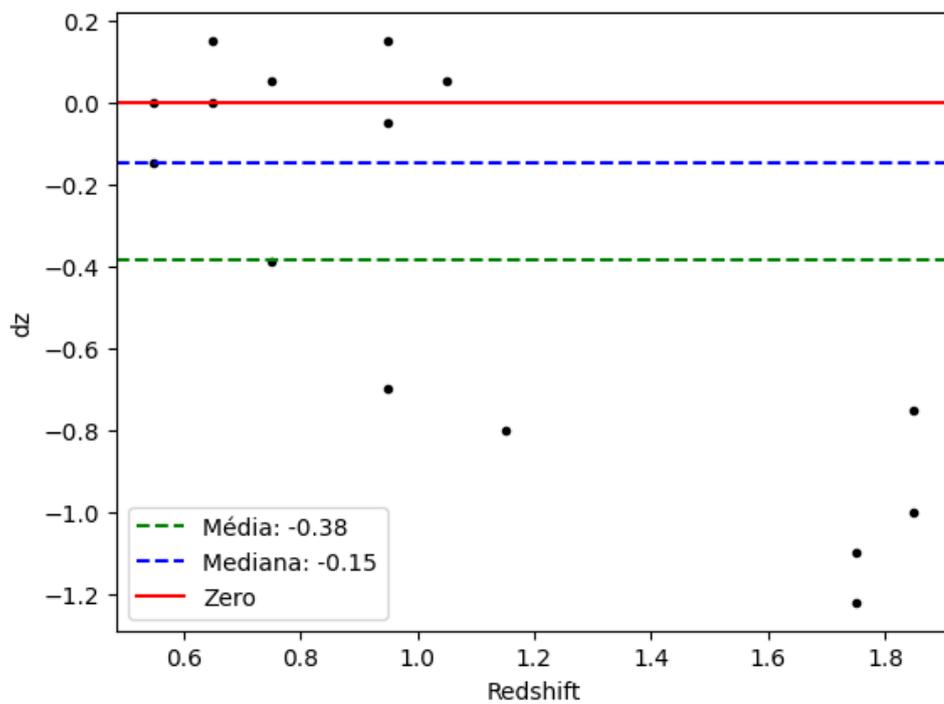


Figura 4.17. Valores de  $dz(z_{F07} - z_{G24})$  versus *redshifts* (Finoguenov *et al.*, 2007).

## 4.4 Resultados do *match* 3D com a literatura

Em uma fase inicial, aplicou-se uma abordagem mais simples com passo de 0,1 sem a interseção entre os *bins* de *redshift*, como por exemplo, uma sequência de *bins* seria: [0,5-0,6], [0,6-0,7], [0,7-0,8] e assim por diante, com passo de 0,1. Os resultados para o caso 3D do passo de 0,1 encontram-se no apêndice.

Na sequência, aplicou-se uma abordagem mais precisa com passo de 0,03 com interseção entre os *bins* de *redshift*, como por exemplo, uma sequência de *bins* seria: [0,5-0,6], [0,53-0,63], [0,56-0,66] e [0,59-0,69] e assim por diante. Obtendo o catálogo previamente denominado G24 na Seção 4.2.

Tendo obtido o catálogo de candidatos a protoaglomerados, na sequência, fez-se o *match* deste catálogo com os catálogos da literatura. Como resultados, as Figuras 4.18, 4.19, 4.20 e 4.21 apresentam os *matches* e os resultados das análises de resíduos de *redshift* entre o G24 com os trabalhos de W11, S12, B11 e F07, respectivamente. Os resíduos dz de baixo valor indicam que há possibilidade de serem as mesmas detecções de objetos encontrados nos catálogos comparados. Nestas figuras, o painel da esquerda refere-se ao *match* e da direita dz em função de z.

Após a comparação com dados da literatura, obteve-se as seguintes quantidades de *matches*: 3 com F07, 9 com W11, 4 com B11 e 26 com S12. Chegou-se assim a um resultado de 26 *matches* totais, resultando em 150 candidatos inéditos para o catálogo G24. Vale ressaltar que um *match* com um catálogo não significa que seja o mesmo para os demais catálogos. Foram obtidos apenas 2 candidatos universais, ou seja, que estiveram presentes em todas as comparações, demonstrando a dificuldade de se obter um catálogo definitivo de aglomerações em uma pequena região do céu.

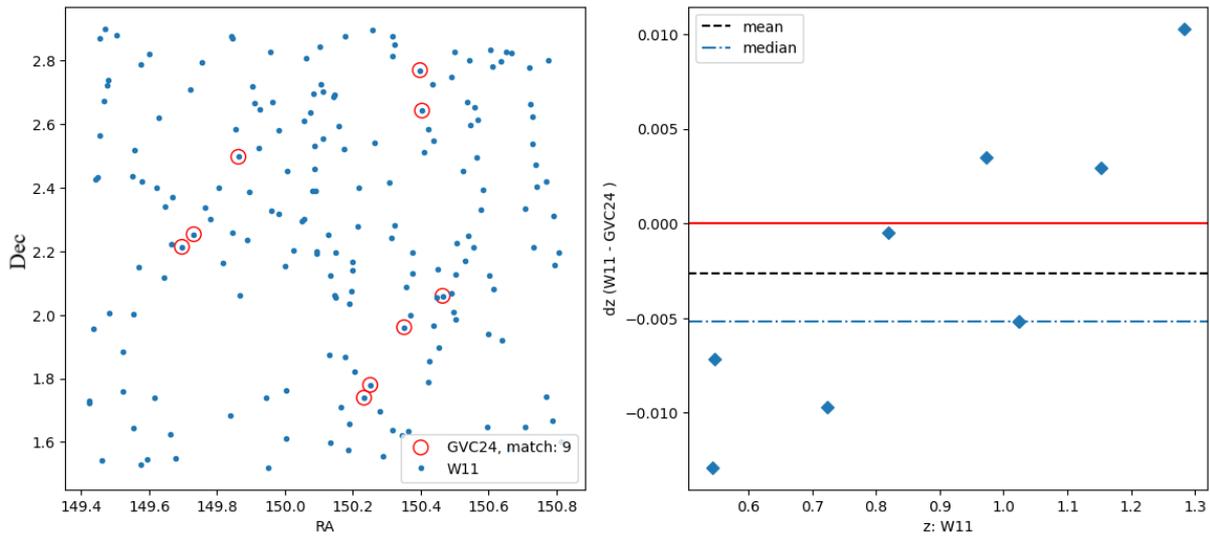


Figura 4.18. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Wen e Han (2011) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os aglomerados de Wen e Han (2011), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

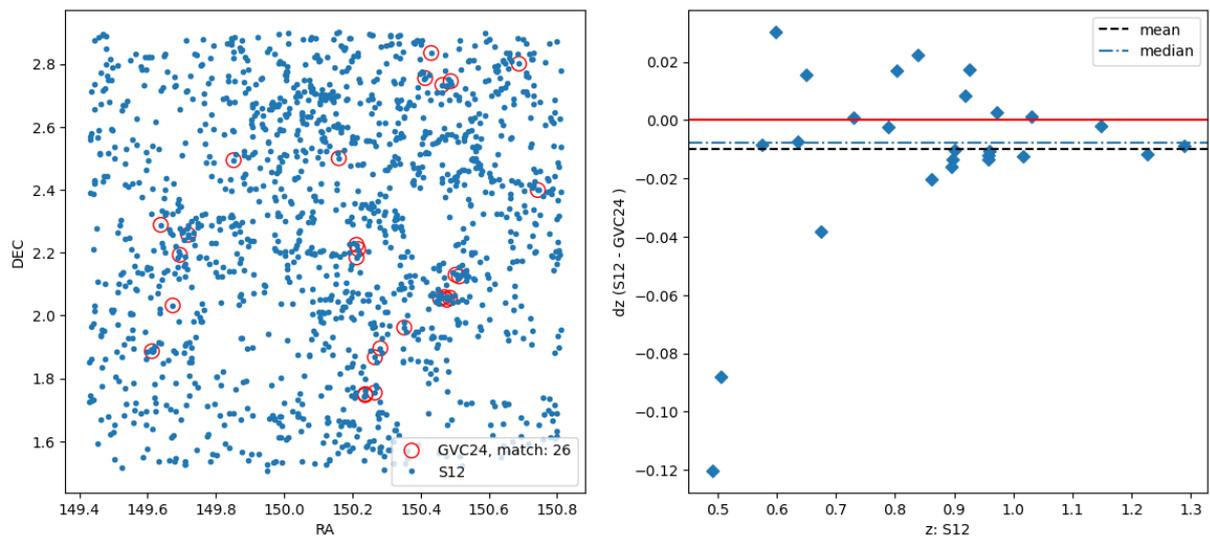


Figura 4.19. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Sochting *et al.* (2012) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os aglomerados de Sochting *et al.* (2012), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

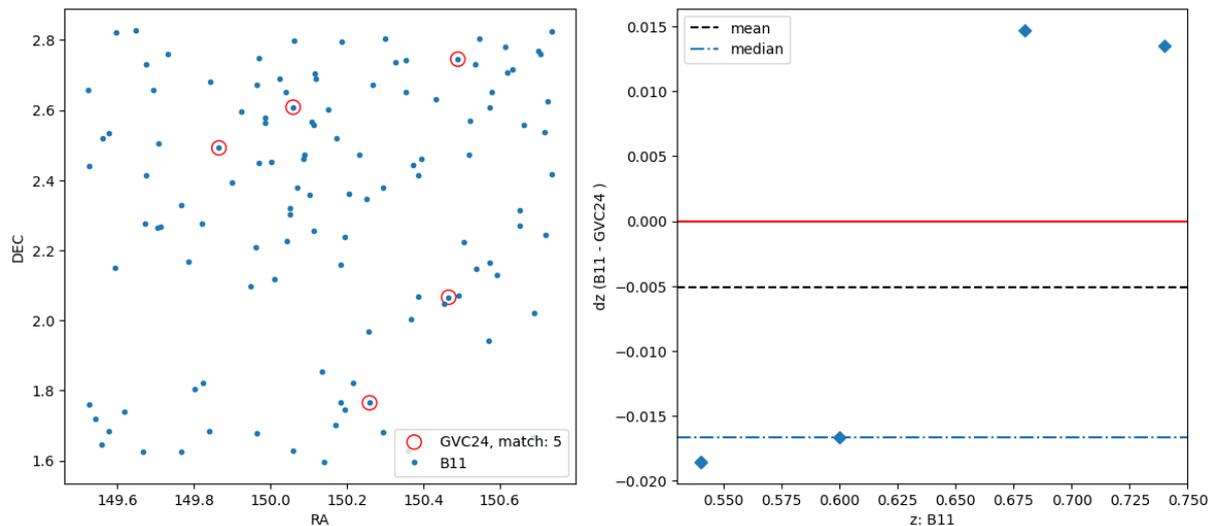


Figura 4.20. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Bellagamba *et al.* (2011) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os aglomerados de Bellagamba *et al.* (2011), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

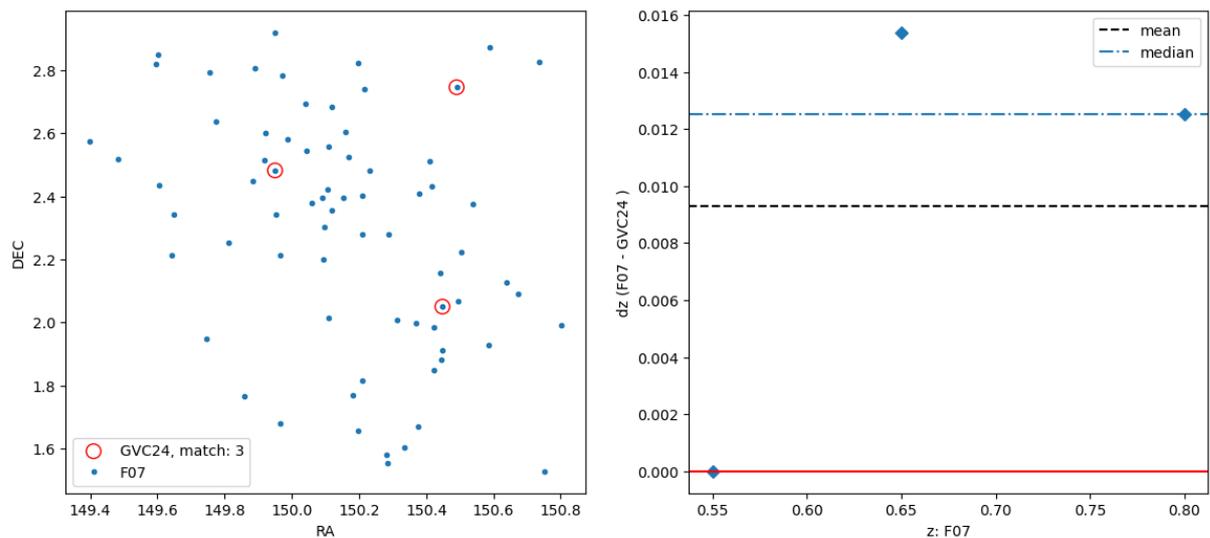


Figura 4.21. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Finoguenov *et al.* (2007) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os aglomerados de Finoguenov *et al.* (2007), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

## 4.5 Validação com catálogos *bootstrap*

Objetivando validar *matches* obtidos com o catálogo S12 e determinar se não foram encontrados de forma aleatória, utilizou-se a rotina *bootstrap* para montar 500 catálogos

*mock* referentes ao G24 e fez-se o *match* entre eles e o catálogo de Sochting *et al.* (2012). Sochting *et al.* (2012) foi escolhido por conter o maior número de aglomerados dentre os trabalhos selecionados para a comparação. Os catálogos foram gerados a partir da amostra de candidatos do presente trabalho, utilizando *bootstrap* (Efron, 1979) para tornar aleatória a sequência de RA e Dec. O histograma presente na Figura 4.22 permite a análise da distribuição de *matches* gerados, sendo assim, pode-se perceber que a maior concentração *match* está em torno de 1 e 2. Os resultados apresentaram mediana e média, respectivamente 2 e 2. Esses resultados confirmam que os 26 *matches* provenientes do catálogo G24 não são aleatórios tendo em vista a grande quantidade de membros do catálogo de Sochting *et al.* (2012). A Figura 4.23 apresenta um resultado típico com 9 *matches*. Na sequência, fez-se a comparação desses catálogos *mock* com os outros catálogos da literatura, porém a quantidade de *match* não passou de 1.

