

Natanael Magalhães Cardoso

Sistema Inteligente para Busca de Objetos Astronômicos por Similaridade Visual usando Aprendizagem Profunda

São Paulo, SP

2024

Natanael Magalhães Cardoso

**Sistema Inteligente para Busca de Objetos
Astronômicos por Similaridade Visual usando
Aprendizagem Profunda**

Trabalho de conclusão de curso apresentado ao Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro.

Universidade de São Paulo – USP
Escola Politécnica
Departamento de Engenharia de Computação e Sistemas Digitais (PCS)

Orientador: Prof. Dr. António Mauro Saraiva
Coorientador: Profa. Dra. Cláudia Mendes de Oliveira

São Paulo, SP
2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação-na-publicação

Cardoso, Natanael Magalhães

Sistema Inteligente para Busca de Objetos Astronômicos por Similaridade Visual usando Aprendizagem Profunda / N. M. Cardoso -- São Paulo, 2024.

122 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1.aprendizagem profunda 2.redes neurais 3.mineração de dados 4.bancos de dados 5.galáxias I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

*Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.*

Agradecimentos

Este projeto foi financiado com uma bolsa CAPES/CNPq.

Foram utilizados votos de centenas de milhares de voluntários em seis campanhas no projeto de ciência cidadã Galaxy Zoo.

Os modelos de aprendizagem profunda foram treinados utilizando recursos computacionais do Centro Brasileiro de Pesquisas Físicas (CBPF) e do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo (IAG-USP).

Os Legacy Surveys consistem em três projetos individuais e complementares: o Dark Energy Camera Legacy Survey (DECaLS; ID da proposta nº 2014B-0404; PIs: David Schlegel e Arjun Dey), o Beijing-Arizona Sky Survey (BASS; ID da prop. NOAO nº 2015A-0801; PIs: Zhou Xu e Xiaohui Fan) e o Mayall z-band Legacy Survey (MzLS; ID da prop. nº 2016A-0453; PI: Arjun Dey). DECaLS, BASS e MzLS juntos incluem dados obtidos, respectivamente, no telescópio Blanco, Observatório Interamericano Cerro Tololo, NOIRLab da NSF; o telescópio Bok, Observatório Steward, Universidade do Arizona; e o telescópio Mayall, Observatório Nacional Kitt Peak, NOIRLab. O processamento do pipeline e as análises dos dados foram apoiados pelo NOIRLab e pelo Laboratório Nacional Lawrence Berkeley (LBNL).

Resumo

Com o aumento exponencial na geração de dados astronômicos, impulsionado por levantamentos astronômicos modernos, a análise manual tornou-se inviável, demandando soluções computacionais eficientes. O objetivo principal foi criar uma ferramenta que automatizasse a identificação e a análise de objetos celestes com propriedades visuais semelhantes, promovendo avanços na pesquisa astronômica. O sistema proposto baseia-se em redes neurais convolucionais (CNNs) treinadas com imagens astronômicas, utilizando a transferência de aprendizado para otimizar a extração de representações visuais (também chamadas de *embeddings*). Essas representações são vetores que codificam as características morfológicas e estruturais das galáxias, permitindo a busca e a recuperação de imagens semelhantes com alta precisão. A arquitetura do sistema inclui uma base de dados escalável para o armazenamento e indexação dos vetores de representação, além de uma interface web responsiva que facilita o uso por pesquisadores. Adicionalmente, o sistema incorpora técnicas de pré-processamento, como ajuste de campo de visão angular e aumento de dados, para garantir a qualidade e a representatividade dos conjuntos de treinamento. Os resultados obtidos evidenciam a eficácia do sistema em identificar padrões visuais entre galáxias, com avaliações robustas utilizando métricas como precisão média (mAP) e matrizes de confusão. Exemplos de buscas realizadas mostraram a aplicabilidade prática do sistema, que é capaz de entregar resultados relevantes mesmo em cenários desafiadores. Além disso, a interface gráfica desenvolvida provou-se intuitiva e eficiente, permitindo que usuários com diferentes níveis de expertise interajam com o sistema e explorem grandes volumes de dados de forma acessível. Por fim, o trabalho destaca também a importância da interdisciplinaridade entre ciência da computação e astronomia. A integração de técnicas de visão computacional, ciência de dados e práticas de engenharia de software resultou em uma solução que atende às demandas científicas e técnicas do domínio astronômico.

Palavras-chave: aprendizagem profunda. visão computacional. astronomia.

Abstract

With the exponential increase in astronomical data generation, driven by modern astronomical surveys, manual analysis has become unfeasible, demanding efficient computational solutions. The main objective was to create a tool that would automate the identification and analysis of celestial objects with similar visual properties, promoting advances in astronomical research. The proposed system is based on convolutional neural networks (CNNs) trained with astronomical images, using transfer learning to optimize the extraction of visual representations (also called embeddings). These representations are vectors that encode the morphological and structural characteristics of galaxies, allowing the search and retrieval of similar images with high precision. The system architecture includes a scalable database for storing and indexing representation vectors, as well as a responsive web interface that facilitates use by researchers. Additionally, the system incorporates preprocessing techniques, such as angular field of view adjustment and data augmentation, to ensure the quality and representativeness of the training sets. The results obtained demonstrate the effectiveness of the system in identifying visual patterns between galaxies, with robust evaluations using metrics such as mean precision (mAP) and confusion matrices. Examples of searches carried out showed the practical applicability of the system, which is capable of delivering relevant results even in challenging scenarios. In addition, the developed graphical interface proved to be intuitive and efficient, allowing users with different levels of expertise to interact with the system and explore large volumes of data in an accessible way. Finally, the work also highlights the importance of interdisciplinarity between computer science and astronomy. The integration of computer vision techniques, data science, and software engineering practices resulted in a solution that meets the scientific and technical demands of the astronomical domain.

Keywords: deep learning. computer vision. astronomy.

Listas de ilustrações

Figura 1 – Esquema de classificação de galáxias Hubble	24
Figura 2 – Cobertura espacial e espectral do levantamento Galex	30
Figura 3 – Cobertura espacial e espectral do levantamento Legacy	31
Figura 4 – Cobertura espacial e espectral do levantamento UKiDSS	31
Figura 5 – Cobertura espacial e espectral do levantamento VHS	32
Figura 6 – Cobertura espacial e espectral do levantamento WISE	32
Figura 7 – Fator de correção (η) ajustado por inspeção visual	35
Figura 8 – Ajuste do campo de visão angular para r_{mag} entre 8 e 13	37
Figura 9 – Ajuste do campo de visão angular para r_{mag} entre 14 e 19	38
Figura 10 – Diagrama de sequência da arquisição das imagens	39
Figura 11 – Pipeline de produção do catálogo de referência	41
Figura 12 – Distribuição da magnitude na banda r	41
Figura 13 – Distribuição do campo de visão angular	42
Figura 14 – Pipeline de produção do catálogo de referência	43
Figura 15 – Área de cobertura espacial do conjunto de inferência	44
Figura 16 – Fluxograma do treinamento do modelo	48
Figura 17 – Aumento artificial dos dados	49
Figura 18 – Treinamento do modelo	57
Figura 19 – Matrizes de confusão para cada questão no conjunto de teste . .	61
Figura 20 – Matrizes de confusão para cada questão no conjunto de teste (continuação)	62

Lista de tabelas

Tabela 1 – Número de objetos por campanha	30
Tabela 2 – Intervalos de busca dos hiperparâmetros otimizados com o MOTPE	58
Tabela 3 – Estrutura da matriz de confusão para classificação binária	59
Tabela 4 – Avaliação do modelo no conjunto de teste	60

List of abbreviations and acronyms

ADQL	Astronomical Data Query Language (Linguagem de Consulta de Dados Astronômicos)
API	Application Programming Interface (Interface de Programação de Aplicativos)
CBIR	Content-Based Image Retrieval (Recuperação de Imagem Baseada em Conteúdo)
CNN	Convolutional Neural Network (Rede Neural Convolucional)
CS	Citizen Science (Ciência Cidadã)
CV	Computer Vision (Visão Computacional)
Dec	Declination (Declinação, latitude)
DL	Deep Learning (Aprendizagem Profunda)
DR	Data Release (Liberação dos Dados)
FITS	Flexible Image Transport System (Sistema de Transporte de Imagem Flexível)
FLOPS	Floating-point Operations Per Second (Operações de Ponto flutuante Por Segundo)
FOV	Field of View (Campo de Visão)
GUI	Graphical User Interface (Interface Gráfica de Usuário)
IVOA	International Virtual Observatory Alliance (Aliança Internacional de Observatórios Virtuais)
LSST	Legacy Survey of Space and Time (Levantamento Legado do Espaço e do Tempo)
ML	Machine Learning (Aprendizagem de Máquina)
RA	Right Ascension (Ascenção Reta, longitude)
SDSS	Sloan Digital Sky Survey

S-PLUS	Southern Photometric Local Universe Survey (Levantamento Fotométrico do Universo Local do Sul)
SQL	Structured Query Language (Linguagem de Consulta Estruturada)
VO	Virtual Observatory (Observatório Virtual)

Listade símbolos

α Ascenção Reta (longitude)

δ Declinação (latitude)

R_e Raio efetivo da galáxia

r_{mag} Magnitude na banda r

Sumário

I Ontologia e Classificação Morfológica	21
1 Classificação de objetos	23
1.1 Introdução	23
1.2 Representação Morfológica por Ontologia	24
1.3 Métodos de Classificação Morfológica	25
1.3.1 Anotação Manual por Especialistas	25
1.3.2 Ciência Cidadã	26
1.3.3 Métodos Computacionais	27
2 Conjuntos de Dados	29
2.1 Metodologia	29
2.1.1 Aquisição dos Votos do GalaxyZoo	29
2.1.2 Seleção dos Levantamentos Astronômicos	30
2.1.2.1 Galex	30
2.1.2.2 Legacy Survey	31
2.1.2.3 United Kingdom Digital Sky Survey	31
2.1.2.4 Vista Hemisphere Survey	31
2.1.2.5 Wide-field Infrared Survey Explorer	31
2.1.3 Ajuste do Campo de Visão Angular	31
2.1.4 Aquisição das Imagens	36
2.2 Conjunto de Referência	40
2.2.1 Pipeline de Dados	40
2.2.2 Descrição	40
2.2.2.1 Conjuntos de Treinamento, Validação e Teste	40
2.2.2.2 Conjunto de Inferência	42
2.2.3 Pipeline de Dados	42
2.2.4 Descrição	43
3 Classificação automática	45
3.1 Introdução	45
3.1.1 Redes Neurais	46
3.1.2 Aprendizado Profundo em Visão Computacional	46
3.2 Modelo de Aprendizagem Profunda	47
3.2.1 Preparação das Características	47
3.2.1.1 Preparação das Características de Entrada	48
3.2.1.2 Preparação dos Rótulos	48
3.2.2 Aumento Artificial de Dados	49

3.2.3	Função de Custo	50
3.2.3.1	Distribuição Multinomial	50
3.2.3.2	Distribuição de Dirichlet	51
3.2.3.3	Distribuição Dirichlet-Multinomial	52
3.2.3.4	Função de Custo e Função de Perda	52
3.2.4	Hiperparâmetros	53
3.2.5	Arquitetura de Rede Neural	55
3.2.5.1	Visual Geometry Group Networks	55
3.2.5.2	InceptionResNetV2	55
3.2.5.3	EfficientNet	55
3.2.5.4	DenseNet	56
3.2.5.5	Escolha da Arquitetura	56
3.2.6	Treinamento	56
3.2.7	Métricas de Avaliação da Predição dos Votos	58
3.2.7.1	Matriz de Confusão	58
3.2.7.2	Acurácia	59
3.2.7.3	Precisão	59
3.2.7.4	Revocação	59
3.2.7.5	F1-Score	60
3.3	Avaliação do Modelo no Conjunto de Teste	60
Referências	65

Parte I

Ontologia e Classificação Morfológica

1

Classificação de objetos

Visão geral do capítulo

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1.1 Introdução

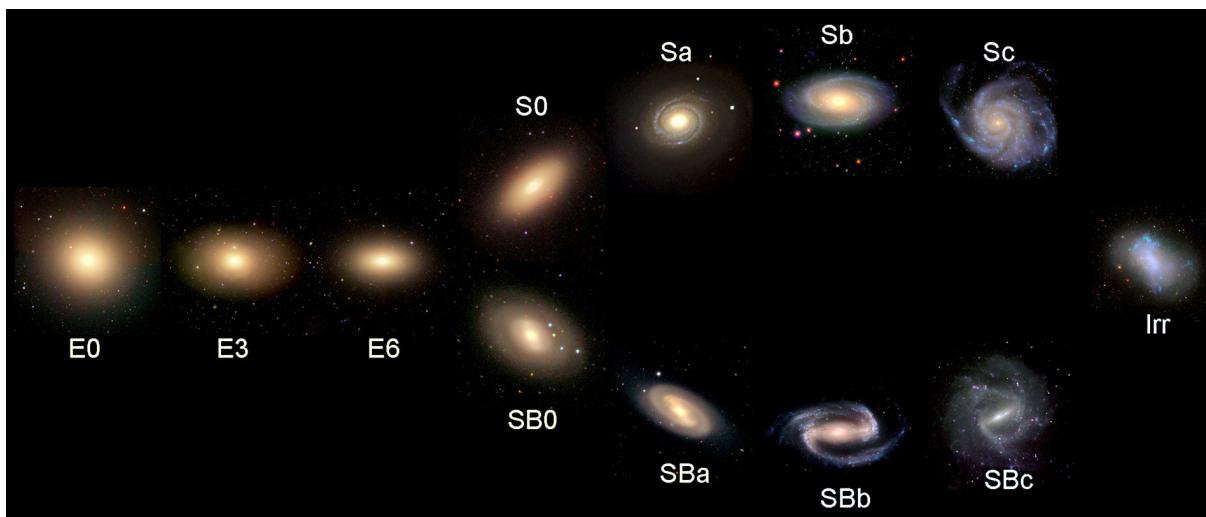
Na astronomia, morfologia refere-se ao estudo das formas e das estruturas observáveis dos objetos celestes, como galáxias, nebulosas, estrelas e aglomerados estelares (BUTA, 2011). A análise morfológica busca descrever e categorizar as características visuais desses objetos, identificando padrões estruturais que possam fornecer percepções sobre seus processos de formação e evolução (STEINMETZ; NAVARRO, 2002). Em particular, as galáxias são frequentemente classificadas com base em sua morfologia. Essas classificações refletem características como a distribuição de estrelas, a presença de braços espirais e a concentração de massa no núcleo, o que está diretamente associado à dinâmica interna e ao histórico evolutivo de cada galáxia (BERGH, 1998).

A classificação morfológica é um sistema de categorização utilizado em astronomia para descrever e agrupar galáxias com base em suas características estruturais e visuais. Esse processo é fundamentado em características observáveis, como a forma, a distribuição de luminosidade, a presença de braços espirais ou um núcleo central proeminente (STEINMETZ; NAVARRO, 2002). A classificação

morfológica permite organizar galáxias em grupos que compartilham propriedades físicas e evolutivas semelhantes, facilitando o estudo de sua formação, dinâmica e história evolutiva. Esse sistema é especialmente valioso na astronomia observational, onde padrões morfológicos podem ser correlacionados com fenômenos físicos específicos, como fusões e interações gravitacionais entre galáxias (BERGH, 1998).

Historicamente, o primeiro sistema de classificação morfológica foi desenvolvido por Hubble (1926), que introduziu o “diagrama de diapasão” para agrupar galáxias em três principais categorias: elípticas, espirais e irregulares. Esse esquema básico de Hubble, também chamado de sequência de Hubble, organizava as galáxias de acordo com sua forma e estrutura, identificando uma possível progressão evolutiva entre os tipos. As galáxias elípticas, por exemplo, são caracterizadas por uma distribuição suave e arredondada de luz e pouca formação estelar, enquanto as espirais possuem braços bem definidos e núcleos brilhantes, indicando regiões de intensa formação de estrelas. As galáxias irregulares, por sua vez, apresentam formas menos definidas e são frequentemente associadas a perturbações gravitacionais.

Figura 1 – Esquema de classificação de galáxias Hubble



Fonte: European Southern Observatory (ESO). Disponível em: <https://supernova.eso.org/exhibition/images/1015_fork-1920/>

1.2 Representação Morfológica por Ontologia

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus

et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

1.3 Métodos de Classificação Morfológica

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1.3.1 Anotação Manual por Especialistas

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc.

Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

No entanto, embora preciso, este método é ineficiente por conta da disparidade entre a capacidade de classificações manuais feitas por especialistas e a quantidade de objetos no Universo.

1.3.2 Ciência Cidadã

Um método de classificação mais eficiente é a utilização da ciência cidadã, que é um campo interdisciplinar emergente que envolve a participação ativa do público geral em tarefas de pesquisas científicas com finalidade de produzir novos conhecimentos para a ciência e para a sociedade (VOHLAND et al., 2021), especialmente na coleta, categorização, transcrição e análise de dados científicos (SILVERTOWN, 2009; BONNEY et al., 2014). Este modelo de pesquisa tem ganhado relevância em diversos campos da ciência, principalmente devido ao aumento do acesso a tecnologias digitais e à internet, que facilitam a comunicação e a organização entre cidadãos e cientistas (LEMMENS et al., 2021).

