



SISTEMA INTELIGENTE PARA BUSCA DE OBJETOS ASTRONÔMICOS A PARTIR DA SIMILARIDADE VISUAL

Iniciação Científica

Relatório Parcial

Natanael Magalhães Cardoso

Orientadora: Profa. Cláudia Mendes de Oliveira

**Universidade de São
Paulo**

Instituto de Astronomia,
Geofísica e Ciências Atmosféricas

Departamento de Astronomia



UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE ASTRONOMIA, GEOFÍSICA E CIÊNCIAS ATMOSFÉRICAS

Departamento de Astronomia



Sistema Inteligente para Busca de Objetos Astronômicos a partir da Similaridade Visual

Natanael Magalhães Cardoso*, , Profa. Cláudia Mendes de Oliveira†

I N F O

* Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica, Universidade de São Paulo.

† Departamento de Astronomia, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo.

 nauxmac@gmail.com

R E S U M O

O aumento exponencial na geração de dados astronômicos, impulsionado por levantamentos modernos demanda soluções computacionais eficientes. O objetivo deste trabalho é criar uma ferramenta que automatize a identificação e a análise de objetos celestes com propriedades visuais semelhantes. O sistema baseia-se em redes neurais convolucionais (CNNs) treinadas com imagens para a extração de representações visuais (também chamadas de *embeddings*), que 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. Exemplos de buscas realizadas mostraram a aplicabilidade prática do sistema, que é capaz de entregar resultados relevantes mesmo em cenários complexos.

Palavras-chave: astronomia extra-galática, visão computacional, redes neurais convolucionais, deep learning, similarity learning.

SUMÁRIO

1	Introdução	8
2	Aspectos Conceituais	10
2.1	Levantamentos Astronômicos	10
2.1.1	Sloan Digital Sky Survey	11
2.1.2	Southern Photometric Local Universe Survey	12
2.1.3	Legacy Surveys	14
2.2	Padrões e Protocolos na Astronomia	15
2.3	Morfologia	17
2.3.1	Raio Efetivo	17
2.3.2	Elipticidade	18
2.3.3	Brilho e Magnitude	19
2.3.4	Classificação Morfológica	19
2.4	Ciência Cidadã	20
2.5	Aprendizagem Profunda	21
2.5.1	Redes Neurais	22
2.5.2	Aprendizado Profundo em Visão Computacional	22
2.5.3	Recuperação de Imagens Baseado em Conteúdo	23
2.6	Sistemas de Informação	24
2.6.1	Bancos de Dados Relacionais	24
2.6.2	API RESTful	25
2.6.3	Aplicações Web Modernas	26

3 Especificação dos Requisitos	28
3.1 Requisitos Funcionais	28
3.1.1 Extração de Características Visuais	28
3.1.2 Base de Dados para Armazenamento e Consulta de Representações Visuais	28
3.1.3 Interface de Consulta para o Usuário	29
3.2 Requisitos Não Funcionais	29
3.2.1 Desempenho	30
3.2.2 Usabilidade e Experiência do Usuário	30
3.2.3 Interoperabilidade	30
3.2.4 Eficiência e Otimização de Recursos	31
3.2.5 Precisão e Consistência dos Resultados	31
4 Conjuntos de Dados	32
4.1 Aquisição dos Votos do GalaxyZoo	32
4.2 Ajuste do Campo de Visão Angular	33
4.3 Aquisição das Imagens	40
4.4 Descrição dos Conjuntos de Dados	42
4.4.1 Conjuntos de Treinamento, Validação e Teste	42
4.4.2 Conjunto de Inferência	43
5 Modelo de Aprendizagem Profunda	44
5.1 Preparação das Características	45
5.1.1 Preparação das Características de Entrada	45
5.1.2 Preparação dos Rótulos	46
5.2 Aumento Artificial de Dados	46
5.3 Função de Custo	48

5.3.1	Distribuição Multinomial	48
5.3.2	Distribuição de Dirichlet	49
5.3.3	Distribuição Dirichlet-Multinomial	50
5.3.4	Função de Custo e Função de Perda	50
5.4	Hiperparâmetros	51
5.5	Arquitetura de Rede Neural	52
5.5.1	Visual Geometry Group Networks	52
5.5.2	InceptionResNetV2	53
5.5.3	EfficientNet	53
5.5.4	DenseNet	53
5.5.5	Escolha da Arquitetura	54
5.6	Treinamento	54
5.7	Métricas de Avaliação da Predição dos Votos	56
5.7.1	Matriz de Confusão	56
5.7.2	Acurácia	57
5.7.3	Precisão	57
5.7.4	Revocação	57
5.7.5	F1-Score	58
5.8	Métricas de Avaliação da Similaridade	58
5.8.1	Distância do Cosseno	58
5.8.2	Produto Interno	59
5.8.3	Distância L1	59
5.8.4	Distância L2	60
5.8.5	Métrica de Desempenho da Busca	60
6	Sistema de Informação	61
6.1	Base de Dados	61

6.1.1	Tecnologias Utilizadas	62
6.1.2	Modelo de Dados	63
6.1.3	Armazenamento e Indexação das Coordenadas	65
6.1.4	Armazenamento e Indexação dos Vetores	66
6.2	Backend	67
6.2.1	Tecnologias Utilizadas	67
6.2.2	Integrações	68
6.2.3	Procedimento de Busca no Backend	68
6.3	Frontend	70
6.3.1	Tecnologias Utilizadas	70
6.3.2	Arquitetura	71
6.3.3	Procedimento de Busca no Frontend	72
6.4	Implementação e Implantação	72
6.4.1	Sistema de Controle de Versão	74
6.4.2	DevOps	74
7	Resultados	76
7.1	Avaliação do Modelo no Conjunto de Teste	76
7.2	Avaliação da Recuperação de Imagens	77
7.3	Avaliação do Sistema de Informação	77
7.3.1	Performance	78
7.3.2	Interface Gráfica	78
8	Conclusão	92
8.1	Sumário	92
8.2	Lições Aprendidas	93
8.3	Contribuição Científica	93

8.4 Trabalhos Futuros	94
--	-----------

1. INTRODUÇÃO

Com o avanço das tecnologias de captura e armazenamento de imagens em larga escala, levantamentos astronômicos modernos, como o Sloan Digital Sky Survey (SDSS; Seção 2.1.1) e o DESI Legacy Survey (Seção 2.1.3), têm produzido catálogos de imagens da ordem de dezenas de bilhões de objetos celestes. Logo, torna-se crescente a demanda por ferramentas eficientes e precisas para a análise de grandes volumes de dados astronômicos de forma automatizada.

Além disso, para a próxima década, está previsto o início da operação de novos telescópios projetados para ter uma capacidade de geração de dados jamais vista na astronomia, como o Giant Magellan Telescope (GMT; [FANSON et al., 2022](#)) e o Vera Rubin Observatory – Legacy Survey of Space and Time (LSST; [IVEZIĆ et al., 2019](#)). Esse aumento exponencial na quantidade de dados causarão novos desafios à pesquisa em astronomia, uma vez que a análise manual tornar-se-á inviável e métodos convencionais de classificação e busca não conseguem lidar com a complexidade e a diversidade das características visuais dos objetos celestes.

Por isso, o desenvolvimento de um sistema de busca baseado em similaridade visual é fundamental para identificar e classificar objetos com propriedades morfológicas e estruturais semelhantes, como galáxias espirais, elípticas e irregulares, que possuem características específicas associadas a sua formação e evolução. Esses aspectos são fundamentais para entender a estrutura do universo e os processos físicos que governam a formação das galáxias. O sistema de busca por similaridade visual, desenvolvido neste trabalho, permite que os astrônomos realizem comparações entre objetos visualmente similares de forma automática, facilitando o estudo de fenômenos astronômicos complexos e a identificação de padrões em grandes amostras de dados. Essa capacidade de análise comparativa é particularmente valiosa para estudos de lentes gravitacionais, interações galácticas e evolução cósmica.

O objetivo deste projeto é desenvolver um sistema inteligente de busca por similaridade visual que permita a pesquisadores realizar consultas visuais rápidas e precisas em grandes bancos de dados astronômicos, utilizando uma forma automática de inspeção visual. Esse sistema deve ser capaz de identificar objetos morfológicamente semelhantes com base em

suas propriedades visuais, promovendo uma análise detalhada e comparativa dos dados astronômicos

Inicialmente, na etapa de concepção, realizou-se uma investigação das principais características necessárias para que o sistema tivesse o impacto científico almejado com a especificação dos requisitos funcionais (Seção 3.1) e não funcionais (Seção 3.2).

Em seguida, na etapa de desenvolvimento, realizou-se uma análise dos levantamentos astronômicos disponíveis para selecionar os dados mais relevantes para o treinamento e avaliação do sistema (Seção 4). Esses levantamentos foram escolhidos devido à sua abrangência, qualidade e disponibilidade pública. Em seguida, as imagens foram baixadas (Seção 4.3) com o devido ajuste do campo de visão angular (Seção 4.2). O conjunto de dados foi segmentado por características físicas, assegurando que as amostras de treinamento, validação e teste fossem representativas e com menos viés possíveis (Seção 4.4).

Posteriormente, foi implementado o modelo de aprendizado profundo, desenvolvido com arquiteturas modernas (Seção 5.5). O treinamento do modelo foi otimizado por meio de técnicas como aumento de dados (Seção 5.2) e ajuste automático de hiperparâmetros (Seção 5.6). Além disso, uma função de custo customizada (Seção 5.3) foi projetada para atender aos requisitos específicos dos dados de treinamento utilizado.

Desta forma, foi feita a integração do modelo desenvolvido com um sistema de informação escalável. Para isso, foi desenvolvido um backend, responsável pelo gerenciamento das consultas de similaridade visual (Seção 6.2) e pela indexação eficiente dos embeddings gerados pelo modelo e de dados georreferenciados (Seção 6.1.1). O frontend, desenvolvido com uma framework de programação reativa e com gerenciamento de estados declarativo, foi projetado com foco na usabilidade, oferecendo uma interface intuitiva (Seção 6.3). A arquitetura do sistema foi complementada pelo uso de bancos de dados relacionais otimizados e pipelines de DevOps (Seção 6.4.2) para implantação e manutenção contínuas.

Por fim, o sistema foi validado por meio de experimentos quantitativos e qualitativos, utilizando métricas de avaliação de modelos (Seção 7.1) e de sistemas de recuperação de imagens (Seção 7.2). Além disso, testes em cenários reais foram conduzidos para demonstrar a aplicabilidade prática do sistema, incluindo buscas por objetos celestes específicos em grandes volumes de dados (Seção 7.2).

2. ASPECTOS CONCEITUAIS

2.1. LEVANTAMENTOS ASTRONÔMICOS

Um levantamento astronômico consiste em uma pesquisa sistemática de grandes áreas do céu com o intuito de coletar dados sobre objetos e fenômenos celestes (SCHNEIDER, 2014, p. 40-42). Estes levantamentos abrangem observações em diferentes faixas do espectro eletromagnético e buscam identificar e catalogar astros como estrelas, galáxias, aglomerados e nebulosas, fornecendo uma base de dados ampla e acessível para estudos detalhados (MICKAELIAN, 2016). A finalidade dos levantamentos é construir um mapa celeste que permita compreender a distribuição e as propriedades dos corpos celestes em diversas escalas, auxiliando na análise estatística de populações estelares e galácticas e na identificação de estruturas de larga escala, como filamentos e vazios cósmicos (BAHCALL, 1995; BALEISIS et al., 1998; JARRETT, 2004).

O grande volume de dados gerado por esses levantamentos astronômicos representa um desafio significativo. Os avanços em tecnologia de sensoriamento e armazenamento permitiram que telescópios modernos capturassem bilhões de objetos celestes em detalhes, gerando petabytes de dados que precisam ser processados, organizados e armazenados (GRAFE, 1993; ALEXANDER S. SZALAY et al., 2000). Este volume de dados impõe desafios em termos de infraestrutura de armazenamento, processamento e análise, além de exigir métodos eficientes de organização e recuperação de informações. Para enfrentar esses desafios, foram desenvolvidos os chamados observatórios virtuais (QUINN et al., 2004), que integram e centralizam os dados provenientes de diversos levantamentos. Estes observatórios permitem o acesso distribuído aos dados e promovem a interoperabilidade entre diferentes conjuntos de dados, facilitando a análise conjunta de observações realizadas em diferentes regiões do espectro e por distintos instrumentos (TAGHIZADEH-POPP et al., 2020).

A existência de múltiplos levantamentos astronômicos fotométricos se justifica pela necessidade de observar o universo em diferentes comprimentos de onda, uma vez que cada faixa do espectro revela características específicas dos objetos astronômicos. Por exemplo, a radiação ultravioleta é útil para identificar regiões de formação estelar intensa,

enquanto o infravermelho é empregado na observação de regiões obscurecidas por poeira interestelar (SCHNEIDER, 2014, p. 28-31). Assim, levantamentos em faixas distintas do espectro eletromagnético fornecem dados complementares que permitem uma visão mais completa dos processos físicos no cosmos. A variabilidade na profundidade (capacidade de observar objetos mais distantes e fracos) e na resolução espacial também diversifica os objetivos dos levantamentos. Nas próximas subseções, serão apresentadas breves descrições dos levantamentos utilizados no desenvolvimento deste projeto.

2.1.1 Sloan Digital Sky Survey

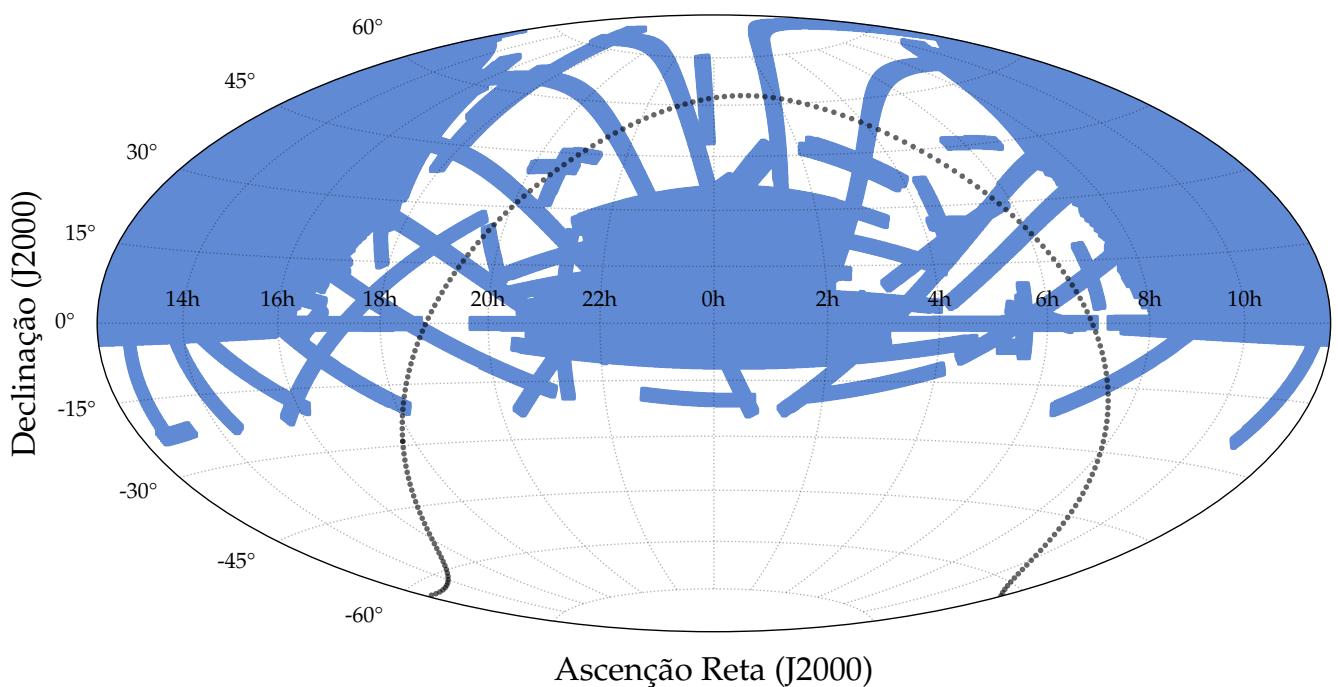


Figura 1: Área de cobertura do levantamento SDSS

O Sloan Digital Sky Survey¹ (SDSS; YORK et al., 2000) é um dos levantamentos astronômicos mais influentes e abrangentes, sendo pioneiro no uso de técnicas digitais para mapear o céu noturno. Iniciado em 2000, o SDSS empregou um telescópio dedicado de 2,5 metros (JAMES E. GUNN et al., 2006) equipado com uma poderosa câmera de dispositivo de carga acoplada (CCD; J. E. GUNN et al., 1998) para mapear o céu do norte com detalhes sem precedentes, cobrindo uma área de cerca de 14.000 graus quadrados, principalmente no hemisfério celestial norte. A Fig. 1 mostra a área de cobertura da décima oitava liberação de dados (DR18; ALMEIDA et al., 2023). Ao longo de suas diferentes fases, o

¹<https://sdss.org>

levantamento produziu dados detalhados de milhões de estrelas, galáxias e quasares, contribuindo significativamente para o avanço de áreas como cosmologia, formação de galáxias e estrutura de larga escala do universo.

O SDSS foi um dos primeiros levantamentos a realizar observações em múltiplas bandas fotométricas, abrangendo cinco faixas do espectro eletromagnético: u, g, r, i e z (FUKUGITA et al., 1996), com comprimentos de onda centrados em 3543 Å, 4770 Å, 6231 Å, 7625 Å e 9134 Å, respectivamente. Essas bandas cobrem desde o ultravioleta próximo até o infravermelho próximo, permitindo uma caracterização abrangente das propriedades físicas e químicas dos objetos observados. Esse mapeamento multiespectral fornece dados que possibilitam a estimativa de parâmetros astrofísicos, como temperatura e metalicidade, e facilita a classificação e estudo da evolução dos objetos celestes (HOGG et al., 2001).

O pioneirismo do SDSS como levantamento astronômico digital se destaca pela implementação de uma infraestrutura de dados digital chamada SkyServer² (ALEXANDER S SZALAY et al., 2002) e de técnicas automatizadas para a aquisição e processamento das imagens (HOGG et al., 2001) e dos espectros (SMEE et al., 2013). Os dados obtidos pelo SDSS foram disponibilizados ao público por meio de lançamentos periódicos, estabelecendo um novo padrão de acessibilidade e transparência na astronomia observacional. Esse levantamento digital não apenas tornou-se uma referência na organização e disseminação de dados, mas também promoveu o desenvolvimento de observatórios virtuais e inspirou a criação de novos levantamentos digitais em diferentes faixas espetrais, solidificando a transição para a era da astronomia baseada em big data (TAGHIZADEH-POPP et al., 2020).

2.1.2 Southern Photometric Local Universe Survey

Em contraste com o SDSS, o levantamento astronômico brasileiro Southern Photometric Local Universe Survey³ (S-PLUS; MENDES DE OLIVEIRA et al., 2019) é um mapeamento do hemisfério celestial sul utilizando uma abordagem multibanda altamente detalhada. O S-PLUS utiliza o telescópio T80-Sul, localizado no Observatório de Cerro Tololo, no Chile, para observar uma área de aproximadamente 8.000 graus quadrados, como mostra a Fig. 2. Este levantamento cobre, portanto, uma ampla região do céu meridional, permitindo um

²<https://skyserver.sdss.org>

³<https://splus.iag.usp.br>

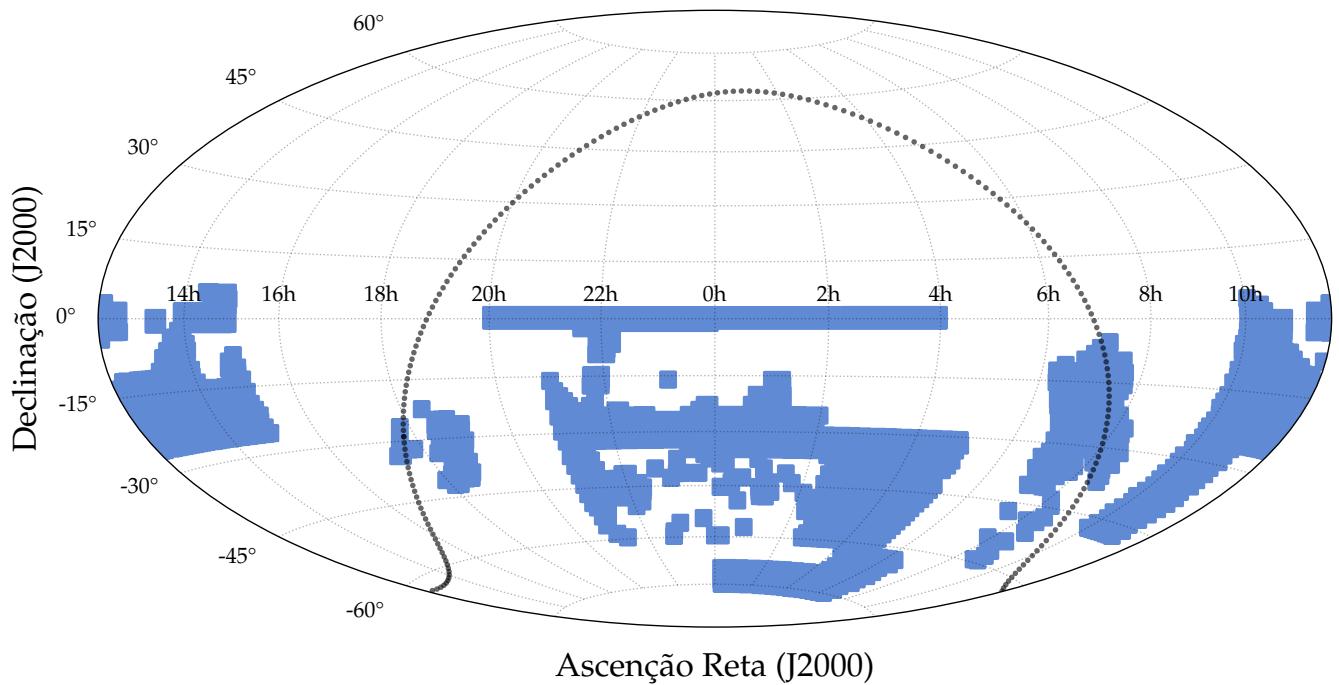


Figura 2: Área de cobertura do levantamento S-PLUS

estudo profundo de diversos objetos celestes, incluindo estrelas, galáxias e aglomerados de galáxias.

Um dos principais diferenciais do S-PLUS é a sua cobertura em doze bandas fotométricas, que incluem cinco bandas similares às do levantamento SDSS (u , g , r , i , z) e sete bandas estreitas (J0378, J0395, J0410, J0430, J0515, J0660, J0861) com comprimentos de onda variando entre 3780 Å e 8610 Å. Essa configuração de bandas permite uma análise espectral rica, capturando detalhes como a presença de linhas de emissão e absorção específicas, que são cruciais para a determinação precisa de parâmetros físicos e químicos dos objetos observados. Por exemplo, as bandas estreitas foram escolhidas para registrar características específicas como o traço do cálcio, magnésio e oxigênio, o que facilita a análise da composição estelar e a identificação de processos de formação estelar. Como levantamento astronômico digital, o S-PLUS disponibiliza seus dados para a comunidade científica em um formato acessível e padronizado através da plataforma S-PLUS Cloud (<https://splus.cloud>), promovendo a interoperabilidade com outros levantamentos e incentivando o desenvolvimento de estudos complexos em astronomia observational e cosmologia.