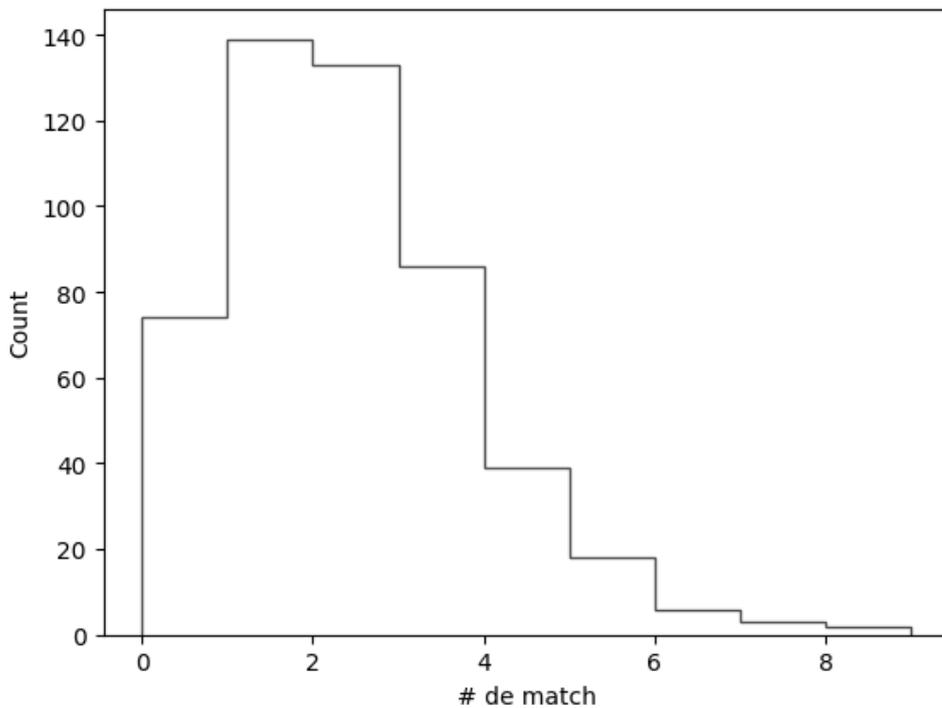


Figura 4.22. Histograma da comparação entre os 500 catálogos *mock* gerados com o catálogo de Sochting *et al.* (2012) (S12).

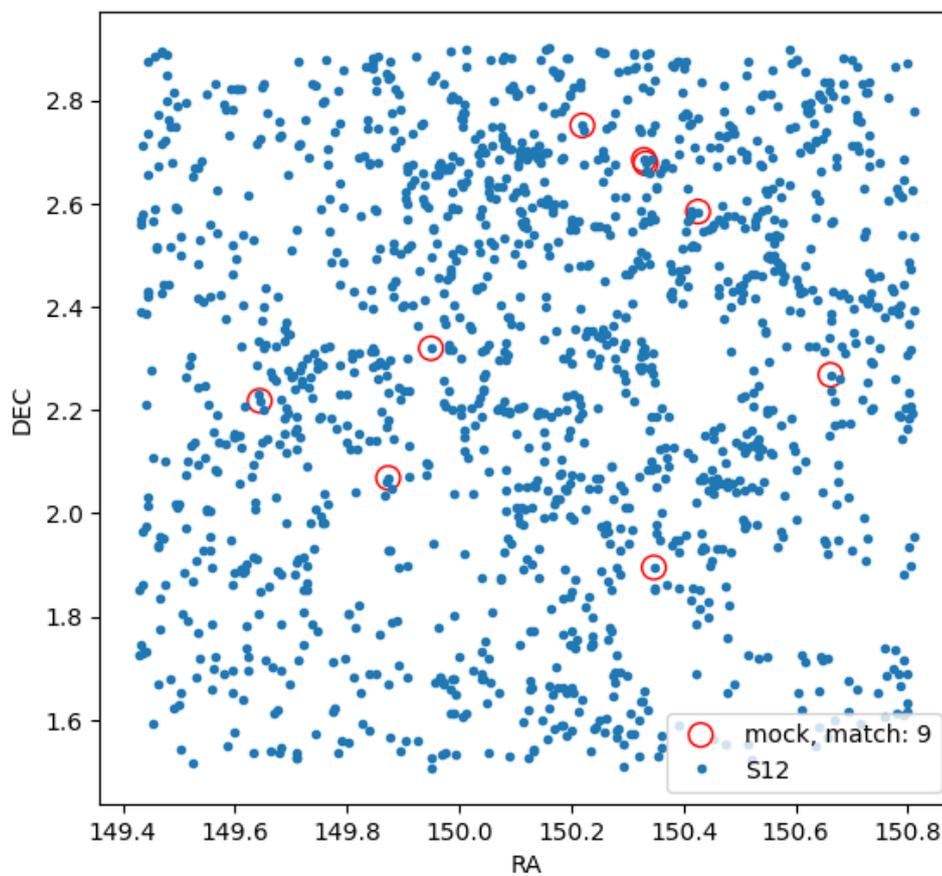


Figura 4.23. Resultados da comparação entre o catálogo *mock* com o catálogo de Sochting *et al.* (2012) (S12). O círculo vermelho representa os *matches* e os pontos azuis são dados de S12



# Capítulo 5

## Resultados e análises dos candidatos

### 5.1 Análise dos candidatos inéditos

Para todos os 150 candidatos inéditos do G24, com o objetivo de caracteriza-los, foram feitas análises como: *k-correction*, gráficos de *contour*, estudada a riqueza de cada candidato, estudada a distribuição da razão entre galáxias vermelhas e azuis pela magnitude  $i$ , foi feito um estudo de CMD dos candidatos com e sem *k-correction*, histogramas de distribuição de *redshift* e estudo da morfologia. O Apêndice B apresenta painéis dos 150 candidatos apresentados no presente trabalho.

O estudo da população das galáxias membro dos candidatos inéditos foi feito utilizando uma rotina de GMM (Pedregosa *et al.*, 2011) e a correção das magnitudes foi feita utilizando uma rotina de *k-Correction* (Blanton e Roweis, 2007).

Com relação ao *k-correction*, pode-se observar o seu efeito a partir da análise da diferença entre as Figuras 5.1 e 5.2 que apresentam o GMM de um candidato sem e com *k-Correction*, respectivamente.

A Figura 5.3 apresenta 4 casos típicos de distribuição bimodal das populações de galáxias para os IDs 23, 38, 39 e 42. A análise dos dados obtidos neste trabalho, por exemplo, ver Figura 5.3, permite concluir que a distribuição de cores em *rest-frame* dessas galáxias é bimodal majoritariamente abaixo de  $z < 1$  (Bell *et al.*, 2004). Essa bimodalidade permite uma definição independente de modelo de galáxias vermelhas do tipo *early-type* e galáxias azuis do tipo *late-type* em qualquer *redshift* dado.

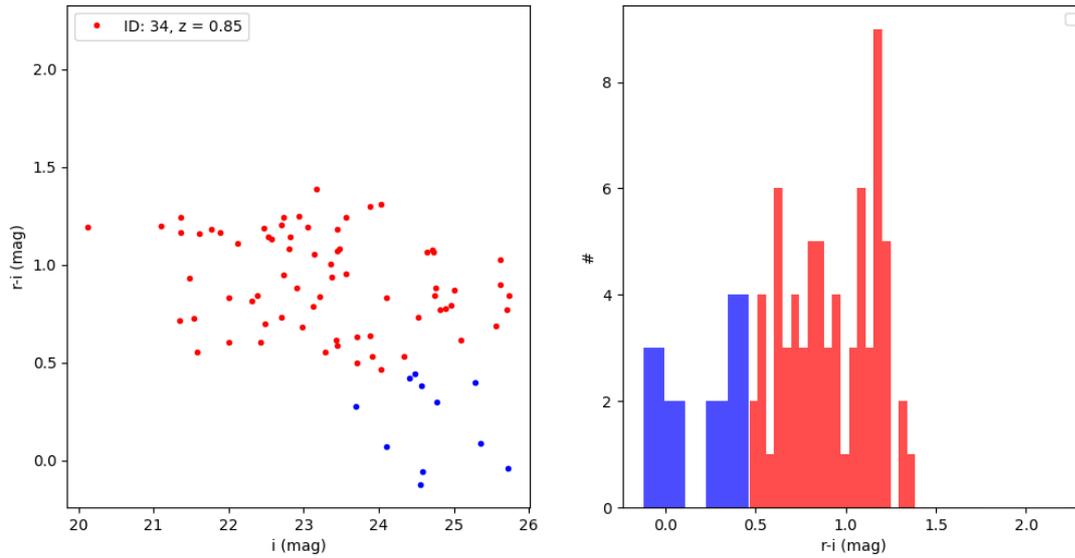


Figura 5.1. O painel da esquerda apresenta dados de  $r-i$  (mag) em função de  $i$  (mag) e o painel da direita a distribuição de  $r-i$  (mag) para o caso em que não foi aplicada a  $k$ -correction. Os dados são referentes ao candidato de ID 34 e *redshift* 0,85 do catálogo proposto neste trabalho.

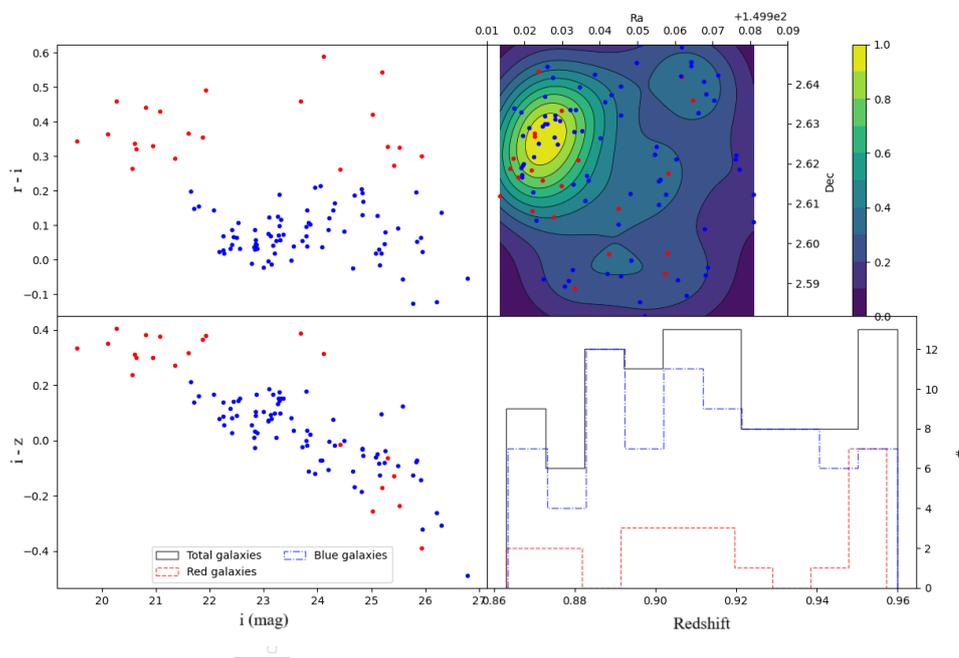


Figura 5.2. Os painéis superior e inferior da esquerda apresentam dados de  $r-i$  (mag) e  $i-z$  (mag) em função de  $i$  (mag), respectivamente. Os painéis superior e inferior da direita apresentam o *countour plot* do candidato e o histograma de *redshift* das galáxias membros. Os dados são referentes ao candidato de ID 34 e *redshift* 0,85 do catálogo proposto neste trabalho.