O GalaxyZoo¹ (LINTOTT et al., 2008) é um projeto de ciência cidadã lançado em 2007, cujo objetivo é classificar morfológicamente galáxias utilizando imagens obtidas por grandes levantamentos astronômicos, como o Sloan Digital Sky Survey (SDSS). Os participantes, voluntários de diferentes formações e níveis de conhecimento científico, analisam imagens de galáxias e fornecem informações sobre suas características, como forma espiral, elíptica ou irregular, e a presença de estruturas específicas, como barras centrais. Essa colaboração massiva permitiu a classificação de milhões de galáxias em um curto período, superando em eficiência o que seria possível com equipes científicas tradicionais.

O impacto científico do GalaxyZoo é expressivo. Além de fornecer uma base de dados robusta e de alta qualidade para a pesquisa astronômica, o projeto

¹ <<https://galaxyzoo.org>>

gerou avanços no entendimento da formação e evolução de galáxias, incluindo a relação entre morfologia e ambiente. Os resultados têm sido utilizados para treinar modelos de aprendizado de máquina, permitindo a automação de tarefas de classificação em levantamentos futuros. O sucesso do GalaxyZoo também inspirou o desenvolvimento de outras plataformas de ciência cidadã, consolidando seu papel como ferramenta metodológica na pesquisa científica.

Para os voluntários, o GalaxyZoo oferece uma oportunidade de engajamento direto com a ciência, promovendo aprendizado e um senso de contribuição para descobertas científicas relevantes. Muitos participantes relatam um aumento na compreensão de conceitos astronômicos e motivação para explorar outras áreas da ciência. Além disso, a plataforma fomenta uma comunidade global de entusiastas que colaboram ativamente, mostrando como iniciativas bem estruturadas podem democratizar o acesso e a participação no progresso científico.

1.3.3 Métodos Computacionais

Os métodos computacionais atuais se baseiam nos métodos anteriores (Seções [1.3.1](#) e [1.3.2](#)) para treinamento de modelos.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Este será o método explorado por este trabalho.

* * *

2

Conjuntos de Dados

Visão geral do capítulo

Este capítulo detalha o método de produção dos conjuntos de dados necessários para implementação da classificação automática proposta neste trabalho. A Seção 2.1 descreve a metodologia utilizada para

2.1 Metodologia

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.1.1 Aquisição dos Votos do GalaxyZoo

Os votos do GalaxyZoo (Seção 1.3.2) são utilizados como rótulos no treinamento supervisionado do modelo de aprendizagem profunda. Para composição desses rótulos, foram utilizadas sete campanhas do GalaxyZoo: GalaxyZoo 1 ([LINTOTT et al., 2011](#)), GalaxyZoo 2 ([WILLETT et al., 2013; HART et al., 2016](#)), GalaxyZoo DECaLS ([WALMSLEY et al., 2021](#)), GalaxyZoo DESI ([WALMSLEY et al., 2023](#)), GalaxyZoo Hubble ([WILLETT et al., 2017](#)), GalaxyZoo Candels ([SIMMONS et al., 2017](#)) e GalaxyCruisers 1 (??).

Após a aquisição de todos os catálogos de votos, os catálogos foram concatenados e foi feita uma correlação com o próprio catálogo em um raio de 8 arcsec para detecção dos objetos repetidos. Essa etapa de limpeza de dados é essencial,

pois um mesmo objeto poderia estar em campanhas distintas do GalaxyZoo e, ao concatenar catálogos de diferentes campanhas, apareceriam objetos duplicados. Isso poderia causar viés na avaliação do modelo, pois, o mesmo objeto poderia aparecer no conjunto de treinamento e de teste/validação simultaneamente.

A Tabela 1 mostra as estatísticas finais por campanha, após a limpeza dos dados.

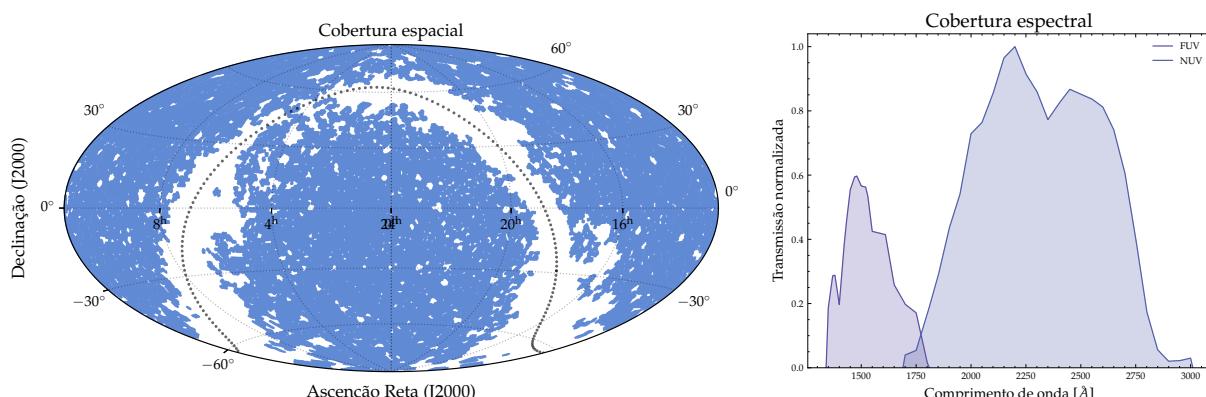
Tabela 1 – Número de objetos por campanha

Campanha	# galáxias	# alternativas
GalaxyZoo 1	93.121	6
GalaxyZoo 2	191.098	30
GalaxyZoo DESI	397.954	68
GalaxyZoo Hubble	92.851	40
GalaxyZoo Candels	35.287	32

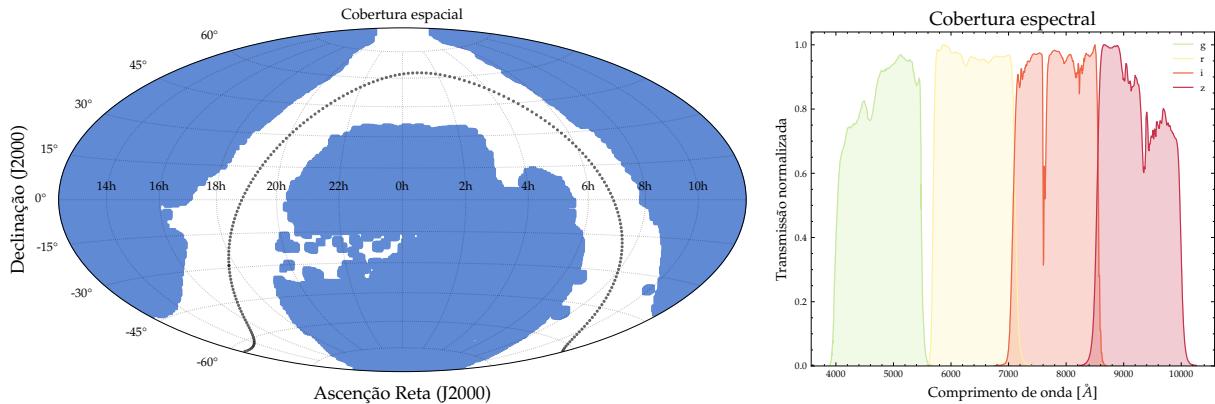
2.1.2 Seleção dos Levantamentos Astronômicos

2.1.2.1 Galex

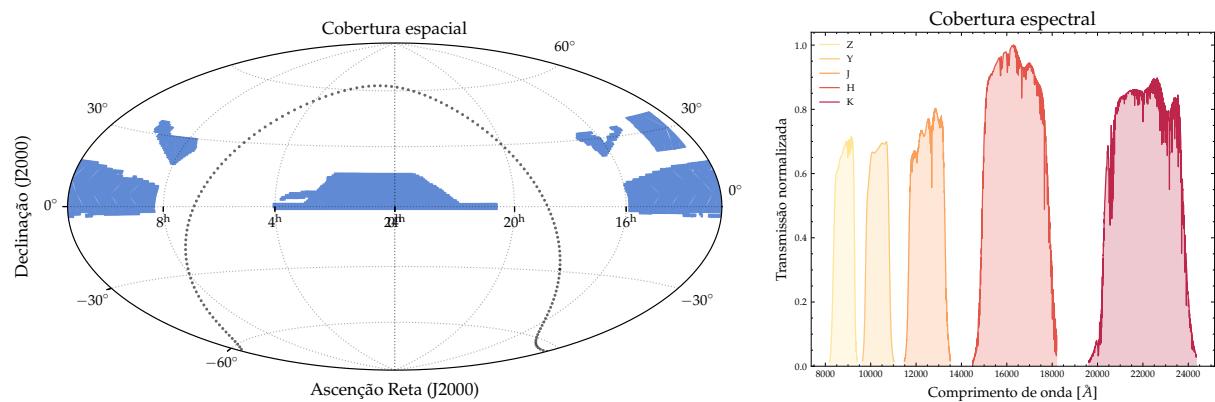
Figura 2 – Cobertura espacial e espectral do levantamento Galex



A área de cobertura espacial do levantamento é mostrada no painel à esquerda na região em azul. A cobertura espectral é mostrada no painel à direita com as curvas de transmissão total dos dois filtros NUV e FUV em função do comprimento de onda.

Figura 3 – Cobertura espacial e espectral do levantamento Legacy

A área de cobertura espacial do levantamento é mostrada no painel à esquerda na região em azul. A cobertura espectral é mostrada no painel à direita com as curvas de transmissão total dos quatro filtros g, r, i e z em função do comprimento de onda.

Figura 4 – Cobertura espacial e espectral do levantamento UKiDSS

A área de cobertura espacial do levantamento é mostrada no painel à esquerda na região em azul. A cobertura espectral é mostrada no painel à direita com as curvas de transmissão total dos cinco filtros Z, Y, J, H e K em função do comprimento de onda.

2.1.2.2 Legacy Survey

2.1.2.3 United Kingdom Digital Sky Survey

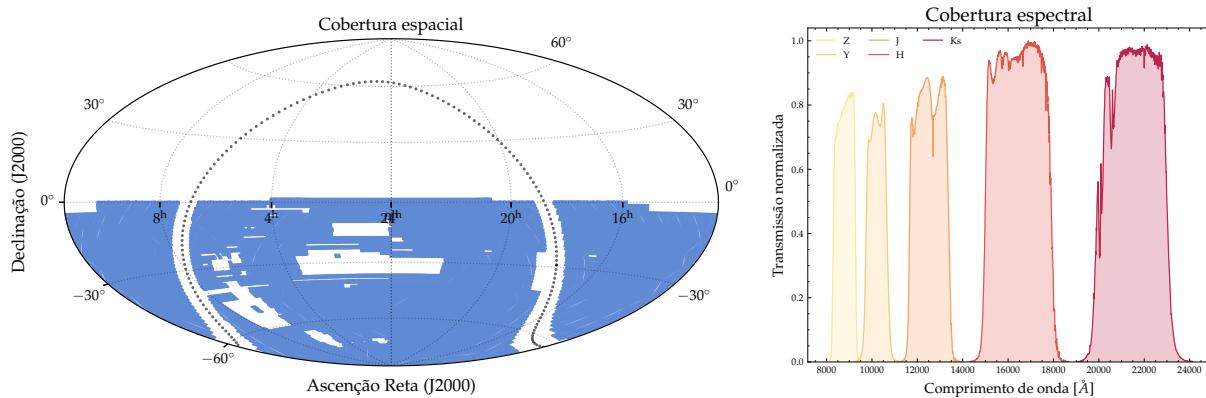
2.1.2.4 Vista Hemisphere Survey

2.1.2.5 Wide-field Infrared Survey Explorer

2.1.3 Ajuste do Campo de Visão Angular

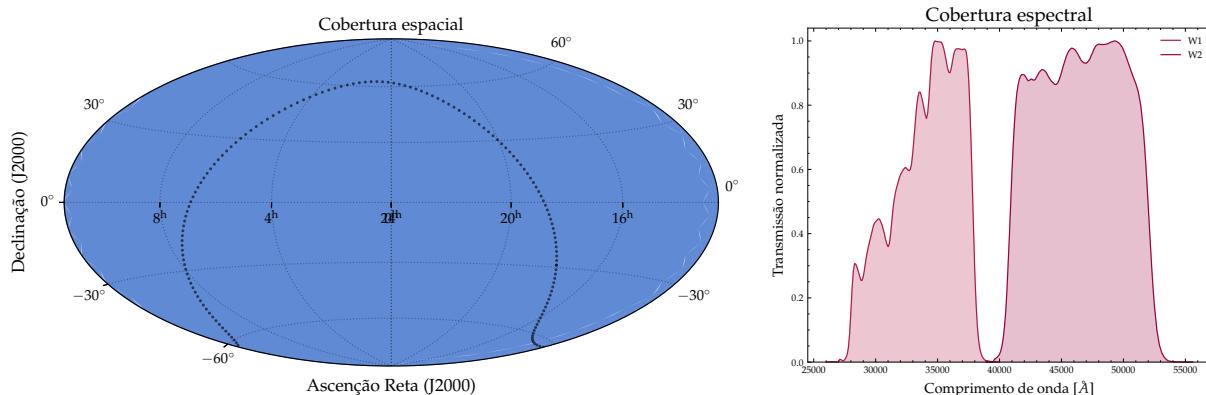
O campo de visão angular (em inglês, field of view – FoV) é a dimensão física da região observada. Essa dimensão é calculada em unidades angulares, pois todas as observações do céu são projetadas na superfície da esfera celeste (Seção

Figura 5 – Cobertura espacial e espectral do levantamento VHS



A área de cobertura espacial do levantamento é mostrada no painel à esquerda na região em azul. A cobertura espectral é mostrada no painel à direita com as curvas de transmissão total dos cinco filtros Z, Y, J, H e Ks em função do comprimento de onda.

Figura 6 – Cobertura espacial e espectral do levantamento WISE



A área de cobertura espacial do levantamento é mostrada no painel à esquerda na região em azul. A cobertura espectral é mostrada no painel à direita com as curvas de transmissão total dos dois filtros W1 e W2 em função do comprimento de onda.

??).

O ajuste do FoV em imagens de galáxias é essencial para o treinamento de modelos baseado em imagens, pois garante que a área da galáxia ocupe uma proporção ideal da imagem. Este ajuste é importante porque afeta diretamente a qualidade das features extraídas, as quais são fundamentais para que o modelo aprenda as características morfológicas das galáxias de forma precisa e consistente. Quando o FoV é ajustado para que a galáxia preencha adequadamente a imagem, o modelo consegue focar nas características principais do objeto, como forma, brilho e estrutura, independentemente de variações no tamanho e na magnitude entre diferentes galáxias.

Se o FoV for proporcionalmente muito pequeno, parte da galáxia pode ser cortada, o que resulta na perda de informações essenciais e introduz inconsistências

nas features extraídas, pois o modelo passa a “ver” apenas uma parte do objeto. Isso pode levar a erros de classificação ou ao mau desempenho na identificação de padrões morfológicos. Por outro lado, se o FoV for muito grande, a galáxia ocupa uma fração pequena da imagem, e o modelo acaba focando mais no céu ao redor, que é irrelevante para o aprendizado das características morfológicas. Neste segundo caso, a quantidade de pixels dedicados ao fundo dilui a presença da galáxia, dificultando a extração de features relevantes e potencialmente introduzindo ruído nos dados de treinamento.

Foram usadas duas abordagens distintas para estimar automaticamente o FoV para cada galáxia: uma usando o raio efetivo (Seção ??) e outra usando a elipticidade complexa (Seção ??). Em ambos os casos, os valores são corrigidos por uma função η que depende da magnitude da galáxia (Seção ??) na banda r. Essa função foi ajustada empiricamente com base na inspeção visual. As medidas de raio efetivo, elipticidade complexa e magnitude foram obtidas do catálogo fotométrico do Legacy Survey (Seção ??) acessadas a partir do AstroDataLab (<<https://datalab.noirlab.edu>>) pelo protocolo TAP (Seção ??).