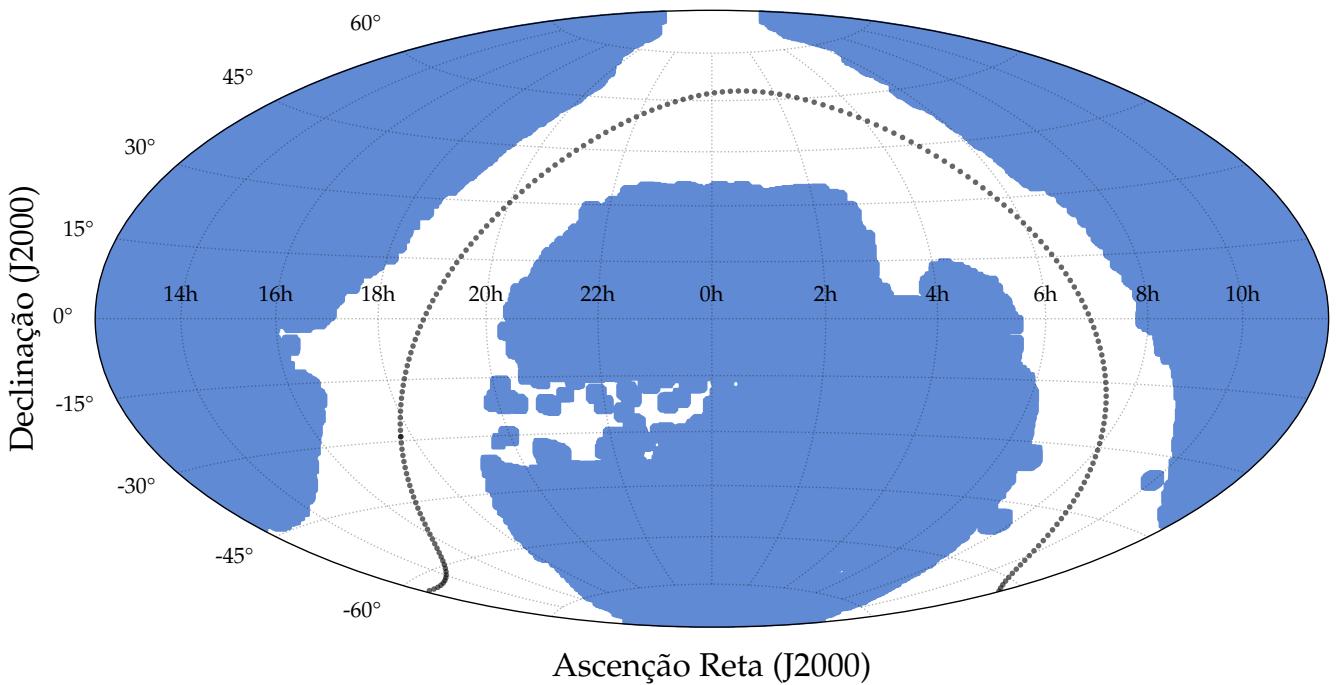


Figura 3: Área de cobertura do levantamento DESI Legacy Survey

2.1.3 Legacy Surveys

Os Legacy Surveys⁴ (DEY et al., 2019) consistem em três projetos individuais e complementares: o Dark Energy Camera Legacy Survey (DECaLS), o Beijing-Arizona Sky Survey (BASS) e o Mayall z-band Legacy Survey (MzLS). Juntos, formam um dos maiores projetos de mapeamento digital do céu, projetado para fornecer uma base fotométrica abrangente para o projeto Dark Energy Spectroscopic Instrument (DESI). O Legacy Survey cobre aproximadamente 14.000 graus quadrados do céu. Esse levantamento oferece uma cobertura profunda e em múltiplas bandas fotométricas, incluindo as bandas g , r e z , com comprimentos de onda centrados em aproximadamente 4750 Å, 6300 Å e 9200 Å, respectivamente. Essas bandas, juntamente com sua grande profundidade, permitem a detecção de objetos mais fracos e distantes, possibilitando estudos detalhados da distribuição de galáxias e outras estruturas em grandes escalas.

A profundidade dos Legacy Surveys é um dos fatores que o diferencia de outros levantamentos, como o SDSS e o S-PLUS. Com exposições projetadas para alcançar uma profundidade mais elevada, os Legacy Surveys são capazes de detectar objetos com brilho mais tênue, o que é fundamental para o mapeamento detalhado da estrutura de larga

⁴<https://legacysurvey.org>

escala do universo. Em comparação, o SDSS foi um dos primeiros levantamentos digitais a mapear o céu com precisão em cinco bandas, mas com uma profundidade moderada, voltada para observações mais próximas. Já o S-PLUS, com suas doze bandas, fornece uma caracterização espectral mais detalhada de cada objeto, mas com um foco em uma área específica do céu no hemisfério celestial sul e uma profundidade relativamente menor em comparação com os Legacy Surveys.

Assim, enquanto o SDSS estabeleceu o padrão para levantamentos digitais de grande escala, e o S-PLUS trouxe um detalhamento espectral fino com bandas estreitas, os Legacy Surveys combinam uma ampla cobertura com alta profundidade, posicionando-se como uma base essencial para o estudo da cosmologia e da estrutura de larga escala do universo. A integração desses levantamentos complementares permite uma análise mais completa do céu, tanto em termos de área quanto de profundidade e detalhamento espectral, promovendo avanços em diversas áreas da astronomia e da astrofísica.

2.2. PADRÕES E PROTOCOLOS NA ASTRONOMIA

Os padrões e protocolos são fundamentais para a pesquisa em astronomia, principalmente devido ao crescente volume e à diversidade dos dados astronômicos disponíveis. Como abordado na seção anterior, cada levantamento astronômico tem um propósito específico, observando fenômenos físicos distintos (a partir de escolhas tomadas na construção do equipamento, como os comprimentos de onda observados, a quantidade de filtros e a resolução da câmera) em diferentes regiões do céu. Nesse sentido, a padronização é essencial para garantir que fontes de dados distintas sejam acessíveis e compatíveis entre si. Ao adotar padrões, torna-se possível integrar informações de diferentes observatórios e levantamentos, facilitando a reutilização de dados e promovendo estudos comparativos em larga escala. Isso cria um ecossistema em que dados de múltiplos projetos podem ser analisados de maneira conjunta, ampliando a possibilidade de descobertas científicas que dependem da integração de informações heterogêneas.

Os protocolos da International Virtual Observatory Alliance⁵ (IVOA; QUINN et al., 2004) desempenham um papel crucial na astronomia, viabilizando a interoperabilidade, o acesso e a utilização de dados astronômicos provenientes de diferentes observatórios e

⁵<https://ivoa.net>

levantamentos. Criados para unificar e padronizar o compartilhamento e a manipulação de dados, esses protocolos compõem uma estrutura abrangente e organizada que facilita o intercâmbio de informações em um ambiente colaborativo global. Com a crescente quantidade e diversidade dos dados astronômicos, o IVOA estabelece normas que garantem que informações oriundas de várias fontes possam ser integradas e acessadas de maneira consistente, independente do formato ou origem, fomentando avanços na pesquisa astronômica.

O conjunto de protocolos IVOA é estruturado em diversas camadas, que funcionam de forma integrada para compor o ambiente de serviço do Observatório Virtual (VO; DOWLER, EVANS et al., 2021). A *Camada de Protocolos de Acesso a Dados* é uma das principais e inclui o Data Access Layer Interface (DALI; DOWLER, DEMLEITNER et al., 2017), que define os princípios gerais de interação com serviços de dados, o Table Access Protocol (TAP; DOWLER, GUY RIXON et al., 2019) para acesso a tabelas, e os protocolos especializados como o Simple Cone Search (SCS; PLANTE et al., 2008), o Simple Spectral Access Protocol (SSAP; TODY et al., 2012) e o Simple Image Access Protocol (SIAP; DOWLER, FRANÇOIS BONNAREL et al., 2015), que facilitam o acesso a catálogos, espectros e imagens, respectivamente. Essas ferramentas possibilitam que os usuários recuperem informações específicas em diferentes bases de dados astronômicas. Por outro lado, a *Camada de Serviços* gerencia a integração de serviços, com o uso de ferramentas como o Universal Worker Service (UWS; HARRISON e G. RIXON, 2016) para coordenar tarefas assíncronas em servidores de dados.

Além dos protocolos de acesso, a *Camada de Formatos* define como os dados devem ser representados e trocados. Os formatos VOTable (OCHSENBEIN et al., 2019), usado para tabelas de dados, e FITS (WELLS e GREISEN, 1979), adotado tanto para imagens e tabelas, garantem que os dados sejam armazenados de forma padronizada, permitindo sua compatibilidade e compreensão entre diferentes sistemas. Já os *Modelos de Dados* (MICHEL et al., 2023) especificam a estrutura e o significado dos dados, fornecendo uma semântica comum para que diferentes sistemas interpretem corretamente o conteúdo dos catálogos e tabelas. Por exemplo, os modelos incluem descrições de propriedades como fluxos, coordenadas e características espetrais, facilitando comparações e agregações de dados de diferentes fontes.

A *Camada de Protocolos Semânticos* inclui padrões como Unified Content Descrip-

tors (UCD; [CECCONI et al., 2023](#)), que proporciona uma nomenclatura padronizada para descrever o conteúdo dos dados, e o VOUnit ([GRAY et al., 2023](#)), que define unidades de medida de forma consistente. Esses protocolos semânticos são essenciais para assegurar que os dados sejam interpretados corretamente por pesquisadores e softwares, evitando ambiguidades e promovendo uma compreensão clara dos conjuntos de dados. Por outro lado, o uso de linguagens de consulta, como a Astronomical Data Query Language (ADQL; [MANTELET et al., 2023](#)), oferece uma sintaxe padronizada para a realização de consultas avançadas em bancos de dados astronômicos, permitindo filtragens e extrações complexas diretamente nos repositórios de dados do VO.

A interação entre essas camadas cria um ambiente de serviço VO em que cada componente desempenha um papel específico para garantir que os dados possam ser acessados, analisados e interpretados de forma integrada. Essa integração e padronização é fundamental para a operação automatizada de um sistema inteligente para descoberta em astronomia.

2.3. MORFOLOGIA

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 e 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 ([VAN DEN BERGH, 1998](#)).

2.3.1 Raio Efetivo

O raio efetivo de uma galáxia é uma medida fundamental em astronomia que caracteriza a distribuição espacial de sua luminosidade. Em termos práticos, o raio efetivo é definido como a distância do centro da galáxia até o ponto onde metade da luz total emitida galáxia está concentrada. Essa métrica é amplamente utilizada para descrever a extensão aparente das galáxias e fornece informações valiosas sobre sua estrutura interna e dinâmica.

O cálculo do raio efetivo depende da análise da distribuição de brilho da galáxia, sendo comumente obtido por meio de perfis de brilho de superfície que representam a intensidade luminosa em função da distância ao centro da galáxia. Para galáxias elípticas, esses perfis de brilho são tipicamente modelados por uma distribuição de Sérsic, que se ajusta bem à sua estrutura concentrada. Já para galáxias espirais, a distribuição de brilho tende a ser mais complexa, com um núcleo brilhante e uma extensão de braços espirais, o que pode exigir ajustes mais detalhados para determinar o raio efetivo.

2.3.2 Elipticidade

A elipticidade complexa é uma métrica amplamente utilizada em estudos de lentes gravitacionais, onde se faz necessária a quantificação precisa da distorção observada nas imagens de objetos distantes devido à ação gravitacional de um objeto massivo interveniente, como uma galáxia ou um aglomerado de galáxias. Essa métrica é essencial para descrever a forma das imagens distorcidas, permitindo a análise da distribuição de massa do objeto que atua como lente. Especificamente, a elipticidade complexa captura a deformação das fontes de luz distantes sob a forma de uma componente elíptica, expressando tanto a intensidade da distorção (ou seja, o alongamento da imagem) quanto a sua orientação no plano do céu.

A métrica de elipticidade complexa é particularmente útil na modelagem e análise de lentes fracas, em que as distorções são sutis e a forma das galáxias de fundo serve como uma sonda estatística para inferir a presença de matéria escura e a estrutura da lente gravitacional. Essa métrica desempenha um papel crucial para estudos que investigam a distribuição de matéria escura e a formação de grandes estruturas no universo. Ao quantificar a distorção induzida por lentes gravitacionais, os astrônomos podem mapear a distribuição de massa em escalas cósmicas, incluindo matéria que não emite luz visível, como a matéria escura. Em estudos de lenteamento fraco, onde os efeitos de distorção são mínimos e requerem uma análise estatística sobre grandes populações de galáxias, a elipticidade complexa oferece uma forma consistente e precisa de capturar e interpretar essas pequenas distorções.

2.3.3 Brilho e Magnitude

O brilho e a magnitude são conceitos fundamentais para quantificar a intensidade de luz emitida ou refletida por objetos celestes, sendo essenciais para a caracterização de estrelas, galáxias e outros corpos celestes. O brilho refere-se à quantidade total de energia luminosa recebida por unidade de área em um detector, como um telescópio. É uma medida física direta da intensidade de luz e varia inversamente com o quadrado da distância do objeto, sendo comumente expresso em termos de fluxo, que é a quantidade de luz recebida por unidade de área por unidade de tempo. A medição do brilho permite aos astrônomos determinar características intrínsecas dos objetos, como temperatura e tamanho, e é um dos parâmetros centrais para estudar a evolução estelar e a estrutura galáctica.

A magnitude, por outro lado, é uma medida que descreve o brilho percebido dos objetos astronômicos em escala logarítmica, introduzida para lidar com a vasta gama de intensidades luminosas observadas no céu. A escala de magnitude é inversa, de forma que objetos mais brilhantes têm magnitudes menores, enquanto objetos mais fracos têm magnitudes maiores. A magnitude aparente quantifica o brilho de um objeto visto da Terra, enquanto a magnitude absoluta define o brilho que o objeto teria se estivesse a uma distância padrão de 10 parsecs. Essa distinção permite que astrônomos comparem o brilho intrínseco de diferentes objetos, independentemente de sua distância da Terra.

2.3.4 Classificação Morfológica

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 e 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 observacional, onde padrões morfológicos podem ser correlacionados com fenômenos físicos específicos, como fusões e interações gravitacionais entre galáxias ([VAN DEN 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.

A importância da classificação morfológica na astronomia reside na sua capacidade de revelar padrões fundamentais sobre a evolução das galáxias e do universo como um todo. A morfologia galáctica não apenas reflete o estado atual de uma galáxia, mas também oferece pistas sobre sua história dinâmica, como colisões passadas, interações com outras galáxias e a presença de matéria escura. Por meio da classificação morfológica, os astrônomos podem investigar como as diferentes populações de galáxias variam ao longo do tempo cósmico, relacionando suas características estruturais com fatores como a densidade do ambiente e a quantidade de gás disponível para a formação estelar. Assim, a classificação morfológica se consolida como um método essencial para compreender a formação e evolução das galáxias, contribuindo para um entendimento mais profundo sobre a estrutura e a dinâmica do universo.

2.4. CIÊNCIA CIDADÃ

A ciência cidadã 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 (BONNEY et al., 2014; SILVERTOWN, 2009). 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).

Na astronomia, o GalaxyZoo⁶ (C. J. 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 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.

2.5. APRENDIZAGEM PROFUNDA

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

⁶<https://galaxyzoo.org>

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.

2.5.1 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 et al. \(1986\)](#), que possibilitou o treinamento de redes com múltiplas camadas ocultas, abrindo caminho para a evolução do campo.

2.5.2 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, C. LINTOTT, GERON et al., 2022), medicina (BELAID e LOUDINI, 2020; FUJIOKA et al., 2020; KRIEGSMANN et al., 2020) e agronomia (ABAS et al., 2018; P. ZHANG et al., 2020). Isso se iniciou com as primeiras arquiteturas convolucionais profundas, como AlexNet (Krizhevsky et al., 2012) e VGG (Simonyan e Zisserman, 2014), passando por Inception (Szegedy, W. Liu et al., 2014), ResNet (He et al., 2015) e EfficientNet (Tan e Le, 2019), até chegar nos modelos baseados em transformadores (Vaswani et al., 2017), como Vision Transformer (Dosovitskiy et al., 2020) e Swin Transformer (Z. 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.

2.5.3 Recuperação de Imagens Baseado em Conteúdo

A recuperação de imagem baseada em conteúdo (CBIR) é uma tarefa de grande relevância na área de visão computacional, que consiste em buscar imagens em um banco de dados utilizando características visuais extraídas automaticamente, em vez de depender de metadados ou descrições textuais. O objetivo principal dessa abordagem é identificar imagens visualmente similares à consulta fornecida pelo usuário, considerando atributos como cor, textura, forma ou padrões visuais específicos.

Inicialmente, as técnicas de CBIR baseavam-se em métodos tradicionais de extração de características. Algoritmos como histogramas de cor, transformadas para análise de textura (como Gabor) e detectores de bordas (como Canny) eram amplamente utilizados.

Essas abordagens geravam descritores manuais, que eram comparados utilizando métricas de similaridade, como distância euclidiana ou similaridade do cosseno. Apesar de funcionais, esses métodos eram limitados pela incapacidade de capturar características semânticas mais complexas, o que frequentemente resultava em baixa precisão em bancos de dados heterogêneos.

Com os avanços em aprendizado profundo, a utilização de redes neurais convolucionais (Convolutional Neural Networks, CNNs) transformou a abordagem de CBIR. As CNNs possuem a capacidade de aprender representações hierárquicas diretamente dos dados, extraíndo características visuais de baixo nível (como bordas) e abstrações de alto nível (como formas e contextos). Essas representações visuais extraídas são codificadas em forma de vetores, também chamados de *embeddings*. Por isso, modelos de aprendizado profundo tornaram-se ferramentas comuns para gerar tais representações visuais das imagens, que podem ser armazenadas e comparadas eficientemente. Essa abordagem não apenas melhora a precisão da recuperação, mas também permite lidar com variações de escala, rotação e iluminação das imagens.

As implementações mais avançadas de CBIR empregam técnicas como aprendizado de métrica e redes neurais siamêsas para otimizar diretamente a tarefa de similaridade visual. Além disso, combinações de embeddings extraídos de CNNs com informações adicionais, como metadados ou descrições textuais, têm sido exploradas para sistemas mais robustos e contextualmente relevantes. Em aplicações complexas, como astronomia ou medicina, onde os bancos de dados são massivos e as diferenças sutis entre imagens são críticas, essas técnicas modernas têm mostrado eficácia superior, tornando a CBIR uma ferramenta essencial para análise de dados e descoberta científica.

2.6. SISTEMAS DE INFORMAÇÃO

2.6.1 Bancos de Dados Relacionais

Um banco de dados relacional (CODD, 1970; 1979) é um sistema de armazenamento e organização de dados que utiliza o modelo relacional para estruturar as informações em tabelas, ou relações, permitindo um acesso eficiente e seguro aos dados armazenados. Nesse modelo, cada tabela é composta por linhas e colunas, onde as linhas representam registros (ou tuplas) e as colunas representam os atributos (ou campos) que descrevem as

características de cada registro. Essa organização de dados possibilita o estabelecimento de relações entre as tabelas por meio de chaves primárias e chaves estrangeiras, facilitando a associação de informações de diferentes tabelas com base em valores comuns.

A principal característica do modelo relacional é a independência dos dados, que permite que as informações sejam manipuladas sem que o usuário precise conhecer detalhes do armazenamento físico. Os dados são acessados e manipulados por meio da Structured Query Language (SQL; [CHAMBERLIN e BOYCE, 1974](#)), uma linguagem de consulta declarativa que permite a execução de operações como inserção, atualização, exclusão e recuperação de dados.

A integridade e a consistência dos dados em um banco de dados relacional são asseguradas pelo uso de restrições e transações. As restrições, como as de chave primária, chave estrangeira e unicidade, garantem que os dados mantêm-se válidos de acordo com as regras estabelecidas para cada relação. As transações, por sua vez, permitem que múltiplas operações sejam executadas de forma atômica, garantindo que, em caso de falha, o banco de dados possa retornar a um estado consistente, preservando a integridade das informações.

Os bancos de dados relacionais são amplamente utilizados em diversas aplicações, desde sistemas de gestão empresarial até plataformas de e-commerce e redes sociais, devido à sua capacidade de organizar e manipular grandes volumes de dados de maneira eficaz. Entre os sistemas de gerenciamento de bancos de dados relacionais (SGBDs) mais populares estão o MySQL, PostgreSQL, Oracle e SQL Server, cada um oferecendo características e funcionalidades específicas para atender às necessidades de diferentes tipos de aplicações.

2.6.2 API RESTful

Uma Interface de Programação de Aplicação (API) Representational State Transfer (RESTful) são interfaces de comunicação amplamente utilizadas no desenvolvimento de sistemas de informação baseados em tecnologias web. Fundamentadas no protocolo HTTP, essas APIs seguem princípios arquiteturais que definem interações padronizadas e escaláveis entre sistemas, utilizando métodos como GET, POST, PUT e DELETE para realizar operações relacionadas a recursos. Os recursos são identificados por URLs e geralmente representados em formatos como JSON ou XML, o que facilita a interoperabilidade

entre diferentes plataformas e linguagens de programação.

No contexto de um sistema de informação baseado em tecnologias web, as APIs RESTful desempenham um papel central ao possibilitar a comunicação entre o backend e o frontend, além de integrar serviços externos. Por exemplo, em um sistema de recuperação de imagens por similaridade visual, a API pode gerenciar solicitações de busca enviadas pelo usuário, consultar bancos de dados para encontrar imagens relevantes e retornar os resultados no formato apropriado. Sua estrutura modular e desacoplada permite que diferentes componentes do sistema evoluam de forma independente, garantindo flexibilidade para adaptações futuras.

2.6.3 Aplicações Web Modernas

Aplicações web modernas são sistemas dinâmicos e interativos projetados para fornecer experiências de usuário fluidas, muitas vezes comparáveis às de aplicações desktop. Essas aplicações aproveitam arquiteturas baseadas em tecnologias web avançadas, como APIs RESTful e frameworks de frontend, para oferecer funcionalidades ricas e responsivas. Diferentemente de aplicações web tradicionais, onde a interação frequentemente exige recarregamentos completos da página, as modernas utilizam técnicas como renderização dinâmica e atualizações assíncronas para otimizar a interação e reduzir a latência percebida pelo usuário.

Um dos pilares no desenvolvimento de aplicações web modernas é o uso de bibliotecas de frontend como React⁷. O React facilita a construção de interfaces de usuário por meio de componentes reutilizáveis e uma abordagem declarativa. Componentes encapsulam lógica e apresentação, permitindo que desenvolvedores criem aplicações modulares e escaláveis. Além disso, o modelo de virtual DOM do React optimiza as atualizações na interface, garantindo alto desempenho, mesmo em aplicações com grande complexidade de interface.

O React também integra-se de maneira eficiente com tecnologias modernas, como gerenciamento de estado global (por exemplo, Redux ou Context API) e ferramentas de roteamento (React Router), essenciais para construir Single Page Applications (SPAs). Essas SPAs são particularmente úteis em sistemas de informação interativos, como painéis de controle e plataformas de visualização de dados, oferecendo uma experiência de

⁷<https://react.dev>

navegação sem interrupções. Combinado a isso, o vasto ecossistema de bibliotecas e suporte comunitário do React acelera o desenvolvimento e facilita a implementação de recursos avançados, como gráficos interativos, animações e integração com APIs externas.