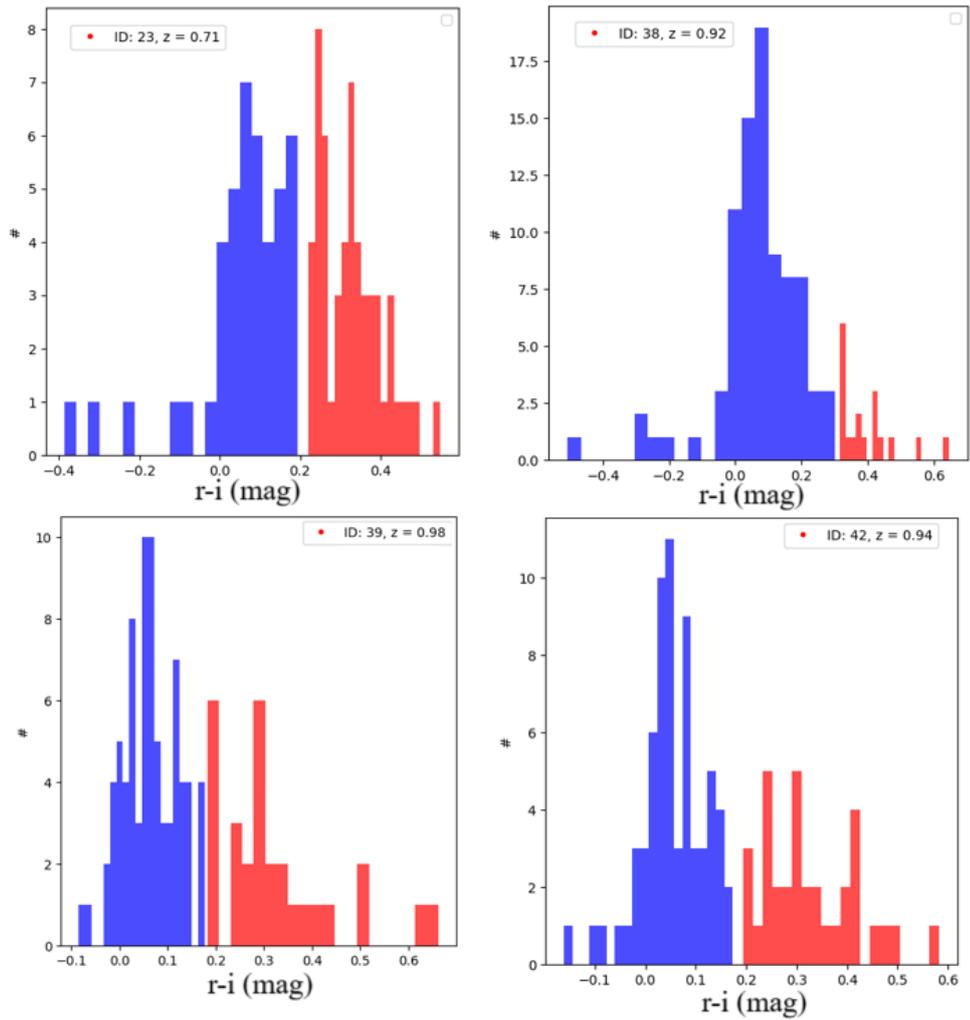


Figura 5.3. Exemplos típicos de distribuição bimodal dos resultados obtidos neste trabalho para os IDs 23, 38, 39 e 42.

Na Figura 5.4 foi apresentada a distribuição da razão de galáxias vermelhas por azul em função do *redshift* para todo o domínio da busca. Pode-se observar que a densidade numérica cresce com o aumento do *redshift*. Desta forma, pode-se concluir que as cores do pico azul tornam-se mais avermelhadas em direção ao presente, e a densidade numérica de galáxias azuis luminosas diminuiu em  $z < 1$ . Essas mesmas características foram observadas no estudo apresentado por Bell *et al.* (2004). Esse resultado apresenta características comuns em estudos de busca por galáxias com sequências vermelhas. Esse tipo de resultados é, geralmente, encontrado em ambientes de aglomerados de galáxias.

Com relação à riqueza<sup>1</sup> dos candidatos, foi observado que para menores valores de *redshift* os dados apresentaram maior riqueza enquanto para maiores *redshift* obteve-se menor riqueza, algo esperado pela função de seleção que perde galáxias com a distância. Como exemplo, pode-se ver a Figura 5.5 para valores de  $z = 0,71$  e  $z = 1,73$ . Para todos os 176 candidatos foram construídos gráficos de *contour* mostrando a concentração de

<sup>1</sup>não calculamos uma riqueza no sentido clássico do termo, apenas o número de galáxias membro

galáxias para cada candidato a aglomerado encontrado. Na Figura 5.6, pode-se observar regiões com alta e baixa densidade de galáxias. Os painéis superior e inferior apresentam o *countour plot* dos candidatos de ID 23 *redshift* 0,71 e 34 *redshift* 0,85, respectivamente.

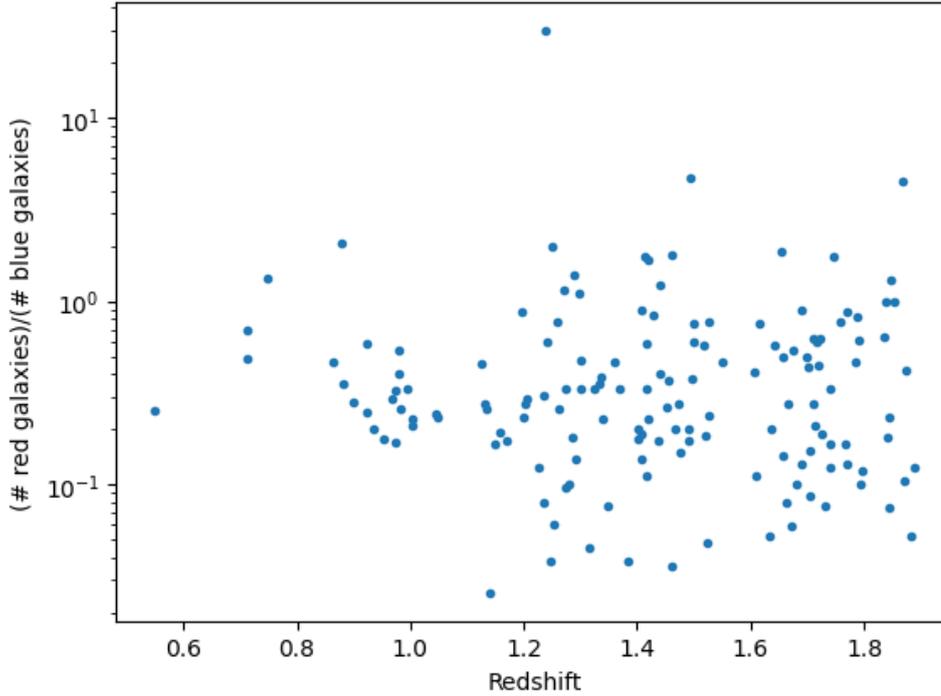


Figura 5.4. Relação de quantidade de galáxias vermelhas por azuis na faixa de *redshift* estudada.

Dentre os 150 candidatos inéditos, fez-se um levantamento daqueles que apresentaram múltiplos núcleos (MN), distorções no Diagrama Cor-Magnitude (CMD), uma grande diferença entre populações de galáxias azuis e vermelhas (P) e candidatos que apresentaram concentração de galáxias nas bordas do pixel (C). Na análise, considerou-se múltiplos núcleos os candidatos que apresentaram duas ou mais sobredensidades. Com relação às distorções no CMD, foram considerados inadequados aqueles candidatos que apresentaram galáxias de populações de cor vermelhas e azuis misturadas. A Tabela 5.1 apresenta dados da caracterização dos IDs denominados como: Característica (MN, CMD, P e C) e Situação (Mantido ou Eliminado). Ao final da análise, decidiu-se manter os IDs que apresentaram MN, P e C e eliminar os IDs com CMD distorcido. Foram encontrados 24 candidatos apresentando MN, 9 apresentando distorções no CMD, 30 que apresentaram P e 10 que apresentaram C. Após essa análise, o número de candidatos passou de 150 para 141. Nesse ponto essa escolha de corte é ligeiramente arbitrária e poderia ser mais estrita, mas de ponto de vista de estudo das populações estelares acreditamos que os problemas no CMD são os piores para a análise.

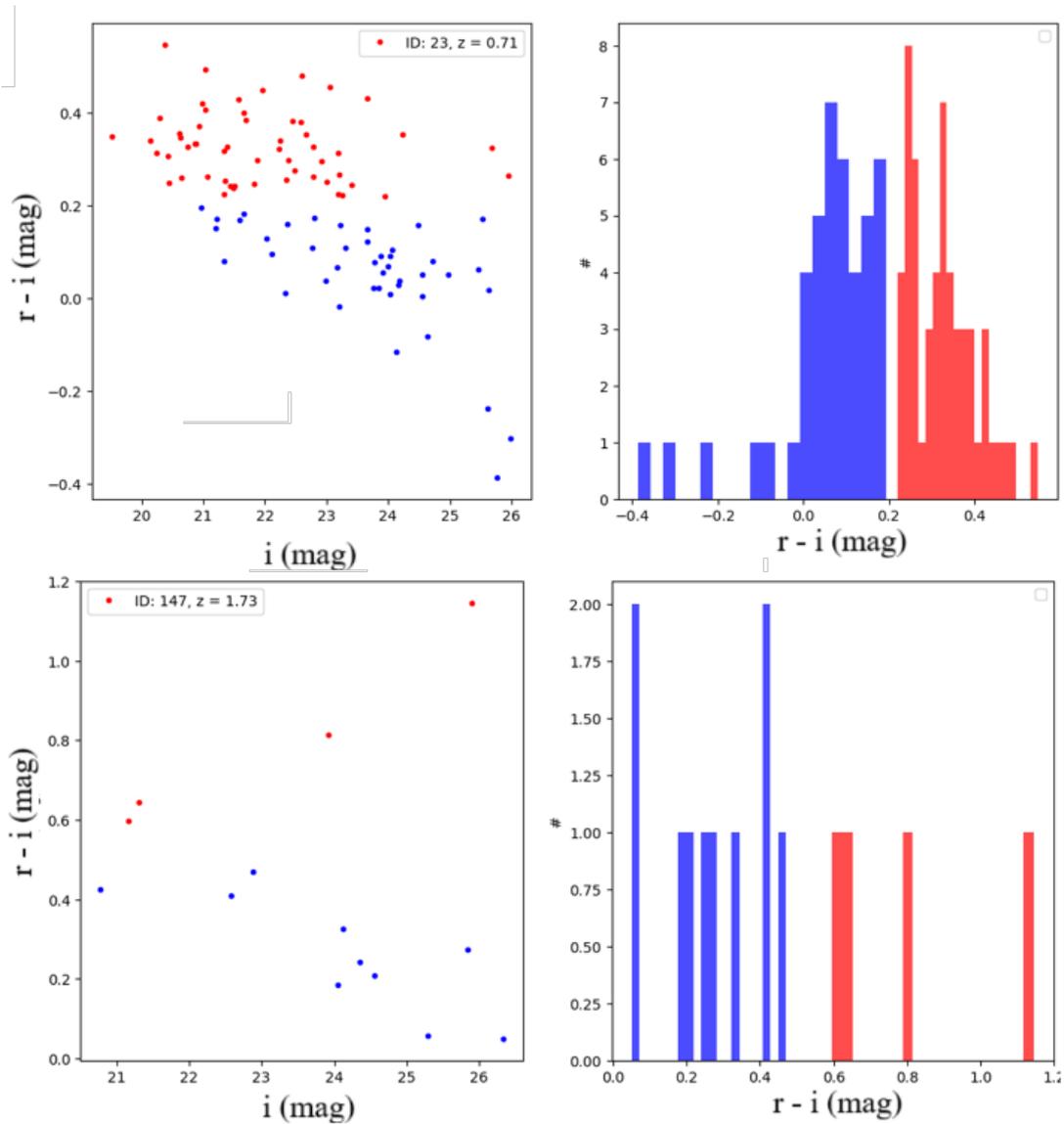


Figura 5.5. O painel superior da esquerda apresenta dados de  $r-i$  (mag) em função de  $i$  (mag) e o painel superior da direita a distribuição de  $r-i$  (mag) para o caso em que foi aplicada a  $k$ -correction. Os dados são referentes ao candidato de ID 23 e *redshift* 0,71 do catálogo proposto neste trabalho. O painel inferior da esquerda apresenta dados de  $r-i$  (mag) em função de  $i$  (mag) e o painel inferior da direita a distribuição de  $r-i$  (mag) para o caso em que foi aplicada a  $k$ -correction. Os dados são referentes ao candidato de ID 147 e *redshift* 1,73 do catálogo proposto neste trabalho.

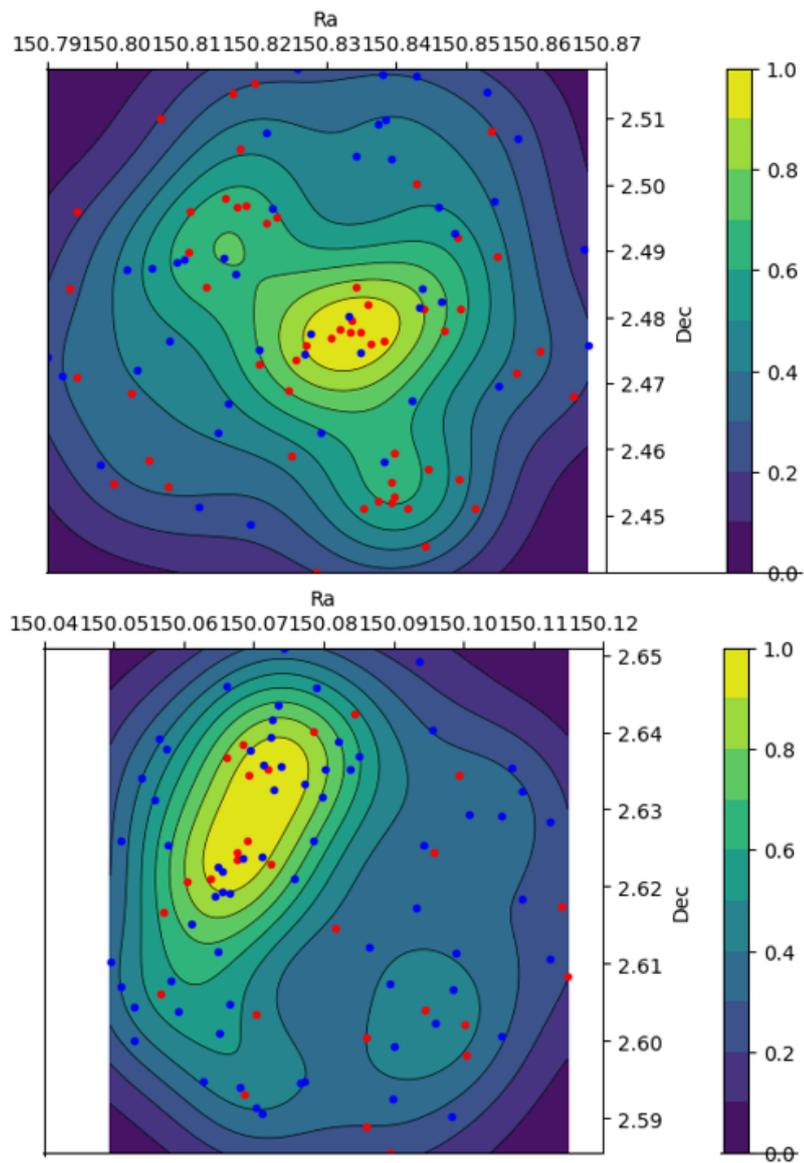


Figura 5.6. Os painéis superior e inferior apresentam o *contour plot* dos candidatos de ID 23 *redshift* 0,71 e 34 *redshift* 0,85, respectivamente, de *redshift* das galáxias membro.