A estimativa do FoV a partir do raio efetivo foi inspirada pelo método desenvolvido por Walmsley et al. (2021). Neste caso, consideramos que a região vista na figura deve conter o dobro do raio efetivo (diâmetro efetivo) corrigido por uma função η (explicada mais adiante) que depende da magnitude (r_{mag}), como mostra a eq. (2.1).

$$\text{fov}_c = 2 \cdot r_e \cdot \eta(r_{\text{mag}}) \quad (2.1)$$

A segunda abordagem consiste na estimativa do FoV a partir da elipticidade expressa pelo número complexo da eq. (2.2). O catálogo fotométrico do Legacy Survey (Seção ??) fornece os valores das componentes real (ϵ_1) e imaginária (ϵ_2).

$$\epsilon = \frac{a - b}{a + b} e^{2i\phi} = \epsilon_1 + i\epsilon_2 \quad (2.2)$$

A partir dos valores das componentes ϵ_1 e ϵ_2 , é calculado o módulo da elipticidade ($|\epsilon|$), conforme a eq. (2.3).

$$|\epsilon| = \sqrt{\epsilon_1^2 + \epsilon_2^2} \quad (2.3)$$

Com o valor do módulo, os valores dos semi-eixos maior e menor (a e b , respectivamente) e do ângulo (ϕ) da elipse são calculados pelas eqs. (2.4) e (2.5).

$$\frac{b}{a} = \frac{1 - |\epsilon|}{1 + |\epsilon|} \quad (2.4)$$

$$\phi = \frac{1}{2} \arctan \frac{\epsilon_2}{\epsilon_1} \quad (2.5)$$

O próximo passo é calcular a caixa delimitadora (*bounding box*) da elipse. Para fazer isso, usamos as equações paramétricas da elipse, conforme as eqs. (2.6) e (2.7).

$$x(t) = x_0 + \frac{a}{2} \cos t \cos \phi - \frac{b}{2} \sin t \sin \phi \quad (2.6)$$

$$y(t) = y_0 + \frac{b}{2} \sin t \cos \phi + \frac{a}{2} \cos t \sin \phi \quad (2.7)$$

Então, expressamos o parâmetro t em termos os semi-eixos e do ângulo, como as eqs. (2.8) e (2.9).

$$t_x = \arctan \left(-\frac{b \tan \phi}{a} \right) \quad (2.8)$$

$$t_y = \arctan \left(\frac{b}{a \tan \phi} \right) \quad (2.9)$$

Então, considerando o ponto (x_0, y_0) como as coordenadas do centro da imagem, em pixels, obtemos as quatro coordenadas da caixa delimitadora da elipse x_1, x_2, y_1 e y_2 , em pixels, como mostra a eq. (2.10).

$$x_1 = x(t_x) \quad x_2 = x(t_x + \pi) \quad y_1 = y(t_y) \quad y_2 = y(t_y + \pi) \quad (2.10)$$

A caixa delimitadora da elipse é um retângulo. No entanto, queremos que as imagens das galáxias sejam quadradas, então calculamos o lado do quadrado ℓ , conforme a eq. (2.11).

$$\ell = \max(|x_2 - x_1|, |y_2 - y_1|) \quad (2.11)$$

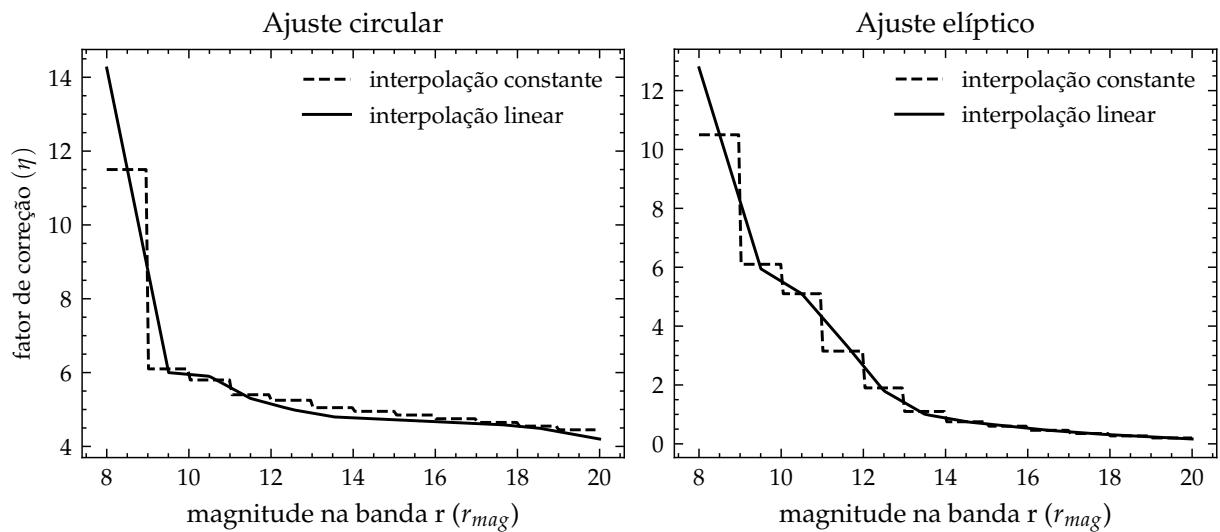
Finalmente, calculamos o FoV pela eq. (2.12)

$$\text{fov}_e = \ell \cdot \eta(r_{\text{mag}}) \quad (2.12)$$

É notável que um valor η , usado como fator de correção, foi usado em ambas as abordagens. Foi necessário porque é necessário escalar as medidas de raio efetivo e elipticidade para que seja possível usá-las como estimativa do FoV. No entanto, o valor depende da magnitude: para objetos mais brilhantes, deve-se multiplicar por um valor maior que para objetos menos brilhantes.

Desse modo, foi considerado um intervalo de r_{mag} entre 8 e 20 e foram amostradas 500 galáxias para cada intervalo de 1 mag (12 amostras). Sendo que, para os intervalos mais brilhantes, foram obtidas menos de 500 galáxias. Para cada intervalo de magnitude, foram atribuídos valores de ajuste para ambas as abordagens. Assim, a função η foi empiricamente determinada por inspeção visual de amostras aleatórias de galáxias em intervalos de magnitude, como mostra a Fig. 7. Por critério de documentação, a inspeção visual realizada está disponível na página <<https://nmcardoso.github.io/ls-stamps>>, contendo todos os ajustes para todas as galáxias inspecionadas.

Figura 7 – Fator de correção (η) ajustado por inspeção visual



O processo de ajuste do FoV é exemplificado nas figs. 8 e 9. Nelas, são mostrados os ajustes de 12 galáxias, uma para cada intervalo de magnitude, tanto para o ajuste circular (primeiro método) quanto para o ajuste elíptico (segundo método). No ajuste circular, a circunferência indica o raio efetivo da galáxia escalonado pela função η . Analogamente, no ajuste elíptico, a elipse é obtida pelos parâmetros calculados nas eqs. (2.4) e (2.5) escalonados pela função η , e o retângulo tracejado indica a caixa delimitadora da elipse, cujos limites são obtidos da eq. (2.10). Em ambas as colunas, o quadrado branco indica o FoV calculado em cada caso, de acordo com as eqs. (2.1) e (2.12). As colunas “corte circular” e “corte elíptico” mostram as figuras finais após o ajuste pelo método do raio efetivo e da elipsidate complexa, respectivamente. A região vista nessas colunas é equivalente à região delimitada pelos quadrados brancos nas colunas de ajuste. Para comparação

entre os métodos, ambas as colunas de ajuste possuem o mesmo FoV, definido manualmente para cada intervalo de magnitude.

Analisando os ajustes com o método do raio efetivo, notamos que há uma tendência dessa abordagem subestimar o FoV de galáxias grandes, distorcidas ou de sistemas de galáxias, como mostra os Objetos 3 ao 6 da Fig. 8 e o Objeto 7 da Fig. 9, além de sobre-estimar o FoV de galáxias menos brilhantes, como os Objetos 10 e 11 da Fig. 9. Por outro lado, o método da elipse pareceu mais consistente para determinação dos FoVs e por isso foi o método escolhido.

2.1.4 Aquisição das Imagens

A aquisição dos pequenos recortes de imagens astronômicas centrados em objetos específicos (stamps) do Legacy Survey por meio do serviço Hips2Fits¹ ([BONNAREL et al., 2000](https://alasky.cds.unistra.fr/hips-image-services/hips2fits)), do Strasbourg Astronomical Data Center² (CDS), envolve um processo meticuloso de coleta de imagens ajustadas para um grande número de galáxias. Considerando que os campos de visão (FOVs) de cada objeto foram previamente calculados para garantir que a área da galáxia ocupe uma proporção ideal na imagem, a aquisição pode então ser automatizada, eficiente e escalável. Para realizar essa aquisição em larga escala, que pode envolver milhões de galáxias, é fundamental utilizar uma arquitetura de paralelização eficiente, como o modelo de produtor-consumidor, combinada com mecanismos de controle de taxa de requisição, mostrado na Fig. 10.

Inicialmente, o sistema é projetado em duas camadas principais: a camada de produtores, responsáveis por gerar as requisições de imagens para cada galáxia com seu respectivo FOV, e a camada de consumidores, encarregada de processar essas requisições e armazenar os stamps retornados. No caso de uma grande quantidade de objetos, é necessário implementar várias threads que operem de forma concorrente para maximizar o desempenho e a eficiência do sistema. Os produtores criam as requisições HTTP para o Hips2Fits, especificando os parâmetros de coordenadas do objeto (como RA e Dec) e o FOV calculado. Esses parâmetros garantem que a imagem adquirida seja adequada à análise morfológica planejada.

A arquitetura paralela é gerenciada com o uso de um semáforo, um mecanismo de controle de concorrência que limita o número de requisições simultâneas feitas ao servidor Hips2Fits. Esse semáforo é configurado de acordo com a taxa de requisição máxima permitida pelo servidor, evitando sobrecarga e bloqueios temporários ou permanentes impostos pelo controle de acesso ao servidor. O

¹ [<https://alasky.cds.unistra.fr/hips-image-services/hips2fits>](https://alasky.cds.unistra.fr/hips-image-services/hips2fits)

² [<https://cds.unistra.fr>](https://cds.unistra.fr)

Figura 8 – Ajuste do campo de visão angular para r_{mag} entre 8 e 13

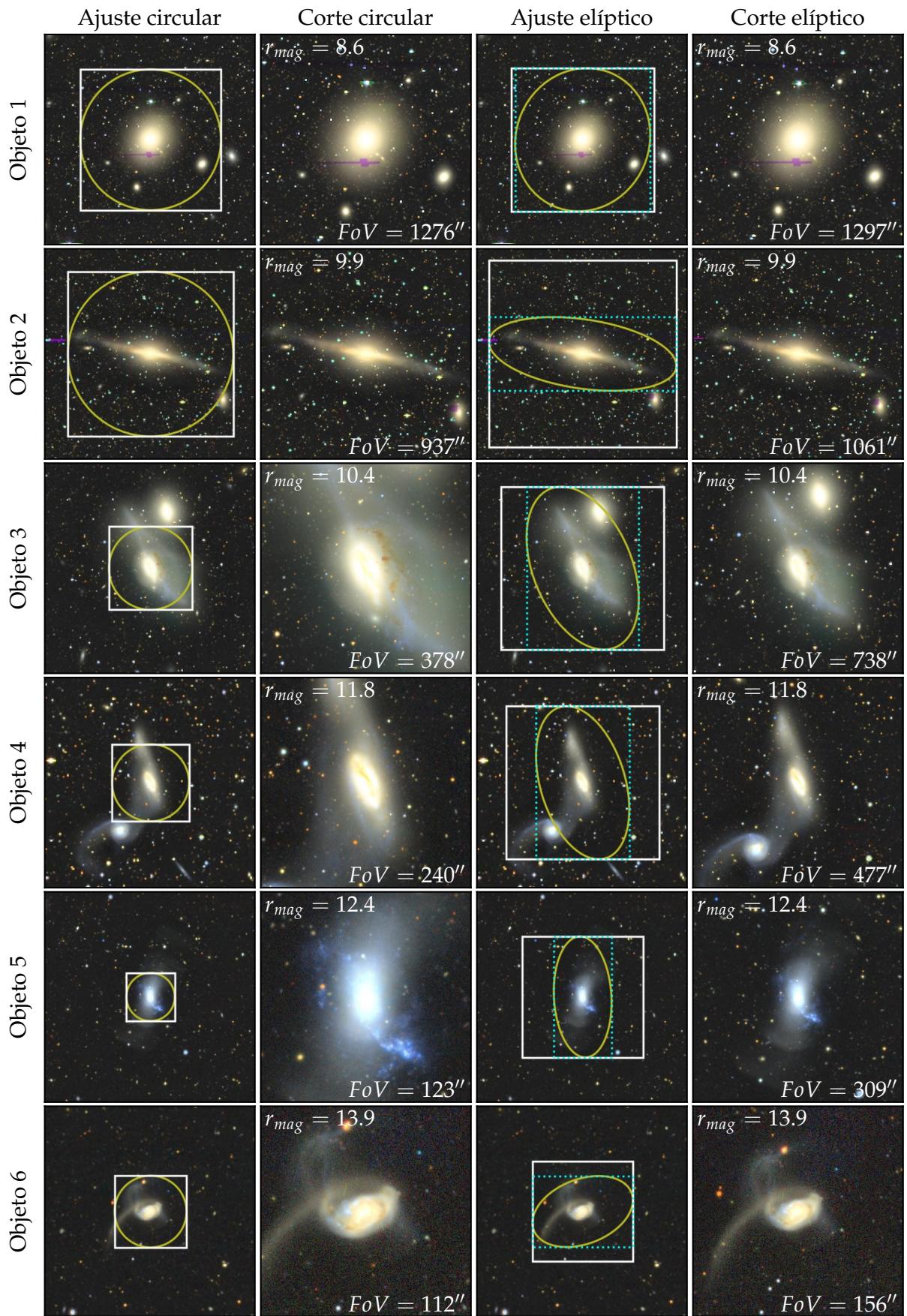


Figura 9 – Ajuste do campo de visão angular para r_{mag} entre 14 e 19

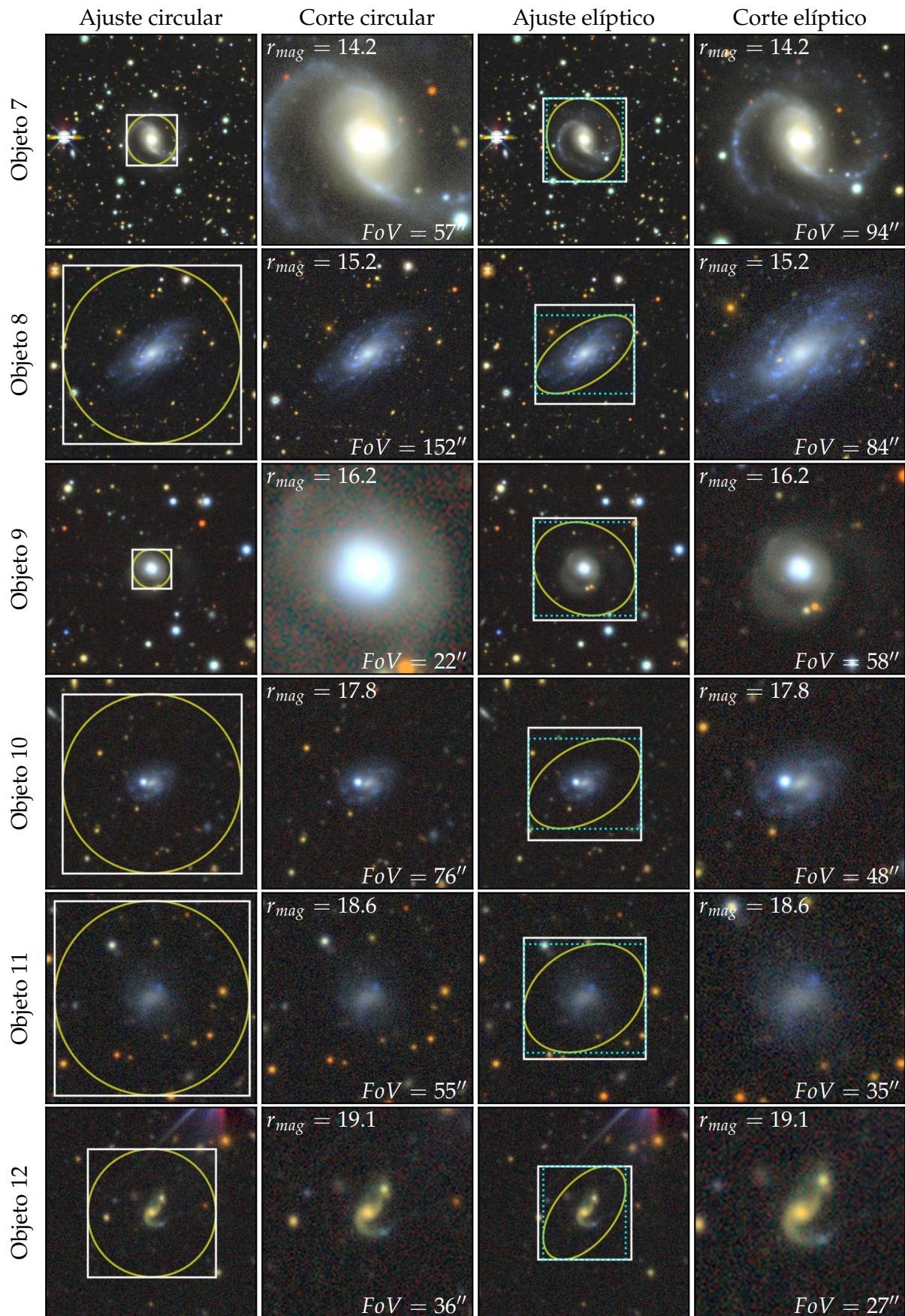
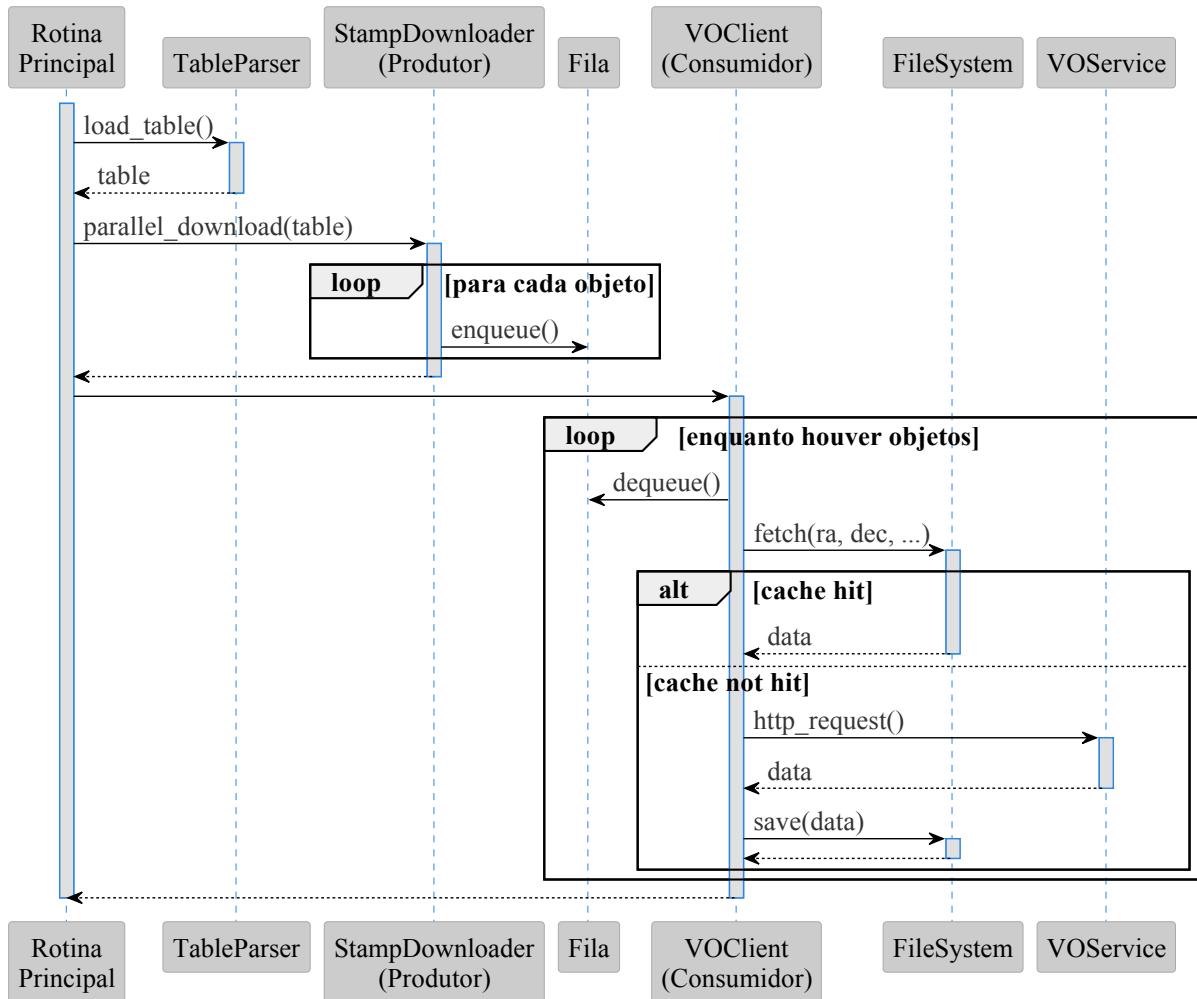