3. ESPECIFICAÇÃO DOS REQUISITOS

A especificação dos requisitos é uma parte fundamental da etapa de concepção do sistema inteligente desenvolvido. Aqui serão abordadas das definições dos requisitos funcionais (Seção 3.1) e não funcionais (Seção 3.2).

3.1. REQUISITOS FUNCIONAIS

A especificação dos requisitos funcionais de um sistema inteligente para busca de objetos astronômicos por similaridade visual envolve uma série de funcionalidades essenciais que visam a eficiência, precisão e usabilidade do sistema. Esse sistema é composto por um modelo de aprendizagem profunda baseado em redes neurais convolucionais (CNNs) para extração de características visuais (*embeddings*; Seção 2.5.3) de objetos astronômicos, um banco de dados otimizado para armazenar e buscar *embeddings* de milhões de objetos, e uma interface gráfica acessível via webapp, que permite ao usuário realizar consultas por similaridade visual. A seguir, são descritos os principais requisitos funcionais do sistema.

3.1.1 Extração de Características Visuais

Um requisito funcional primário do sistema é a capacidade de extrair características visuais dos objetos astronômicos com precisão e eficiência, por meio de um modelo de aprendizagem profunda baseado em CNNs. Esse modelo deve ser capaz de processar as imagens de entrada, normalmente provenientes de grandes levantamentos astronômicos, e gerar um vetor de características (*embedding*) que capture as propriedades morfológicas e fotométricas essenciais do objeto. Esse processo de extração deve ser automatizado e otimizado para processar grandes volumes de dados, permitindo o cálculo e atualização de *embeddings* de novos objetos astronômicos conforme necessário.

3.1.2 Base de Dados para Armazenamento e Consulta de Representações Visuais

O sistema requer uma base de dados especializada para o armazenamento eficiente de milhões de *embeddings*, permitindo o armazenamento de vetores de características

em uma estrutura que suporte operações rápidas de consulta e comparação. Esta base de dados deve ser otimizada para consultas de similaridade, utilizando técnicas de indexação espacial, como árvores KD ou técnicas de hashing para vetores de alta dimensão, visando minimizar o tempo de resposta em consultas complexas. Um requisito funcional essencial é a capacidade de realizar buscas por similaridade no banco de dados, retornando os *embeddings* mais próximos de um vetor de consulta fornecido pelo usuário. Adicionalmente, o banco de dados deve permitir atualizações contínuas, com a inserção de novos *embeddings* e a remoção de registros desatualizados, garantindo que o sistema mantenha uma representação atualizada do universo observado.

3.1.3 Interface de Consulta para o Usuário

O sistema deve incluir uma interface gráfica de usuário acessível por uma aplicação web, composto por uma camada de back-end e uma de front-end, que permite a realização de consultas de forma interativa e intuitiva. No front-end, o usuário deve poder fazer selecionar um objeto astronômico de referência a partir do nome ou coordenada para iniciar a busca por similaridade. Além disso, deve haver a possibilidade de ajustar parâmetros de consulta, como a sensibilidade ou alcance da similaridade, para refinar os resultados. Os resultados devem ser exibidos em uma lista ou galeria de imagens similares, com informações detalhadas sobre cada objeto encontrado, incluindo coordenadas, magnitude e outras propriedades físicas relevantes.

No back-end, os requisitos incluem a implementação de uma API que gerencie as solicitações de busca, comunique-se com o banco de dados para realizar a consulta dos *embeddings* e retorne os resultados ao front-end. A API deve também realizar a pré-processamento de imagens enviadas pelo usuário, preparando-as para serem passadas pelo modelo de CNN para extração do embedding. Este processo inclui etapas como normalização de brilho e ajuste de escala, garantindo que as imagens de entrada estejam padronizadas.

3.2. REQUISITOS NÃO FUNCIONAIS

Os requisitos não funcionais de um sistema inteligente para busca de objetos astronômicos por similaridade visual são fundamentais para garantir a qualidade, desempenho e usabilidade do sistema, complementando os requisitos funcionais ao definir critérios para

atributos como escalabilidade, desempenho, usabilidade, segurança e interoperabilidade. Este sistema é composto por um modelo de aprendizagem profunda, uma base de dados para armazenar *embeddings* de milhões de objetos e uma interface gráfica acessível por um webapp, que permite ao usuário realizar consultas de busca de forma interativa.

3.2.1 Desempenho

O sistema deve ser projetado para oferecer alta performance tanto na extração quanto na consulta dos *embeddings*. Dado o grande volume de dados astronômicos, o tempo de resposta para uma consulta por similaridade visual precisa ser minimizado, idealmente abaixo de um segundo por consulta para garantir a fluidez do sistema. Este requisito implica otimizações na base de dados, como o uso de indexação especializada para busca de similaridade em alta dimensão (por exemplo, KD-trees ou hashing vetorial). O tempo de processamento do modelo de CNN para extração de *embeddings* também deve ser otimizado, garantindo que as inferências sejam feitas em tempo hábil, permitindo a inclusão de novos objetos em uma janela de tempo viável para atualizações.

3.2.2 Usabilidade e Experiência do Usuário

A interface gráfica do sistema deve ser intuitiva, responsiva e fácil de navegar. Para isso, o webapp precisa fornecer uma experiência de usuário consistente e fluida, com uma apresentação visual clara dos resultados de busca e opções de filtro de consulta. Além disso, o sistema deve incluir uma documentação clara e completa, auxiliando o usuário a entender como fazer consultas, interpretar resultados e utilizar todas as funcionalidades oferecidas. Elementos como feedback em tempo real, indicadores de carregamento e mensagens de erro claras são essenciais para melhorar a experiência do usuário e aumentar a usabilidade.

3.2.3 Interoperabilidade

Como o sistema lida com dados astronômicos que podem vir de várias fontes e bancos de dados, a interoperabilidade é um requisito essencial para a integração e o processamento desses dados. O sistema deve ser compatível com padrões internacionais de dados astronômicos, como os protocolos do IVOA (Seção 2.2), que garantem consistência e integração com outros sistemas e bases de dados astronômicas. A interoperabilidade

facilita a incorporação de dados externos e amplia as possibilidades de uso e aplicação do sistema, permitindo que os usuários explorem uma maior variedade de informações de forma integrada.

3.2.4 Eficiência e Otimização de Recursos

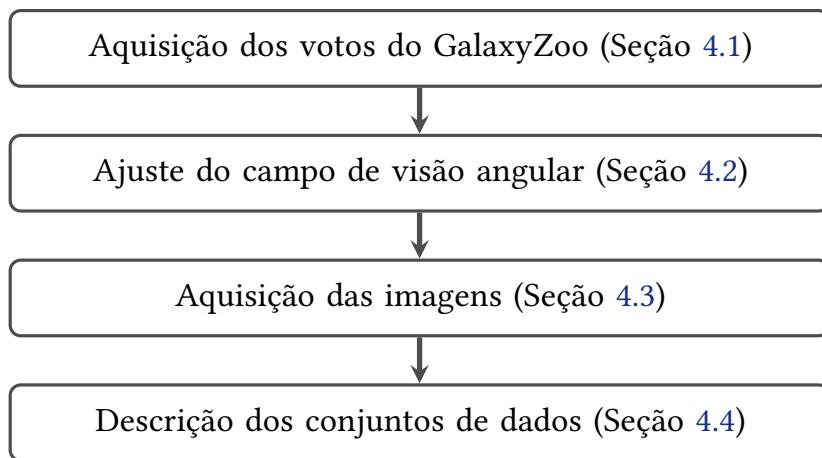
Dada a intensidade computacional envolvida na extração de *embeddings* e na execução de consultas de similaridade, o sistema deve ser eficiente no uso de recursos, tanto em termos de processamento quanto de armazenamento. O uso de técnicas de compressão para *embeddings* e otimizações no armazenamento de dados, como a escolha de formatos compactos e estruturas de dados adequadas, são essenciais para reduzir o consumo de armazenamento. No processamento, a implementação de técnicas de paralelização e uso de GPUs pode reduzir o tempo de inferência do modelo e melhorar a eficiência em consultas em grande escala.

3.2.5 Precisão e Consistência dos Resultados

O sistema deve fornecer resultados de busca por similaridade visual com alta precisão e consistência. Para isso, o modelo de CNN utilizado para extração de características deve ser rigorosamente treinado e validado, com controle de qualidade nas inferências geradas. A precisão é especialmente importante em astronomia, onde os pesquisadores dependem de comparações visuais precisas para análise morfológica e classificações. Assim, é essencial que o sistema mantenha a consistência dos resultados de consultas realizadas em diferentes momentos, evitando variações significativas nos *embeddings* gerados para o mesmo objeto.

4. CONJUNTOS DE DADOS

Nessa seção é abordado o método de aquisição dos dados para o treinamento supervisionado do modelo de aprendizagem profunda (Seção 5) e da base de dados de consulta (Seção 6.1). Os rótulos são provenientes do projeto de ciência cidadã Galaxy Zoo (Seção 2.4) e as imagens de entrada são obtidas do DESI Legacy Survey (Seção 2.1.3). A Fig. 4 descreve a sequência de passos realizada para a obtenção desses dados.



■ **Figura 4:** Fluxograma da aquisição de dados

4.1. AQUISIÇÃO DOS VOTOS DO GALAXYZOO

Os votos do GalaxyZoo (Seção 2.4) são utilizados como rótulos no treinamento supervisionado do modelo de aprendizagem profunda. Para composição desses rótulos, foram utilizadas 5 campanhas do GalaxyZoo: GalaxyZoo 1 ([C. LINTOTT et al., 2011](#)), GalaxyZoo 2 ([HART et al., 2016](#); [WILLETT, C. J. LINTOTT et al., 2013](#)), GalaxyZoo DESI ([WALMSLEY, GÉRON et al., 2023](#); [WALMSLEY, C. LINTOTT, GÉRON et al., 2021](#)), GalaxyZoo Hubble ([WILLETT, GALLOWAY et al., 2017](#)) e GalaxyZoo Candels ([SIMMONS et al., 2017](#)).

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

4.2. 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.

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 2.3.1) e outra usando a elipticidade complexa (Seção 2.3.2). Em ambos os casos, os valores são corrigidos por uma função η que depende da magnitude da galáxia (Seção 2.3.3) 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 2.1.3) acessadas a partir do AstroDataLab (<https://datalab.noirlab.edu>) pelo protocolo TAP (Seção 2.2).

A estimativa do FoV a partir do raio efetivo foi inspirada pelo método desenvolvido por [WALMSLEY, C. LINTOTT, GÉRON 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. (1).

$$\text{fov}_c = 2 \cdot r_e \cdot \eta(r_{\text{mag}}) \quad (1)$$

A segunda abordagem consiste na estimativa do FoV a partir da elipticidade expressa pelo número complexo da eq. (2). O catálogo fotométrico do Legacy Survey (Seção 2.1.3) fornece os valores das componentes real (ϵ_1) e imaginária (ϵ_2).

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

A partir dos valores das componentes ϵ_1 e ϵ_2 , é calculado o módulo do elipticidade ($|\epsilon|$), conforme a eq. (3).

$$|\epsilon| = \sqrt{\epsilon_1^2 + \epsilon_2^2} \quad (3)$$

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. (4) e (5).

$$\frac{b}{a} = \frac{1 - |\epsilon|}{1 + |\epsilon|} \quad (4)$$

$$\phi = \frac{1}{2} \arctan \frac{\epsilon_2}{\epsilon_1} \quad (5)$$

O próximo passo é calcular a caixa delimitadora (*bounding box*) da elipse. Para fazer isso, usamos as equações paramétricas da elipse, conforme as eqs. (6) e (7).

$$x(t) = x_0 + \frac{a}{2} \cos t \cos \phi - \frac{b}{2} \sin t \sin \phi \quad (6)$$

$$y(t) = y_0 + \frac{b}{2} \sin t \cos \phi + \frac{a}{2} \cos t \sin \phi \quad (7)$$

Então, expressamos o parâmetro t em termos os semi-eixos e do ângulo, como as eqs. (8) e (9).

$$t_x = \arctan \left(-\frac{b \tan \phi}{a} \right) \quad (8)$$

$$t_y = \arctan \left(\frac{b}{a \tan \phi} \right) \quad (9)$$

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. (10).

$$x_1 = x(t_x) \quad x_2 = x(t_x + \pi) \quad y_1 = y(t_y) \quad y_2 = y(t_y + \pi) \quad (10)$$

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. (11).

$$\ell = \max(|x_2 - x_1|, |y_2 - y_1|) \quad (11)$$

Finalmente, calculamos o FoV pela eq. (12)

$$\text{fov}_e = \ell \cdot \eta(r_{\text{mag}}) \quad (12)$$

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. 5. 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.

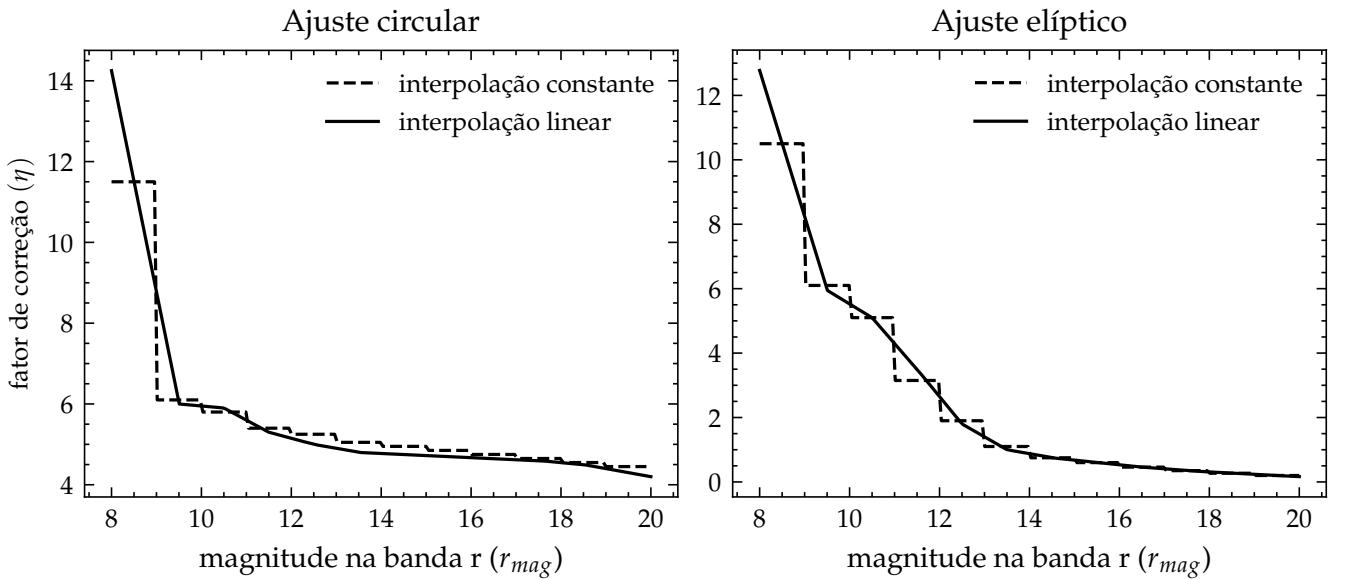
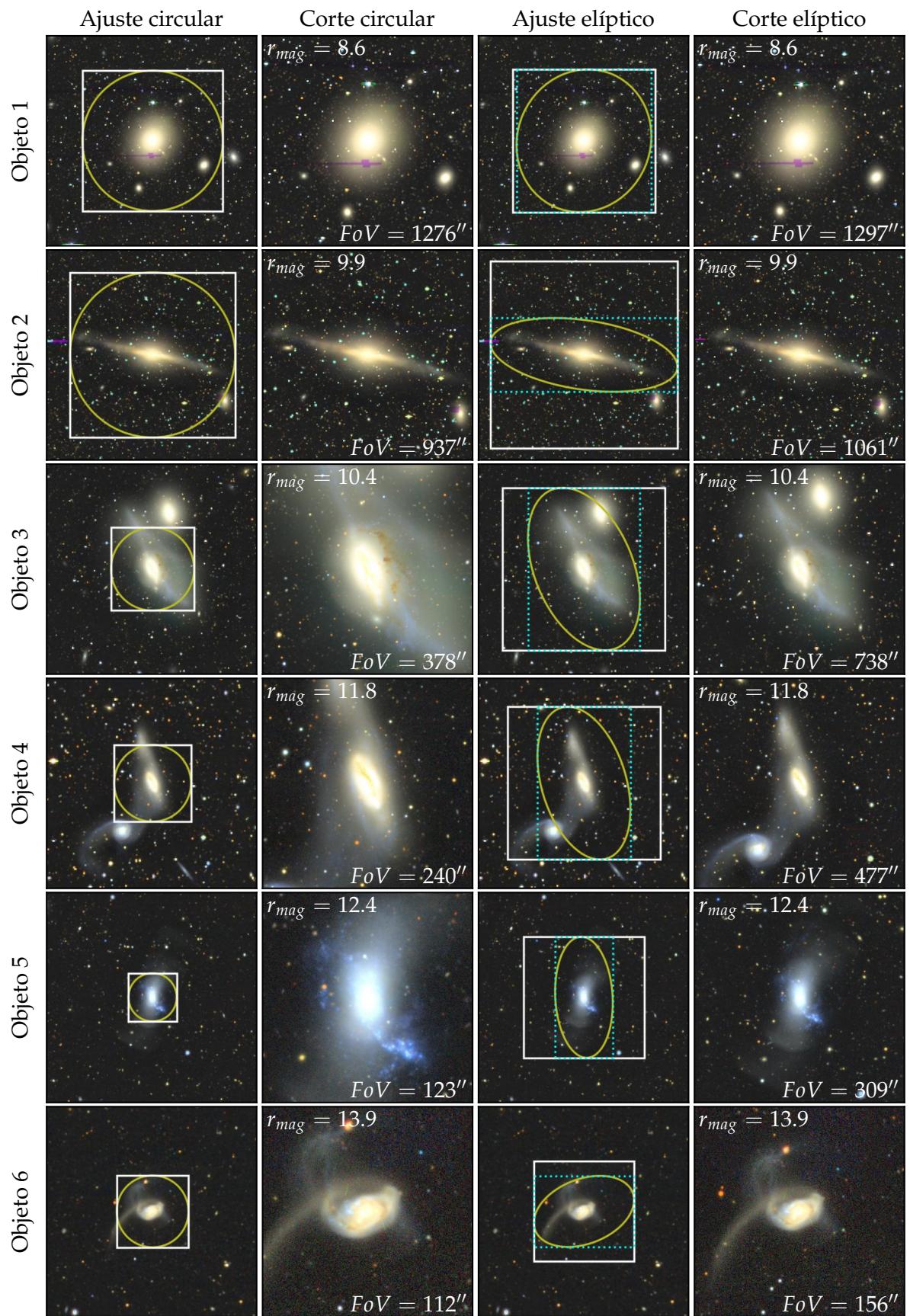


Figura 5: Fator de correção (η) ajustado por inspeção visual

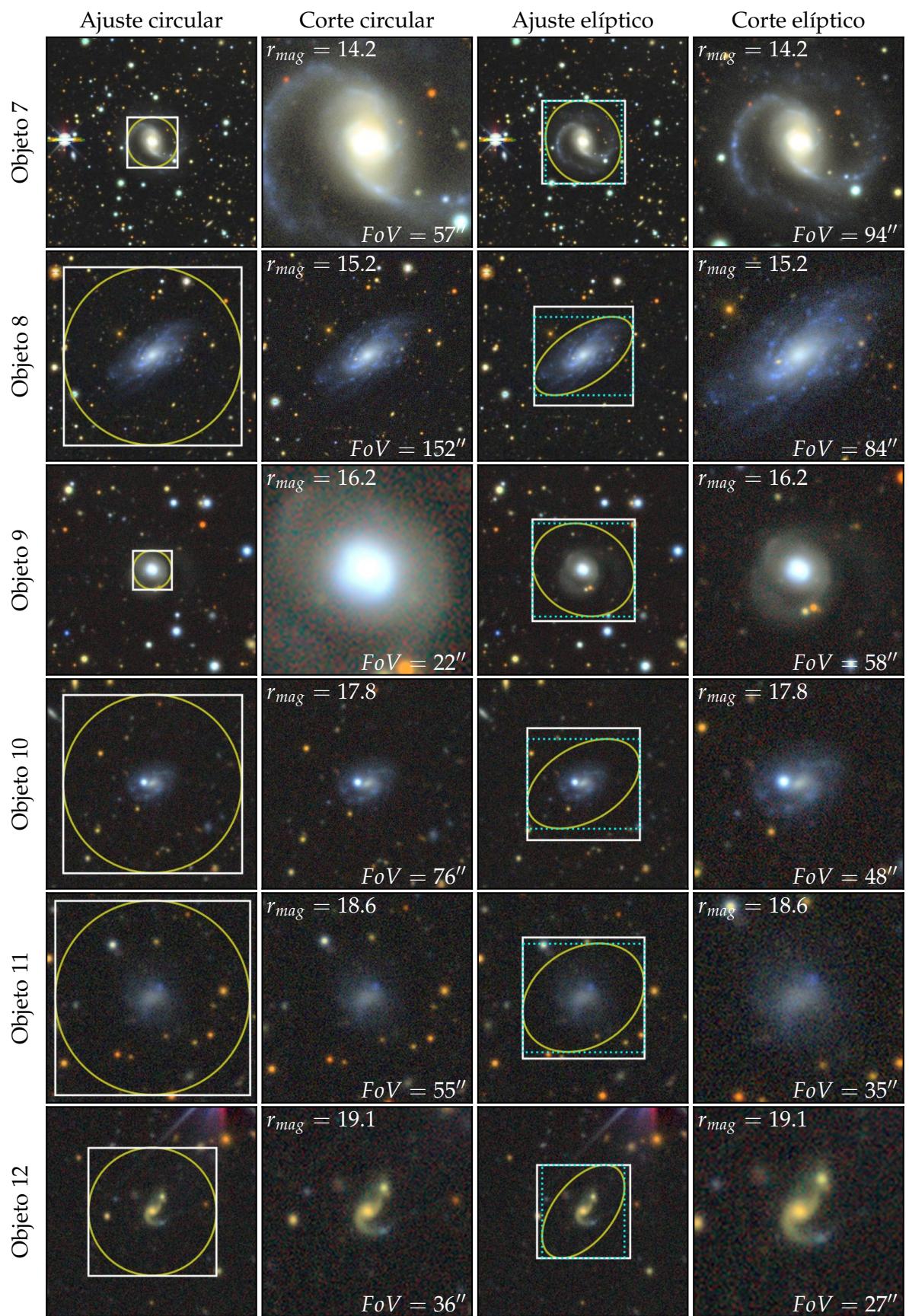
O processo de ajuste do FoV é exemplificado nas figs. 6 e 7. Nas, são mostrados os ajustes de 12 galáxias, uma para cada intervalo de magnitude, tanto para o ajuste circular (primeiro método) quanto para o ajuste elíptico (segundo método). No ajuste circular, a circunferência indica o raio efetivo da galáxia escalonado pela função η . Analogamente,

no ajuste elíptico, a elipse é obtida pelos parâmetros calculados nas eqs. (4) e (5) escalonados pela função η , e o retângulo tracejado indica a caixa delimitadora da elipse, cujos limites são obtidos da eq. (10). Em ambas as colunas, o quadrado branco indica o FoV calculado em cada caso, de acordo com as eqs. (1) e (12). As colunas “corte circular” e “corte elíptico” mostram as figuras finais após o ajuste pelo método do raio efetivo e da elipsidade 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. 6 e o Objeto 7 da Fig. 7, além de sobre-estimar o FoV de galáxias menos brilhantes, como os Objetos 10 e 11 da Fig. 7. Por outro lado, o método da elipse pareceu mais consistente para determinação dos FoVs e por isso foi o método escolhido.