ID	Característica	Situação	ID	Característica	Situação
19	MN	Mantido	92	P	Mantido
31	P	Mantido	94	P	Mantido
32	P	Mantido	95	P	Mantido
35	CMD/MN	Eliminado	107	P	Mantido
36	CMD	Eliminado	110	P	Mantido
39	MN	Mantido	111	P	Mantido
40	MN	Mantido	113	MN	Mantido
45	P	Mantido	114	P	Mantido
43	CMD	Eliminado	115	P	Mantido
44	CMD	Eliminado	117	MN/P	Mantido
46	CMD	Eliminado	118	MN	Mantido
49	C	Mantido	119	MN	Mantido
50	C	Mantido	121	P	Mantido
52	P	Mantido	124	MN	Mantido
53	MN	Mantido	125	MN	Mantido
54	P	Mantido	128	C	Mantido
55	MN	Mantido	132	P/C	Mantido
56	P	Mantido	135	C	Mantido
57	CMD	Eliminado	136	MN/P	Mantido
64	CMD	Eliminado	137	P	Mantido
70	CMD	Eliminado	139	P	Mantido
73	MN	Mantido	140	C	Mantido
74	CMD	Eliminado	143	MN	Mantido
75	P	Mantido	149	MN	Mantido
76	P	Mantido	150	C	Mantido
77	MN	Mantido	152	MN/P	Mantido
79	P	Mantido	154	C	Mantido
81	P	Mantido	160	MN	Mantido
83	P	Mantido	161	P	Mantido
84	P	Mantido	164	P	Mantido
86	P	Mantido	165	MN	Mantido
87	P	Mantido	168	MN/C	Mantido
89	MN	Mantido	169	MN/C	Mantido
90	P	Mantido	170	MN/C	Mantido

Tabela 5.1: Dados de ID, Característica e Situação de candidatos que apresentaram múltiplos núcleos (MN), distorções no Diagrama Cor-Magnitude (CMD), grande diferença entre populações de galáxias azuis e vermelhas (P) e candidatos que apresentam uma concentração de galáxias na extremidade do pixel (C).

A Figura 5.7 apresenta um exemplo típico de um candidato que apresentou MN, com dados de  $r-i$  (mag) e  $i-z$  (mag) em função de  $i$  (mag) e *countour plot* do candidato e o histograma de *redshift* das galáxias membros. Os dados são referentes ao candidato de ID 19 e *redshift* 0,75 do catálogo G24. Já a Figura 5.8 apresenta um exemplo típico de um candidato que apresentou problema no CMD, com dados de  $r-i$  (mag) em função de  $i$  (mag) e distribuição de  $r-i$  (mag). Os dados são referentes ao candidato de ID 35 e *redshift* 0,97 do catálogo proposto neste trabalho.

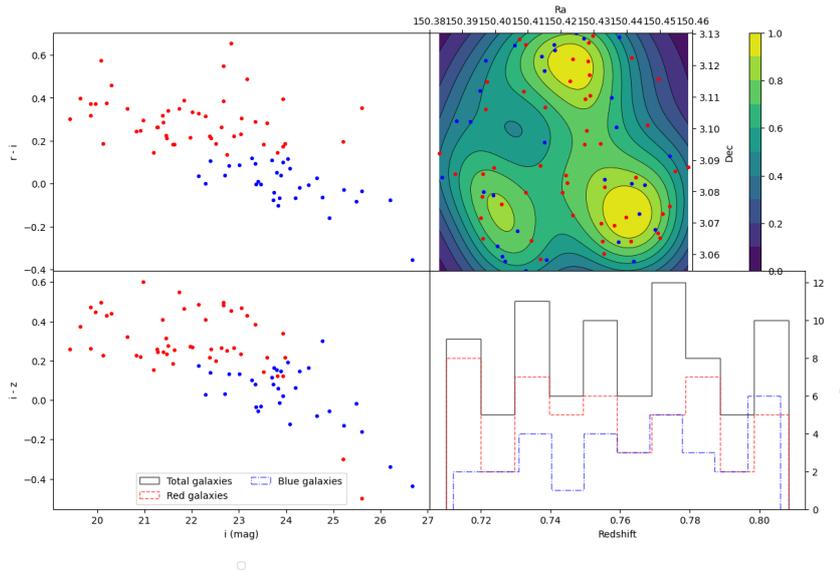


Figura 5.7. Exemplo típico de um candidato que apresentou MN. Os painéis superior e inferior da esquerda apresentam dados de  $r-i$  (mag) e  $i-z$  (mag) em função de  $i$  (mag), respectivamente. Os painéis superior e inferior da direita apresentam o *countour plot* do candidato e o histograma de *redshift* das galáxias membros. Os dados são referentes ao candidato de ID 19 e *redshift* 0,75 do catálogo proposto neste trabalho.

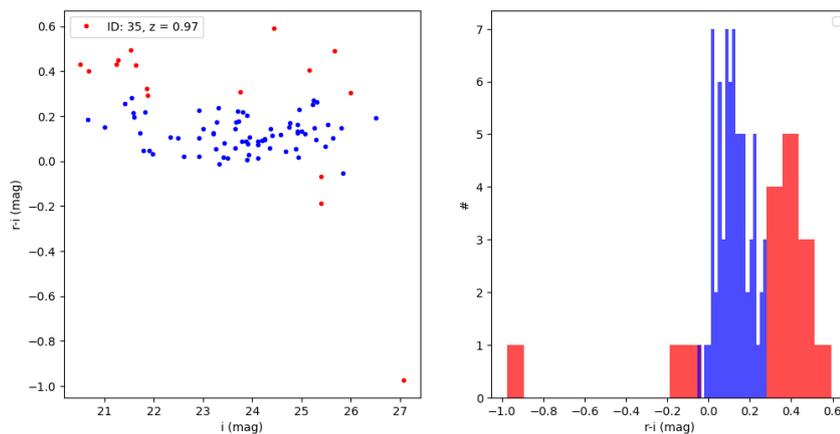


Figura 5.8. Exemplo típico de um candidato que apresentou problema no CMD. O painel da esquerda apresenta dados de  $r-i$  (mag) em função de  $i$  (mag) e o painel da direita apresenta a distribuição de  $r-i$  (mag). Os dados são referentes ao candidato de ID 35 e *redshift* 0,97 do catálogo proposto neste trabalho.

## 5.2 Morfologia das galáxias

É importante determinar a morfologia das galáxias, por fornecer pistas valiosas sobre a sua origem e formação. A morfologia abrange uma variedade de características, como a presença de barras, braços espirais e núcleos ativos. Desta forma, compreender a origem, o tipo, a frequência e a densidade morfológica são de suma importância. No entanto, a determinação dessas propriedades requer dados de boa qualidade e classificação confiável durante o processo de construção de grandes catálogos para estudos posteriores.

Objetivando determinar a morfologia dos candidatos a protoaglomerados de galáxias, foram estudados o comportamento da variável BDF `fracdev` em função do *redshift*. Esta variável indica a contribuição das componentes do bojo e do disco no perfil de brilho das galáxias, sendo detalhado a seguir o seu significado.

É possível atribuir uma relação matemática entre o parâmetro BDF `fracdev` e o modelo de de Vaucouleurs dependendo do contexto e dos dados observacionais na classificação e modelagem de galáxias (Vincent e Ryden, 2005). Para este parâmetro, existem três valores típicos: 0, 0,5 e 1,0, que podem ter implicações na interpretação da estrutura e formação das galáxias. O valor de BDF `fracdev` = 0,5, pode indicar que metade da luminosidade observada da galáxia é atribuída ao componente de de Vaucouleurs e a outra metade a outros componentes, como discos exponenciais (Rodríguez e Padilla, 2013). Desta forma, BDF `fracdev` = 0,5 indica que o perfil de brilho da galáxia é uma combinação equilibrada entre o perfil de de Vaucouleurs e outros perfis. Por exemplo, BDF `fracdev` = 0,5 pode sugerir que a galáxia possui tanto uma componente central densa e concentrada (perfil de Vaucouleurs) quanto, por exemplo, um disco exponencial (modelos exponenciais). Valores de BDF `fracdev` = 0,0 e igual a BDF `fracdev` = 1,0 podem indicar que as galáxias são do tipo espirais e elípticas, respectivamente.

A Figura 5.9 apresenta dados para o conjunto total de galáxias do COSMOS e a Figura 5.10 os dados para as galáxias contidas nos candidatos a aglomerados do G24. Para os dados presentes no catálogo do COSMOS e no catálogo G24, os resultados apresentaram comportamentos semelhantes, ver Figuras 5.9 e 5.10.

O BDF `fracdev` é um parâmetro que tem como função definir o tipo de perfil de brilho que classifica as galáxias em elípticas (bojo) e espirais (disco). As galáxias elípticas (*Early-Type*) possuem um perfil de luz com uma dependência do tipo de Vaucouleurs (BDF `fracdev` = 1), onde o brilho varia de forma abrupta com a distância. Já galáxias espirais (*Late-Type*) possuem uma variação de brilho com dependência exponencial (BDF `fracdev` = 0), onde o brilho superficial decai de maneira suave ao longo do raio.

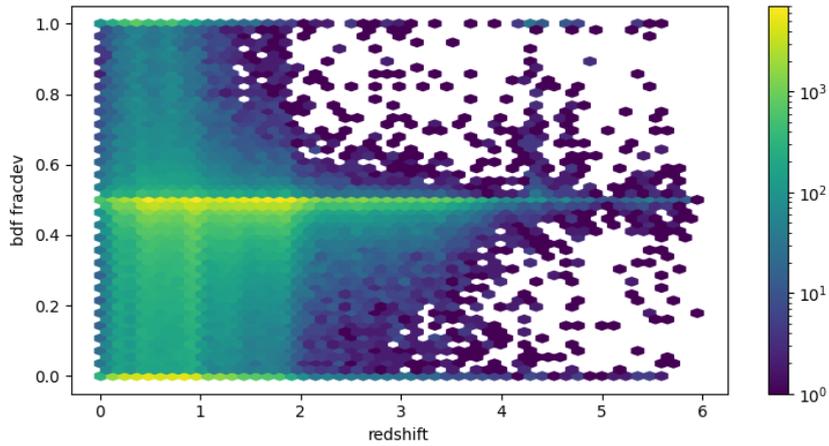


Figura 5.9. Dados do parâmetro BDF `fracdev` em função do *redshift* para o catálogo de galáxias do COSMOS.

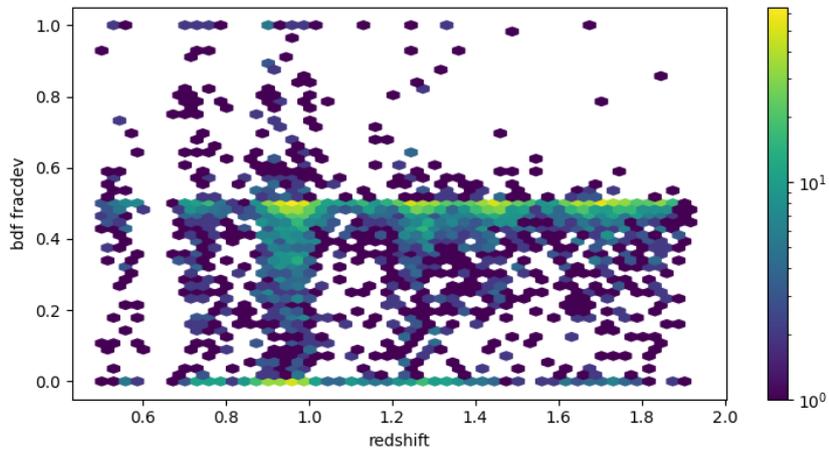


Figura 5.10. Dados do parâmetro BDF `fracdev` em função do *redshift* para o catálogo de galáxias de G24 (catálogo de candidatos encontrados neste trabalho).

A Figura 5.11 apresenta dados de BDF `fracdev` em função do *redshift*, onde as colunas representam respectivamente as cores g-r (painel superior), r-i (painel central) e i-z (painel inferior). Pode-se observar que para BDF `fracdev` entre 0,5 e 1,0 a população de galáxias é, predominantemente, vermelha. Na região de  $0,8 < z < 1,0$  e para valores de BDF `fracdev` inferiores a 0,5, existe uma alta concentração de galáxias de população azul.