Figura 10 – Diagrama de sequência da arquisição das imagens



semáforo atua para que um número limitado de threads possa realizar requisições ao mesmo tempo; assim, ao alcançar o limite, novas requisições aguardam até que uma thread finalize seu processo e libere o semáforo para outra requisição.

Esse modelo de produtor-consumidor com controle de taxa de requisição não apenas respeita os limites de acesso do Hips2Fits, mas também assegura uma coleta eficiente e escalável. Como cada thread pode solicitar e processar imagens independentemente, o tempo total de execução é significativamente reduzido, permitindo a aquisição dos stamps para milhões de galáxias em um tempo viável. Essa abordagem de paralelização e controle de acesso torna o sistema robusto para grandes volumes de dados, possibilitando o processamento em lotes de imagens astronômicas de forma a atender à demanda de análise em larga escala na pesquisa astronômica.

2.2 Conjunto de Referência

2.2.1 Pipeline de Dados

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

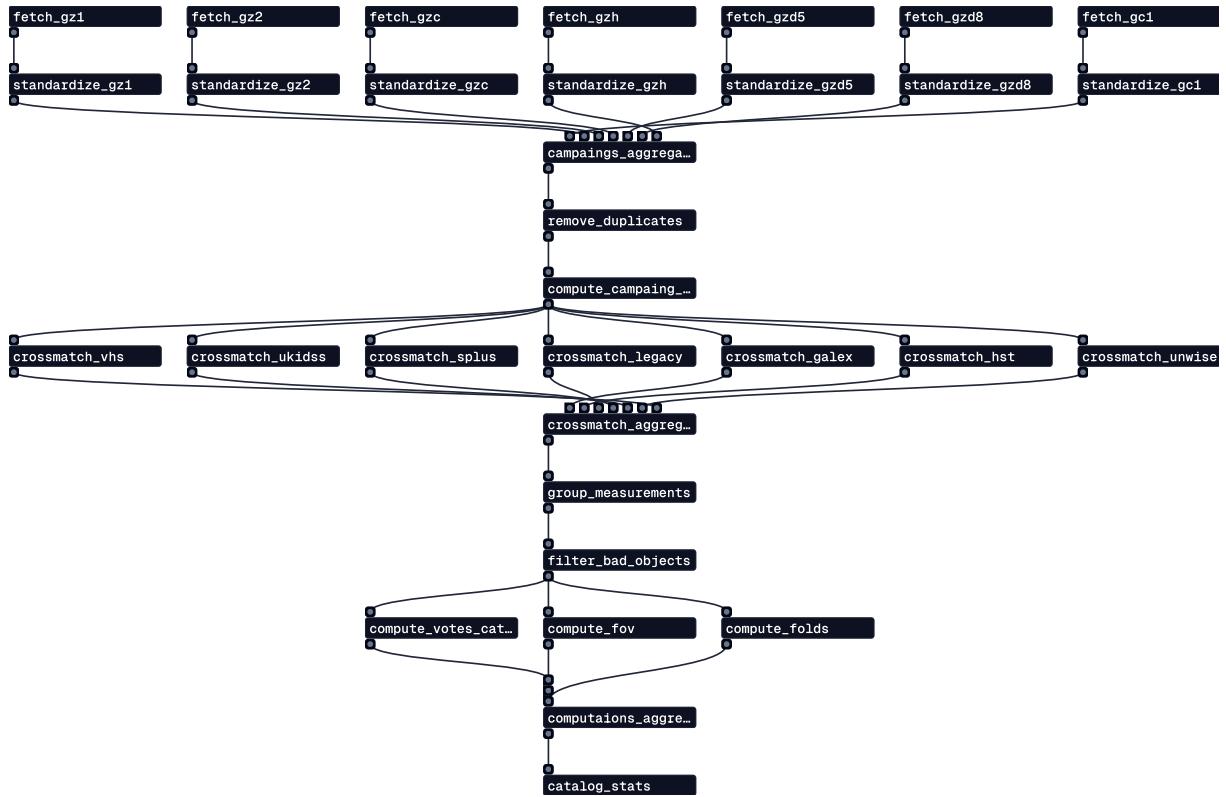
Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

2.2.2 Descrição

Nesta Seção, é feita uma análise, de forma agregada, sobre as propriedades físicas das galáxias dos conjuntos de dados para garantir a coerência no treinamento do modelo.

2.2.2.1 Conjuntos de Treinamento, Validação e Teste

Garantir que os conjuntos de treinamento, validação e teste possuam distribuições similares de magnitude (Fig. 12) e campo de visão angular (Fig. 13) é crucial para o treinamento do modelo. Distribuições consistentes asseguram que o modelo seja exposto a dados representativos durante o treinamento, prevenindo vieses que poderiam comprometer sua capacidade de generalizar para novos dados. Magnitudes diferentes podem indicar variações no brilho dos objetos, influenciando os padrões visuais extraídos pelos modelos, enquanto campos de visão angulares distintos podem alterar o contexto espacial e os detalhes observados. Desequilíbrios entre esses conjuntos podem levar a discrepâncias na avaliação,

Figura 11 – Pipeline de produção do catálogo de referência

onde métricas de desempenho no conjunto de teste não refletem a eficácia real do modelo em aplicações práticas.

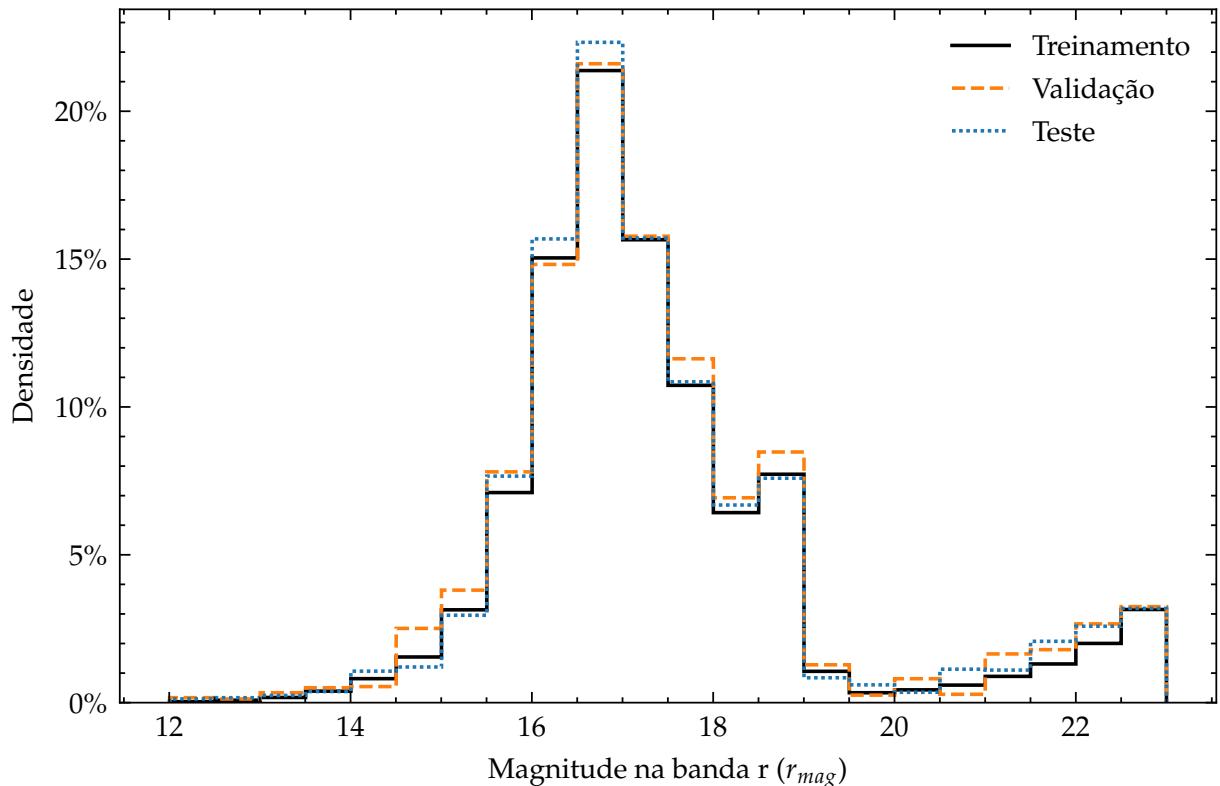
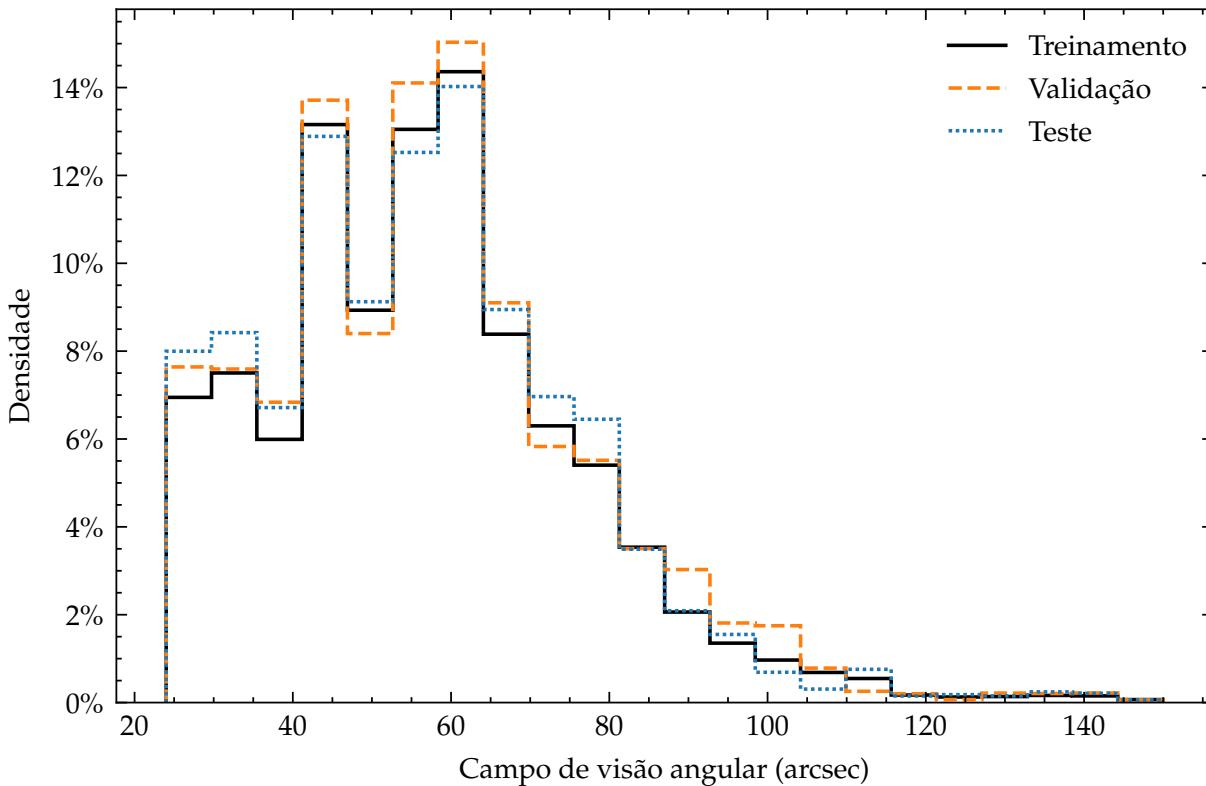
Figura 12 – Distribuição da magnitude na banda r

Figura 13 – Distribuição do campo de visão angular



2.2.2.2 Conjunto de Inferência

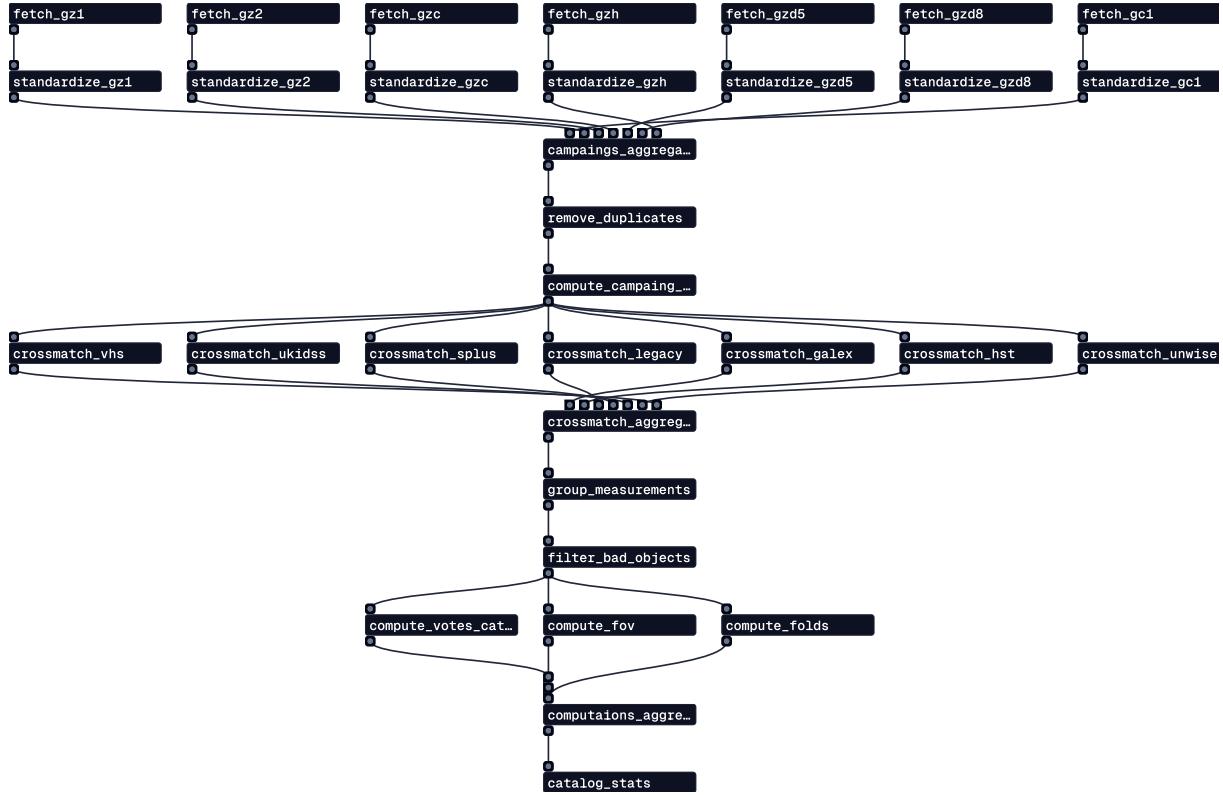
2.2.3 Pipeline de Dados

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et

magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Figura 14 – Pipeline de produção do catálogo de referência