■ **Figura 6:** Ajuste do campo de visão angular para r_{mag} entre 8 e 13



■ **Figura 7:** Ajuste do campo de visão angular para r_{mag} entre 14 e 19

4.3. AQUISIÇÃO DAS IMAGENS

A aquisição dos pequenos recortes de imagens astronômicas centrados em objetos específicos (stamps) do Legacy Survey por meio do serviço Hips2Fits⁸ (F. BONNAREL et al., 2000), do Strasbourg Astronomical Data Center⁹ (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. 8.

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 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

⁸<https://alasky.cds.unistra.fr/hips-image-services/hips2fits>

⁹<https://cds.unistra.fr>

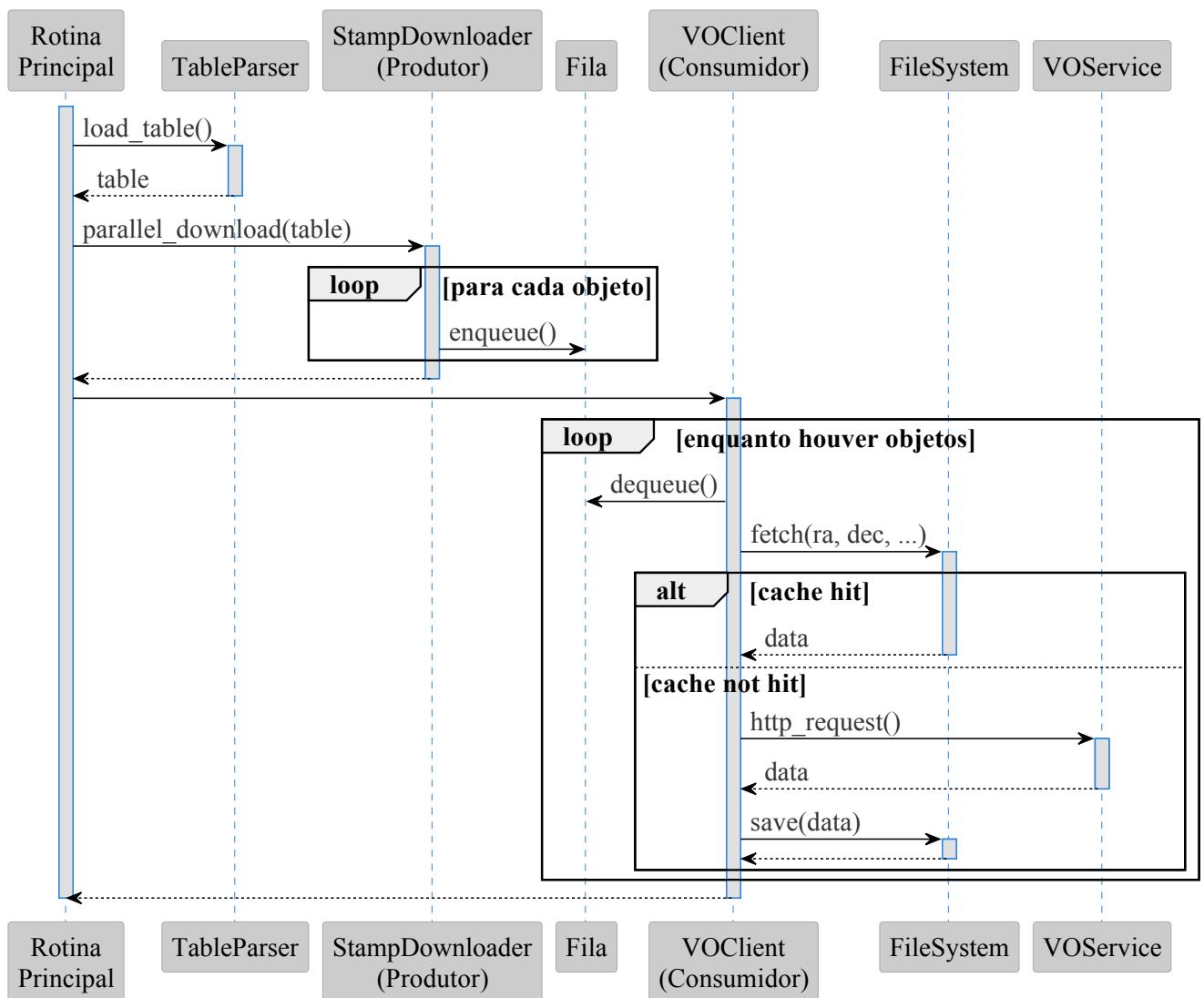


Figura 8: Diagrama de sequência da arquisição das imagens

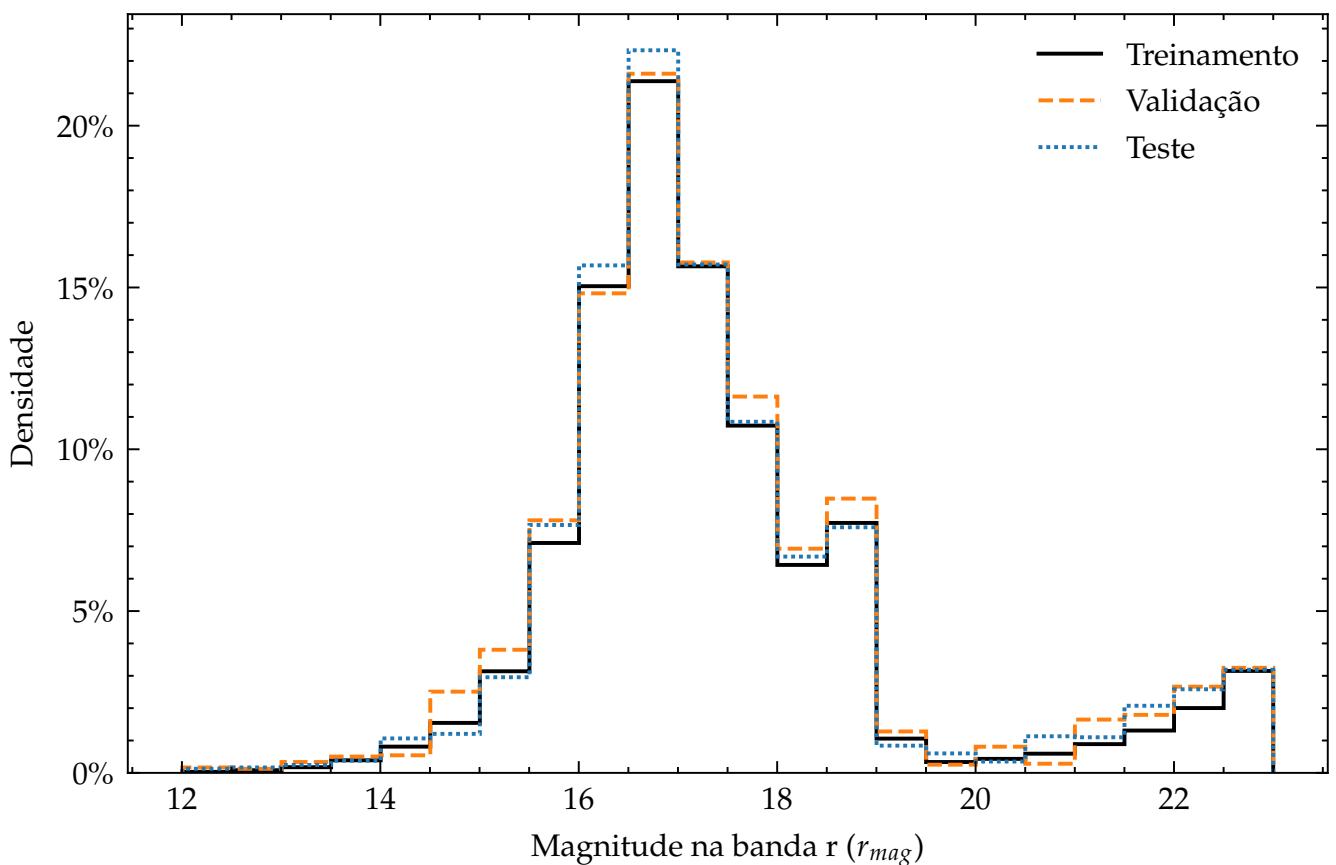
respeita os limites de acesso do Hips2Fits, mas também assegura uma coleta eficiente e escalável. Como cada thread pode solicitar e processar imagens independentemente, o tempo total de execução é significativamente reduzido, permitindo a aquisição dos stamps para milhões de galáxias em um tempo viável. 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.

4.4. DESCRIÇÃO DOS CONJUNTOS DE DADOS

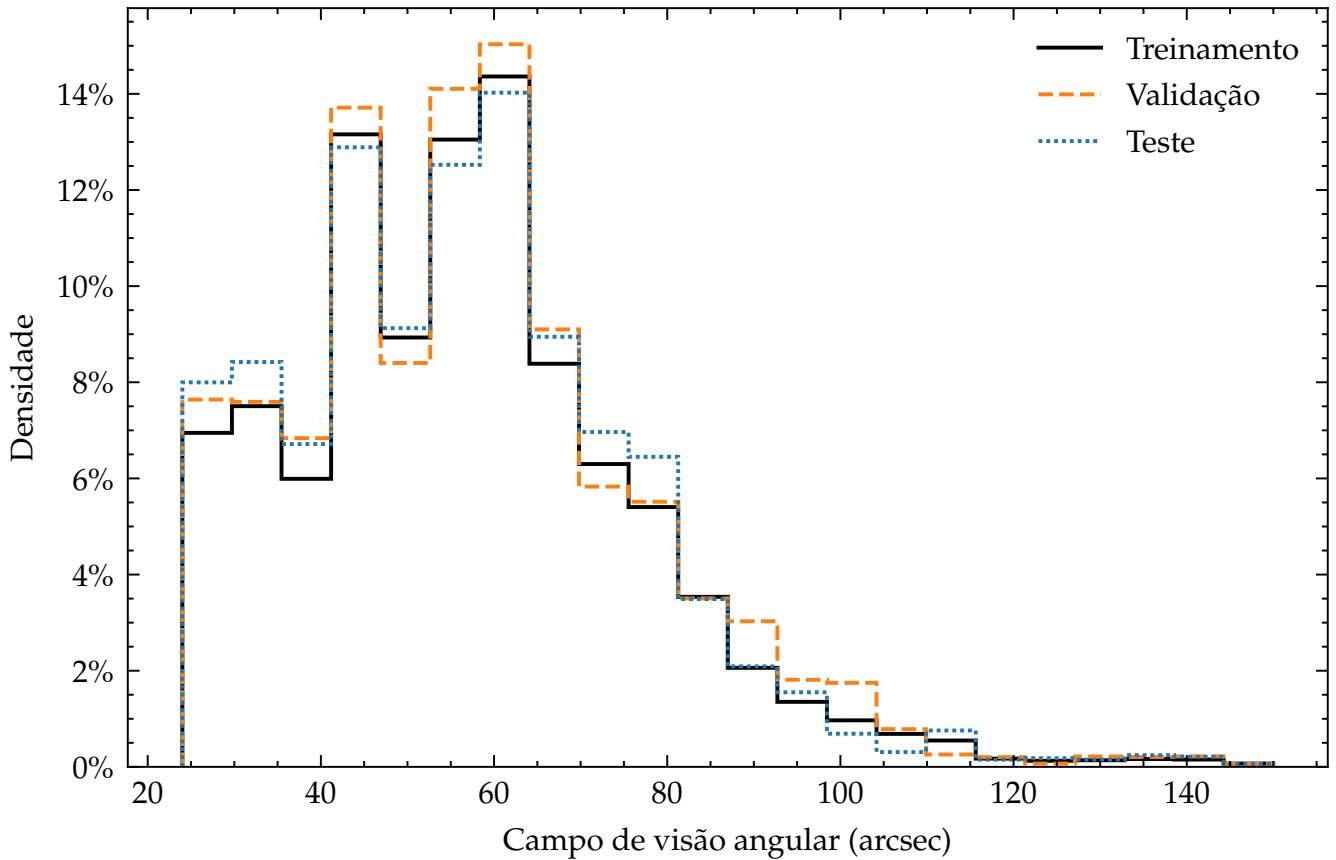
Nesta Seção, é feita uma análise, de forma agregada, sobre as propriedades físicas das galáxias dos conjuntos de dados para garantir a coerência no treinamento do modelo.

4.4.1 Conjuntos de Treinamento, Validação e Teste

Garantir que os conjuntos de treinamento, validação e teste possuam distribuições similares de magnitude (Fig. 9) e campo de visão angular (Fig. 10) 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, onde métricas de desempenho no conjunto de teste não refletem a eficácia real do modelo em aplicações práticas.



■ **Figura 9:** Distribuição da magnitude na banda r



■ **Figura 10:** Distribuição do campo de visão angular

4.4.2 Conjunto de Inferência

O conjunto de inferência é formado por aproximadamente 8 milhões de objetos. A Fig. 11 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 transitório, 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

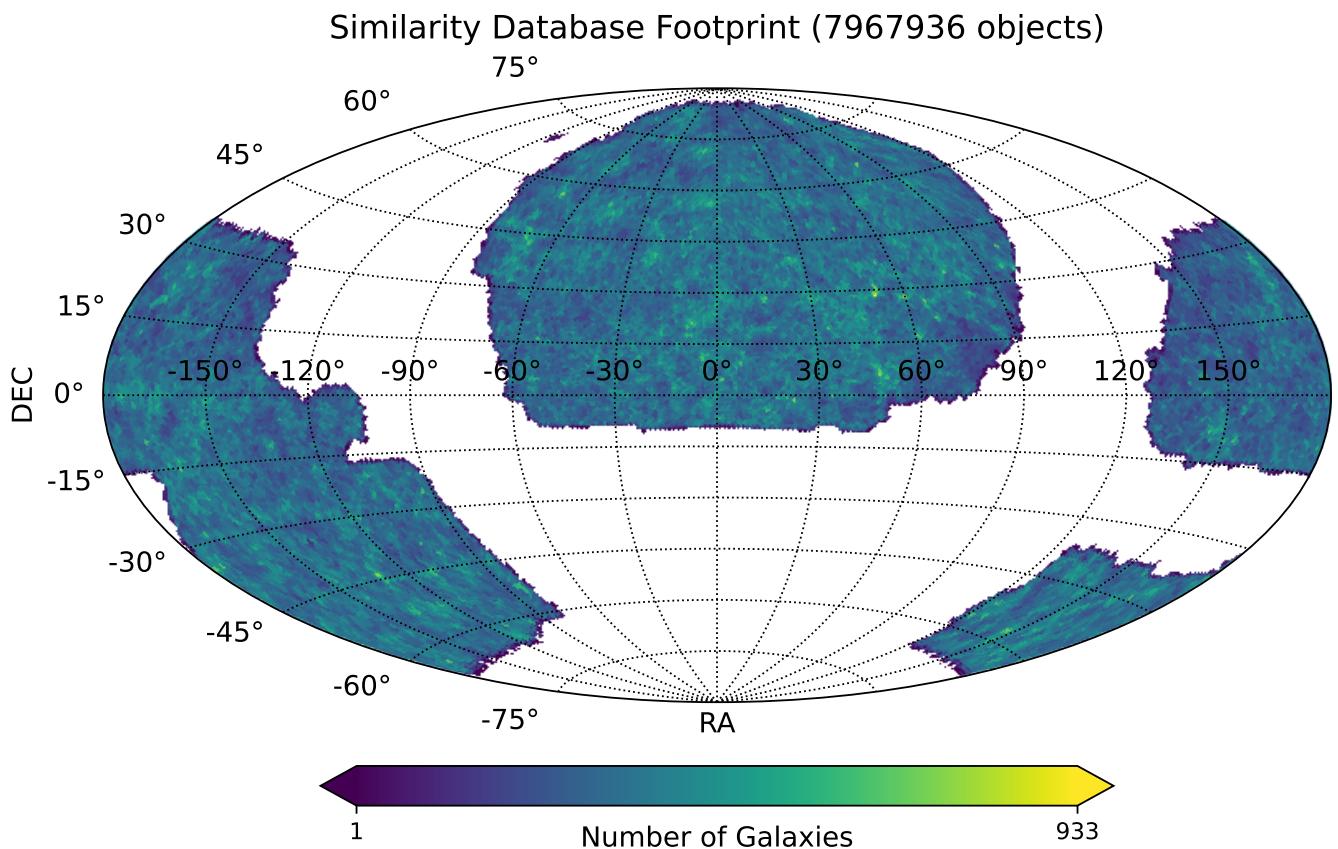


Figura 11: Área de cobertura espacial do conjunto de inferência

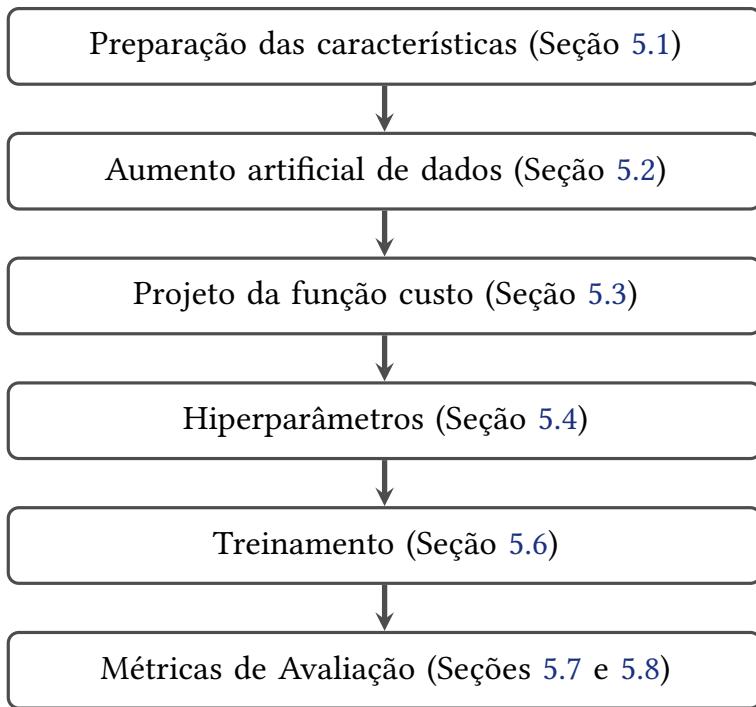
e complementares, à medida que novos modelos de aprendizado profundo e técnicas de análise de dados são desenvolvidos. Essa durabilidade é particularmente valiosa em um contexto científico onde as demandas computacionais e as abordagens metodológicas evoluem rapidamente, garantindo que o conjunto de 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.

5. 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 metamodelo.



■ **Figura 12:** Fluxograma do treinamento do modelo

5.1. PREPARAÇÃO DAS CARACTERÍSTICAS

A preparação dos dados é a etapa intuitivamente subsequente, após a aquisição (Seção 4). Nessa subseção, é discutido o processamento e codificação das características de entrada (Seção 5.1.1) e dos rótulos (Seção 5.1.2).

5.1.1 Preparação das Características de Entrada

As características de entrada do modelo são as imagens do Legacy Survey DR10 (Seção 2.1.3), obtidas pelo método descrito na Seção 4.3. 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

4.2.

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).

5.1.2 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 valore de referênci para a estimativa das contagens de votos durante o treinamento do modelo.

5.2. AUMENTO ARTIFICIAL DE DADOS

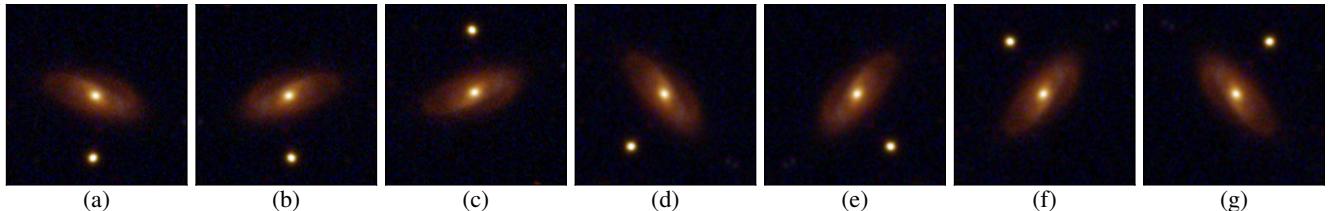


Figura 13: Exemplo do aumento artificial de dados em uma imagem original, mostrada no painel (a). Os painéis (b), (c), (d), (e), (f) e (g) contêm os resultados da eq. (16) substituindo M por diferentes combinações das transformações das eqs. (13), (14) e (15). Em (b) $M = V$, em (c) $M = H$, em (d) $M = R(30^\circ)$, em (e) $M = VR(30^\circ)$, em (f) $M = HR(30^\circ)$ e em (g) $M = HVR(30^\circ)$.

Aumento artificial de dados (YAEGER et al., 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. (13), (14) e (15) definem as transformações usadas.

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (13)$$

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (14)$$

$$V = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

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. (13), (14) e (15), 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. (13), (14) e (15) em cada píxel da imagem original usando a equação (16), 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.

$$\begin{bmatrix} x^* \\ y^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} M \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (16)$$

Além disso, ainda é aplicada uma interpolação bilinear como *anti-aliasing* (Bovik, 2009; McREYNOLDS e BLYTHE, 2005). Durante o treinamento da rede, novas imagens de entrada são geradas a cada época a partir da transformação das imagens originais. A Figura 13 mostra a imagem original, no painel (a), e diversas transformações, nos demais painéis, aplicadas substituindo M da equação (16) por combinações (multiplicação matricial) das transformações das eqs. (13), (14) e (15). 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 (CHOLLET, 2017; SIMARD et al., 2003). Frequentemente são relatados bons resultados com o uso desta técnica (II e DINC, 2020; MITTAL et al., 2019; NONAKA e SEITA, 2020), principalmente quando existe grande similaridade entre as classes.

5.3. FUNÇÃO DE CUSTO

5.3.1 Distribuição Multinomial

A distribuição multinomial (FORBES et al., 2010e) é uma generalização da distribuição binomial ([b](#)) 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. (17), onde $x_i \geq 0$ representa a contagem de ocorrências para cada categoria i .

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} \prod_{i=1}^k p_i^{x_i} \quad (17)$$

A distribuição multinomial (SEBER, 2015) é amplamente utilizada em aplicações de aprendizado de máquina e estatística, como na modelagem da transmissão de Dengue (L. WANG e M. ZHANG, 2025), na predição de cotação de mercado (NEVASALMI, 2020) e na análise de tráfego aéreo (OLIVEIRA TORRES et al., 2023). Uma de suas principais aplicações é em modelagem de eventos categóricos (KIBRIYA et al., 2004; LUO et al., 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 (JIANG et al., 2016; KALCHEVA et al., 2023), amplamente empregado em problemas de classificação de texto (KANG et al., 2005; ODEH et al., 2022), 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.