A Figura 5.12 apresenta o histograma do número de galáxias de G24 em função do BDF `fracdev` para a totalidade dos dados no painel da esquerda e com zoom na faixa de 0,47 e 0,53 no painel da direita. Pode-se observar uma alta concentração de galáxias apresentando valores de BDF `fracdev` em torno de 0,5. Com o objetivo de visualizar com mais detalhes outros valores de BDF `fracdev`, foi construída a Figura 5.13 eliminando galáxias com valor de BDF `fracdev` =  $0,5^{+0,005}_{-0,01}$  e valores de  $r(mag) > 24$ , representado no painel da direita da Figura 2.12.

Como pode-se perceber, um elevado número de galáxias concentra-se na faixa de

$\text{BDF fracdev} = 0,5_{-0,01}^{+0,005}$ . Este fato deve-se por  $\text{BDF fracdev}$  estar centrado em 0,5. Objetos fracos apresentam  $\text{BDF fracdev}$  mal definidos, sendo atribuídos valores próximos a 0,5 como *prior* do modelo. As Figuras 5.14 e 5.15 apresentam cortes na variável  $\text{BDF fracdev}$ . Comparando as Figuras 5.10 e 5.14 e o painel superior e inferior da Figura 5.15, pode-se perceber a diferença na densidade de pontos na região próxima a 0,5.

A Figura 5.16 apresenta os candidatos do catálogo G24 em diferentes cores para  $\text{BDF fracdev}$  em função do *redshift* para o catálogo de galáxias de G24 após o corte em  $\text{BDF fracdev} = 0,5_{-0,01}^{+0,005}$  e valores de  $r(\text{mag}) > 24$ .

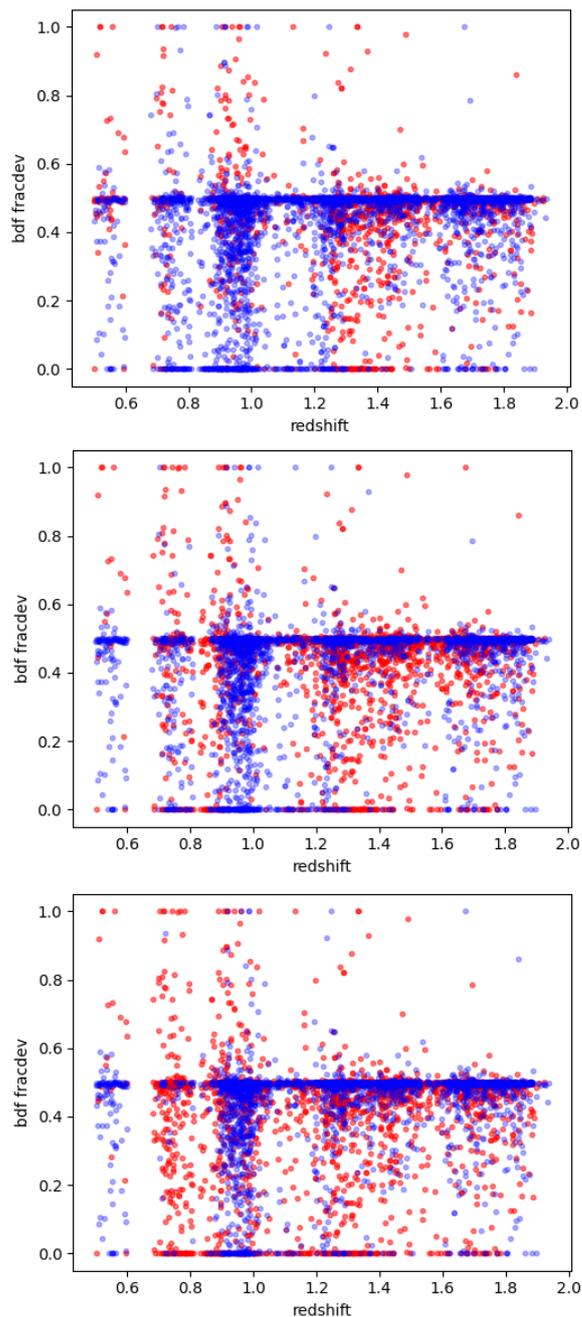


Figura 5.11. Divisão em duas populações nas cores azul e vermelha referente aos dados do parâmetro BDF *fracdev* em função do *redshift* para o catálogo de galáxias G24. As colunas representam respectivamente as cores g-r (painel superior), r-i (painel central) e i-z (painel inferior).

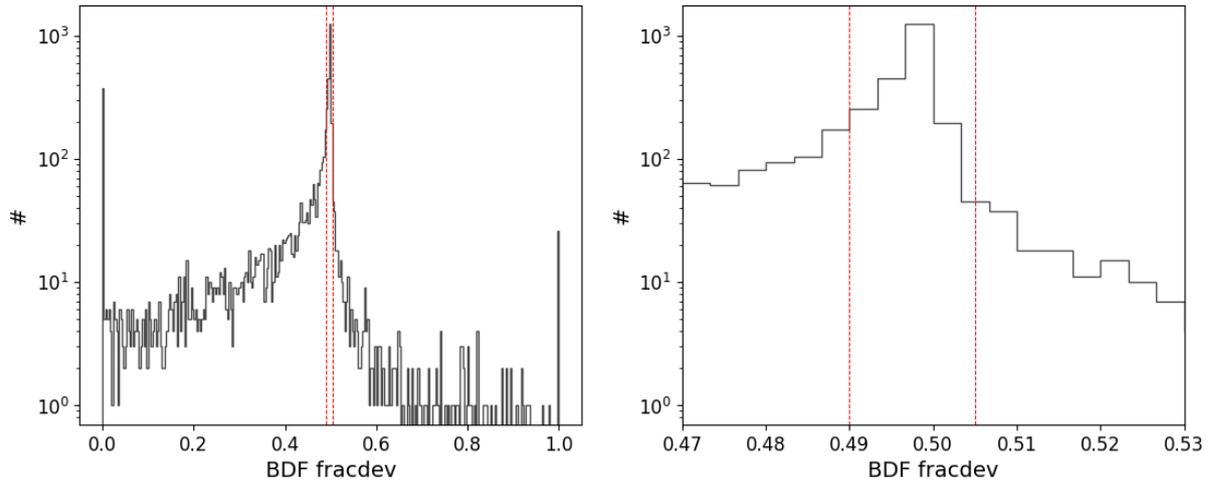


Figura 5.12. Histogramas do número de galáxias de G24 em função do BDF *fracdev* para a totalidade dos dados. Painel da esquerda apresenta a representação normal, e o da direita apresenta um zoom na faixa de 0,47 a 0,53. As linhas vermelhas verticais apresentam a faixa do possível corte.

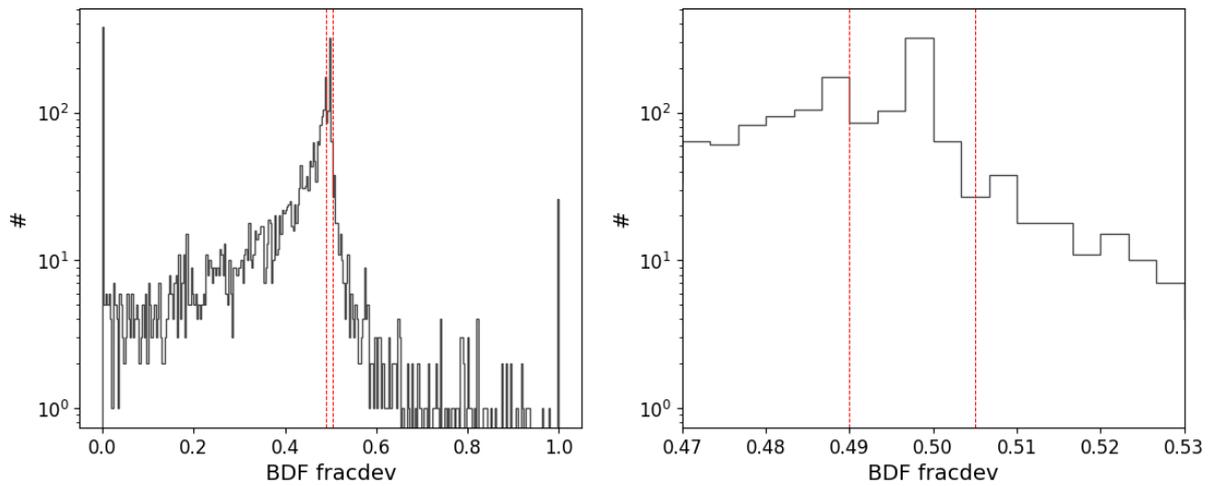


Figura 5.13. Histogramas do número de galáxias de G24 em função do BDF *fracdev* com corte dos dados na região  $\text{BDF } \textit{fracdev} = 0,5^{+0,005}_{-0,01}$  e valores de  $r(\textit{mag}) > 24$ . Painel da esquerda apresenta a representação normal, e o da direita apresenta um zoom na faixa de 0,47 a 0,53. As linhas vermelhas verticais apresentam a faixa do corte.

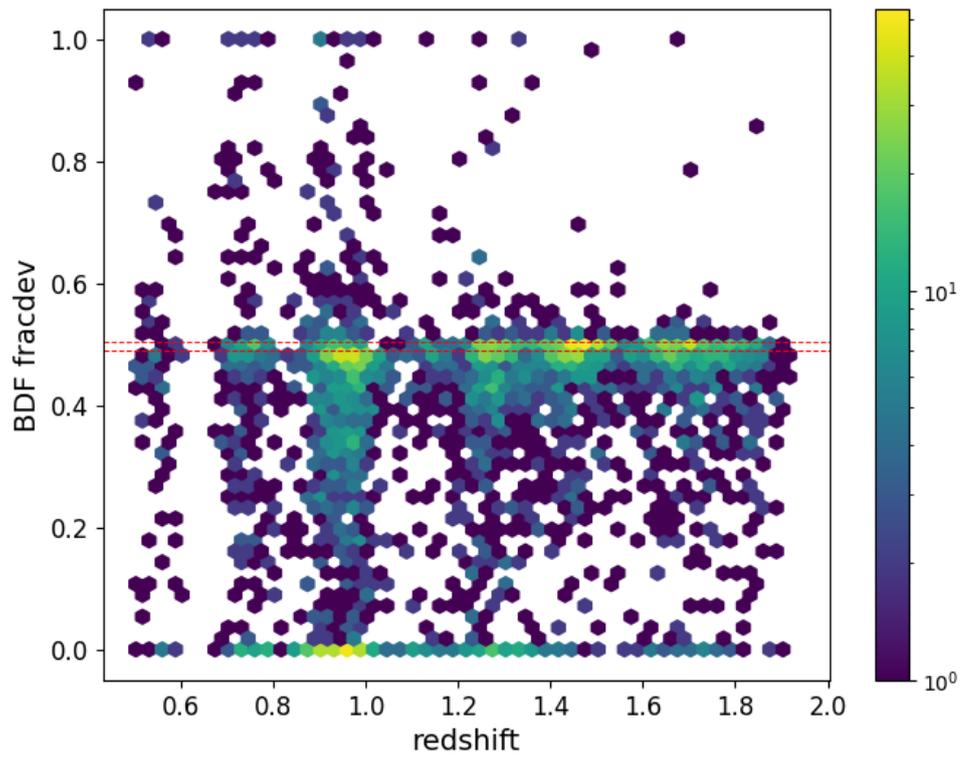


Figura 5.14. Dados do parâmetro  $BDF\ fracdev$  em função do  $redshift$  para o catálogo de galáxias de G24 após o corte na faixa de  $BDF\ fracdev = 0,5^{+0,005}_{-0,01}$  e valores de  $r(mag) > 24$ .

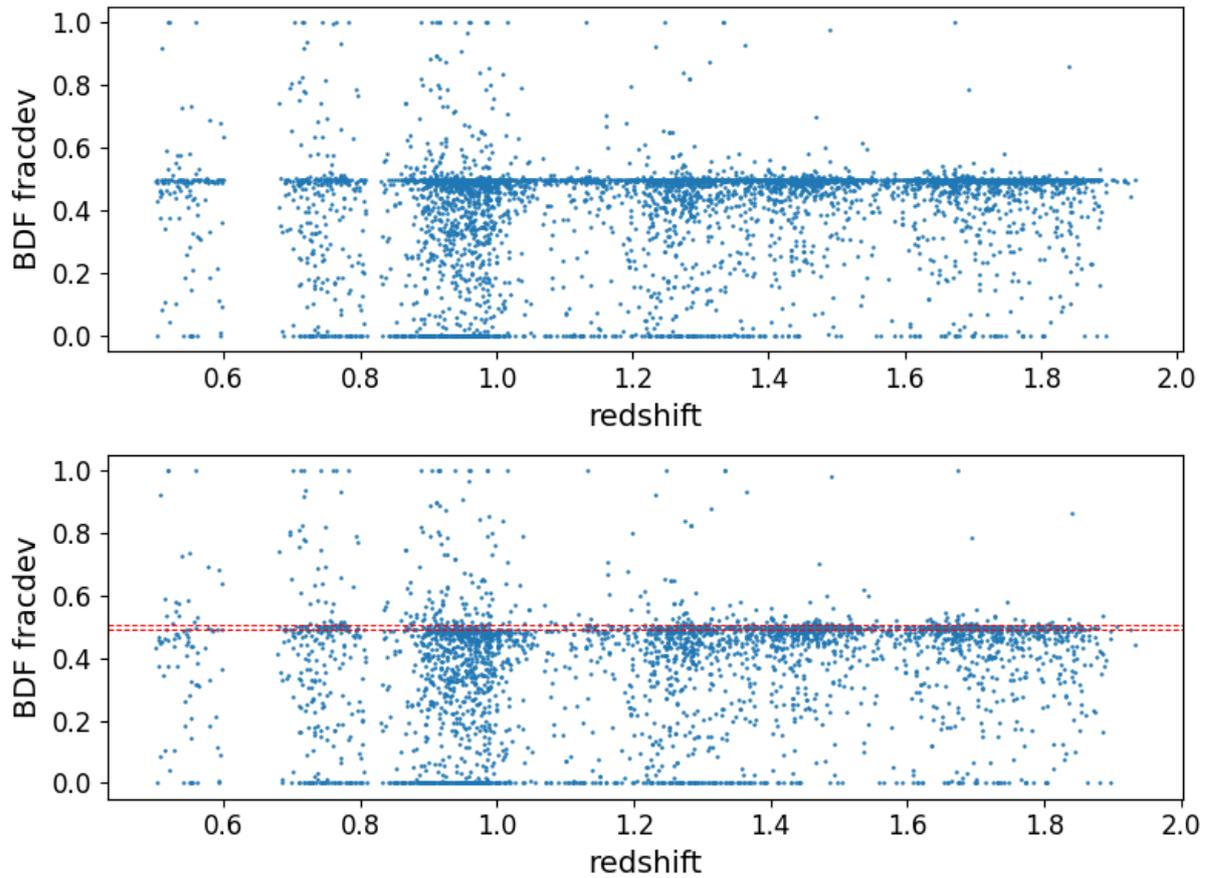


Figura 5.15. BDF *fracdev* em função do *redshift* para o catálogo de galáxias de G24. O painel superior apresenta dados sem corte e o painel inferior apresenta os dados após o corte em  $BDF \text{ fracdev} = 0,5^{+0,005}_{-0,01}$  e valores de  $r(mag) > 24$ .