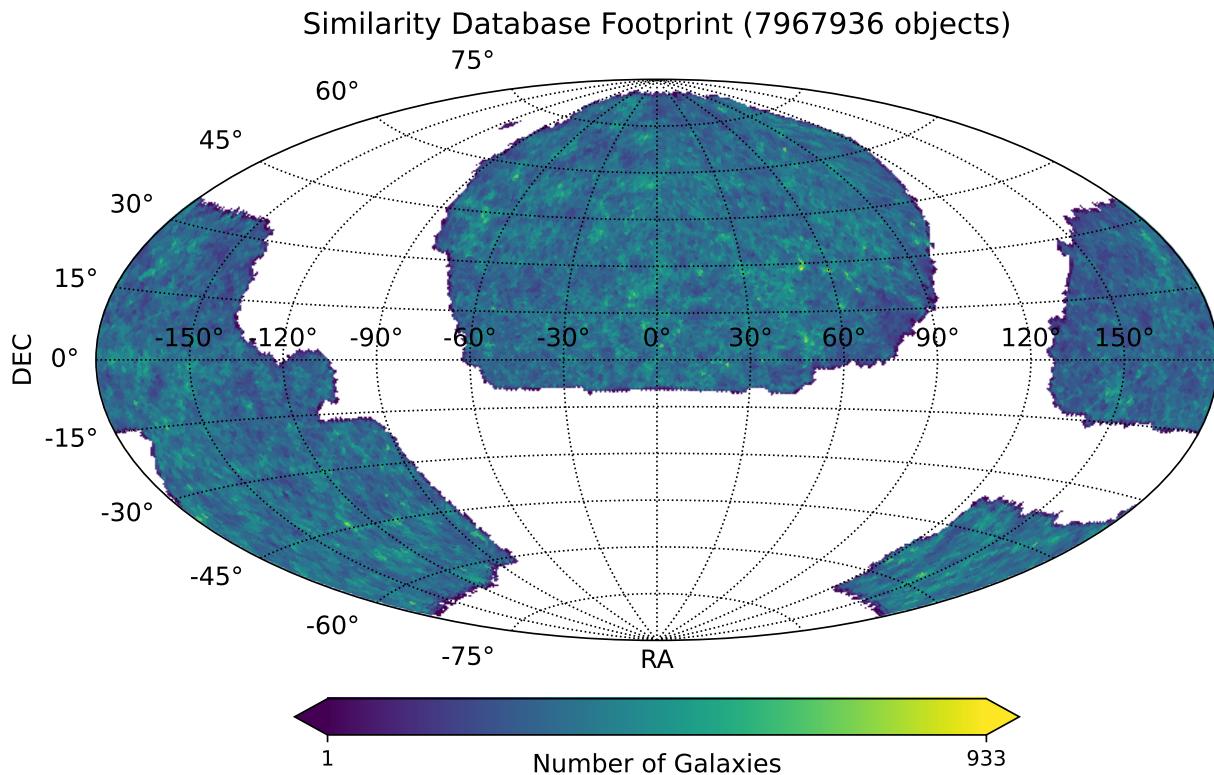


2.2.4 | Descrição

O conjunto de inferência é formado por aproximadamente 8 milhões de objetos. A Fig. 15 mostra a distribuição espacial das galáxias, em que as cores codificam a densidade de galáxias em uma determinada região, conforme mostrado na escala.

Os conjuntos de treinamento e inferência utilizados neste trabalho, composto por 400 mil e 8 milhões de imagens obtidas do DESI Legacy Survey, respectivamente, possuem uma longa durabilidade científica devido às características intrínsecas do levantamento e ao tipo de observação realizada. Como um levantamento de campo profundo e não transitente, o Legacy Survey foca na captura de objetos celestes distantes e estáticos, como galáxias e aglomerados de galáxias, que não apresentam variações significativas em escalas temporais humanas. Esses objetos, frequentemente dominados por matéria escura, são fundamentais para o estudo da evolução cósmica e da distribuição de massa no universo em larga escala. A natureza estável dessas observações permite que os dados permaneçam relevantes por décadas, fornecendo uma base sólida para análises contínuas e com-

Figura 15 – Área de cobertura espacial do conjunto de inferência



plementares, à medida que novos modelos de aprendizado profundo e técnicas de análise de dados são desenvolvidos. Essa durabilidade é particularmente valiosa em um contexto científico onde as demandas computacionais e as abordagens metodológicas evoluem rapidamente, garantindo que o conjunto de inferência continue a contribuir para avanços na pesquisa astronômica e cosmológica.

Esta análise encerra a seção de descrição dos dados. A seguir, serão detalhados os métodos de aprendizado profundo utilizados para o treinamento do modelo.

* * *

3

Classificação automática

Visão geral do capítulo

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.1 Intrrodução

A aprendizagem profunda, ou deep learning, é uma subárea da inteligência artificial (IA) que envolve o uso de redes neurais artificiais com múltiplas camadas para modelar e processar dados complexos. Inspirada inicialmente na estrutura e no funcionamento do cérebro humano, essa abordagem visa reproduzir, em um ambiente computacional, a capacidade do cérebro de identificar padrões e inferir informações a partir de grandes volumes de dados. O termo “profunda” refere-se ao uso de várias camadas ocultas na rede neural, que permite que modelos de aprendizagem profunda extraiam hierarquias de características complexas, aumentando a precisão e a capacidade do modelo em resolver problemas como reconhecimento de fala, visão computacional e processamento de linguagem natural.

3.1.1 Redes Neurais

As redes neurais artificiais são modelos computacionais inspirados no funcionamento do cérebro humano, constituídos por unidades denominadas “neurônios artificiais” interligados em uma estrutura de camadas. Esses modelos buscam simular a capacidade do cérebro de reconhecer padrões, processar informações e realizar tarefas complexas de forma autônoma. A estrutura básica de uma rede neural é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada camada possui neurônios conectados por pesos que são ajustados durante o processo de treinamento, permitindo que a rede aprenda a realizar tarefas específicas.

O desenvolvimento das redes neurais artificiais remonta aos anos 1940, com os trabalhos de [McCulloch e Pitts \(1943\)](#), que propuseram o primeiro modelo de neurônio artificial. Nos anos 1960, o perceptron, desenvolvido em [Rosenblatt \(1960\)](#), trouxe um avanço significativo, sendo capaz de resolver problemas lineares simples. No entanto, o perceptron era limitado a problemas linearmente separáveis. Duas décadas depois, ocorreu um significativo avanço que permitiu que redes neurais resolvessem problemas não-lineares: a introdução do algoritmo de retropropagação de erro, proposto em [Rumelhart, Hinton e Williams \(1986\)](#), que possibilitou o treinamento de redes com múltiplas camadas ocultas, abrindo caminho para a evolução do campo.

3.1.2 Aprendizado Profundo em Visão Computacional

A partir dos anos 2000, o aumento da capacidade computacional e o acesso a grandes volumes de dados catalisaram o desenvolvimento das redes neurais profundas, ou aprendizagem profunda. As redes neurais profundas (ou deep learning) são redes com múltiplas camadas ocultas, que permitem a extração de representações complexas e hierárquicas dos dados. Modelos profundos, como as redes neurais convolucionais (CNNs) e as redes neurais recorrentes (RNNs), provaram ser extremamente eficazes em tarefas de reconhecimento de imagem, processamento de linguagem natural e até mesmo na descoberta de novas partículas na física, evidenciando seu potencial em diversas áreas científicas.

Durante a última década, a aprendizagem profunda conquistou um notório sucesso em visão computacional devido sua capacidade de processamento de dados complexos, exercendo um impacto profundo nas pesquisas científicas, especialmente naquelas baseadas em grandes volumes de dados, como astronomia ([BOM et al., 2021; WALMSLEY et al., 2022](#)), medicina ([BELAID; LOUDINI, 2020](#); [KRIEGSMANN et al., 2020; FUJIOKA et al., 2020](#)) e agronomia ([ZHANG; YANG; LI, 2020; ABAS et al., 2018](#)). Isso se iniciou com as primeiras arquiteturas con-

volucionais profundas, como AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) e VGG (SIMONYAN; ZISSERMAN, 2014), passando por Inception (SZE-GEDY et al., 2014), ResNet (HE et al., 2015) e EfficientNet (TAN; LE, 2019), até chegar nos modelos baseados em transformadores (VASWANI et al., 2017), como Vision Transformer (DOSOVITSKIY et al., 2020) e Swin Transformer (LIU et al., 2021), além de arquiteturas híbridas como Multi-axis Vision Transformer (TU et al., 2022) e Fast Vision Transformer (VASU et al., 2023). Todos esses modelos dependem de um grande volume de dados para serem treinados, como ImageNet (DENG et al., 2009), MS-COCO (LIN et al., 2014) ou até maiores. Isso nos leva a entender que o avanço da capacidade preditiva dos modelos de visão computacional é, essencialmente, sustentado pelo aumento da complexidade do modelo e, consequentemente, por um grande conjunto de treinamento.

Este capítulo aborda o desenvolvimento do sistema inteligente proposto. Inicialmente, são discutidas a abordagem de preparação dos conjuntos de dados (Seção 2) e o método de aprendizagem profunda (Seção 3.2), essenciais para desenvolver um modelo para busca por objetos astronômicos. Em seguida, é detalhada a abordagem para a construção do sistema de informação (Seção ??).

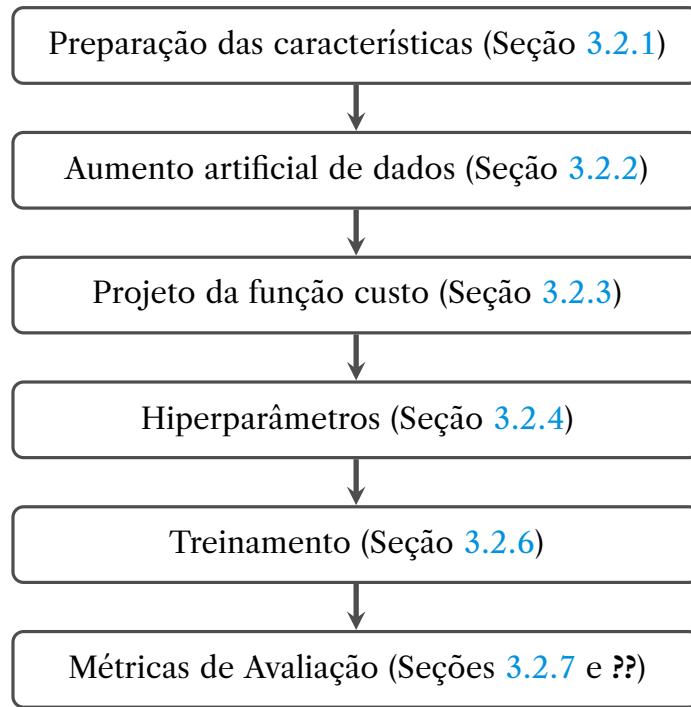
3.2 | Modelo de Aprendizagem Profunda

Neste capítulo, será explicado sobre a preparação dos dados e como fizemos o aumento artificial dos dados para obter melhores resultados na avaliação dos modelos. Em seguida, descrevemos as redes convolucionais utilizadas: VGG, Inception Resnet, EfficientNet e DenseNet. Introduzimos o conceito de (*Ensemble*) e descrevemos as técnicas usadas anteriormente e que fundamentaram nossas escolhas. Em seguida, apresentamos as principais definições das redes e parâmetros utilizados neste trabalho e por fim detalhamos como foram feitas as modelagens e treinamentos dos classificadores e do nosso meta-modelo.

3.2.1 | Preparação das Características

A preparação dos dados é a etapa intuitivamente subsequente, após a aquisição (Seção 2). Nessa subseção, é discutido o processamento e codificação das características de entrada (Seção 3.2.1.1) e dos rótulos (Seção 3.2.1.2).

Figura 16 – Fluxograma do treinamento do modelo



3.2.1.1 Preparação das Características de Entrada

As características de entrada do modelo são as imagens do Legacy Survey DR10 (Seção ??), obtidas pelo método descrito na Seção 2.1.4. Uma etapa importante da seleção dessas características é o ajuste do campo de visão (FoV) de cada imagem. Esse ajuste determina a porcentagem que a galáxia ocupa na imagem. Visto que os píxeis referentes às galáxias são as únicas características de interesse, esse ajuste é fundamental para garantir o desempenho do modelo. Como o FoV deve ser especificado previamente, no momento da aquisição da imagem, essa etapa de ajuste do FoV já foi discutida na Seção 2.1.3.

O primeiro pré-processamento feito é a soma pixel a pixel no eixo dos canais RGB, produzindo uma imagem em níveis de cinza, com um único canal. Dessa forma, é mantida apenas a informação estrutural de cada galáxia. Em seguida, é feita uma normalização usando a média e o desvio padrão do lote (BatchNormalization).

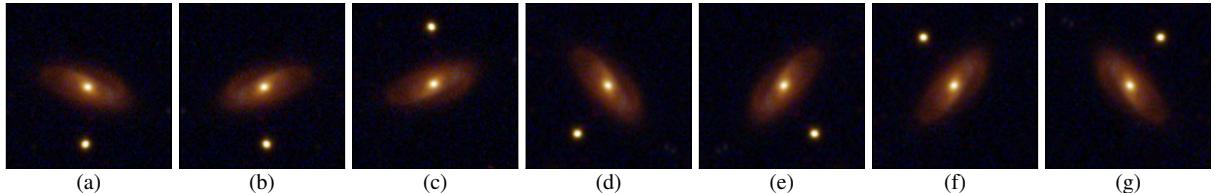
3.2.1.2 Preparação dos Rótulos

Os rótulos são formados pelas contagens das respostas dos voluntários do GalaxyZoo para cada galáxia. Isto é, os votos são agregados em relação a cada galáxia. As contagens de votos de cada alternativa são arranjadas na forma de vetor

e os votos de diferentes campanhas são concatenados em um único vetor de votos, que é considerado o valor de referência para a estimativa das contagens de votos durante o treinamento do modelo.

3.2.2 Aumento Artificial de Dados

Figura 17 – Aumento artificial dos dados



Exemplo do aumento artificial de dados em uma imagem original, mostrada no painel (a). Os painéis (b), (c), (d), (e), (f) e (g) contêm os resultados da eq. (3.4) substituindo M por diferentes combinações das transformações das eqs. (3.1), (3.2) e (3.3). Em (b) $M = V$, em (c) $M = H$, em (d) $M = R(30^\circ)$, em (e) $M = VR(30^\circ)$, em (f) $M = HR(30^\circ)$ e em (g) $M = HVR(30^\circ)$.

Aumento artificial de dados (YAEGER; LYON; WEBB, 1996) é a aplicação de transformações afins nas imagens do conjunto de treinamento, por exemplo rotação, reflexão, translação e mudança de escala. As matrizes das eqs. (3.1), (3.2) e (3.3) definem as transformações usadas.

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.1)$$

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.2)$$

$$V = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.3)$$

onde $R(\theta)$ é a transformação rotação por um ângulo θ , H é a transformação reflexão horizontal e V é a transformação reflexão vertical, nas eqs. (3.1), (3.2) e (3.3), respectivamente.

Seja M a matriz das transformações combinadas, (x, y) a coordenada do píxel da imagem original e (x^*, y^*) a coordenada transformada do píxel, as transformações nas imagens são feitas remapeando as coordenadas dos píxeis originais

aplicando uma combinação das matrizes das eqs. (3.1), (3.2) e (3.3) em cada píxel da imagem original usando a equação (3.4), onde (t_x, t_y) é a coordenada do centro da imagem e as matrizes ao redor de M são as matrizes translação. Isso é feito para que a transformação M tenha o centro da imagem como ponto de simetria.

$$\begin{bmatrix} x^* \\ y^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} M \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.4)$$

Além disso, ainda é aplicada uma interpolação bilinear como *anti-aliasing* (MCREYNOLDS; BLYTHE, 2005; BOVIK, 2009). Durante o treinamento da rede, novas imagens de entrada são geradas a cada época a partir da transformação das imagens originais. A Figura 17 mostra a imagem original, no painel (a), e diversas transformações, nos demais painéis, aplicadas substituindo M da equação (3.4) por combinações (multiplicação matricial) das transformações das eqs. (3.1), (3.2) e (3.3). Tais transformações não mudam a interpretação da classe da imagem original, pois o espaço visual é invariante a elas. Logo, o objetivo de aplicar estas transformações nas imagens de entrada da rede é deixar que o algoritmo infira tal invariância, criando, assim, uma “noção” do espaço visual, o que resulta no aumento do potencial de generalização da rede (SIMARD; STEINKRAUS; PLATT, 2003; CHOLLET, 2017). Frequentemente são relatados bons resultados com o uso desta técnica (NONAKA; SEITA, 2020; II; DINC, 2020; MITTAL et al., 2019), principalmente quando existe grande similaridade entre as classes.

3.2.3 Função de Custo

3.2.3.1 Distribuição Multinomial

A distribuição multinomial (FORBES et al., 2010e) é uma generalização da distribuição binomial (FORBES et al., 2010b) que descreve o comportamento de um experimento discreto onde cada tentativa resulta em um de k possíveis resultados mutuamente exclusivos.

Supondo que um experimento é repetido n vezes, e que a probabilidade de cada resultado específico i , para $i = 1, 2, \dots, k$, seja p_i , com $\sum_{i=1}^k p_i = 1$. A distribuição multinomial caracteriza a probabilidade conjunta de observar cada um dos k resultados um número específico de vezes (x_1, x_2, \dots, x_k) , onde $\sum_{i=1}^k x_i = n$. A função de probabilidade associada é definida pela eq. (3.5), onde $x_i \geq 0$ representa a contagem de ocorrências para cada categoria i .