5.3.2 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 (a) 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. (18), onde $B(\vec{\alpha})$ é a função beta multivariada, definida na eq. (19), e $\Gamma(\cdot)$ representa a função gama (d). 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.

$$P(\vec{X}|\vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{i=1}^k X_i^{\alpha_i-1} \quad (18)$$

$$B(\vec{\alpha}) = \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^k \alpha_i\right)} \quad (19)$$

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 et al., 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 et al., 2009).

5.3.3 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 5.3.1) e da distribuição Dirichlet (Seção 5.3.2). 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. (20), 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.

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{\Gamma\left(\sum_{i=1}^k \alpha_i\right)}{\Gamma\left(\sum_{i=1}^k \alpha_i + n\right)} \prod_{i=1}^k \frac{\Gamma(x_i + \alpha_i)}{\Gamma(\alpha_i) x_i!} \quad (20)$$

5.3.4 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. (21).

$$\mathcal{L}_q = \int \text{Multi}(\vec{k}|\vec{\rho}, N) \text{Dirichlet}(\vec{\rho}|\vec{\alpha}) d\vec{\rho} \quad (21)$$

Já o valor da perda de um exemplo é a soma das perdas de cada questão q , conforme a eq. (22).

$$\log \mathcal{L} = \sum_q \mathcal{L}_q(\vec{k}_q, N_q, \vec{f}_q^w) \quad (22)$$

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. (23).

$$J = \sum_i \log \mathcal{L}_i \quad (23)$$

5.4. 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 e BA, 2017), NAdam (DOZAT, 2015), RAdam (L. LIU et al., 2020) e RMSprop (HINTON et al., 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).

5.5. ARQUITETURA DE REDE NEURAL

5.5.1 Visual Geometry Group Networks

A família de arquiteturas Visual Geometry Group Networks (VGGNets) (SIMONYAN e 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 escalamento 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 e LOUDINI, 2020).

5.5.2 InceptionResNetV2

A arquitetura InceptionResNetV2 (SZEGEDY, IOFFE et al., 2016) usa os blocos Inception, que são convoluções fatorizadas, introduzidos em (SZEGEDY, W. LIU 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).

5.5.3 EfficientNet

A arquitetura EfficientNet (TAN e 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 (P. ZHANG et al., 2020), eletrocardiogramas (NONAKA e SEITA, 2020) e cristalização de proteínas (II e DINC, 2020).

5.5.4 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 (Z. LI et al., 2020), classificação de músicas (SOLANO-ROJAS et al., 2020), câncer de mama (X. LI et al., 2020) e esclerose múltipla (S.-H. WANG e Y.-D. ZHANG, 2020).

5.5.5 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.

5.6. 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; Y. LI et al., 2021), mostrada na Fig. 14, 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; L. LI et al., 2016), disponíveis na biblioteca Optuna¹⁰ (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. 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

¹⁰<https://optuna.org>

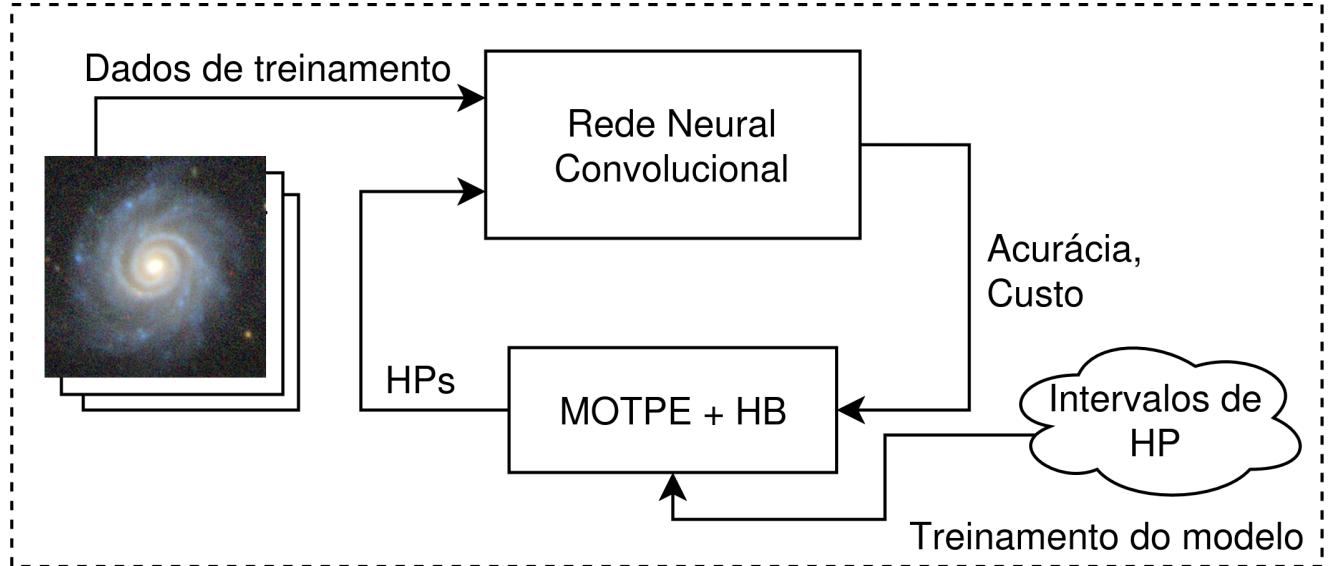


Figura 14: Treinamento do modelo

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}

5.7. 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.

5.7.1 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
- Verdadeiro Negativo (VN): O modelo prevê uma instância negativa corretamente
- Falso Positivo (FP): O modelo prevê uma instância positiva, mas, na verdade, é

Tabela 3: Estrutura da matriz de confusão para classificação binária

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

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.

5.7.2 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. 24.

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}} \quad (24)$$

5.7.3 Precisão

A precisão avalia a proporção de exemplos classificados como positivos que são realmente positivos, conforme a eq. (25), sendo particularmente importante em cenários onde o custo de falsos positivos é alto.

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}}. \quad (25)$$

5.7.4 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. (26), sendo principalmente útil em contextos onde a identificação de todos os exemplos positivos é crucial

$$\text{Revocação} = \frac{\text{VP}}{\text{VP} + \text{FN}}. \quad (26)$$

5.7.5 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. (27). 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.

$$F1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (27)$$

5.8. MÉTRICAS DE AVALIAÇÃO DA SIMILARIDADE

A recuperação de imagens baseada em embeddings gerados por modelos de aprendizado de máquina envolve a comparação de vetores em um espaço latente para determinar a semelhança ou dissimilaridade entre as representações das imagens. Para isso, são amplamente utilizadas métricas de distância ou similaridade, como similaridade do cosseno (Seção 5.8.1), produto interno (Seção 5.8.2), distância L1 (Seção 5.8.3) e distância L2 (Seção 5.8.4). Essas métricas capturam diferentes aspectos das relações entre os embeddings e são escolhidas com base nas características do espaço latente e nos objetivos do sistema. Ademais, nessa Seção, também é definida a métrica que quantifica a capacidade preditiva da tarefa de recuperação por similaridade visual na Seção 5.8.5.

5.8.1 Distância do Cosseno

A similaridade do cosseno mede o ângulo entre dois vetores no espaço vetorial, desconsiderando suas magnitudes. Ela é amplamente usada em tarefas de recuperação de imagens e processamento de linguagem natural, pois é invariável à escala dos vetores, o que é útil em embeddings que capturam informações relativas às direções no espaço latente.

Formalmente, ela é definida pela eq. (28), onde $\mathbf{u} \cdot \mathbf{v}$ é o produto interno dos vetores

\mathbf{u} e \mathbf{v} , e $\|\mathbf{u}\|$ e $\|\mathbf{v}\|$ são suas normas. Valores de $\cos(\theta)$ variam entre -1 e 1, onde 1 indica vetores perfeitamente alinhados (paralelos), -1 indica vetores opostos, e 0 indica vetores ortogonais. Já a similaridade do cosseno é mais parecida com outras métricas de distância, pois parte do princípio que coisas próximas possuem valores de distância pequenos, variando de 0 a 2. Quanto mais próximo de zero, mais similar.

$$\text{Similaridade do Cosseno} = 1 - \cos(\theta) = 1 - \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \quad (28)$$

5.8.2 Produto Interno

O produto interno é uma métrica simples que mede a projeção de um vetor sobre outro, conforme a eq. (29), onde u_i e v_i são as componentes correspondentes dos vetores \mathbf{u} e \mathbf{v} . Diferentemente da similaridade do cosseno, o produto interno considera tanto a direção quanto a magnitude dos vetores, o que pode ser útil em espaços latentes onde a magnitude dos embeddings também carrega informações relevantes.

$$\text{Produto Interno} = \mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n u_i v_i \quad (29)$$

5.8.3 Distância L1

A distância L1, também conhecida como distância de Manhattan, mede a soma das diferenças absolutas entre as componentes correspondentes de dois vetores. Ela é expressa na eq. (30).

Essa métrica é útil em espaços onde as dimensões são interpretadas como contribuições independentes, e variações locais são importantes. A distância L1 tende a ser mais robusta a outliers do que a distância L2, pois considera contribuições lineares de cada dimensão sem elevar as diferenças ao quadrado.

$$\text{Distância L1} = \|\mathbf{u} - \mathbf{v}\|_1 = \sum_{i=1}^n |u_i - v_i| \quad (30)$$

5.8.4 Distância L2

A distância L2, ou distância Euclidiana, é a métrica mais comum para medir a similaridade entre vetores. Ela calcula a raiz quadrada da soma dos quadrados das diferenças entre as componentes correspondentes dos vetores, conforme a eq. (31).

Esta métrica é amplamente utilizada devido à sua interpretação geométrica intuitiva como a distância direta entre dois pontos em um espaço multidimensional. No entanto, ela é sensível a outliers, pois penaliza diferenças grandes mais severamente do que a distância L1.

$$\text{Distância L2} = \|\mathbf{u} - \mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n (u_i - v_i)^2} \quad (31)$$

5.8.5 Métrica de Desempenho da Busca

A métrica mAP (mean Average Precision) é amplamente utilizada para avaliar o desempenho de sistemas de recuperação de informações, incluindo a recuperação de imagens baseada em embeddings gerados por modelos de aprendizado de máquina. Essa métrica combina precisão e revocação em um único valor, fornecendo uma medida robusta da capacidade de um modelo em recuperar itens relevantes em tarefas onde a ordenação dos resultados importa.

O cálculo do mAP começa pela definição da *Average Precision* (AP), que mede a precisão média em diferentes níveis de revocação para um conjunto de resultados ordenados. Em um cenário típico de recuperação, para cada consulta, o sistema retorna uma lista ordenada de itens, e o objetivo é avaliar como os itens relevantes estão distribuídos ao longo dessa lista. Dessa forma, para uma consulta q , a AP é definida pela eq. (32), onde N é o número total de itens retornados; $P(k)$ é a precisão até a posição k na lista de resultados; e $rel(k)$ é uma função indicadora que vale 1 se o item na posição k for relevante e 0 caso contrário.

$$AP(q) = \frac{\sum_{k=1}^N P(k) \cdot rel(k)}{\text{Total de itens relevantes}} \quad (32)$$

Intuitivamente, a AP calcula a média das precisões em cada ponto onde um item relevante é encontrado, ponderando pela relevância. Isso recompensa sistemas que

colocam itens relevantes nas primeiras posições, refletindo a importância da ordenação.

A métrica mAP é obtida como a média da AP sobre todas as consultas Q , conforme a eq. (33), onde $|Q|$ é o número total de consultas. Esse valor sintetiza o desempenho global do sistema, considerando tanto a qualidade das previsões individuais quanto a consistência do modelo em diferentes consultas.

$$mAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \quad (33)$$

O mAP é mais informativo que métricas como acurácia ou precisão isolada, pois considera explicitamente a ordenação dos resultados e é adequado para sistemas de recuperação onde o número de itens relevantes por consulta pode variar. Além disso, o mAP não é influenciado por classes desbalanceadas, pois avalia separadamente a qualidade da recuperação para cada consulta.

6. SISTEMA DE INFORMAÇÃO

Nesta seção, será descrito desenvolvimento do sistema de informação, esboçado na Fig. 15, composto por um banco de dados (Seção 6.1), um backend (Seção 6.2) e um frontend (Seção 6.3), que, juntos, são responsáveis pelo armazenamento, manipulação e recuperação dos dados.

6.1. BASE DE DADOS

A base de dados desempenha um papel fundamental no sistema inteligente desenvolvido, pois ela tem a função de armazenar as representações visuais, geradas a partir do modelo de aprendizagem profunda (Seção 5) e delimitadas pelo conjunto de inferência (Seção 4.4.2), com o propósito de viabilizar a busca rápida no céu por similaridade visual em uma grande área do céu. Esta funcionalidade foi especificada durante a concepção do projeto como o requisito funcional de armazenamento e recuperação de representações visuais (Seção 3.1.2). A seguir, são discutidas as tecnologias (Seção 6.1.1), o modelo de dados (Seção 6.1.2) e as otimizações (Seções 6.1.3 e 6.1.4) utilizadas do desenvolvimento da base de dados.

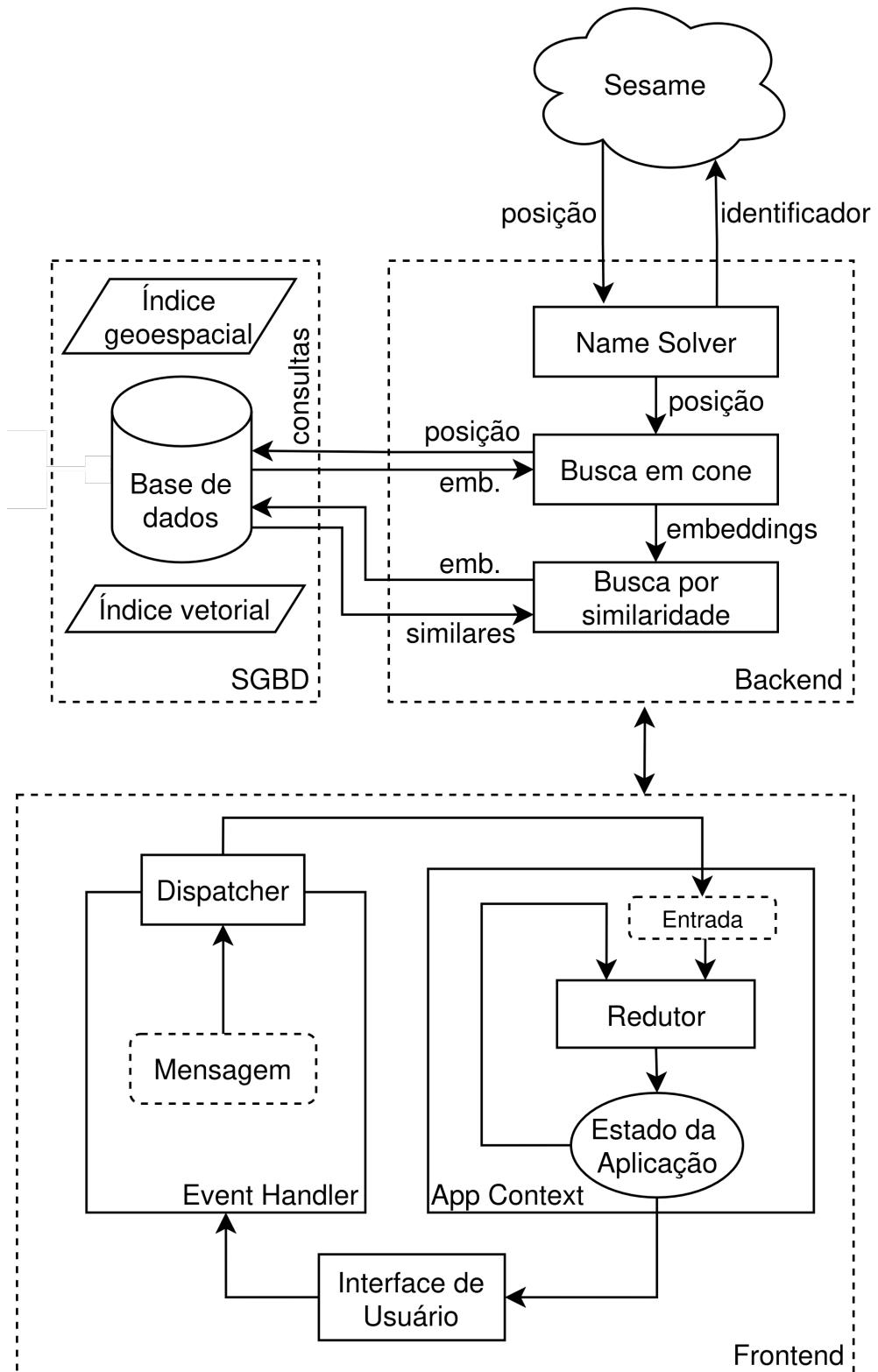


Figura 15: Diagrama do Sistema de Informação

6.1.1 Tecnologias Utilizadas

Uma parte fundamental no projeto de uma base de dados é a escolha do Sistema Gerenciador de Banco de Dados (SGBD), essa escolha está fortemente relacionada com o

tipo de dado a ser armazenado e os tipos de buscas a serem realizadas. Nesse sentido, para atender os requisitos da aplicação desenvolvida, as especificações do SGBD são listadas a seguir.

- **Armazenamento de vetores:** as características visuais extraídas pelo modelo de aprendizagem profunda (Seção 5) são representadas na forma de vetores. Logo, é essencial que o SGBD consiga manusear esse tipo de dado;
- **Busca eficiente de vetores:** a recuperação de objetos similares é a busca por objetos com menor distância do vetor de referência. Nesse sentido, o SGBD precisa realizar operações com distância de vetores além de ter um mecanismo de indexação que otimize a busca por distância;
- **Busca por posição:** além das representações visuais, a base de dados deve armazenar metadados dos objetos, como a posição. A busca por posição é fundamental para integrar o sistema como outros serviços web astronômicos, como Sesame, Simbad e Vizier, pois ela permite correlacionar informações de um mesmo objeto proveniente de fontes diferentes.

Com base nas especificações apresentadas e levando em consideração as tecnologias do estado-da-arte para a tarefa desenvolvida (PAN et al., 2024), o PostgreSQL¹¹ foi escolhido como SGDB, pois sua flexibilidade e extensibilidade permitem as funcionalidades de busca por posição e vetorial, discutidas nas Seções 6.1.3 e 6.1.4, respectivamente.

6.1.2 Modelo de Dados

A definição de um modelo de dados eficiente é fundamental para a integração das diversas partes do sistema. A Fig. 16 mostra o diagrama entidade-relacionamento da base de dados projetada para o armazenamento e a recuperação dos objetos por similaridade visual.

As Tabelas 4 e 5 informam o tipo de dado e uma breve descrição de cada coluna das tabelas mostradas na Fig. 16.

A indexação de colunas é um mecanismo fundamental de otimização de consultas em bases de dados relacionais, representando os dados por uma estrutura de dados auxiliar, projetada para garantir mais desempenho em tipos específicos de problemas. Na sequência, são mostradas as técnicas de indexação para as colunas `ra` e `dec` (Seção 6.1.3)

¹¹<https://postgresql.org>

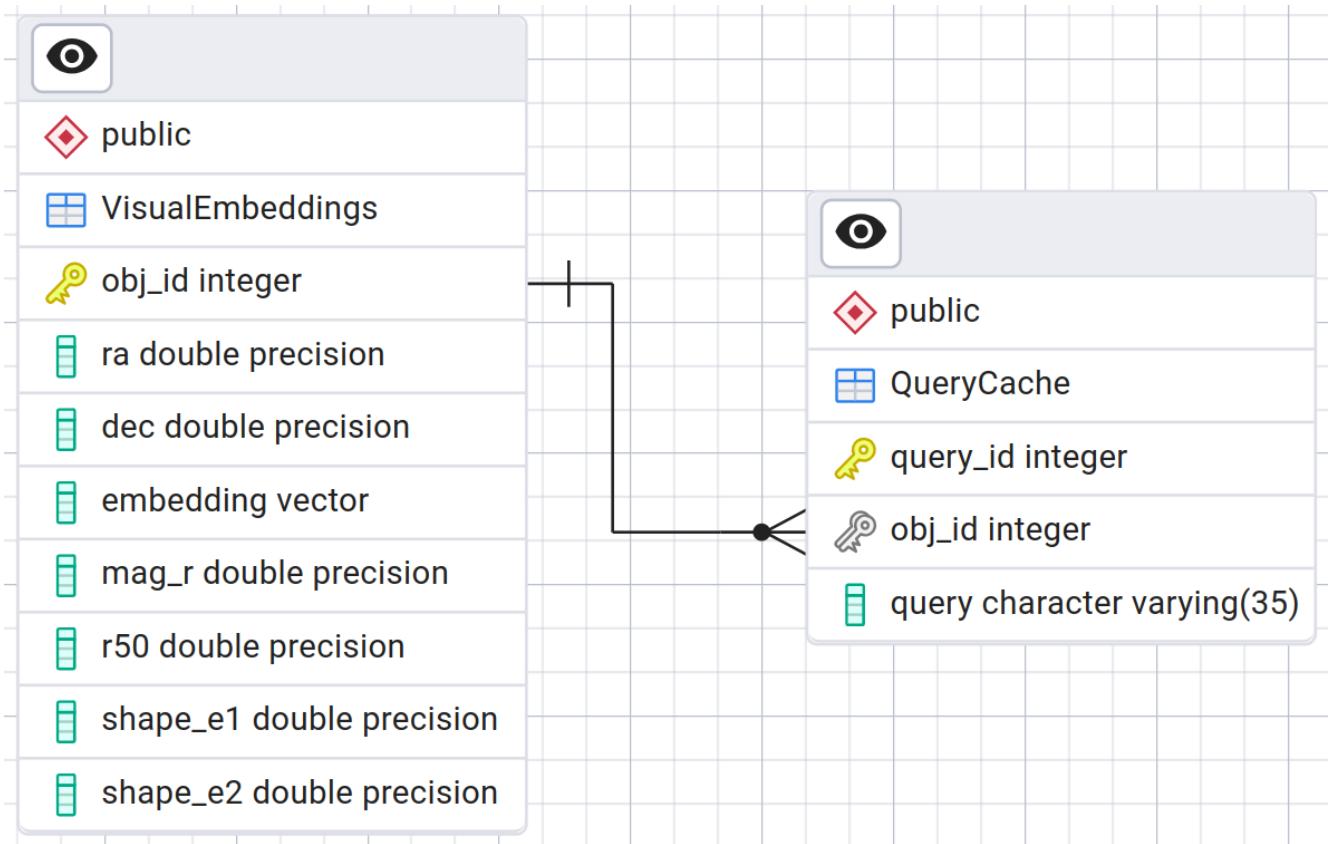


Figura 16: O diagrama entidade-relacionamento mostra duas tabelas: `VisualEmbeddings` (esquerda) e `QueryCache` (direita). A primeira armazena as representações visuais e algumas propriedades físicas de todas as galáxias do conjunto de inferência (Seção 4.4.2). Já a segunda é um cache que associa um termo de busca à chave primária da tabela anterior, tendo a função de evitar requisição ao sistema Sesame e a correlação por coordenadas.

e para a coluna `embedding` (Seção 6.1.4).