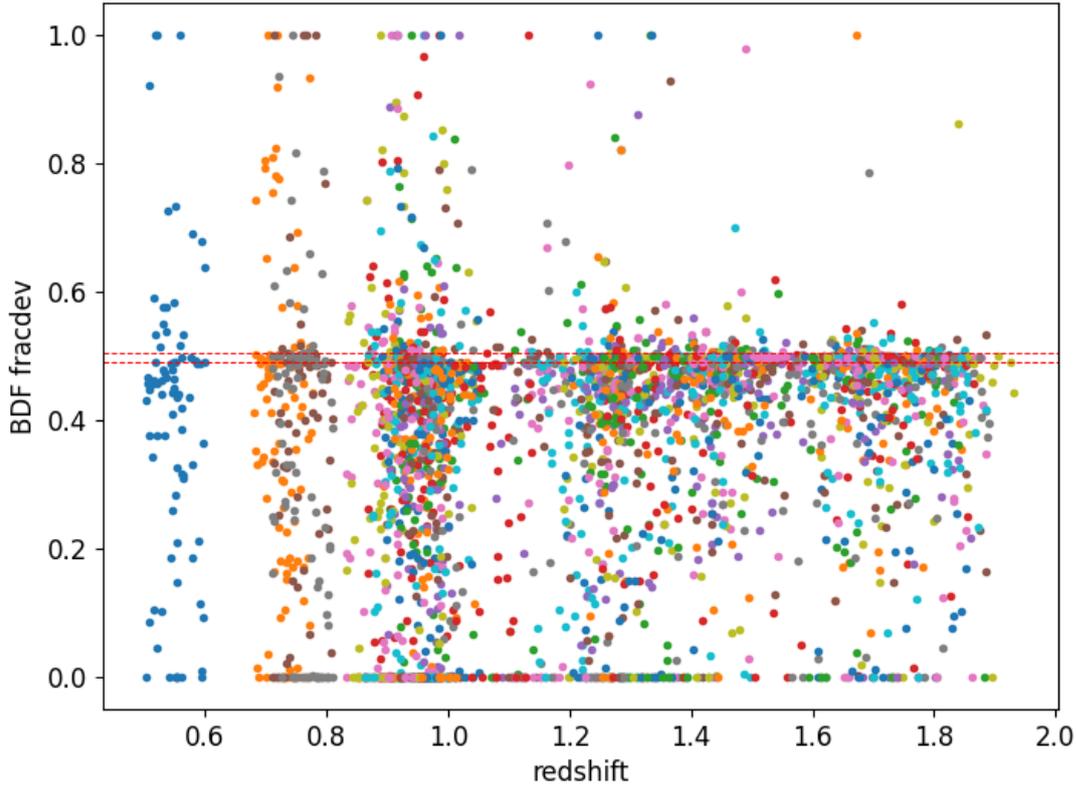


Figura 5.16. BDF `fracdev` em função do `redshift` para o catálogo de galáxias de G24 após o corte em  $\text{BDF } \text{fracdev} = 0,5^{+0,005}_{-0,01}$  e valores de  $r(\text{mag}) > 24$ . Os candidatos do catálogo são representados em cores diferentes.

A análise morfológica baseada no parâmetro BDF `fracdev` permitiu verificar para o catálogo G24 uma alta concentração de galáxias onde não é possível definir se a galáxia é espiral ou elíptica, apresentando valores de BDF `fracdev` em torno de 0,5. As galáxias elípticas que apresentam BDF `fracdev` = 1, tem maior ocorrência em baixos redshift enquanto as espirais apresentam ocorrência em toda faixa de `redshift` estudada. Pode-se observar na Figura 5.11 que a população de galáxias azuis está localizada, majoritariamente, em valores BDF `fracdev` < 0,5 (Vincent e Ryden, 2005).

# Capítulo 6

## Conclusões

Neste trabalho, foi feita uma busca por protoaglomerados de galáxias a partir do catálogo dos *Deep Fields* do *Dark Energy Survey* (DF DES). A busca foi feita com um algoritmo escrito em linguagem Python<sup>1</sup> baseado fundamentalmente na busca por regiões densas por meio da técnica de sobredensidades de galáxias. Foi utilizado o código para executar buscas em áreas de alta densidade de galáxias em fatias de *redshift*, onde há pontos em que galáxias estão concentradas. Os pontos densos foram considerados candidatos a aglomerados e todas as galáxias da região foram consideradas como parte do candidato.

O algoritmo desenvolvido foi validado com base no trabalho de Narwal *et al.* (2020), onde foi possível reproduzir os resultados dos autores, tendo encontrado uma estrutura no intervalo de Dec de -50 a -48 e RA de -2,5 a 0,0. Após a validação, foi executada a busca para região do COSMOS contida no catálogo dos DF DES para *redshifts* na faixa de 0,5 a 2,0.

Fez-se uma divisão da faixa de *redshift* em *bins*, havendo interseção entre os *bins*. Foram detectados 309 candidatos a aglomerados para a região delimitada. Foi escrito um programa em Python para eliminar candidatos que estivessem em *bins* diferentes. A quantidade final foi de 176 candidatos. Foi estudada a influência do tamanho do pixel do histograma 2D na buscando o maior número de *matches* com a literatura. Concluiu-se que o melhor caso foi o que considerava o pixel igual a 2 Mpc. Para verificar se os *matches* encontrados não eram aleatórios, montou-se catálogos *mock* usando a técnica de *bootstrap* constatando a eficiência da técnica ao se encontrar poucos *matches* no *mock*. Fez-se o *match* 3D com os catálogos da literatura, chegando a 26 candidatos comuns a algum catálogo (levando em consideração os dados de Wen e Han (2011), Bellagamba *et al.* (2011) e Sochting *et al.* (2012)) e apenas 2 em comum com todos os catálogos, um problema conhecido na literatura, mesmo que os autores reportem alta completeza e pureza em seus respectivos catálogos.

Os resultados compõe-se de um total de 176 candidatos encontrados na região do

---

<sup>1</sup>base de todo o desenvolvimento aqui nesse trabalho

COSMOS na faixa de *redshift* de 0,5 - 2,0. Após a comparação com dados da literatura, chegou-se a um resultado final de 150 candidatos inéditos. Como o objetivo de caracterizar os candidatos inéditos foram feitas análises com *k-correction*, gráficos de *contour* para aferir sua distribuição espacial, a distribuição da razão entre galáxias vermelhas e azuis separadas pelo algoritmo GMM. Pode-se observar, que a distribuição de cores em *rest-frame* de galáxias em  $z < 1$  é, majoritariamente, bimodal (Bell *et al.*, 2004).

Dentre os 150 candidatos inéditos do G24, decidiu-se manter 22 candidatos que apresentaram múltiplos núcleos (MN), 30 candidatos que apresentaram uma desproporção elevada entre as populações de galáxias azuis e vermelhas (P) e 12 candidatos que foram localizados na extremidade do pixel (C). Foram eliminados os 9 candidatos que apresentaram distorções no Diagrama Cor-Magnitude (CMD). A quantidade final de candidatos passou de 150 para 141.

Com relação ao estudo de morfologia, os resultados mostraram que o catálogo de candidatos a protoaglomerados apresentou uma diminuição das galáxias do tipo elíptica com a aumento do *redshift*. Para a variável BDF **fracdev**, pode-se observar que com o aumento do *redshift* a população de galáxias elípticas diminui e observou-se que entre 0,5 e 1,0 a população de galáxias foi predominantemente vermelha. Para a região de  $0,8 < \textit{redshift} < 1,0$  e para valores de BDF **fracdev** inferiores a 0,5, existe uma alta concentração de galáxias de população azul. Para a região de  $1,0 < \textit{redshift} < 2,0$  os valores de BDF **fracdev** são, predominantemente, inferiores a 0,5, indicando uma predominância de galáxias espirais (Vincent e Ryden, 2005, Rodríguez e Padilla, 2013). A correlação entre a separação de cores do GMM e a classificação morfológica do BDF **fracdev** é muito tranquilizadora.

Para alcançar a confirmação dos candidatos precisaríamos de observações espectroscópicas com telescópios como SOAR e Gemini, a depender da magnitude e distância do candidato, mas isso está além do escopo do tempo do mestrado e as dependências com lançamentos de dados do DES. O plano original era usar dados completos com os seis anos do DES, mas esses não foram liberados a tempo deste trabalho, portanto usamos os dados do ano 3, não tão profundos quanto os dados completos (Hartley *et al.*, 2022).

Esperamos demonstrar que com novos levantamentos ópticos profundos, como o LSST (Ivezić *et al.*, 2019), deve ser relativamente simples encontrar aglomerações no universo distante que possam ensinar sobre a formação e evolução de estruturas.

## Referências Bibliográficas

- Abbott, T. M. C., Abdalla, F. B., Allam, S., et al., 2018, “The Dark Energy Survey: Data Release 1”, *Astrophysical Journal, Supplement*, 239(2):18. doi: 10.3847/1538-4365/aae9f0.
- Abbott, T. M. C., Allam, S., Andersen, P., et al., 2019, “First Cosmology Results using Type Ia Supernovae from the Dark Energy Survey: Constraints on Cosmological Parameters”, *Astrophys. J. Lett.*, 872(2):L30. doi: 10.3847/2041-8213/ab04fa.
- Abbott, T. M. C., Adamów, M., Aguena, M., et al., 2021, “The Dark Energy Survey Data Release 2”, *Astrophys. J. Suppl.*, 255(2):20. doi: 10.3847/1538-4365/ac00b3.
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al., 2022, “The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package”, *apj*, 935(2):167. doi: 10.3847/1538-4357/ac7c74.
- Bamford, S. P., Nichol, R. C., Baldry, I. K., et al., 2009, “Galaxy Zoo: the dependence of morphology and colour on environment\*”, *Mon. Not. Roy. Astron. Soc.*, v. 393, n. 4 (mar.), pp. 1324–1352. doi: 10.1111/j.1365-2966.2008.14252.x.
- Bell, E. F., Wolf, C., Meisenheimer, K., et al., 2004, “Nearly 5000 Distant Early-Type Galaxies in COMBO-17: A Red Sequence and Its Evolution since  $z \sim 1$ ”, *Astrophys. J.*, v. 608, n. 2 (jun.), pp. 752–767. doi: 10.1086/420778.
- Bellagamba, F., Maturi, M., Hamana, T., et al., 2011, “VizieR Online Data Catalog: Galaxy clusters in the COSMOS field (Bellagamba+, 2011)”, *VizieR Online Data Catalog*, art. J/MNRAS/413/1145.
- Blanton, M. R., Roweis, S., 2007, “K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared”, *Astron. J.*, v. 133, n. 2 (fev.), pp. 734–754. doi: 10.1086/510127.
- Bleem, L. E., Stalder, B., de Haan, T., et al., 2015, “Galaxy Clusters Discovered via the Sunyaev-Zel’dovich Effect in the 2500-Square-Degree SPT-SZ Survey”, *Astrophys. J. Suppl.*, 216(2):27. doi: 10.1088/0067-0049/216/2/27.