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!}{x_1!x_2!\dots x_k!} \prod_{i=1}^k p_i^{x_i} \quad (3.5)$$

A distribuição multinomial (SEBER, 2015) é amplamente utilizada em aplicações de aprendizado de máquina e estatística, como na modelagem da transmissão de Dengue (WANG; ZHANG, 2025), na predição de cotação de mercado (NEVASALMI, 2020) e na análise de tráfego aéreo (TORRES; GUTERRES; CELESTINO, 2023). Uma de suas principais aplicações é em modelagem de eventos categóricos (KIBRIYA et al., 2004; LUO; WONG; VONG, 2021). Por exemplo, em tarefas de classificação multiclasse, onde o objetivo é prever a probabilidade de uma amostra pertencer a uma entre várias categorias, a distribuição multinomial é usada para modelar a saída do modelo. Um exemplo prático é o uso em algoritmos como o Naive Bayes Multinomial (KALCHEVA; MARINOVA; TODOROVA, 2023; JIANG et al., 2016), amplamente empregado em problemas de classificação de texto (ODEH; ODEH; ODEH, 2022; KANG et al., 2005), como análise de sentimentos (SARAVANAN et al., 2023) e categorização de documentos. Neste contexto, as palavras em um documento são tratadas como eventos que seguem uma distribuição multinomial.

3.2.3.2 Distribuição de Dirichlet

A distribuição Dirichlet (FORBES et al., 2010c) é uma distribuição de probabilidade multivariada contínua definida no simplex k -dimensional, sendo uma generalização da distribuição Beta (FORBES et al., 2010a) para múltiplas dimensões. Ela é frequentemente utilizada para modelar distribuições de probabilidade sobre proporções, em que cada elemento do vetor de variáveis aleatórias representa uma fração de um todo e, portanto, suas somas totalizam 1.

Dada uma variável aleatória $\vec{X} = (X_1, X_2, \dots, X_k)$ que segue uma distribuição Dirichlet com parâmetro de concentração $\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k)$ para uma quantidade k de categorias, a função de densidade de probabilidade é definida pela eq. (3.6), onde $B(\vec{\alpha})$ é a função beta multivariada, definida na eq. (3.7), e $\Gamma(\cdot)$ representa a função gama (FORBES et al., 2010d). O parâmetro α_i controla a concentração da probabilidade ao redor das categorias: valores maiores de α_i indicam maior densidade de probabilidade associada à categoria correspondente.

$$P(\vec{X}|\vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{i=1}^k X_i^{\alpha_i-1} \quad (3.6)$$

$$B(\vec{\alpha}) = \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^k \alpha_i\right)} \quad (3.7)$$

Em aprendizado de máquina, a distribuição Dirichlet possui aplicações importantes em problemas que envolvem modelagem de distribuições de probabilidade sobre categorias. Um exemplo clássico é na modelagem de misturas probabilísticas, como no algoritmo Latent Dirichlet Allocation (LDA; [BLEI; NG; JORDAN, 2003](#)), amplamente utilizado para análise de tópicos em grandes corpora de texto ([JELODAR et al., 2019](#)). No LDA, a distribuição Dirichlet é usada como uma distribuição a priori para representar a probabilidade de tópicos em documentos e a probabilidade de palavras em tópicos, fornecendo uma forma eficiente de capturar relações semânticas latentes ([CANINI; SHI; GRIFFITHS, 2009](#)).

3.2.3.3 Distribuição Dirichlet-Multinomial

A distribuição Dirichlet-multinomial ([FORBES et al., 2010c](#)) é uma distribuição de probabilidade discreta que combina as propriedades da distribuição multinomial (Seção 3.2.3.1) e da distribuição Dirichlet (Seção 3.2.3.2). Ela é usada para modelar processos em que as probabilidades das categorias em uma distribuição multinomial são, elas mesmas, amostras de uma distribuição Dirichlet. Isto é, a distribuição Dirichlet é usada como distribuição a priori para estimar os parâmetros da distribuição multinomial. Essa estrutura permite incorporar incertezas sobre as probabilidades subjacentes das categorias.

Formalmente, dado um vetor $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ de probabilidades categóricas que segue uma distribuição Dirichlet com parâmetro $\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k)$, e uma distribuição multinomial condicional sobre $\vec{\theta}$, a probabilidade conjunta da distribuição Dirichlet-multinomial é expressa na eq. (3.8), onde $n = \sum_{i=1}^k x_i$ é o número total de experimentos, e x_i representa o número de ocorrências da i -ésima categoria.

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{\Gamma\left(\sum_{i=1}^k \alpha_i\right)}{\Gamma\left(\sum_{i=1}^k \alpha_i + n\right)} \prod_{i=1}^k \frac{\Gamma(x_i + \alpha_i)}{\Gamma(\alpha_i) x_i!} \quad (3.8)$$

3.2.3.4 Função de Custo e Função de Perda

A função de perda (*loss function*) e a função de custo (*cost function*) têm papéis relacionados, mas distintos. A primeira avalia o erro do modelo em uma

única amostra, medindo a discrepância entre a predição do modelo e o valor esperado. Já a segunda, é uma agregação da primeira, geralmente calculada como a média ou soma das perdas sobre todo o conjunto de treinamento, representando o desempenho global do modelo.

A função de perda de uma pergunta q é calculada a partir da distribuição Dirichlet-Multinomial, conforme a eq. (3.9).

$$\mathcal{L}_q = \int \text{Multi}(\vec{k}|\vec{\rho}, N) \text{Dirichlet}(\vec{\rho}|\vec{\alpha}) d\vec{\rho} \quad (3.9)$$

Já o valor da perda de um exemplo é a soma das perdas de cada questão q , conforme a eq. (3.10).

$$\log \mathcal{L} = \sum_q \mathcal{L}_q(\vec{k}_q, N_q, \vec{f}_q^w) \quad (3.10)$$

Existem duas formas recorrentes de agregar as perdas dos exemplos para calcular o custo total do modelo: calcular a soma ou a média. Nesse trabalho, a função de custo é definida como a soma de todas as perdas de cada exemplo do conjunto de dados, conforme eq. (3.11).

$$J = \sum_i \log \mathcal{L}_i \quad (3.11)$$

3.2.4 Hiperparâmetros

Nesta seção descrevemos as definições dos principais conceitos, no contexto de deep learning, que serão úteis para o entendimento dos métodos aqui utilizados. A função de ativação, função de custo, o otimizador, o learning rate, o número de épocas, além do número de camadas dos modelos, são importantes parâmetros responsáveis pela construção do modelo definido a seguir.

Função de ativação

A função de ativação é responsável por adicionar não-linearidade à rede. Sem ela, a saída de uma camada seria apenas uma transformação linear dos dados de entrada e a rede não seria beneficiada pelo empilhamento de diversas camadas lineares, pois isso não aumentaria o espaço de hipóteses. Logo, a função de ativação viabiliza representações mais complexas da rede, uma vez que define a complexidade de um modelo e, consequentemente, sua capacidade de generalização (CHOLLET, 2017). Neste trabalho, a função $\text{ReLU}(x) := \max(0, x)$ é usada nas camadas densas dos classificadores, a

equação tangente hiperbólica é usada nas camadas densas do meta-modelo e a função Softmax ([BRIDLE, 1990](#)) foi usada na última camada, tanto dos classificadores quanto do meta-modelo.

Otimizador

O otimizador é um algoritmo iterativo com objetivo de minimizar a função de custo. Uma escolha típica é o método de gradiente descendente e suas demais variações. Este tipo de algoritmo tem um parâmetro livre relacionado ao passo da iteração conhecido como taxa de aprendizado ou *learning rate*. Neste trabalho foram testados diversos algoritmos considerados como estado-da-arte dos otimizadores como Adam ([KINGMA; BA, 2017](#)), NAdam ([DOZAT, 2015](#)), RAdam ([LIU et al., 2020](#)) e RMSprop ([HINTON; SRIVASTAVA; SWERSKY, 2012](#)).

Número de Épocas

O Número de épocas se referem a quantidade de vezes que o dataset de treino foi utilizado completamente no processo de otimização iterativa da função de custo. Um número de épocas adequado é necessário para que a função de custo seja minimizada.

Tamanho do Batch

O processo de otimização acontece em batches, cada iteração para minimizar a função custo é realizada com um número fixo de amostras, quando todas as amostras de treino são utilizadas se completa uma época.

Unidades de neurônios na última camada

A última camada da rede antes da camada de saída é responsável por condensar toda a informação extraída da rede para o processo de classificação final. Por esta razão, a quantidade de neurônios nessa camada pode ser particularmente sensível para a performance da rede. Neste trabalho utilizamos diferentes valores de neurônios para encontrar a quantidade que pode gerar a melhor performance.

Dropout

Dropout ([SRIVASTAVA et al., 2014](#)) é uma técnica de regularização muito utilizada em redes neurais por seu bom desempenho e baixo custo computacional. Aplicar esta regularização em uma camada consiste em eliminar aleatoriamente uma taxa dos neurônios de saída desta camada durante o treinamento, sendo geralmente escolhido um valor entre 0.2 e 0.5 para esta taxa ([CHOLLET, 2017](#)).

3.2.5 Arquitetura de Rede Neural

3.2.5.1 Visual Geometry Group Networks

A família de arquiteturas Visual Geometry Group Networks (VGGNets) ([SIMONYAN; ZISSERMAN, 2015](#)) foi criada durante a competição de classificação de imagens *Large Scale Visual Recognition Challenge* ([RUSSAKOVSKY et al., 2015](#)). Ela se destaca por estar entre as primeiras redes a adotar, com sucesso, o escalamiento em profundidade (quantidade de camadas) para aumentar o desempenho na classificação de imagens usando redes convolucionais. Neste trabalho, consideramos a arquitetura VGG16, que já foi usada em diversas tarefas de classificação, como a classificação de software malicioso ([REZENDE et al., 2018](#)), de plantas ([ABAS et al., 2018](#)) e de tumores cerebrais ([BELAID; LOUDINI, 2020](#)).

3.2.5.2 InceptionResNetV2

A arquitetura InceptionResNetV2 ([SZEGEDY et al., 2016](#)) usa os blocos Inception, que são convoluções fatorizadas, introduzidos em ([SZEGEDY et al., 2015](#)), motivada pela construção de redes mais profundas com um menor custo computacional e menor overfitting, com a adição de conexões residuais ([HE et al., 2016](#)) motivada pelo problema de dissipação do Gradiente (*vanishing gradients*). Isso permite treinar redes profundas com maior acurácia e mais rápido. Esta arquitetura já foi usada, por exemplo, para classificação de imagens de satélite ([MAHDIANPARI et al., 2018](#)), de ultrasonografia ([FUJIOKA et al., 2020](#)) e de células cancerígenas ([KRIEGSMANN et al., 2020](#)).

3.2.5.3 EfficientNet

A arquitetura EfficientNet ([TAN; LE, 2020](#)) foi desenvolvida como uma resposta à questão de como escalar modelos de convolução. Foram considerados três diferentes aspectos: profundidade, largura e resolução da imagem de entrada. Em vez de dimensionar cada aspecto manualmente, o modelo implementa um escalonamento composto que equilibra os aspectos para obter melhor desempenho, com isso a rede consegue uma alta acurácia usando muito menos parâmetros e operações de ponto flutuante por segundo (*FLOPS*). Esta rede já foi usada na classificação de doenças em vegetais ([ZHANG; YANG; LI, 2020](#)), eletrocardiogramas ([NONAKA; SEITA, 2020](#)) e cristalização de proteínas ([II; DINC, 2020](#)).

3.2.5.4 DenseNet

A rede DenseNet ([HUANG et al., 2018](#)) também usa conexões residuais que conectam cada camada a todas as outras camadas seguintes, o que reduz ainda mais o número de parâmetros na rede sem perda significativa da precisão. O uso desta rede incluem predição do mapa de contato de proteínas ([LI et al., 2020](#)), classificação de músicas ([SOLANO-ROJAS; VILLALÓN-FONSECA; MARÍN-RAVENTÓS, 2020](#)), câncer de mama ([LI et al., 2020](#)) e esclerose múltipla ([WANG; ZHANG, 2020](#)).

3.2.5.5 Escolha da Arquitetura

Todas as arquiteturas convolucionais mencionadas nessa seção foram testadas e a arquitetura que teve melhor desempenho na predição dos votos dos voluntários foi escolhida. O processo de treinamento e avaliação está descrito na seção seguinte.

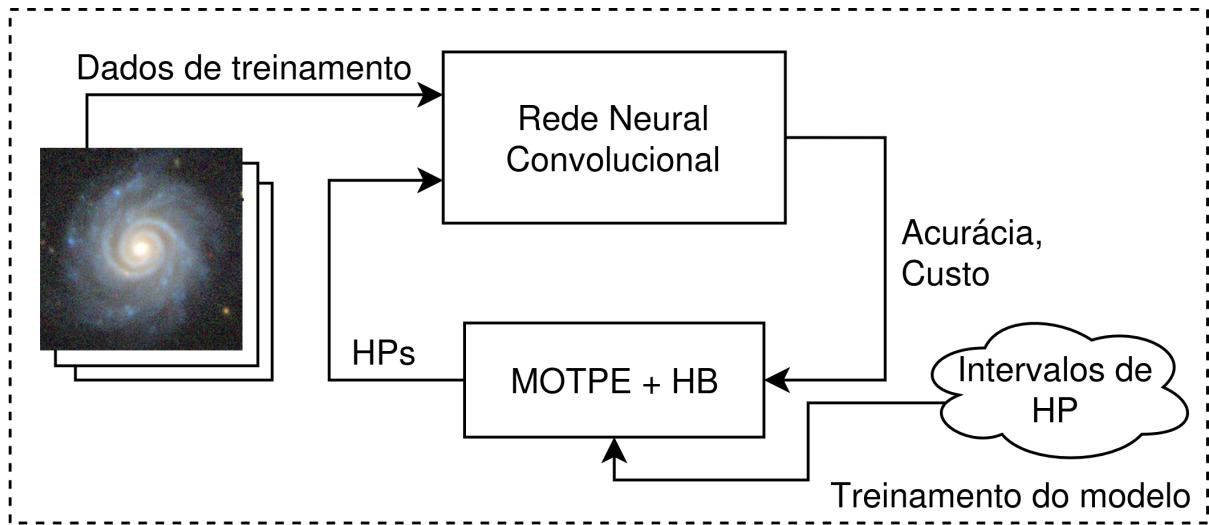
3.2.6 Treinamento

O treinamento de modelos de aprendizado profundo é um processo complexo que envolve a escolha de uma arquitetura adequada e a definição de hiperparâmetros que otimizem o desempenho do modelo em um conjunto de dados específico. A técnica de treinamento automatizado (AutoML; [LI et al., 2021](#)), mostrada na Fig. 18, facilita esse processo de seleção de hiperparâmetros e configurações de treinamento. Neste trabalho, foi utilizada uma abordagem que se mostrou eficaz nesse contexto: os algoritmos Multi-Objective Tree-structured Parzen Estimator (MOTPE; [OZAKI et al., 2020](#)) e HyperBand (HB; [LI et al., 2016](#)), disponíveis na biblioteca Optuna¹ ([AKIBA et al., 2019](#)), para a otimização de múltiplos objetivos simultaneamente.

O treinamento com AutoML utilizando o MOTPE envolve diversas etapas estruturadas, desde a configuração inicial do espaço de busca dos hiperparâmetros até a avaliação iterativa dos modelos. A seguir, descrevem-se as principais etapas do processo:

- 1. Definição do Espaço de Busca dos Hiperparâmetros:** Inicialmente, define-se um espaço de busca para os hiperparâmetros do modelo, que pode incluir a taxa de aprendizado, número de camadas, número de neurônios por camada, taxas de dropout, tipos de ativação, otimizadores, entre outros.

¹ <<https://optuna.org>>

Figura 18 – Treinamento do modelo

Para cada hiperparâmetro, especificam-se limites, distribuições ou valores discretos que serão explorados pelo algoritmo.

- 2. Criação da Função de Objetivo:** A função de objetivo é o núcleo do processo de otimização. Ela recebe uma configuração de hiperparâmetros como entrada, treina e avalia o modelo utilizando essa configuração, e retorna uma ou mais métricas de desempenho que serão otimizadas. No caso do MOTPE, que é multiobjetivo, é possível considerar métricas conflitantes, como minimizar o erro de validação enquanto reduz o tempo de treinamento.
- 3. Treinamento Iterativo com Amostragem Bayesiana:** O MOTPE utiliza um modelo probabilístico para guiar a busca por configurações promissoras. Ele avalia iterativamente as combinações de hiperparâmetros, ajustando o modelo probabilístico com base nos resultados obtidos em iterações anteriores. Isso permite uma exploração mais eficiente do espaço de busca, concentrando-se em áreas com maior probabilidade de melhorar o desempenho.
- 4. Validação e Regularização:** Cada configuração de hiperparâmetros é validada em um conjunto de validação separado para garantir que o modelo não esteja superajustando os dados de treinamento. Além disso, o uso de técnicas como early stopping e regularização (L2 ou dropout) é integrado ao treinamento para evitar overfitting.
- 5. Seleção do Melhor Modelo:** Após um número predefinido de iterações ou ao atingir um critério de parada, o AutoML seleciona a configuração de hiperparâmetros que apresenta o melhor desempenho médio em relação

às métricas otimizadas. No caso de otimização multiobjetivo, a solução é escolhida com base no conceito de fronteira de Pareto.

A Tabela 2 sumariza os hiperparâmetros otimizados pelo MOTPE. Nela são mostrados os intervalos de busca para cada hiperparâmetro, bem como a distribuição utilizada para cada um deles.