Tabela 4: Descrição da tabela `VisualEmbeddings`

Coluna	Tipo	Descrição
<code>obj_id</code>	inteiro	Chave primária do objeto
<code>ra</code>	precisão dupla	Ascenção Reta do obejeto
<code>dec</code>	precisão dupla	Declinação do obejeto
<code>embedding</code>	vetor	Vetor de representações visuais
<code>mag_r</code>	precisão dupla	Magnitude do objeto (Seção 2.3.3)
<code>shape_e1</code>	precisão dupla	Componente real da elipticidade complexa (Seção 2.3.2)
<code>shape_e2</code>	precisão dupla	Componente imaginária da elipticidade complexa (Seção 2.3.2)

Tabela 5: Descrição da tabela QueryCache

Coluna	Tipo	Descrição
query_id	inteiro	Chave primária
obj_id	inteiro	Chave estrangeira referenciando a coluna obj_id da tabela VisualEmbeddings (Tab. 4)
query	string	Termo de busca utilizado para fazer a consulta

6.1.3 Armazenamento e Indexação das Coordenadas

A indexação das coordenadas, definidas pelas colunas `ra` e `dec`, utiliza o mecanismo Quad Tree Cube¹² (Q3C; KOPOSOV e BARTUNOV, 2006), que é uma técnica utilizada na astronomia para otimizar o armazenamento e a recuperação de dados espaciais em bancos de dados de grandes levantamentos astronômicos. Esse mecanismo foi desenvolvido para lidar com a quantidade massiva de dados gerada por observatórios modernos, que capturam informações sobre bilhões de objetos celestes, incluindo sua posição, brilho e outras características. A principal função do Q3C é permitir consultas eficientes em catálogos astronômicos, possibilitando a localização rápida de objetos com base em suas coordenadas celestes.

O Q3C utiliza uma estrutura baseada em árvore quad (ou quadtree) projetada para dividir iterativamente a superfície da esfera celeste em regiões menores, como mostra a Fig. 17. Cada região é identificada por um índice (chamado IPIX), o que permite armazenar e acessar informações de localização de forma hierárquica. No Q3C, a superfície esférica é projetada em um cubo e subdividida em quadrantes, o que facilita a representação de posições celestes e otimiza a indexação espacial. Essa estrutura de dados permite que consultas espaciais, como encontrar todos os objetos dentro de um certo raio de um ponto específico, sejam executadas com alta eficiência, reduzindo a necessidade de examinar todas as entradas do catálogo.

A implementação do Q3C aproveita funções de hashing para transformar as coordenadas esféricas em valores que podem ser facilmente indexados. Esse processo de hash mapeia coordenadas (RA, DEC) para uma representação numérica, permitindo uma indexação que é eficiente tanto para operações de busca por proximidade quanto para

¹²<https://github.com/segasai/q3c>

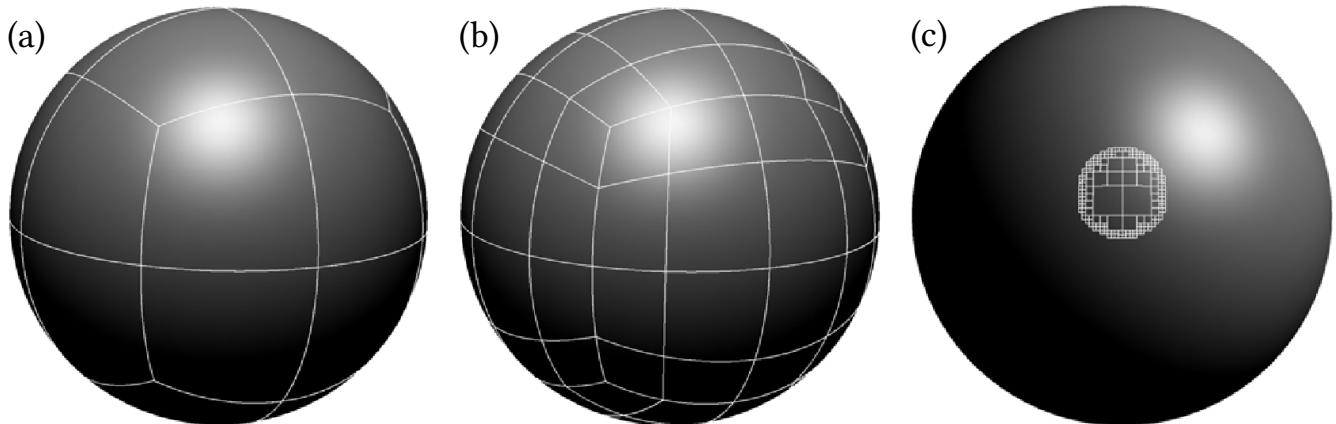


Figura 17: Segmentação da esfera celeste pelo Q3C. Nos painéis (a) e (b), são mostradas segmentações da esfera celeste pelo Q3C para duas quantidades de bits distintas. Já o painel (c) mostra um exemplo de busca usando o Q3C. Fonte: Compilação das Figs. 1 e 2 de .

consultas de intervalo. As consultas são formuladas de maneira que a hierarquia de quadrantes seja explorada, de modo que apenas uma fração dos dados seja analisada para atender à consulta. Além disso, o Q3C pode realizar operações de cruzamento de catálogos, comparando dados de diferentes levantamentos com rapidez.

6.1.4 Armazenamento e Indexação dos Vetores

Embora o PostgreSQL seja tradicionalmente um banco de dados relacional, sua flexibilidade e extensibilidade permitem a implementação de funcionalidades tipicamente associadas a bancos de dados vetoriais (TAIPALUS, 2024), como o armazenamento e a busca eficiente de vetores. Isso é alcançado por meio de extensões, como o pgvector¹³, que integram capacidades vetoriais ao PostgreSQL. Com essa abordagem, é possível armazenar vetores diretamente em tabelas relacionais e realizar consultas por similaridade, transformando o PostgreSQL em uma solução híbrida que combina recursos relacionais e vetoriais.

O mecanismo de indexação de vetores no pgvector é projetado para otimizar a indexação e busca de embeddings de alta dimensionalidade, especialmente aqueles gerados por modelos de aprendizagem profunda. Tais representações são amplamente utilizadas em aplicações como recuperação de informação, recomendação e classificação de dados, onde é necessário medir similaridade entre itens com base em seus embeddings. O pgvector

¹³<https://github.com/pgvector/pgvector>

oferece suporte nativo para armazenar e indexar esses vetores no PostgreSQL, tornando viável a realização de buscas eficientes mesmo em bancos de dados com grandes volumes de dados vetoriais.

O pgvector utiliza o HNSW, uma estrutura de grafos, chamada Hierarchical Navigable Small World Graph (HNSW; [MAL'KOV e YASHUNIN, 2020](#)), que facilita a busca aproximada por similaridade. O HNSW é baseado em uma hierarquia de grafos navegáveis, onde cada nó representa um vetor, e conexões entre nós são criadas com base na similaridade vetorial. Esse grafo permite uma navegação eficiente, utilizando uma busca aproximada que, apesar de não garantir a precisão absoluta, oferece um bom equilíbrio entre desempenho e precisão. O HNSW é particularmente eficaz em cenários com alta dimensionalidade, pois reduz o número de cálculos de similaridade necessários para encontrar o vetor mais próximo, sendo capaz de realizar buscas sublineares em grandes volumes de dados. Em termos de aplicação prática, esse método torna viável a busca por similaridade em bancos de dados com milhões de embeddings, permitindo um tempo de resposta mais rápido.

6.2. BACKEND

O backend refere-se à camada responsável pelo processamento lógico, manipulação de dados e comunicação com bancos de dados e serviços externos. Ele opera nos servidores e é invisível ao usuário final, funcionando como a infraestrutura que sustenta e executa as operações solicitadas pelo frontend. O backend normalmente utiliza frameworks e ferramentas para gerenciar APIs, autenticação, controle de acesso e persistência de dados. Sua principal função é garantir que as informações sejam processadas de forma segura, eficiente e escalável, permitindo que a aplicação ofereça funcionalidades dinâmicas aos usuários.

6.2.1 Tecnologias Utilizadas

O desenvolvimento do backend requer a utilização de tecnologias que garantam desempenho, escalabilidade e flexibilidade. Nesse contexto, a biblioteca Blacksheep¹⁴, implementada em Python, oferece uma solução moderna para construir APIs robustas e eficientes. Trata-se de um framework assíncrono que utiliza o modelo de programação

¹⁴<https://www.neoteroi.dev/blacksheep>

baseado em corrotinas do Python, permitindo lidar de maneira eficiente com operações de entrada e saída (E/S) intensivas, como consultas a bancos de dados de grande escala.

A Blacksheep é projetada sobre o motor de eventos do módulo `asyncio`, proporcionando suporte nativo para o desenvolvimento de APIs RESTful de alta performance. Sua abordagem assíncrona é vantajosa para sistemas que requerem comunicação frequente com bancos de dados, como é o caso deste sistema de recuperação de imagens astronômicas. Além disso, a biblioteca adota um design simples e modular, permitindo personalizar funcionalidades conforme necessário.

Foram utilizados os middlewares e roteamento eficiente da biblioteca Blacksheep, possibilitando a construção de uma API organizada e segura. A biblioteca também integra-se facilmente com o SQLAlchemy, essencial em aplicações astronômicas, que frequentemente demandam operações complexas de busca e agregação de dados.

6.2.2 Integrações

O backend é composto por uma API RESTful responsável por implementar a lógica de consulta por similaridade e permitir o acesso aos dados pelo frontend. Pensando em aumentar a flexibilidade e a experiência de usuário, o sistema permite que a busca seja feito pelo nome oficial da galáxia, apelido ou coordenada. Para garantir essa funcionalidade, é necessária uma interação com o webserviço astronômico *Sesame Name Solver*¹⁵, que é utilizado para resolver o termo de busca enviado pelo usuário, convertendo-o em posição angular da galáxia (RA e Dec).

6.2.3 Procedimento de Busca no Backend

A Fig. 18 mostra o diagrama de sequência do procedimento de busca por similaridade visual do ponto de vista do backend. Nela é mostrada as três etapas da busca: a resolução do nome do objeto, transformando-o em coordenadas; a busca por posição no banco de dados para encontrar os embeddings da galáxia buscada e, por fim, a busca por similaridade.

¹⁵<https://cds.unistra.fr/cgi-bin/Sesame>

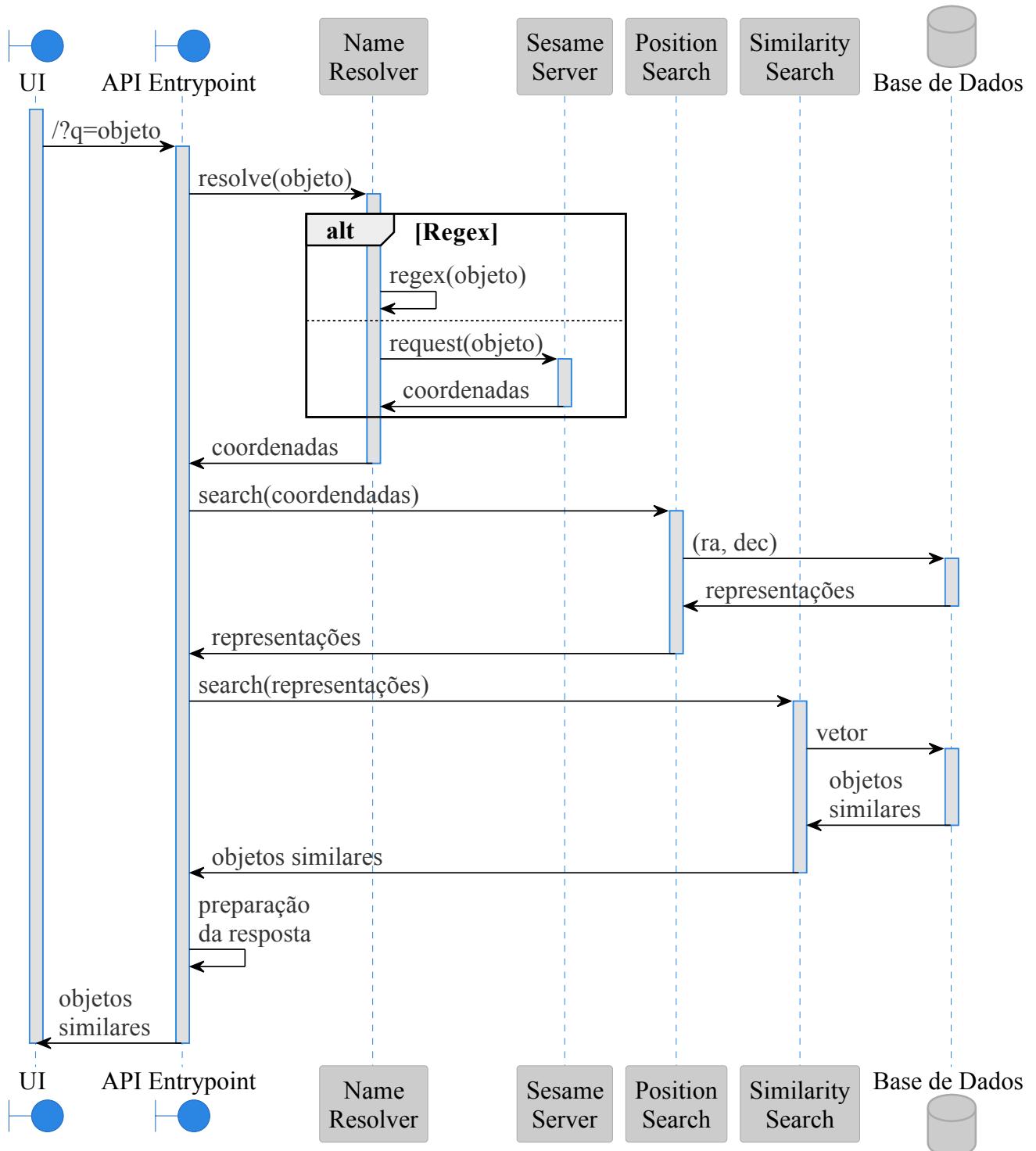


Figura 18: Diagrama de sequência da busca no backend. Primeiramente, o termo de busca é decodificado em coordenadas RA e Dec utilizando o webserviço Sesame. Em seguida, uma busca por posição é feita no banco de dados para obter as representações da galáxia busca. Por fim, é feita uma busca por similaridade para obter as galáxias similares.

6.3. FRONTEND

O frontend refere-se à camada da aplicação interpretada diretamente pelo navegador responsável pela interface visual e interação direta com o usuário. Ele abrange a estrutura, o estilo e o comportamento das páginas ou telas que os usuários visualizam e utilizam. O frontend conecta-se ao back-end por meio de APIs, recebendo e enviando dados para apresentar informações dinâmicas e garantir uma experiência de uso interativa e responsiva.

6.3.1 Tecnologias Utilizadas

O frontend é a camada da aplicação que é executada no navegador. Ele foi desenvolvido em TypeScript¹⁶, que é uma linguagem de programação de código aberto que expande o JavaScript ao adicionar suporte a tipagem estática e outras funcionalidades avançadas de desenvolvimento. O TypeScript é amplamente utilizado para melhorar a robustez e a escalabilidade do código, permitindo a detecção de erros em tempo de compilação e facilitando a manutenção de projetos complexos. As aplicações feitas nesta linguagem são transpiladas para JavaScript para serem interpretadas pelos navegadores.

O frontend foi desenvolvido utilizando o Next.js¹⁷, que é um framework de desenvolvimento web baseado na biblioteca React que permite a criação de aplicações modernas e escaláveis. A principal ferramenta usada é a geração estática (Static Site Generation, SSG), que permite compilar e distribuir o frontend como uma aplicação completamente independente do backend. Além disso, o Next.js fornece uma estrutura que simplifica o desenvolvimento ao incluir funcionalidades integradas, como roteamento automático baseado em arquivos, suporte a APIs serverless, otimização de desempenho e internacionalização.

A biblioteca React¹⁸ revolucionou a forma como interfaces de usuário (UI) são projetadas e construídas no desenvolvimento de aplicativos modernos. Baseada no conceito de componentes reutilizáveis, o React permite que desenvolvedores criem aplicações web e móveis de forma declarativa, eficiente e altamente escalável.

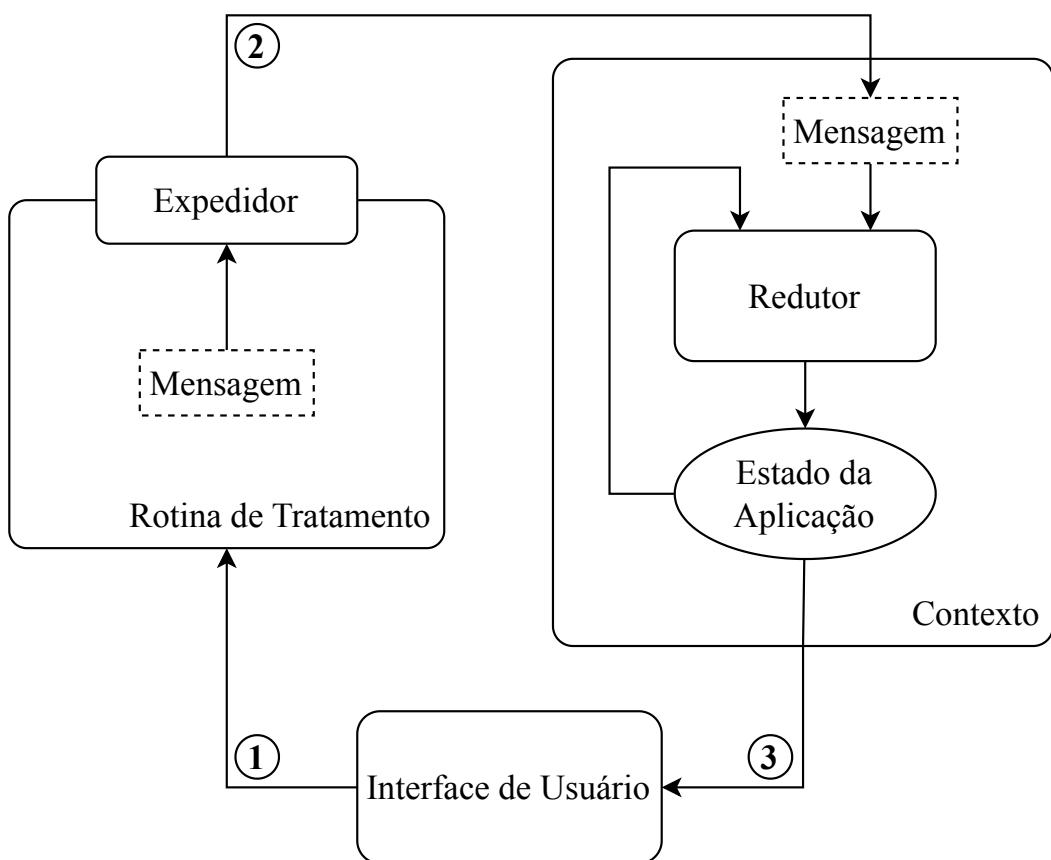
¹⁶<https://typescriptlang.org>

¹⁷<https://nextjs.org>

¹⁸<https://react.dev>

6.3.2 Arquitetura

Um dos principais diferenciais do React é sua abordagem declarativa para o desenvolvimento de UI. Em vez de manipular diretamente o Document Object Model (DOM), os desenvolvedores descrevem como a interface deve se comportar em diferentes estados, e o React cuida de atualizar o DOM de maneira eficiente. Essa abstração é possibilitada pelo Virtual DOM, uma representação leve do DOM real que permite ao React identificar e aplicar somente as alterações necessárias, minimizando a manipulação direta do DOM e melhorando o desempenho de aplicativos, especialmente em sistemas complexos com múltiplos estados dinâmicos.



■ **Figura 19:** Arquitetura do gerenciamento de estado da interface de usuário

Nessa abordagem declarativa, o gerenciamento de estado da aplicação é o mecanismo mais importante. Este mecanismo pode ser visto como uma máquina de estados finita (J. WANG, 2019) que depende do estado atual e uma entrada para produzir o próximo estado, como a máquina MEALY (1955). Este mecanismo é ilustrado no diagrama da Fig. 19 e seus elementos são explicados a seguir.

1. O usuário interage com a *Interface do Usuário* (UI), causando um evento, que é enviado para a *Rotina de Tratamento*. Este evento pode ser um clique do mouse, um pressionamento de tecla, uma seleção de arquivo e assim por diante.
2. A *Rotina de Tratamento* é a função responsável por processar o evento. Como o sistema é desenvolvido usando abordagem declarativa, ao invés de executar alterações diretas na UI, esta função prepara uma mensagem contendo todas as informações necessárias para mudar a aplicação para outro estado e passá-la para o *Expedidor*, responsável por enviar a mensagem para o *Redutor*.
3. O *Redutor* é a função, chamada pelo *Contexto da Aplicação*, responsável por gerar o próximo estado da aplicação. Esta função recebe o estado atual e a mensagem do *Expedidor* para criar um novo estado. A mutação no *Estado da Aplicação* dispara a atualização da UI, realizada pela biblioteca React. O *Redutor* pode ser visto como a função de transição da Máquina Mealy.

Além disso, a arquitetura baseada em componentes do React é outro ponto crucial para seu sucesso. Aplicações são divididas em partes menores e independentes, chamadas componentes, que encapsulam lógica, estilo e estrutura de forma modular. Esses componentes podem ser reutilizados em diferentes partes de um aplicativo ou até mesmo em projetos distintos, promovendo consistência no design e eficiência no desenvolvimento. Além disso, a reutilização de componentes reduz o tempo necessário para implementar novas funcionalidades e simplifica a manutenção do código, uma característica essencial para projetos modernos, que frequentemente envolvem equipes de desenvolvimento distribuídas e de grande escala.

6.3.3 Procedimento de Busca no Frontend

A Fig. 20 mostra o diagrama de sequência do procedimento de busca por similaridade visual do ponto de vista do frontend.

6.4. IMPLEMENTAÇÃO E IMPLANTAÇÃO

Nesta seção, são abordadas as técnicas de desenvolvimento e implantação utilizadas para a disponibilização do sistema de informação desenvolvido como uma aplicação web.