- Bom, C. R., Cortesi, A., Ribeiro, U., et al., 2024, “An extended catalogue of galaxy morphology using deep learning in southern photometric local universe survey data release 3”, *Mon. Not. Roy. Astron. Soc.*, v. 528, n. 3 (mar.), pp. 4188–4208. doi: 10.1093/mnras/stad3956.
- Bonnett, C., Troxel, M. A., Hartley, W., et al., 2016, “Redshift distributions of galaxies in the Dark Energy Survey Science Verification shear catalogue and implications for weak lensing”, *Phys. Rev. D*, 94(4):042005. doi: 10.1103/PhysRevD.94.042005.
- Brammer, G. B., van Dokkum, P. G., Coppi, P., 2008a, “EAZY: A Fast, Public Photometric Redshift Code”, *Astrophys. J.*, v. 686, n. 2 (out.), pp. 1503–1513. doi: 10.1086/591786.
- Brammer, G. B., van Dokkum, P. G., Coppi, P., 2008b, “EAZY: A Fast, Public Photometric Redshift Code”, *Astrophys. J.*, v. 686, n. 2 (out.), pp. 1503–1513. doi: 10.1086/591786.
- Brinch, M., Greve, T. R., Weaver, J. R., et al., 2023, “COSMOS2020: Identification of High-z Protocluster Candidates in COSMOS”, *Astrophys. J.*, 943(2):153. doi: 10.3847/1538-4357/ac9d96.
- Brinch, M., Greve, T. R., Sanders, D. B., et al., 2024, “DEIMOS spectroscopy of  $z = 6$  protocluster candidate in COSMOS - a massive protocluster embedded in a large-scale structure?” *Mon. Not. Roy. Astron. Soc.*, v. 527, n. 3 (jan.), pp. 6591–6615. doi: 10.1093/mnras/stad3409.
- Capak, P. L., Riechers, D., Scoville, N. Z., et al., 2011, “A massive protocluster of galaxies at a redshift of  $z \sim 5.3$ ”, *Nature*, v. 470, n. 7333 (fev.), pp. 233–235. doi: 10.1038/nature09681.
- Cautun, M., van de Weygaert, R., Jones, B. J. T., et al., 2014, “Evolution of the cosmic web”, *Monthly Notices of the Royal Astronomical Society*, v. 441, n. 4 (05), pp. 2923–2973. ISSN: 0035-8711. doi: 10.1093/mnras/stu768. Disponível em: <<https://doi.org/10.1093/mnras/stu768>>.
- Chiang, Y.-K., Overzier, R. A., Gebhardt, K., et al., 2017, “Galaxy Protoclusters as Drivers of Cosmic Star Formation History in the First 2 Gyr”, *Astrophys. J. Lett.*, 844(2):L23. doi: 10.3847/2041-8213/aa7e7b.
- Chilingarian, I. V., Melchior, A.-L., Zolotukhin, I. Y., 2010, “Analytical approximations of K-corrections in optical and near-infrared bands”, *Mon. Not. Roy. Astron. Soc.*, v. 405, n. 3 (jul.), pp. 1409–1420. doi: 10.1111/j.1365-2966.2010.16506.x.

- Dalton, G. B., Caldwell, M., Ward, A. K., et al., 2006, “The VISTA infrared camera”. In: McLean, I. S., Iye, M. (Eds.), *Ground-based and Airborne Instrumentation for Astronomy*, v. 6269, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 62690X, jun. doi: 10.1117/12.670018.
- Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al., 2016, “The Dark Energy Survey: more than dark energy - an overview”, *Mon. Not. Roy. Astron. Soc.*, v. 460, n. 2 (ago.), pp. 1270–1299. doi: 10.1093/mnras/stw641.
- de Souza, R. E., Gadotti, D. A., dos Anjos, S., 2004, “BUDDA: A New Two-dimensional Bulge/Disk Decomposition Code for Detailed Structural Analysis of Galaxies”, *Astrophys. J. Suppl.*, v. 153, n. 2 (ago.), pp. 411–427. doi: 10.1086/421554.
- Diener, C., Lilly, S. J., Ledoux, C., et al., 2015, “A PROTOCLUSTER AT  $z = 2.45$ ”, *The Astrophysical Journal*, v. 802, n. 1 (mar), pp. 31. doi: 10.1088/0004-637X/802/1/31. Disponível em: <<https://dx.doi.org/10.1088/0004-637X/802/1/31>>.
- Dong, C., Lee, K.-G., Ata, M., et al., 2023, “Observational Evidence for Large-scale Gas Heating in a Galaxy Protocluster at  $z = 2.30$ ”, *Astrophys. J. Lett.*, 945(2):L28. doi: 10.3847/2041-8213/acba89.
- Drlica-Wagner, A., Sevilla-Noarbe, I., Rykoff, E. S., et al., 2018, “Dark Energy Survey Year 1 Results: The Photometric Data Set for Cosmology”, *Astrophys. J. Suppl.*, 235(2):33. doi: 10.3847/1538-4365/aab4f5.
- Efron, B., 1979, “Bootstrap Methods: Another Look at the Jackknife”, *The Annals of Statistics*, v. 7, n. 1, pp. 1–26. ISSN: 00905364. Disponível em: <<http://www.jstor.org/stable/2958830>>.
- Emerson, J., McPherson, A., Sutherland, W., 2006, “Visible and Infrared Survey Telescope for Astronomy: Progress Report”, *The Messenger*, v. 126 (dez.), pp. 41–42.
- Farrens, S., Abdalla, F. B., Cypriano, E. S., et al., 2011, “Friends-of-friends groups and clusters in the 2SLAQ catalogue”, *Mon. Not. Roy. Astron. Soc.*, v. 417, n. 2 (out.), pp. 1402–1416. doi: 10.1111/j.1365-2966.2011.19356.x.
- Finoguenov, A., Guzzo, L., Hasinger, G., et al., 2007, “VizieR Online Data Catalog: XMM Clusters of galaxies in COSMOS field (Finoguenov+ 2007)”, *VizieR Online Data Catalog*, art. J/ApJS/172/182. doi: 10.26093/cds/vizier.21720182.
- Flaugher, B., Diehl, H. T., Honscheid, K., et al., 2015, “The Dark Energy Camera”, *Astron. J.*, 150(5):150. doi: 10.1088/0004-6256/150/5/150.

- Gladders, M. D., Yee, H. K. C., 2005, “The Red-Sequence Cluster Survey. I. The Survey and Cluster Catalogs for Patches RCS 0926+37 and RCS 1327+29”, *Astrophys. J. Suppl.*, v. 157, n. 1 (mar.), pp. 1–29. doi: 10.1086/427327.
- Hao, J., McKay, T. A., Koester, B. P., et al., 2010, “A GMBCG Galaxy Cluster Catalog of 55,424 Rich Clusters from SDSS DR7”, *Astrophys. J. Suppl.*, v. 191, n. 2 (dez.), pp. 254–274. doi: 10.1088/0067-0049/191/2/254.
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al., 2020, “Array programming with NumPy”, *Nature*, v. 585, n. 7825 (set.), pp. 357–362. doi: 10.1038/s41586-020-2649-2. Disponível em: <<https://doi.org/10.1038/s41586-020-2649-2>>.
- Hartley, W. G., Choi, A., Amon, A., et al., 2022, “Dark Energy Survey Year 3 Results: Deep Field optical + near-infrared images and catalogue”, *Mon. Not. Roy. Astron. Soc.*, v. 509, n. 3 (jan.), pp. 3547–3579. doi: 10.1093/mnras/stab3055.
- Hubble, E., 1929, “A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae”, *Proceedings of the National Academy of Science*, v. 15, n. 3 (mar.), pp. 168–173. doi: 10.1073/pnas.15.3.168.
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al., 2019, “LSST: From Science Drivers to Reference Design and Anticipated Data Products”, *Astrophys. J.*, 873(2):111. doi: 10.3847/1538-4357/ab042c.
- Jiang, L., Wu, J., Bian, F., et al., 2018, “A giant protocluster of galaxies at redshift 5.7”, *Nature Astronomy*, v. 2 (out.), pp. 962–966. doi: 10.1038/s41550-018-0587-9.
- Kimm, T., Yi, S. K., 2007, “Intrinsic Axis Ratio Distribution of Early-Type Galaxies from the Sloan Digital Sky Survey”, *Astrophys. J.*, v. 670, n. 2 (dez.), pp. 1048–1055. doi: 10.1086/522573.
- Koester, B. P., McKay, T. A., Annis, J., et al., 2007, “A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey”, *Astrophys. J.*, v. 660, n. 1 (maio), pp. 239–255. doi: 10.1086/509599.
- Lahav, O., Naim, A., Buta, R. J., et al., 1995, “Galaxies, Human Eyes, and Artificial Neural Networks”, *Science*, v. 267, n. 5199 (fev.), pp. 859–862. doi: 10.1126/science.267.5199.859.
- Lammers, C., Hill, R., Lim, S., et al., 2022, “Candidate high-redshift protoclusters and lensed galaxies in the Planck list of high-z sources overlapping with Herschel-SPIRE imaging”, *Mon. Not. Roy. Astron. Soc.*, v. 514, n. 4 (ago.), pp. 5004–5023. doi: 10.1093/mnras/stac1555.

- Lazar, I., Kaviraj, S., Martin, G., et al., 2023, “Relaxed blue ellipticals: accretion-driven stellar growth is a key evolutionary channel for low mass elliptical galaxies”, *Mon. Not. Roy. Astron. Soc.*, v. 520, n. 2 (abr.), pp. 2109–2120. doi: 10.1093/mnras/stad224.
- Miller, C. J., Nichol, R. C., Reichart, D., et al., 2005, “The C4 Clustering Algorithm: Clusters of Galaxies in the Sloan Digital Sky Survey”, *Astron. J.*, v. 130, n. 3 (set.), pp. 968–1001. doi: 10.1086/431357.
- Muldrew, S. I., Hatch, N. A., Cooke, E. A., 2018, “Galaxy evolution in protoclusters”, *Mon. Not. Roy. Astron. Soc.*, v. 473, n. 2 (jan.), pp. 2335–2347. doi: 10.1093/mnras/stx2454.
- Narwal, T., Goto, T., Hashimoto, T., et al., 2020, “Discovery of a Supercluster Candidate at  $z \sim 1.1$ ”, *The Astrophysical Journal Letters*, v. 895, n. 2, pp. L23.
- Oke, J. B., Sandage, A., 1968, “Energy Distributions, K Corrections, and the Stebbins-Whitford Effect for Giant Elliptical Galaxies”, *Astrophys. J.*, v. 154 (out.), pp. 21. doi: 10.1086/149737.
- Overzier, R. A., 2016, “The realm of the galaxy protoclusters. A review”, *Astron. Astrophys. Rev.*, 24(1):14. doi: 10.1007/s00159-016-0100-3.
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al., 2011, “Scikit-learn: Machine Learning in Python”, *Journal of Machine Learning Research*, v. 12, pp. 2825–2830.
- Poggianti, B. M., 1997, “K and evolutionary corrections from UV to IR”, *Astron. Astrophys. Suppl.*, v. 122 (maio), pp. 399–407. doi: 10.1051/aas:1997142.
- Reshetnikov, V. P., Marchuk, A. A., Chugunov, I. V., et al., 2023, “Evolution of the spiral structure of galaxies from the HST COSMOS field”, *arXiv e-prints*, art. arXiv:2302.02366. doi: 10.48550/arXiv.2302.02366.
- Ribeiro, A. L. B., Nascimento, R. S., Morell, D. F., et al., 2023, “Late growth of early-type galaxies in low- $z$  massive clusters”, *Mon. Not. Roy. Astron. Soc.*, v. 521, n. 1 (maio), pp. 1221–1232. doi: 10.1093/mnras/stad468.
- Rodríguez, S., Padilla, N. D., 2013, “The intrinsic shape of galaxies in SDSS/Galaxy Zoo”, *Mon. Not. Roy. Astron. Soc.*, v. 434, n. 3 (set.), pp. 2153–2166. doi: 10.1093/mnras/stt1168.
- Sánchez, C., Clampitt, J., Kovacs, A., et al., 2017, “Cosmic voids and void lensing in the Dark Energy Survey Science Verification data”, *Mon. Not. Roy. Astron. Soc.*, v. 465, n. 1 (fev.), pp. 746–759. doi: 10.1093/mnras/stw2745.

- Sánchez, C., Alarcon, A., Bernstein, G. M., et al., 2022, “The Dark Energy Survey Year 3 high redshift sample: Selection, characterization and analysis of galaxy clustering”, *Monthly Notices of the Royal Astronomical Society*, (November).
- Scoville, N., Aussel, H., Brusa, M., et al., 2007, “The Cosmic Evolution Survey (COSMOS): Overview\*”, *The Astrophysical Journal Supplement Series*, v. 172, n. 1 (sep), pp. 1. doi: 10.1086/516585. Disponível em: <<https://dx.doi.org/10.1086/516585>>.
- Sérsic, J. L., 1963, “Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy”, *Boletín de la Asociación Argentina de Astronomía La Plata Argentina*, v. 6 (fev.), pp. 41–43.
- Sochting, I. K., Coldwell, G. V., Clowes, R. G., et al., 2012, “VizieR Online Data Catalog: CES ultra deep catalogue of galaxy structures (Sochting+, 2012)”, *VizieR Online Data Catalog*, art. J/MNRAS/423/2436.
- Toshikawa, J., Kashikawa, N., Overzier, R., et al., 2016, “A Systematic Survey of Proto-clusters at  $z \sim 3-6$  in the CFHTLS Deep Fields”, *Astrophys. J.*, 826(2):114. doi: 10.3847/0004-637X/826/2/114.
- Toshikawa, J., Wuyts, S., Kashikawa, N., et al., 2024, “An enhanced abundance of bright galaxies in protocluster candidates at  $z \sim 3-5$ ”, *Mon. Not. Roy. Astron. Soc.*, v. 527, n. 3 (jan.), pp. 6276–6291. doi: 10.1093/mnras/stad3162.
- Vanderplas, J., Connolly, A., Ivezić, Ž., et al., 2012, “Introduction to astroML: Machine learning for astrophysics”. In: *Conference on Intelligent Data Understanding (CIDU)*, pp. 47–54, oct. doi: 10.1109/CIDU.2012.6382200.
- Vincent, R. A., Ryden, B. S., 2005, “The Dependence of Galaxy Shape on Luminosity and Surface Brightness Profile”, *Astrophys. J.*, v. 623, n. 1 (abr.), pp. 137–147. doi: 10.1086/428765.
- Wen, Z. L., Han, J. L., 2011, “Galaxy Clusters at High Redshift and Evolution of Brightest Cluster Galaxies”, *Astrophys. J.*, 734(1):68. doi: 10.1088/0004-637X/734/1/68.
- Wen, Z. L., Han, J. L., Liu, F. S., 2012, “A Catalog of 132,684 Clusters of Galaxies Identified from Sloan Digital Sky Survey III”, *Astrophys. J. Suppl.*, 199(2):34. doi: 10.1088/0067-0049/199/2/34.
- York, D. G., Adelman, J., Anderson, John E., J., et al., 2000, “The Sloan Digital Sky Survey: Technical Summary”, *Astron. J.*, v. 120, n. 3 (set.), pp. 1579–1587. doi: 10.1086/301513.