Tabela 2 – Intervalos de busca dos hiperparâmetros otimizados com o MOTPE

Hiperparâmetro	Distribuição	Intervalo de Busca
Arquitetura	Categórica	{VGG, IRV2, EffNet, DenseNet}
Taxa de aprendizado	Log-Uniforme	$[10^{-5}, 10^{-1}]$
Número de camadas ocultas	Int-Uniforme	$[0, 5]$
Neurônios por camada	Int-Uniforme	$[32, 512]$
Taxa de dropout	Uniforme	$[0.1, 0.5]$
Otimizador	Categórica	{Adam, NAdam, SGD, RMSprop}
Batch size	Int-Log-Uniforme	$[16, 256]$
Peso da regularização L2	Log-Uniforme	$[10^{-5}, 10^{-2}]$
Função de ativação	Categórica	{ReLU, Leaky ReLU, ELU}

3.2.7 Métricas de Avaliação da Predição dos Votos

As métricas aqui utilizadas baseiam-se no erro ou acerto na associação dos objetos às classes pelo modelo. Para relacionar a probabilidade calculada pelo modelo à classe, é definido um limiar de discriminação, que é a probabilidade mínima para que um exemplar pertença à uma classe. Para o cálculo das métricas, foi considerado um de limiar de 0.5.

3.2.7.1 Matriz de Confusão

A matriz de confusão consiste em um modelo tabular que cruza métricas de predição e referência, sendo comumente usada para avaliar a qualidade preditiva de modelos de aprendizagem de máquina. A Tabela 3 mostra um esquema desta matriz para o caso de classificação binária (apenas duas classes). A análise para mais classes pode ser feita isolando uma classe positiva por vez e reduzindo o problema a multiplas classificações binárias.

A seguir, são definidas cada uma das métricas usadas para montar a matriz de confusão.

- Verdadeiro Positivo (VP): O modelo prevê uma instância verdadeira corretamente

Tabela 3 – Estrutura da matriz de confusão para classificação binária

	Predição	
	Verdadeiro Positivo (VP)	Falso Negativo (FN)
Verdadeiro	Falso Positivo (FP)	Verdadeiro Negativo (VN)

- Verdadeiro Negativo (VN): O modelo prevê uma instância negativa corretamente
- Falso Positivo (FP): O modelo prevê uma instância positiva, mas, na verdade, é negativa. Este é o erro tipo I, que consiste em rejeitar a hipótese nula quando ela é verdadeira.
- Falso Negativo (FN): O modelo prevê uma instância negativa, mas, na verdade, é positiva. Este é o erro tipo II, em estatística, que consiste na falha em rejeitar a hipótese nula quando ela é falsa.

3.2.7.2 Acurácia

A acurácia é uma métrica que mede a proporção de previsões corretas em relação ao total de instâncias avaliadas. Ela é definida como indicado na eq. 3.12.

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}} \quad (3.12)$$

3.2.7.3 Precisão

A precisão avalia a proporção de exemplos classificados como positivos que são realmente positivos, conforme a eq. (3.13), sendo particularmente importante em cenários onde o custo de falsos positivos é alto.

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}}. \quad (3.13)$$

3.2.7.4 Revocação

A revocação, também conhecida como sensibilidade ou taxa de verdadeiros positivos, mede a capacidade do modelo de identificar corretamente os exemplos positivos, conforme a eq. (3.14), sendo principalmente útil em contextos onde a identificação de todos os exemplos positivos é crucial.

$$\text{Revocação} = \frac{\text{VP}}{\text{VP} + \text{FN}}. \quad (3.14)$$

3.2.7.5 F1-Score

O F1-score é a média harmônica entre precisão e revocação, sendo usado como uma métrica de equilíbrio entre essas duas dimensões, conforme a eq. (3.15). Esta métrica é útil quando há uma compensação entre precisão e revocação, como em cenários com classes desbalanceadas, pois fornece uma medida única que leva em conta ambos os aspectos. Valores de $F1$ próximos de 1 indicam que o modelo apresenta um bom equilíbrio entre os dois fatores.

$$F1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (3.15)$$

3.3 Avaliação do Modelo no Conjunto de Teste

A análise quantitativa das métricas de desempenho no conjunto de teste é fundamental para avaliar a capacidade preditiva e de generalização de um modelo de aprendizado profundo. A Tabela 4 sumariza as métricas de acurácia (Seção 3.2.7.2), precisão (Seção 3.2.7.3), revocação (Seção 3.2.7.4) e F1-score (Seção 3.2.7.5) para várias perguntas do GalaxyZoo.

Tabela 4 – Avaliação do modelo no conjunto de teste

Questão	Acurácia	Precisão	Revocação	F1
Spiral winding	0.833771	0.714138	0.717103	0.715594
Bar	0.890751	0.715289	0.780659	0.742811
Edge on bulge	0.943199	0.817953	0.895207	0.851695
Disk edge on	0.973585	0.969010	0.970718	0.969856
Smooth or featured	0.974996	-	-	-
Has spiral arms	0.942789	0.898665	0.859011	0.877276
Galaxy Symetrical	0.980000	0.973438	0.966997	0.970173
Clumpy appearance	0.962085	0.956045	0.958264	0.957141
How rounded	0.952254	0.919239	0.927339	0.923121
Merging	0.981400	0.666607	0.771892	0.703503
Bulge size	0.931864	0.683324	0.734342	0.701055
Spiral arms count	0.954545	0.684281	0.758884	0.716117

Para dar continuidade na avaliação quantitativa no conjunto de teste, a seguir, são mostradas as matrizes de confusão (Seção 3.2.7.1) para questão da Tabela 4.

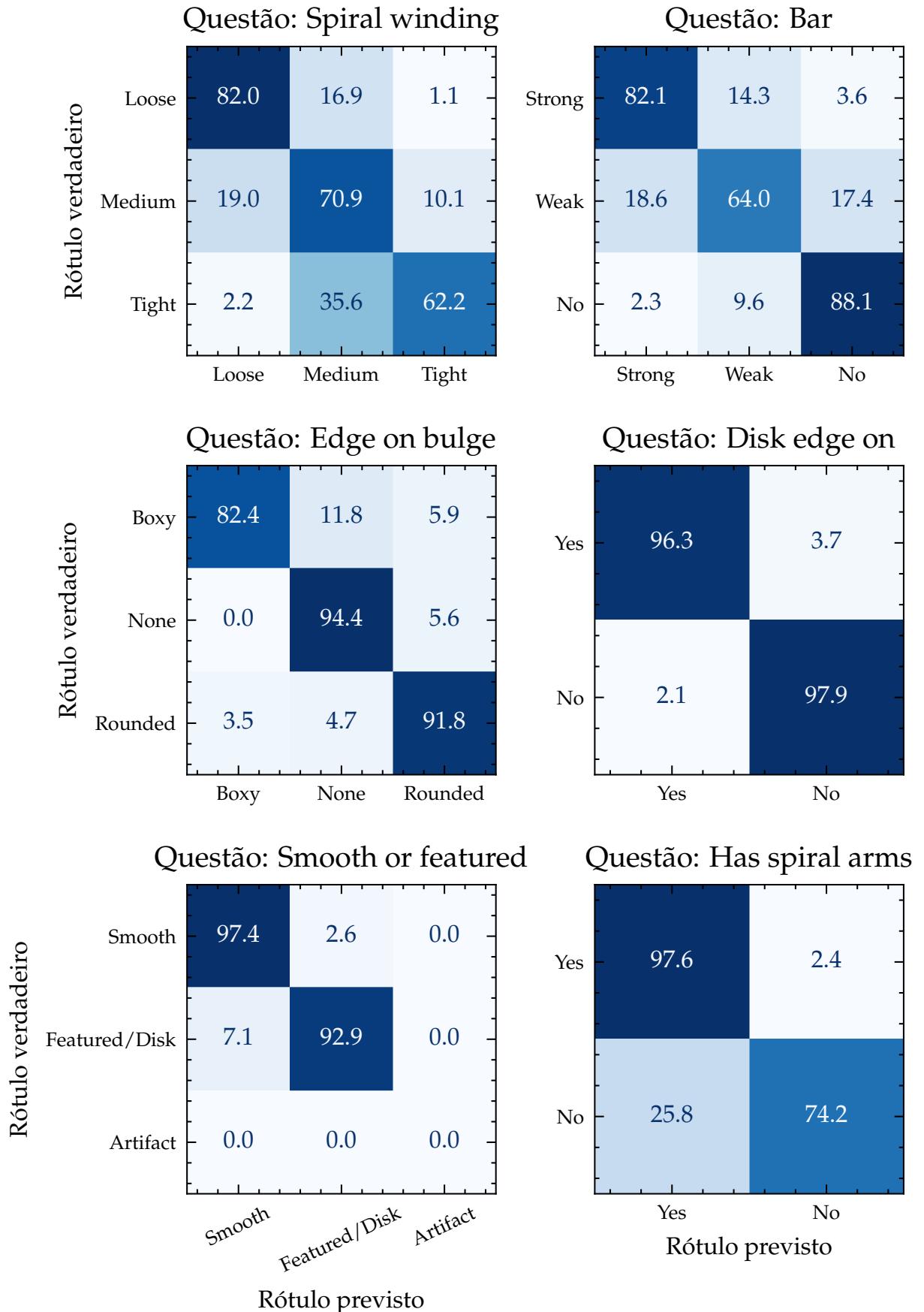
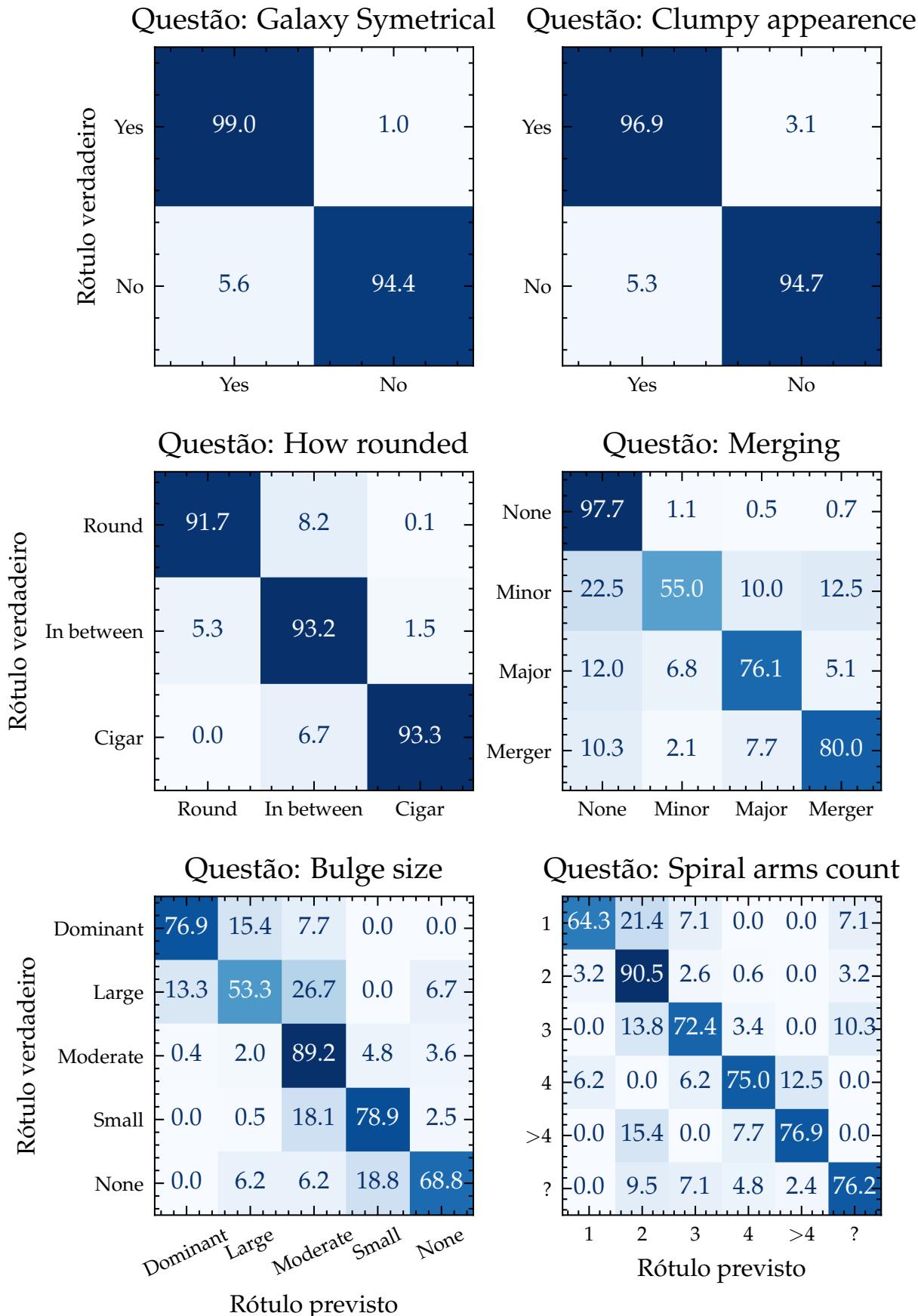
Figura 19 – Matrizes de confusão para cada questão no conjunto de teste

Figura 20 – Matrizes de confusão para cada questão no conjunto de teste (continuação)

Os valores mostrados na Tabela 4 e nas Figs. 19 e 20 mostram a eficiência do modelo em identificar as características morfológicas das galáxias. Essa análise evidencia que o modelo não apenas possui bom desempenho em dados de treinamento, mas também é capaz de generalizar para novos conjuntos de dados, aspecto crucial para aplicações científicas confiáveis.

* * *

Referências

- ABAS, M. A. H. et al. VGG16 for plant image classification with transfer learning and data augmentation. *International Journal of Engineering & Technology*, Science Publishing Corporation, v. 7, n. 4.11, p. 90, out. 2018. Disponível em: <<https://doi.org/10.14419/ijet.v7i4.11.20781>>. Citado 2 vezes nas páginas 46 e 55.
- AKIBA, T. et al. Optuna. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2019. p. 2623–2631. ISBN 9781450362016. Disponível em: <<https://doi.org/10.1145/3292500.3330701>>. Citado na página 56.
- BELAID, O. N.; LOUDINI, M. Classification of brain tumor by combination of pre-trained vgg16 cnn. *Journal of Information Technology Management*, Faculty of Management, University of Tehran, v. 12, n. 2, p. 13–25, 2020. ISSN 2008-5893. Disponível em: <https://jjtm.ut.ac.ir/article_75788.html>. Citado 2 vezes nas páginas 46 e 55.
- BERGH, S. Van den. *Galaxy morphology and classification*. [S.l.]: Cambridge University Press, 1998. Citado 2 vezes nas páginas 23 e 24.
- BLEI, D. M.; NG, A. Y.; JORDAN, M. I. Latent dirichlet allocation. *The Journal of Machine Learning Research*, JMLR.orgPUB6573, v. 3, p. 993–1022, 3 2003. ISSN 1532-4435. Disponível em: <<https://doi.org/10.5555/944919.944937>>. Citado na página 52.
- BOM, C. R. et al. Deep Learning assessment of galaxy morphology in S-PLUS Data Release 1. *Mon. Not. R. Astron. Soc.*, Oxford Academic, v. 507, n. 2, p. 1937–1955, out. 2021. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1093/mnras/stab1981>>. Citado na página 46.
- BONNAREL, F. et al. The aladin interactive sky atlas. *Astronomy and Astrophysics Supplement Series*, EDP Sciences, v. 143, p. 33–40, 4 2000. ISSN 0365-0138. Disponível em: <<http://doi.org/10.1051/aas:2000331>>. Citado na página 36.
- BONNEY, R. et al. Citizen science. Next steps for citizen science. *Science (New York, N.Y.)*, v. 343, n. 6178, p. 1436–7, mar. 2014. ISSN 1095-9203. Disponível em: <<https://doi.org/10.1126/science.1251554>>. Citado na página 26.
- BOVIK, A. C. Basic gray level image processing. In: *The Essential Guide to Image Processing*. Elsevier, 2009. p. 43–68. Disponível em: <<https://doi.org/10.1016/b978-0-12-374457-9.00003-2>>. Citado na página 50.
- BRIDLE, J. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. In: TOURETZKY, D. (Ed.). *Advances in Neural Information Processing Systems*. Morgan-Kaufmann,

1990. v. 2. Disponível em: <<https://proceedings.neurips.cc/paper/1989/file/0336dcbab05b9d5ad24f4333c7658a0e-Paper.pdf>>. Citado na página 54.

BUTA, R. J. *Galaxy Morphology*. 2011. Citado na página 23.

CANINI, K.; SHI, L.; GRIFFITHS, T. Online inference of topics with latent dirichlet allocation. In: DYK, D. van; WELLING, M. (Ed.). *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. PMLR, 2009. v. 5, p. 65–72. Disponível em: <<https://proceedings.mlr.press/v5/canini09a.html>>. Citado na página 52.

CHOLLET, F. *Deep Learning with Python*. 1st. ed. USA: Manning Publications Co., 2017. ISBN 1617294438. Citado 3 vezes nas páginas 50, 53 e 54.

DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. [s.n.], 2009. p. 248–255. Disponível em: <<https://doi.org/10.1109/CVPR.2009.5206848>>. Citado na página 47.

DOSOVITSKIY, A. et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv*, out. 2020. Disponível em: <<https://doi.org/10.48550/arXiv.2010.11929>>. Citado na página 47.

DOZAT, T. Incorporating nesterov momentum into. In: . [S.l.: s.n.], 2015. Citado na página 54.