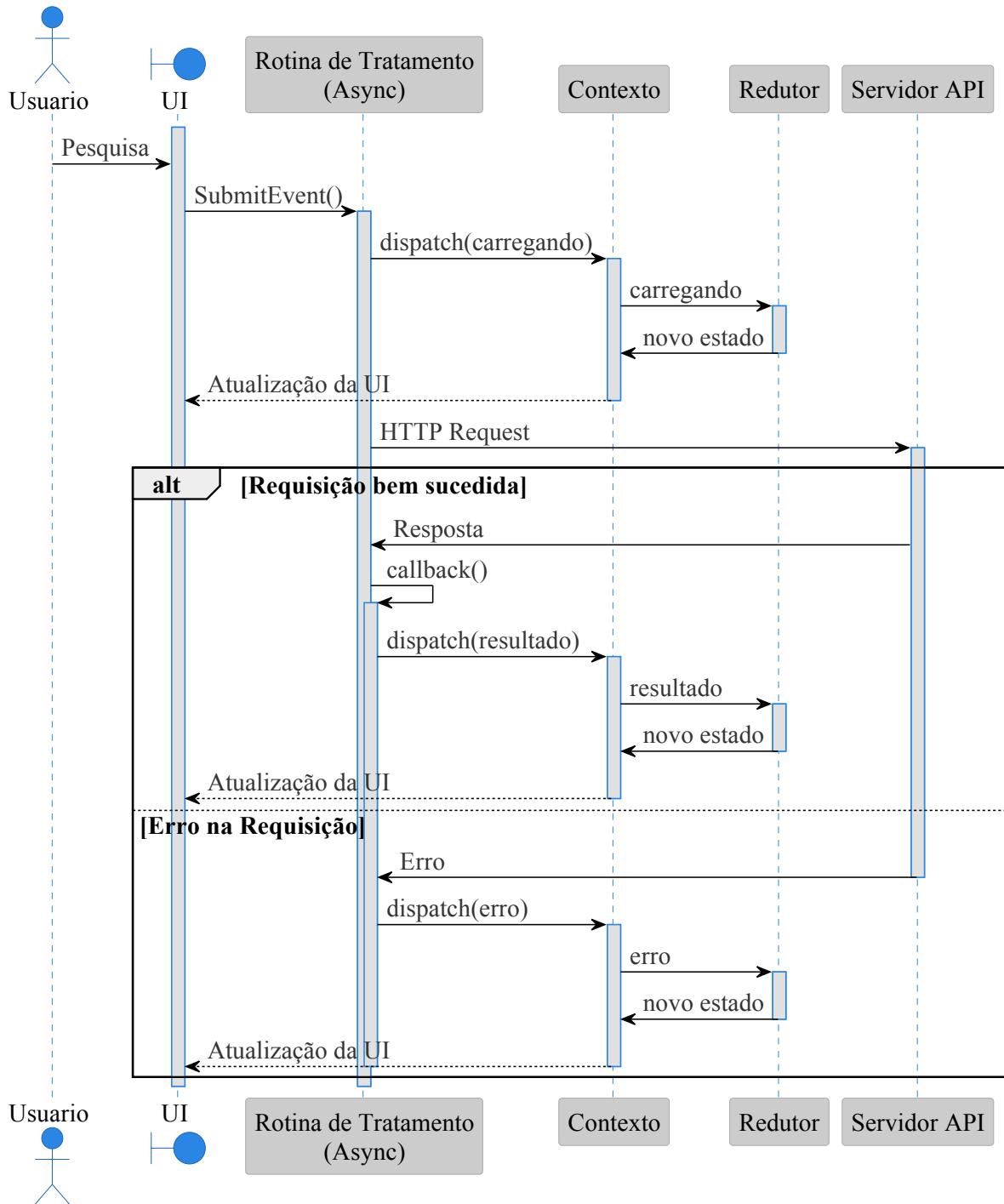


Figura 20: Diagrama de sequência da busca no frontend. Ao interagir com a barra de pesquisa, o usuário engatilha um evento, que é interceptado pelo motor de eventos do navegador, que inicializa a rotina de tratamento registrada para o evento. Esta função despacha uma mensagem, ao contexto da aplicação, para mudar o estado da UI para “carregando” e, em seguida, faz uma requisição HTTP ao servidor. Ao receber a resposta da API, a função callback é responsável por despachar uma mensagem com os resultados ou com erro, causando uma nova atualização da UI.

6.4.1 Sistema de Controle de Versão

O uso combinado do sistema de controle de versão Git e da plataforma GitHub¹⁹ no desenvolvimento de um sistema inteligente para busca de imagens similares em astronomia oferece uma série de vantagens que otimizam a colaboração, a gestão do código e a automação de processos. O Git²⁰, sendo um sistema distribuído, permite que os desenvolvedores rastreiem todas as alterações feitas no código-fonte de forma detalhada, possibilitando reverter modificações indesejadas e garantir a integridade do desenvolvimento. Sua estrutura descentralizada assegura que cada colaborador mantenha uma cópia completa do repositório, aumentando a segurança e a resiliência contra falhas.

A integração com o GitHub amplia essas capacidades, oferecendo uma interface intuitiva e funcionalidades adicionais que promovem o trabalho em equipe. A plataforma facilita a criação de ramificações (branches) para desenvolvimento paralelo, revisão de código por meio de pull requests e resolução de conflitos de forma transparente. Tais recursos são fundamentais em projetos de sistemas inteligentes, onde equipes frequentemente trabalham em componentes diferentes, como algoritmos de similaridade, infraestrutura de backend e interfaces gráficas.

Além disso, o GitHub fornece ferramentas de integração contínua (CI) e entrega contínua (CD), permitindo a execução automática de testes, validações e implantações a cada mudança no código. Isso reduz significativamente o risco de erros e acelera o ciclo de desenvolvimento, o que é crucial em projetos que lidam com grandes volumes de dados astronômicos. Recursos como issues e projetos também facilitam o gerenciamento de tarefas, promovendo organização e transparência.

6.4.2 DevOps

A uso de práticas DevOps e pipelines de integração contínua e entrega contínua (CI/CD) com o GitHub no desenvolvimento do sistema de informação de um sistema inteligente para busca de imagens similares em astronomia oferece vantagens significativas em termos de automação, colaboração e qualidade do software. Este procedimento promove uma forma de integração entre desenvolvimento e operações, facilitando o gerenciamento eficiente de infraestrutura e aplicações.

¹⁹<https://github.com>

²⁰<https://git.org>

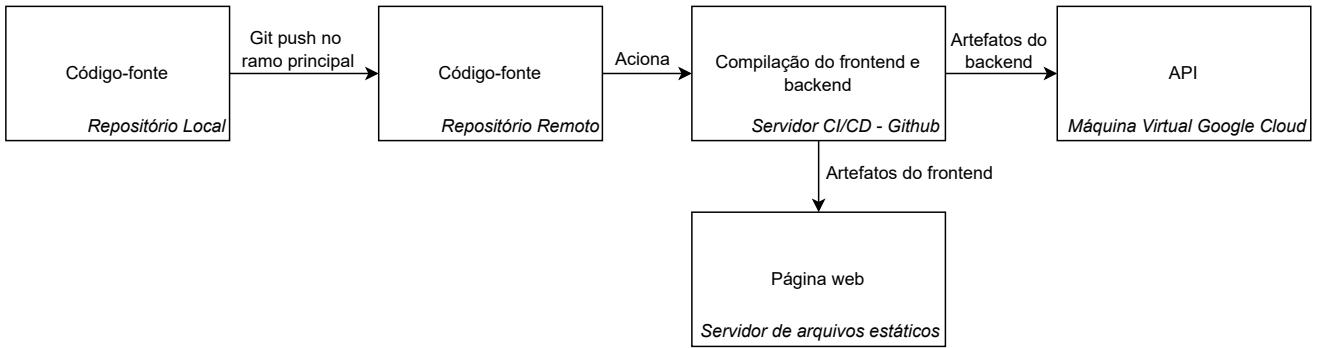


Figura 21: Integração contínua

Como mostra a Fig. 21, as pipelines de CI/CD foram configuradas no GitHub para serem acionadas quando ocorrer alguma atualização no ramo principal, automatizando a execução de validação de código e implantação de novas versões do sistema. A pipeline compila o frontend e o backend, que são despachados automaticamente para servidores distintos. Os arquivos estáticos do frontend (HTML, CSS, JS, imagens) são enviados ao serviço de hospedagem de arquivos estáticos e páginas web do próprio GitHub. Já o backend é enviado à uma máquina virtual do Google Cloud.

Isso resultou em uma maior agilidade na entrega de atualizações. Além disso, a rastreabilidade e transparência proporcionadas pelo GitHub, combinadas com as práticas DevOps, asseguram um fluxo de trabalho organizado e escalável, que suporta tanto a evolução do sistema quanto a colaboração entre múltiplas equipes e pesquisadores.

7. RESULTADOS

A análise detalhada do modelo desenvolvido é fundamental, pois a recuperação precisa de galáxias impacta diretamente na qualidade das análises científicas e na descoberta de fenômenos raros. Além disso, ao medir quantitativamente o desempenho, é possível identificar limitações, realizar ajustes ou refinar o processo de geração dos embeddings do modelo para otimizar sua aplicação em levantamentos astronômicos de larga escala.

Neste capítulo, será feita a avaliação do modelo no conjunto de teste para as tarefas de simulação dos votos (Seção 7.1) e recuperação de imagens a partir das representações visuais (Seção 7.2), por fim, também é feita uma avaliação de usabilidade do sistema (Seção 7.3).

7.1. 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 6 sumariza as métricas de acurácia (Seção 5.7.2), precisão (Seção 5.7.3), revocação (Seção 5.7.4) e F1-score (Seção 5.7.5) para várias perguntas do GalaxyZoo.

Tabela 6: 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 5.7.1) para questão da Tabela 6.

Os valores mostrados na Tabela 6 e nas Figs. 22 e 23 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.

7.2. AVALIAÇÃO DA RECUPERAÇÃO DE IMAGENS

A avaliação quantitativa do desempenho do sistema para a tarefa de recuperação de imagens é feita utilizando a métrica mean Average Precision (mAP; Seção 5.8.5), pois esta medida é crucial para avaliar a capacidade do sistema em priorizar e classificar corretamente imagens relevantes em relação a consultas específicas. A mAP integra a precisão ao longo de diferentes níveis de revocação, proporcionando uma visão detalhada do desempenho global do modelo em recuperar galáxias em grandes bancos de dados.

Os valores de mAP apresentados na Tabela 7 evidenciam a habilidade do modelo em manter alta relevância entre os itens recuperados, mesmo em cenários de grande complexidade ou com classes desbalanceadas. Essa métrica assegura que o sistema não apenas é eficaz em identificar galáxias relevantes, mas também organiza as recuperações de forma a maximizar sua utilidade científica, destacando a robustez e a capacidade de generalização do modelo para aplicações astronômicas.

A seguir, as Figs. 24, 25, 26, 27, 28, 29 e 30 mostram as buscas por similaridade visual para vários tipos de galáxias com a finalidade de exemplificar os resultados das buscas por similaridade visual obtidas pelo sistema desenvolvido.

7.3. AVALIAÇÃO DO SISTEMA DE INFORMAÇÃO

Na avaliação do Sistema de Informação, são considerados aspectos fundamentais para garantir uma alta experiência de usuário, como o tempo gasto para executar cada tarefa (Seção 7.3.1) e os aspectos de interface com o usuário (Seção 7.3.2).

Tabela 7: Avaliação do modelo para tarefa de recuperação de imagens

Questão	mAP
Spiral winding	0.988254
Bar	0.980521
Edge on bulge	0.990124
Disk edge on	0.999986
Smooth or featured	0.987412
Has spiral arms	0.968340
Galaxy Symetrical	0.962513
Clumpy appearance	0.957814
How rounded	0.978536
Merging	0.963656
Bulge size	0.987431
Spiral arms count	0.960259

7.3.1 Performance

A performance relacionada ao tempo de busca do sistema inteligente é fundamental para garantir sua viabilidade prática, especialmente diante do crescente volume de dados provenientes de levantamentos astronômicos de larga escala. Um tempo de busca otimizado garante a eficiência na recuperação de imagens relevantes, permitindo que pesquisadores obtenham resultados em tempo hábil para análises científicas. Isso é particularmente crítico em contextos como o monitoramento de eventos astronômicos transitórios, onde a agilidade na identificação de padrões visuais similares pode influenciar descobertas e reações rápidas.

A Tabela 8 sumariza o tempo gasto em cada etapa da busca. As otimizações implementadas na base de dados (Seções 6.1.3 e 6.1.4) a partir de índices específicos para cada tipo de dado colaboram para o desempenho da busca. No entanto, a maior parte do tempo é gasto na resolução do objeto, não sendo possível mitigá-lo, pois é realizado em um serviço na nuvem externo. Este foi o principal motivo para implementar o sistema de cache.

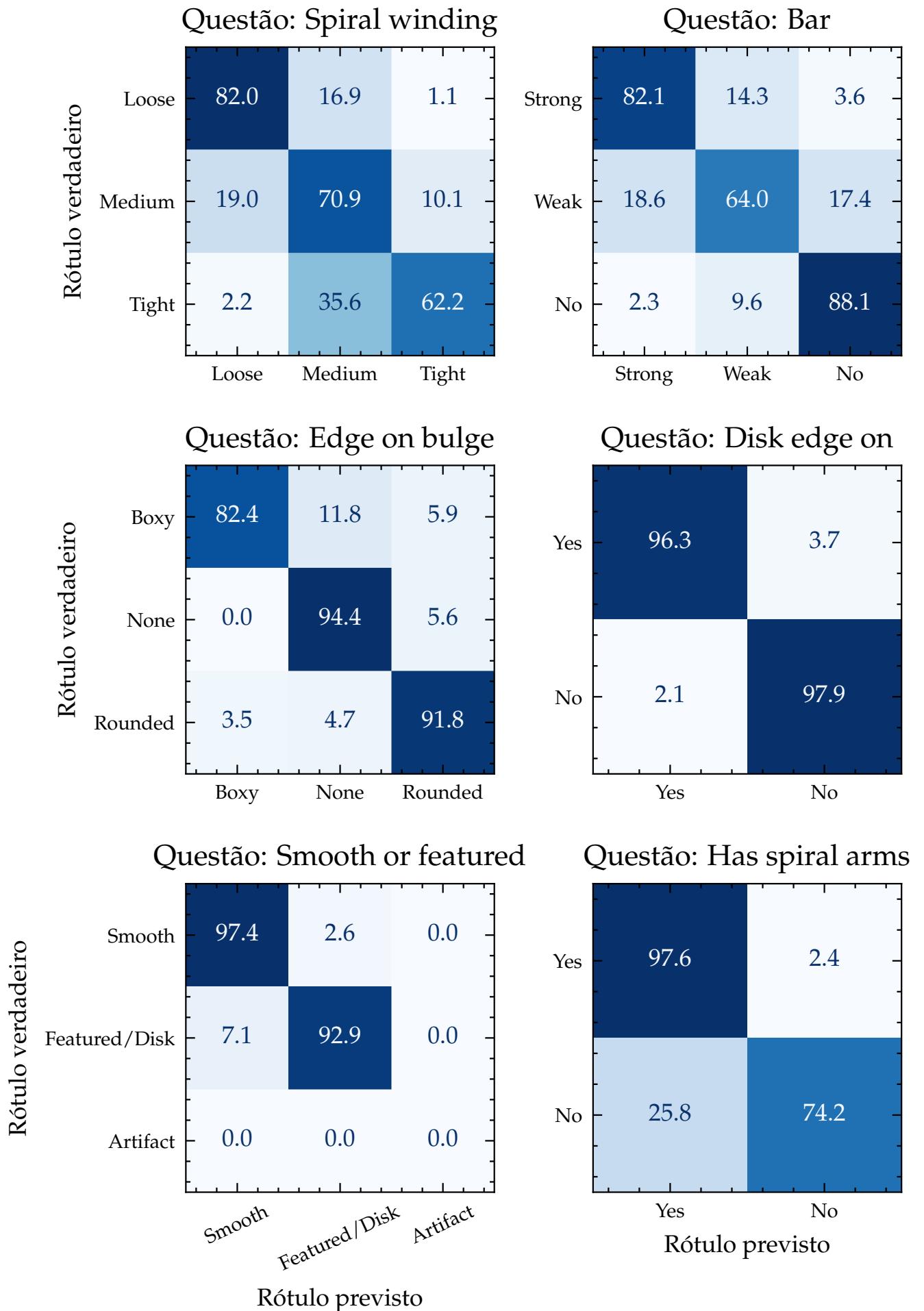
7.3.2 Interface Gráfica

Uma interface bem projetada permite que pesquisadores interajam intuitivamente com o sistema, interpretando os resultados de maneira eficiente, sem a necessidade de conhe-

Tarefa	Tempo gasto (sem cache)	Tempo gasto (com cache)
Resolução do nome	4.3 ± 2.8	0.1 ± 0.0
Busca em cone	1.7 ± 0.8	0.0 ± 0.0
Busca vetorial	2.1 ± 0.9	0.1 ± 0.0

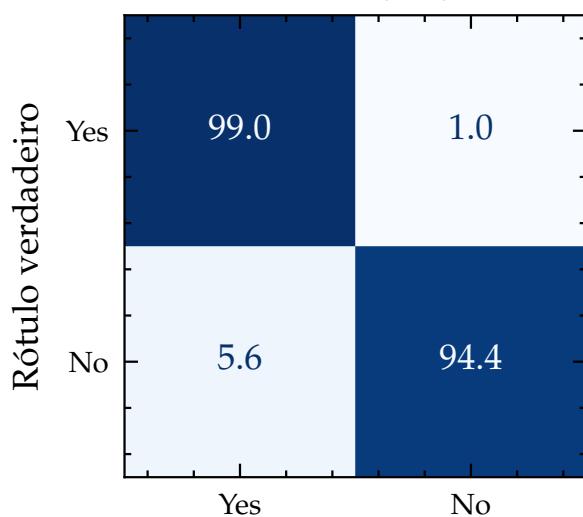
■ **Tabela 8:** Tempo gasto nas tarefas executadas durante a busca

cimentos avançados em programação ou sistemas de busca. Além disso, funcionalidades como visualização interativa dos resultados e apresentação clara de metadados relacionados às imagens recuperadas otimizam o processo de análise e exploração científica. A combinação de usabilidade, eficiência e clareza na interface gráfica garante que o sistema não apenas seja tecnicamente robusto, mas também facilite descobertas e análises dos dados. As Figs. 31, 32 e 33 mostram a interface gráfica do sistema.

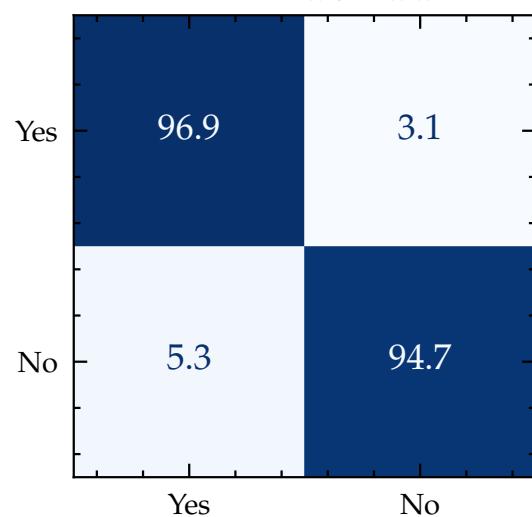


■ Figura 22: Matrizes de confusão para cada questão no conjunto de teste ⁸⁰

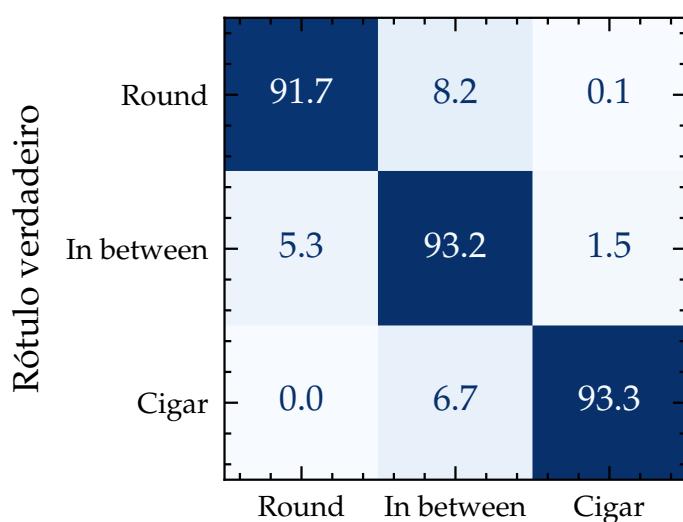
Questão: Galaxy Symetrical



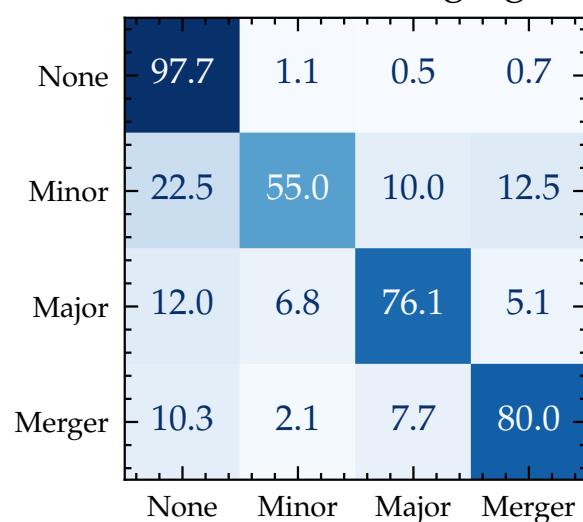
Questão: Clumpy appearance



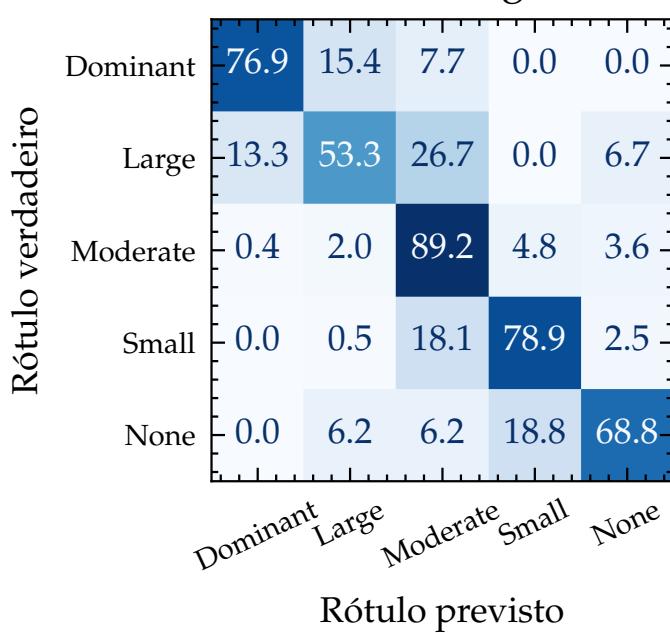
Questão: How rounded



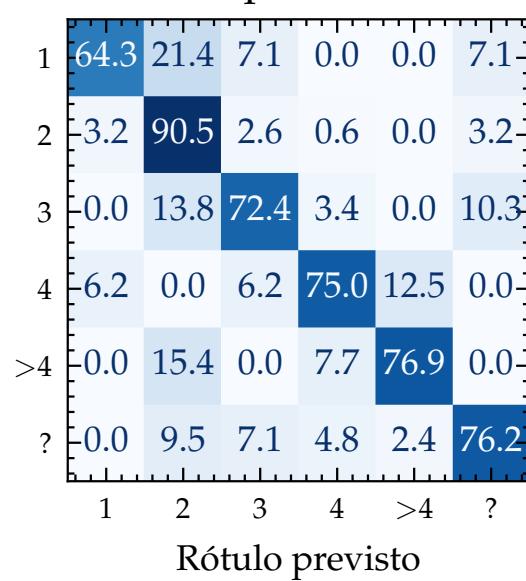
Questão: Merging



Questão: Bulge size



Questão: Spiral arms count

Figura 23: Matrizes de confusão para cada⁸¹ questão no conjunto de teste (continuação)