# Apêndice A

## Busca com passo 0,1, sem interseção

Em uma etapa inicial, por questão de aprendizado de implementação da técnica de busca por sobredensidade de galáxias, foi aplicado um passo de 0,1 para os *bins* de *redshift*, sem interseção entre eles.

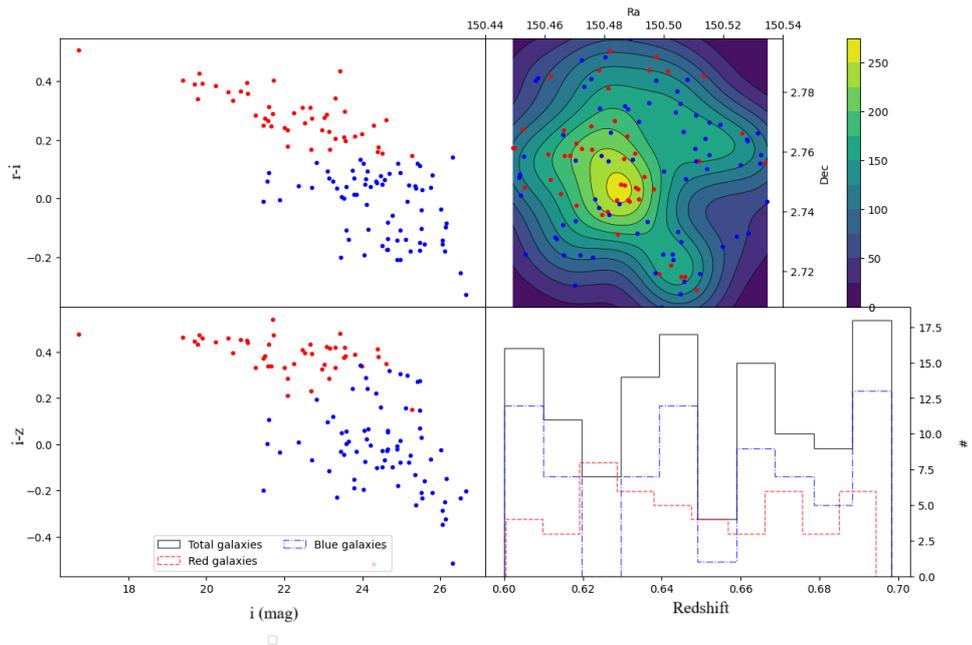


Figura A.1. As imagens superior e inferior da coluna da esquerda apresentam dados de  $r-i$  (mag) e  $i-z$  (mag) em função de  $i$  (mag), respectivamente. As imagens superior e inferior da coluna da direita apresentam o *countour plot* do candidato e o histograma de *redshift* das galáxias membros, referentes ao candidato de ID 12 e *redshift* 0,68.

Tendo escolhido o tamanho do pixel de 2 Mpc, foram analisados os *matches* entre os candidatos obtidos no presente trabalho e da literatura com resultados apresentados nas Figuras A.2, A.3, A.4 e A.5. Pode-se observar nestas figuras os *match* dos resultados deste trabalho com relação a literatura analisada nos painéis da esquerda e as diferenças de  $dz$  nas imagens da direita.

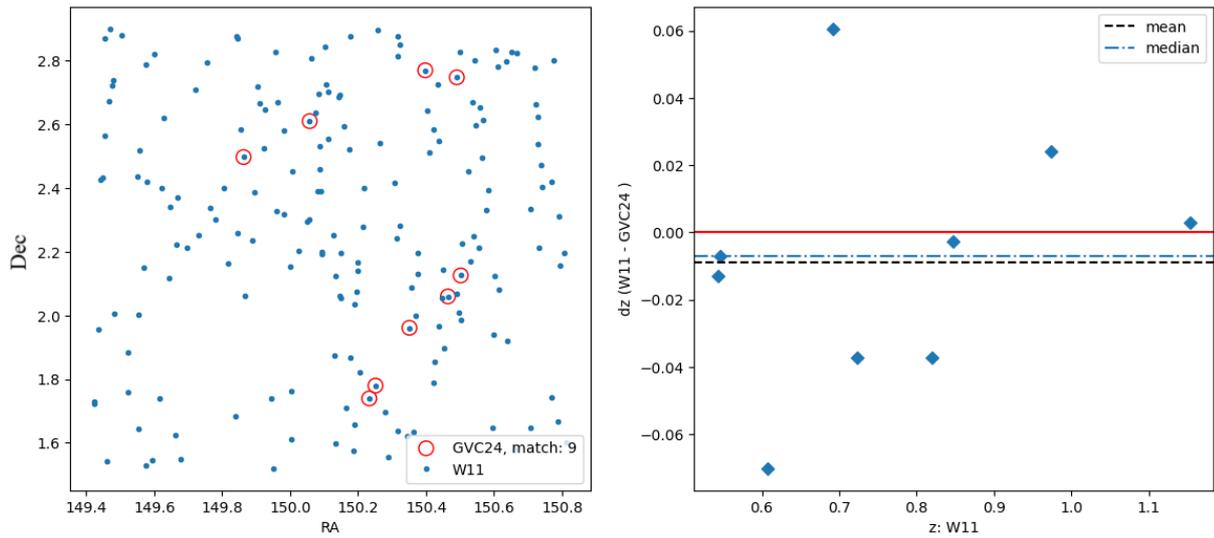


Figura A.2. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Wen e Han (2011) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os *clusters* de Wen e Han (2011), a linha ponto-tracejada azul é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

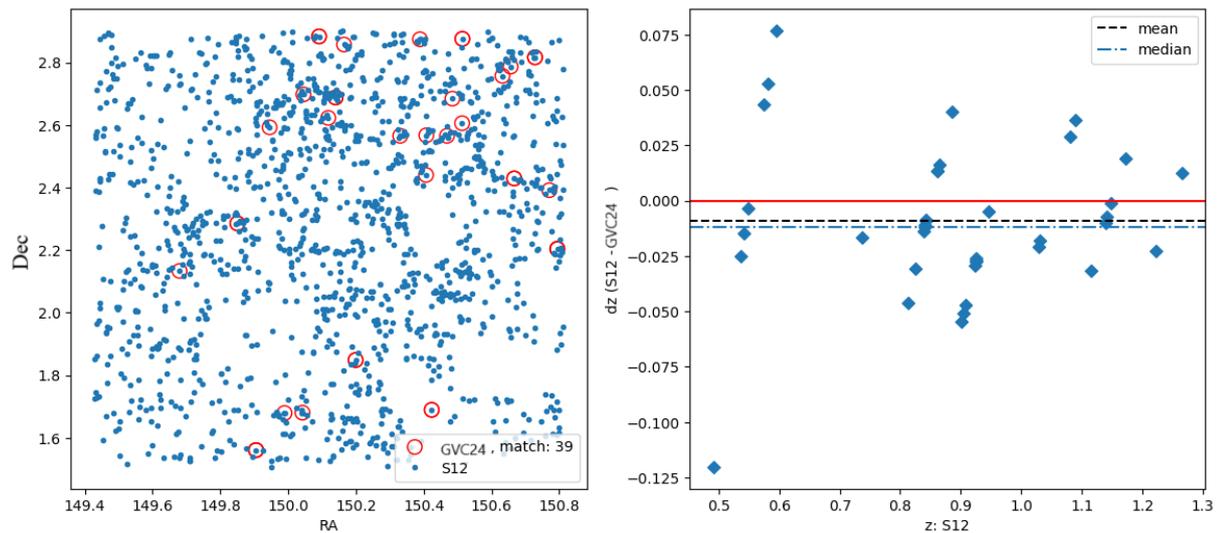


Figura A.3. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Sochting *et al.* (2012) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os *clusters* de Sochting *et al.* (2012), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

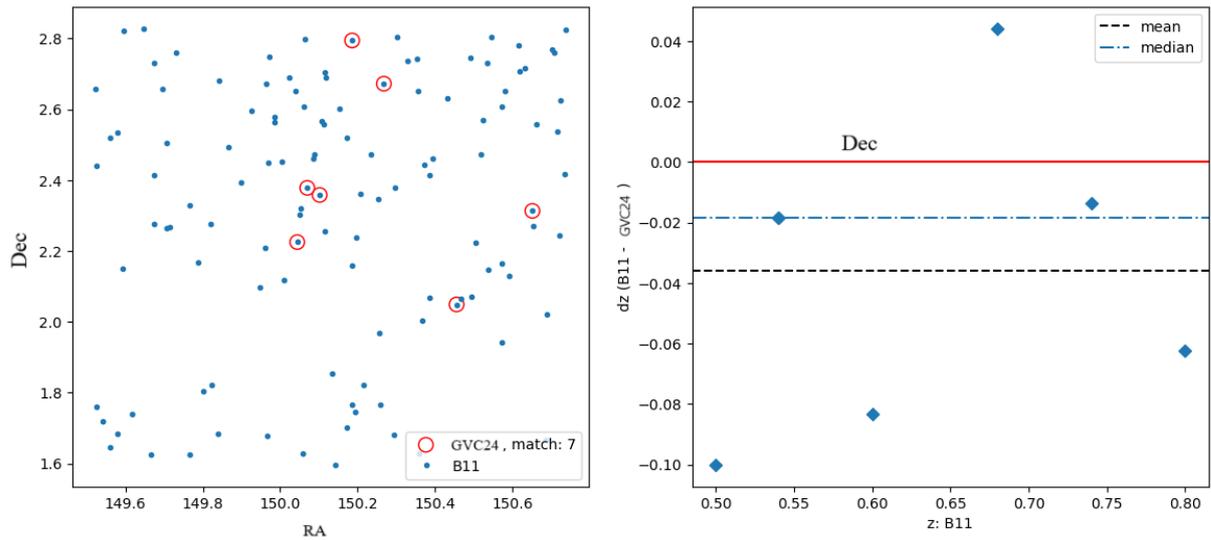


Figura A.4. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Bellagamba *et al.* (2011) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os *clusters* de Bellagamba *et al.* (2011), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .

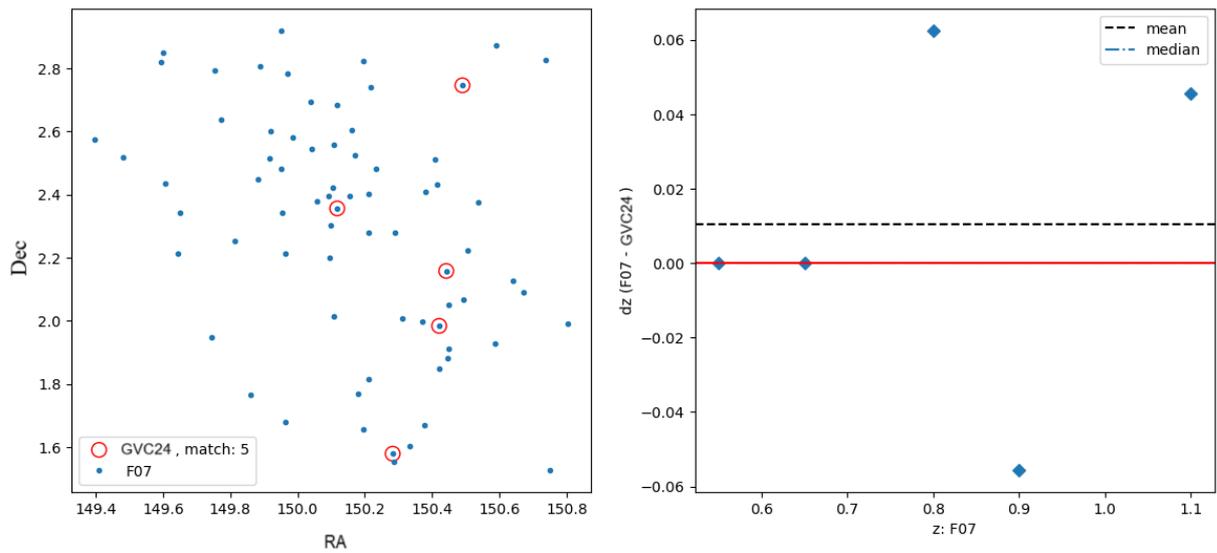


Figura A.5. O painel da direita apresenta os *matches* entre o presente trabalho (círculos vermelhos) e o catálogo de aglomerados de Finoguenov *et al.* (2007) (pontos azuis). O painel da esquerda apresenta  $dz$  vs  $z$  dos candidatos com os *clusters* de Finoguenov *et al.* (2007), a linha ponto-tracejada preta é a mediana de  $dz$ , a linha tracejada preta é a média de  $dz$ .



# Apêndice B

## Dados de GMM e CMD dos candidatos inéditos de G24

Segue o link para acesso de todos os candidatos inéditos presentes neste trabalho:

[https://github.com/GabrielCalcada/Apendice\\_Dissertacao\\_Mestrado/blob/main/Apendice\\_B.pdf](https://github.com/GabrielCalcada/Apendice_Dissertacao_Mestrado/blob/main/Apendice_B.pdf)