FORBES, C. et al. Beta distribution. In: _____. 4. ed. Wiley, 2010. p. 55–61. ISBN 9780470390634. Disponível em: <<https://doi.org/10.1002/9780470627242.ch8>>. Citado na página 51.

FORBES, C. et al. Binomial distribution. In: _____. 4. ed. Wiley, 2010. p. 62–65. ISBN 9780470390634. Disponível em: <<https://doi.org/10.1002/9780470627242.ch9>>. Citado na página 50.

FORBES, C. et al. Dirichlet distribution. In: _____. 4. ed. Wiley, 2010. p. 77–78. ISBN 9780470390634. Disponível em: <<https://doi.org/10.1002/9780470627242.ch13>>. Citado 2 vezes nas páginas 51 e 52.

FORBES, C. et al. Gamma distribution. In: _____. 4. ed. Wiley, 2010. p. 109–113. ISBN 9780470390634. Disponível em: <<https://doi.org/10.1002/9780470627242.ch22>>. Citado na página 51.

FORBES, C. et al. Multinomial distribution. In: _____. 4. ed. Wiley, 2010. p. 135–136. ISBN 9780470390634. Disponível em: <<https://doi.org/10.1002/9780470627242.ch30>>. Citado na página 50.

FUJIOKA, T. et al. Classification of breast masses on ultrasound shear wave elastography using convolutional neural networks. *Ultrasonic Imaging*, SAGE Publications, v. 42, n. 4-5, p. 213–220, jun. 2020. Disponível em: <<https://doi.org/10.1177/0161734620932609>>. Citado 2 vezes nas páginas 46 e 55.

- HART, R. E. et al. Galaxy zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 461, p. 3663–3682, 10 2016. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1093/mnras/stw1588>>. Citado na página 29.
- HE, K. et al. Deep Residual Learning for Image Recognition. *arXiv*, dez. 2015. Disponível em: <<https://doi.org/10.48550/arXiv.1512.03385>>. Citado na página 47.
- HE, K. et al. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [s.n.], 2016. p. 770–778. Disponível em: <<https://doi.org/10.1109/CVPR.2016.90>>. Citado na página 55.
- HINTON, G.; SRIVASTAVA, N.; SWERSKY, K. *On the Variance of the Adaptive Learning Rate and Beyond*. 2012. Citado na página 54.
- HUANG, G. et al. *Densely Connected Convolutional Networks*. 2018. Disponível em: <<https://doi.org/10.48550/arXiv.1608.06993>>. Citado na página 56.
- Hubble, E. P. Extragalactic nebulae. *Astrophysical Journal*, v. 64, p. 321–369, Dec 1926. Citado na página 24.
- II, D. W. E.; DINC, I. Classification of protein crystallization images using EfficientNet with data augmentation. In: *CSBio '20: Proceedings of the Eleventh International Conference on Computational Systems-Biology and Bioinformatics*. ACM, 2020. Disponível em: <<https://doi.org/10.1145/3429210.3429220>>. Citado 2 vezes nas páginas 50 e 55.
- JELODAR, H. et al. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, Springer New York LLC, v. 78, p. 15169–15211, 6 2019. ISSN 1380-7501. Disponível em: <<https://doi.org/10.1007/s11042-018-6894-4>>. Citado na página 52.
- JIANG, L. et al. Structure extended multinomial naive bayes. *Information Sciences*, Elsevier, v. 329, p. 346–356, 2 2016. ISSN 0020-0255. Citado na página 51.
- KALCHEVA, N.; MARINOVA, G.; TODOROVA, M. Comparative analysis of the bernoulli and multinomial naive bayes classifiers for text classification in machine learning. In: *2023 International Conference Automatics and Informatics (ICAI)*. IEEE, 2023. p. 28–31. ISBN 979-8-3503-1291-1. Disponível em: <<https://ieeexplore.ieee.org/document/10339077/>>. Citado na página 51.
- KANG, D.-K. et al. Multinomial event model based abstraction for sequence and text classification. In: _____. Springer, Berlin, Heidelberg, 2005. v. 3607 LNAI, p. 134–148. ISBN 978-3-540-31882-8. Disponível em: <http://doi.org/10.1007/11527862_10>. Citado na página 51.
- KIBRIYA, A. M. et al. Multinomial naive bayes for text categorization revisited. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, Springer, Berlin, Heidelberg, v. 3339, p. 488–499, 2004. ISSN

1611-3349. Disponível em: <https://doi.org/10.1007/978-3-540-30549-1_43>. Citado na página 51.

KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Disponível em: <<https://doi.org/10.48550/arXiv.1412.6980>>. Citado na página 54.

KRIEGSMANN, M. et al. Deep learning for the classification of small-cell and non-small-cell lung cancer. *Cancers*, MDPI AG, v. 12, n. 6, p. 1604, jun. 2020. Disponível em: <<https://doi.org/10.3390/cancers12061604>>. Citado 2 vezes nas páginas 46 e 55.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. v. 25. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>. Citado na página 47.

LEMMENS, R. et al. Citizen science in the digital world of apps. In: _____. Springer International Publishing, 2021. p. 461–474. ISBN 978-3-030-58278-4. Disponível em: <https://doi.org/10.1007/978-3-030-58278-4_23>. Citado na página 26.

LI, L. et al. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *arXiv*, mar. 2016. Disponível em: <<https://doi.org/10.48550/arXiv.1603.06560>>. Citado na página 56.

LI, X. et al. Classification of breast cancer histopathological images using interleaved densenet with senet (idsnet). *PLoS ONE*, v. 15, 2020. Disponível em: <<https://doi.org/10.1371/journal.pone.0232127>>. Citado na página 56.

LI, Y. et al. Automl: A perspective where industry meets academy. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, p. 4048–4049, 8 2021. Disponível em: <<https://doi.org/10.1145/3447548.3470827>>. Citado na página 56.

LI, Z. et al. Protein contact map prediction based on resnet and densenet. *BioMed Research International*, v. 2020, 2020. Disponível em: <<https://doi.org/10.1155/2020/7584968>>. Citado na página 56.

LIN, T.-Y. et al. Microsoft COCO: Common Objects in Context. *arXiv*, maio 2014. Disponível em: <<https://doi.org/10.48550/arXiv.1405.0312>>. Citado na página 47.

LINTOTT, C. et al. Galaxy zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 410, p. 166–178, 1 2011. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1111/j.1365-2966.2010.17432.x>>. Citado na página 29.

LINTOTT, C. J. et al. Galaxy zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 389, p. 1179–1189, 9 2008.

- ISSN 13652966. Disponível em: <<https://dx.doi.org/10.1111/j.1365-2966.2008.13689.x>>. Citado na página 26.
- LIU, L. et al. *On the Variance of the Adaptive Learning Rate and Beyond*. 2020. Disponível em: <<https://doi.org/10.48550/arXiv.1908.03265>>. Citado na página 54.
- LIU, Z. et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv*, mar. 2021. Disponível em: <<https://doi.org/10.48550/arXiv.2103.14030>>. Citado na página 47.
- LUO, J.; WONG, C.-M.; VONG, C.-M. Multinomial bayesian extreme learning machine for sparse and accurate classification model. *Neurocomputing*, Elsevier, v. 423, p. 24–33, 1 2021. ISSN 09252312. Disponível em: <<https://doi.org/10.1016/j.neucom.2020.09.061>>. Citado na página 51.
- MAHDIANPARI, M. et al. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing*, MDPI AG, v. 10, n. 7, p. 1119, jul. 2018. Disponível em: <<https://doi.org/10.3390/rs10071119>>. Citado na página 55.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, Kluwer Academic Publishers, v. 5, n. 4, p. 115–133, dez. 1943. ISSN 1522-9602. Disponível em: <<https://doi.org/10.1007/BF02478259>>. Citado na página 46.
- MCREYNOLDS, T.; BLYTHE, D. Antialiasing. In: *Advanced Graphics Programming Using OpenGL*. Elsevier, 2005. p. 169–184. Disponível em: <<https://doi.org/10.1016/b978-155860659-3.50012-3>>. Citado na página 50.
- MITTAL, A. et al. Data augmentation based morphological classification of galaxies using deep convolutional neural network. *Earth Science Informatics*, Springer Science and Business Media LLC, v. 13, n. 3, p. 601–617, dez. 2019. Disponível em: <<https://doi.org/10.1007/s12145-019-00434-8>>. Citado na página 50.
- NEVASALMI, L. Forecasting multinomial stock returns using machine learning methods. *The Journal of Finance and Data Science*, Elsevier, v. 6, p. 86–106, 11 2020. ISSN 24059188. Disponível em: <<https://doi.org/10.1016/j.jfds.2020.09.001>>. Citado na página 51.
- NONAKA, N.; SEITA, J. Electrocardiogram classification by modified EfficientNet with data augmentation. In: *2020 Computing in Cardiology Conference (CinC)*. Computing in Cardiology, 2020. Disponível em: <<https://doi.org/10.22489/cinc.2020.063>>. Citado 2 vezes nas páginas 50 e 55.
- ODEH, A. H.; ODEH, M.; ODEH, N. Using multinomial naive bayes machine learning method to classify, detect, and recognize programming language source code. *Proceedings - 2022 23rd International Arab Conference on Information Technology, ACIT 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. Disponível em: <<https://doi.org/10.1109/ACIT57182.2022.9994117>>. Citado na página 51.

OZAKI, Y. et al. Multiobjective tree-structured parzen estimator for computationally expensive optimization problems. In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. ACM, 2020. p. 533–541. ISBN 9781450371285. Disponível em: <<https://doi.org/10.1145/3377930.3389817>>. Citado na página 56.

REZENDE, E. et al. Malicious software classification using VGG16 deep neural network's bottleneck features. In: *Advances in Intelligent Systems and Computing*. Springer International Publishing, 2018. p. 51–59. Disponível em: <https://doi.org/10.1007/978-3-319-77028-4_9>. Citado na página 55.

ROSENBLATT, F. Perceptron simulation experiments. *Proceedings of the IRE*, v. 48, n. 3, p. 301–309, 1960. Disponível em: <<https://doi.org/10.1109/JRPROC.1960.287598>>. Citado na página 46.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, Nature Publishing Group, v. 323, p. 533–536, out. 1986. ISSN 1476-4687. Disponível em: <<https://doi.org/10.1038/323533a0>>. Citado na página 46.

RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015. Disponível em: <<https://doi.org/10.1007/s11263-015-0816-y>>. Citado na página 55.

SARAVANAN, U. K. et al. Multinomial naive bayes based machine learning analysis of twitter sentiment. *Proceedings of the 2nd International Conference on Edge Computing and Applications, ICECAA 2023*, Institute of Electrical and Electronics Engineers Inc., p. 429–434, 2023. Disponível em: <<https://doi.org/10.1109/ICECAA58104.2023.10212150>>. Citado na página 51.

SEBER, G. A. F. Multinomial distribution. In: _____. Springer International Publishing, 2015. p. 181–188. ISBN 978-3-319-21930-1. Disponível em: <https://doi.org/10.1007/978-3-319-21930-1_12>. Citado na página 51.

SILVERTOWN, J. A new dawn for citizen science. *Trends in ecology & evolution*, v. 24, n. 9, p. 467–71, set. 2009. ISSN 0169-5347. Disponível em: <<https://doi.org/10.1016/j.tree.2009.03.017>>. Citado na página 26.

SIMARD, P. Y.; STEINKRAUS, D.; PLATT, J. C. Best practices for convolutional neural networks applied to visual document analysis. In: *Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2*. USA: IEEE Computer Society, 2003. (ICDAR '03), p. 958. ISBN 0769519601. Disponível em: <<https://doi.org/10.5555/938980.939477>>. Citado na página 50.

SIMMONS, B. D. et al. Galaxy zoo: Quantitative visual morphological classifications for 48 000 galaxies from candels. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 464, p. 4420–4447, 2 2017. ISSN 13652966. Disponível em: <<https://doi.org/10.1093/mnras/stw2587>>. Citado na página 29.

- SIMONYAN, K.; ZISSERMAN, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*, set. 2014. Disponível em: <<https://doi.org/10.48550/arXiv.1409.1556>>. Citado na página 47.
- SIMONYAN, K.; ZISSERMAN, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2015. Disponível em: <<https://doi.org/10.48550/arXiv.1409.1556>>. Citado na página 55.
- SOLANO-ROJAS, B.; VILLALÓN-FONSECA, R.; MARÍN-RAVENTÓS, G. Alzheimer's disease early detection using a low cost three-dimensional densenet-121 architecture. In: *Lecture Notes in Computer Science*. Springer International Publishing, 2020. p. 3–15. Disponível em: <https://doi.org/10.1007/978-3-030-51517-1_1>. Citado na página 56.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, n. 56, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Citado na página 54.
- STEINMETZ, M.; NAVARRO, J. F. The hierarchical origin of galaxy morphologies. *New Astronomy*, Elsevier, v. 7, n. 4, p. 155–160, 2002. Citado na página 23.
- SZEGEDY, C. et al. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. 2016. Disponível em: <<https://doi.org/10.48550/arXiv.1602.07261>>. Citado na página 55.
- SZEGEDY, C. et al. Going Deeper with Convolutions. *arXiv*, set. 2014. Disponível em: <<https://doi.org/10.48550/arXiv.1409.4842>>. Citado na página 47.
- SZEGEDY, C. et al. Going deeper with convolutions. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [s.n.], 2015. p. 1–9. Disponível em: <<https://doi.org/10.1109/CVPR.2015.7298594>>. Citado na página 55.
- TAN, M.; LE, Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv*, maio 2019. Disponível em: <<https://doi.org/10.48550/arXiv.1905.11946>>. Citado na página 47.
- TAN, M.; LE, Q. V. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. 2020. Disponível em: <<https://doi.org/10.48550/arXiv.1905.11946>>. Citado na página 55.
- TORRES, G. de O.; GUTERRES, M. X.; CELESTINO, V. R. R. Legal actions in brazilian air transport: A machine learning and multinomial logistic regression analysis. *Frontiers in Future Transportation*, Frontiers Media SA, v. 4, p. 1070533, 4 2023. ISSN 2673-5210. Disponível em: <<https://doi.org/10.3389/ffutr.2023.1070533>>. Citado na página 51.
- TU, Z. et al. MaxViT: Multi-Axis Vision Transformer. *arXiv*, abr. 2022. Disponível em: <<https://doi.org/10.48550/arXiv.2204.01697>>. Citado na página 47.
- VASU, P. K. A. et al. FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization. *arXiv*, mar. 2023. Disponível em: <<https://doi.org/10.48550/arXiv.2303.14189>>. Citado na página 47.

VASWANI, A. et al. Attention Is All You Need. *arXiv*, jun. 2017. Disponível em: <<https://doi.org/10.48550/arXiv.1706.03762>>. Citado na página 47.

VOHLAND, K. et al. Editorial: The science of citizen science evolves. In: _____. Springer International Publishing, 2021. p. 1–12. ISBN 978-3-030-58278-4. Disponível em: <https://doi.org/10.1007/978-3-030-58278-4_1>. Citado na página 26.

WALMSLEY, M. et al. Galaxy zoo desi: Detailed morphology measurements for 8.7m galaxies in the desi legacy imaging surveys. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 526, p. 4768–4786, 10 2023. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1093/mnras/stad2919>>. Citado na página 29.

WALMSLEY, M. et al. Galaxy Zoo DECaLS: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies. *Mon. Not. R. Astron. Soc.*, Oxford Academic, v. 509, n. 3, p. 3966–3988, jan. 2022. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1093/mnras/stab2093>>. Citado na página 46.

WALMSLEY, M. et al. Galaxy zoo decals: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 509, p. 3966–3988, 12 2021. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1093/mnras/stab2093>>. Citado 2 vezes nas páginas 29 e 33.

WANG, L.; ZHANG, M. Statistical modeling of dengue transmission dynamics with environmental factors. *Computational Statistics & Data Analysis*, North-Holland, v. 203, p. 108080, 3 2025. ISSN 01679473. Disponível em: <<https://doi.org/10.1016/j.csda.2024.108080>>. Citado na página 51.

WANG, S.-H.; ZHANG, Y.-D. Densenet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification. *ACM Trans. Multimedia Comput. Commun. Appl.*, Association for Computing Machinery, New York, NY, USA, v. 16, n. 2s, jun. 2020. ISSN 1551-6857. Disponível em: <<https://doi.org/10.1145/3341095>>. Citado na página 56.

WILLETT, K. W. et al. Galaxy zoo: morphological classifications for 120 000 galaxies in hst legacy imaging. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 464, p. 4176–4203, 2 2017. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1093/mnras/stw2568>>. Citado na página 29.

WILLETT, K. W. et al. Galaxy zoo 2: detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, Oxford Academic, v. 435, p. 2835–2860, 11 2013. ISSN 0035-8711. Disponível em: <<https://dx.doi.org/10.1093/mnras/stt1458>>. Citado na página 29.

YAEGER, L.; LYON, R.; WEBB, B. Effective training of a neural network character classifier for word recognition. In: *Proceedings of the 9th International Conference on Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 1996. (NIPS'96), p. 807–813. Citado na página 49.

ZHANG, P.; YANG, L.; LI, D. EfficientNet-b4-ranger: A novel method for greenhouse cucumber disease recognition under natural complex environment. *Computers and Electronics in Agriculture*, Elsevier BV, v. 176, p. 105652, set. 2020. Disponível em: <<https://doi.org/10.1016/j.compag.2020.105652>>. Citado 2 vezes nas páginas 46 e 55.