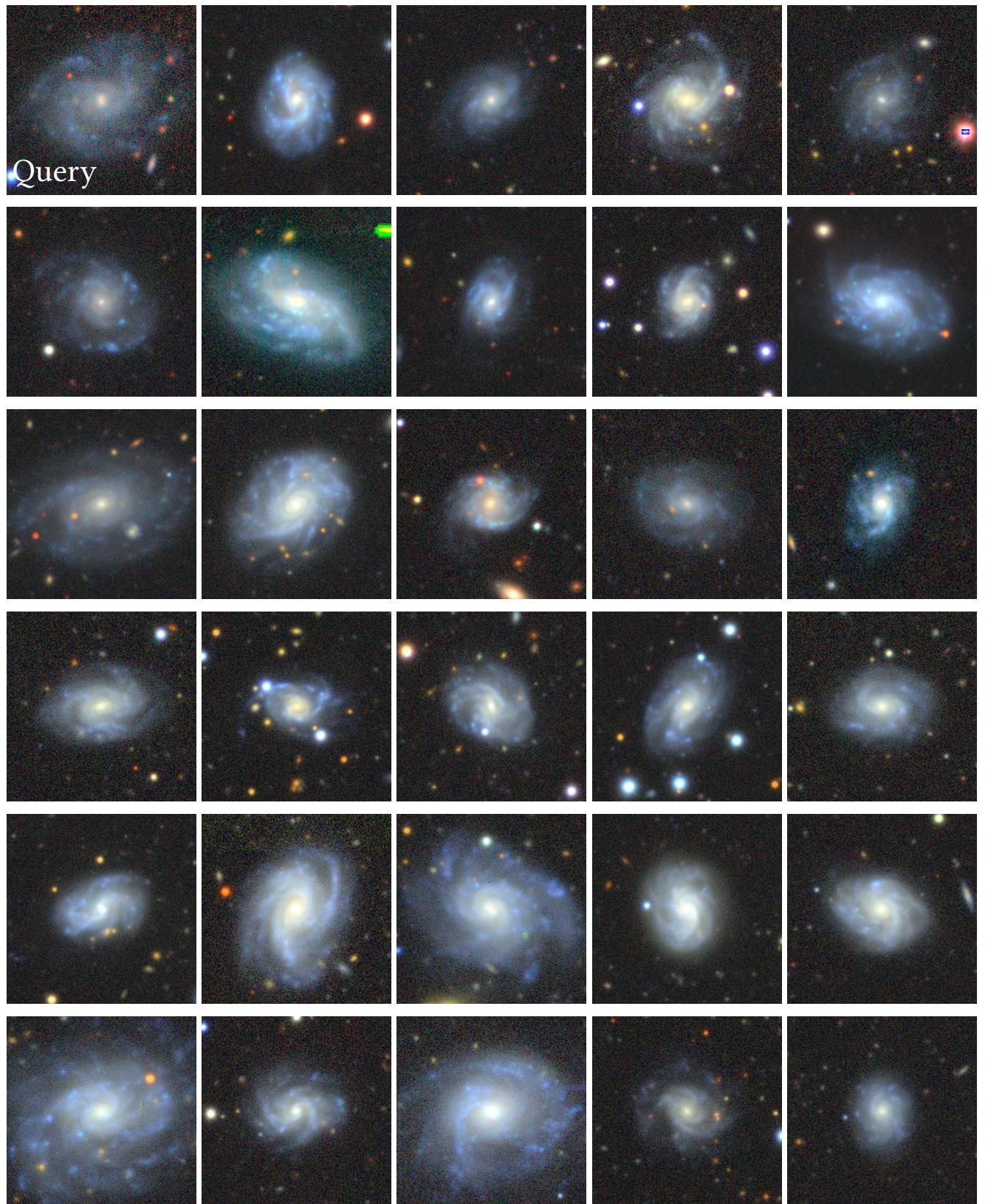


Figura 24: UGC 9010 é uma galáxia espiral com orientação face-on (face para o observador). O primeiro painel (Query) mostra a imagem da galáxia UGC 9010, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

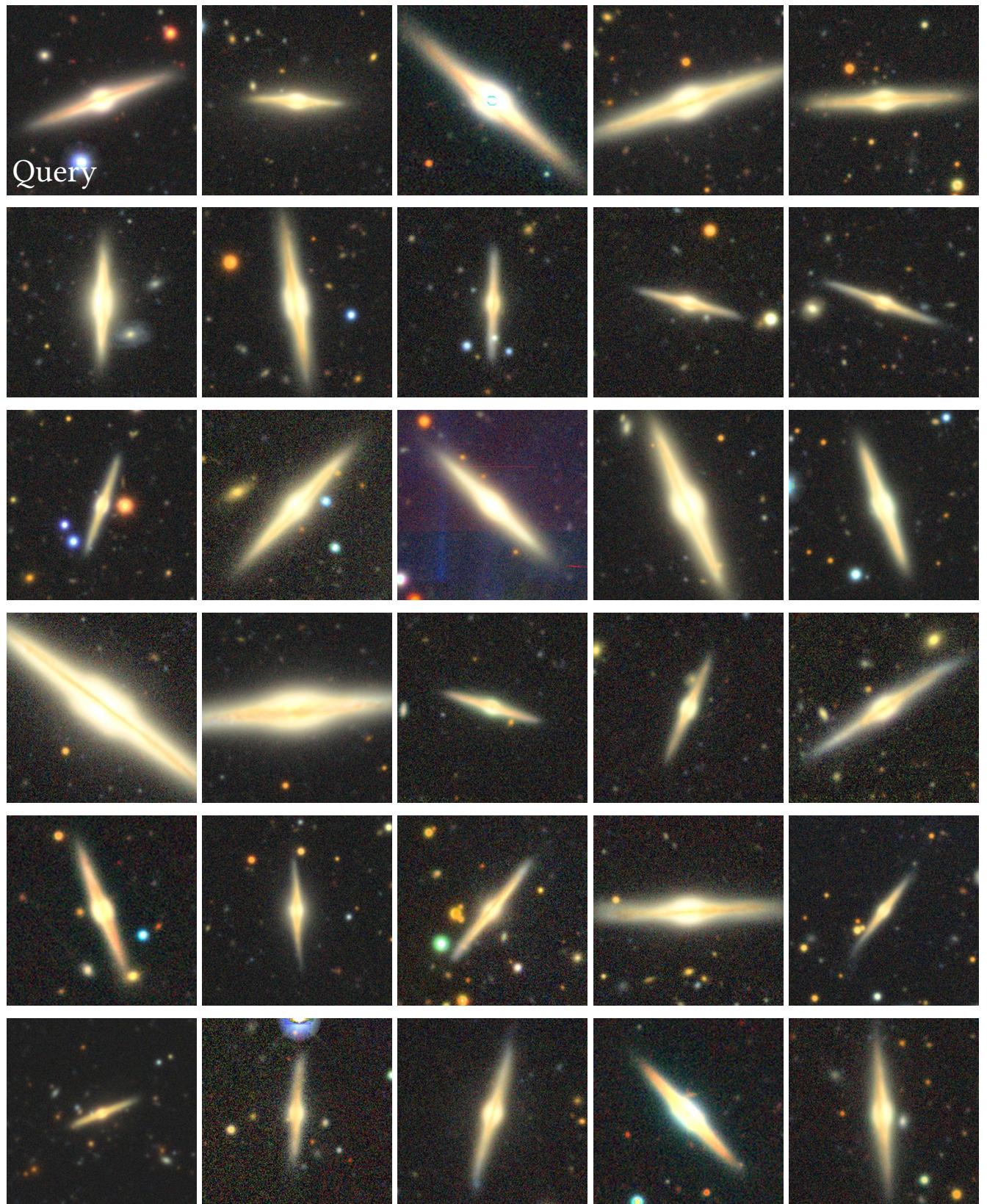


Figura 25: NGC 1043 é uma galáxia espiral com orientação edge-on (borda para o observador). O primeiro painel (Query) mostra a imagem da galáxia NGC 1043, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

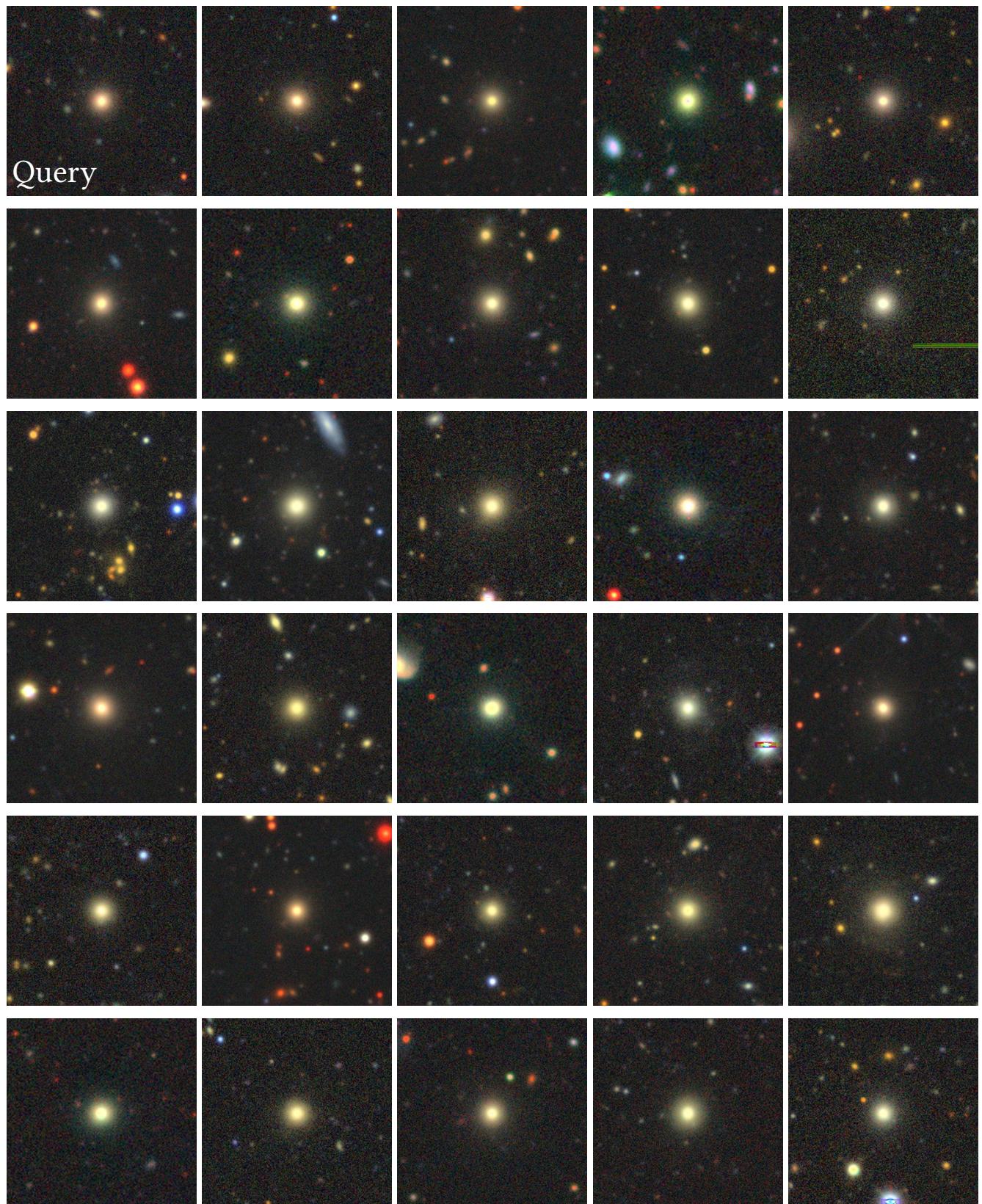


Figura 26: J082404.58+315023.3 é uma galáxia compacta, sua ocorrência na natureza é extremamente rara. O primeiro painel (Query) mostra a imagem da galáxia J082404.58+315023.3, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

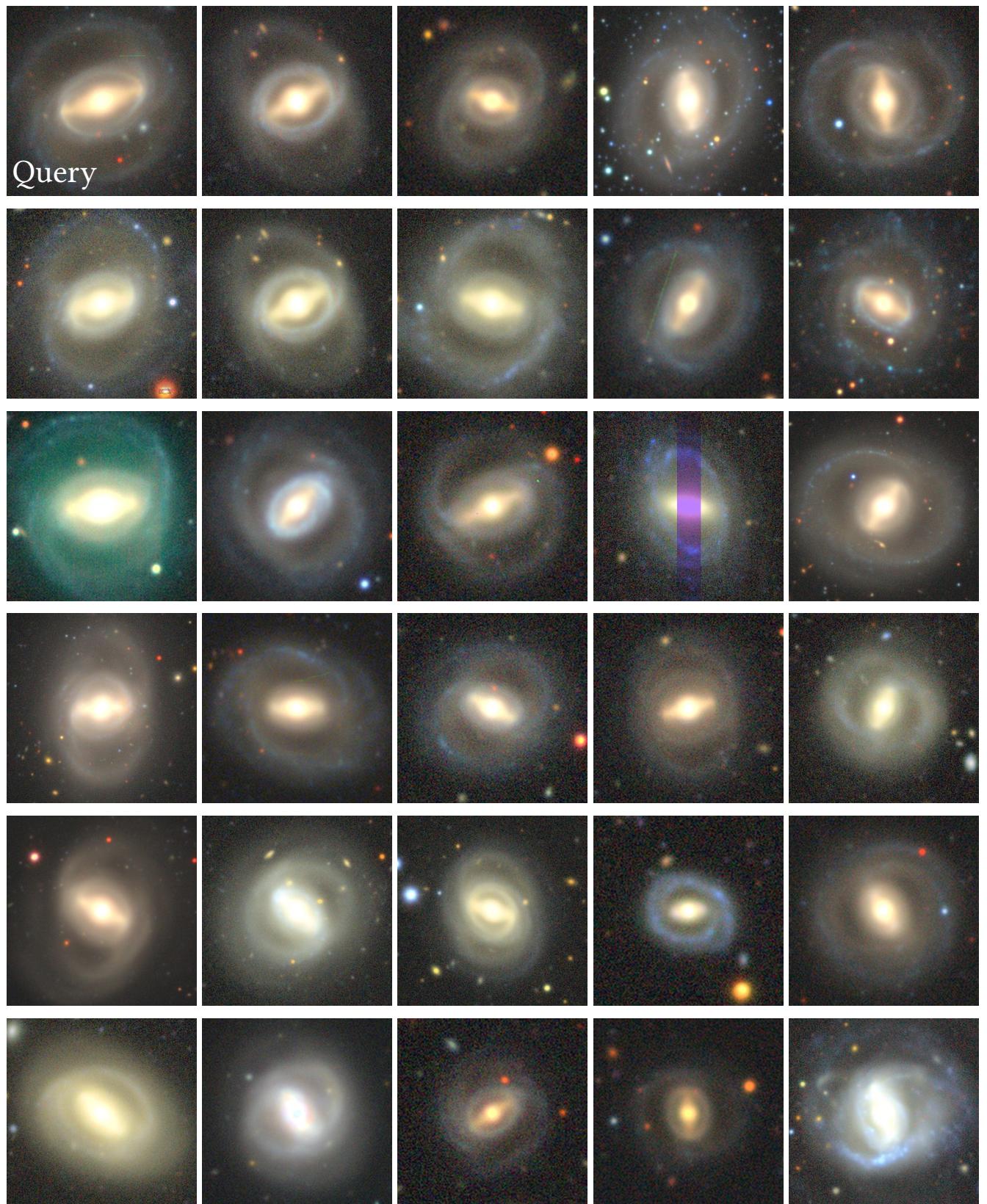


Figura 27: UGC 1698 é uma espiral barrada, o sistema foi capaz de recuperar galáxias muito parecidas. O primeiro painel (Query) mostra a imagem da galáxia UGC 1698, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

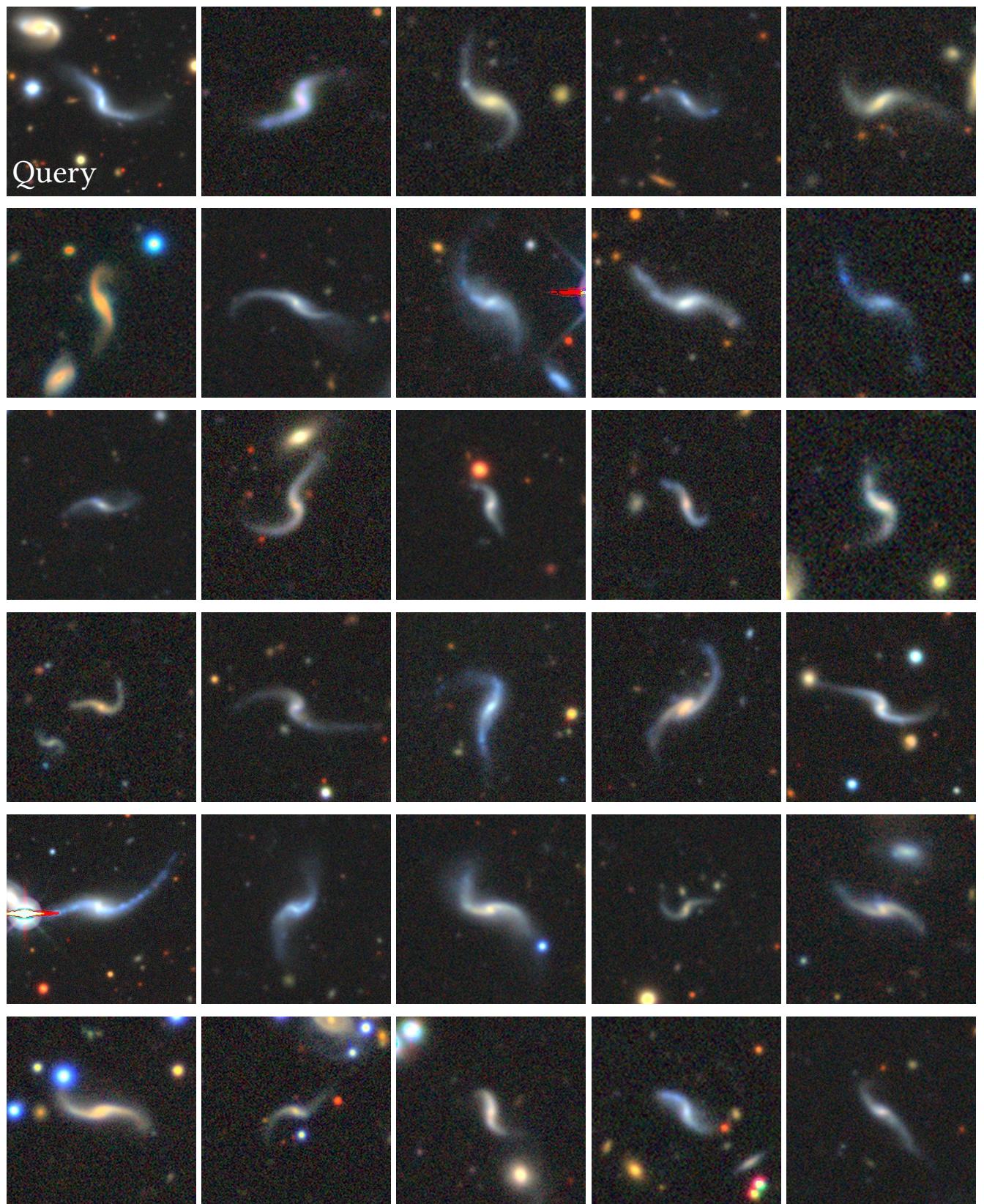


Figura 28: J060430.95-470755.7 é uma galáxia com enrolamento dos braços espirais aberdos. O primeiro painel (Query) mostra a imagem da galáxia J060430.95-470755.7, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

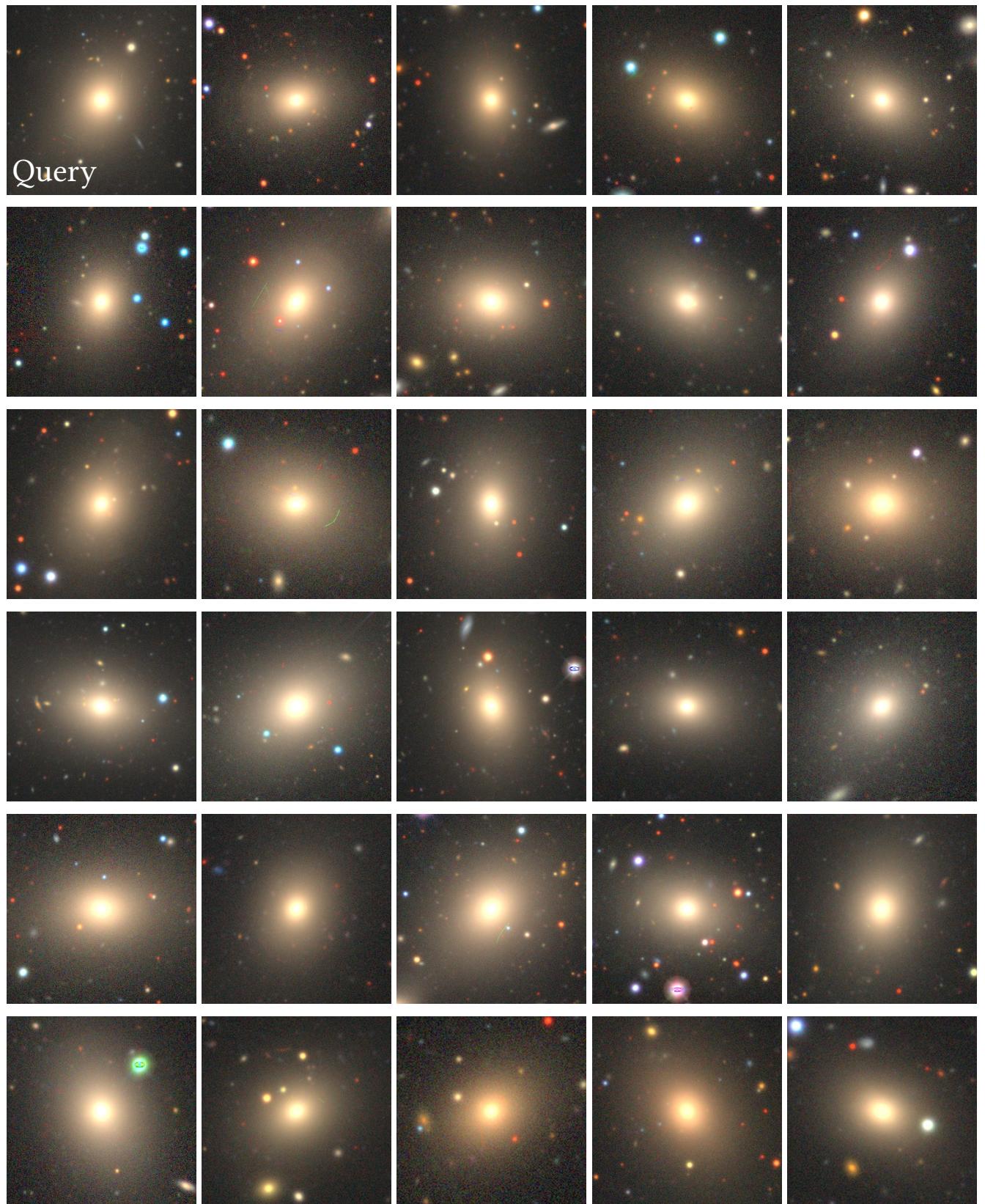


Figura 29: UGC 767 é uma galáxia elíptica. O primeiro painel (Query) mostra a imagem da galáxia UGC 767, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

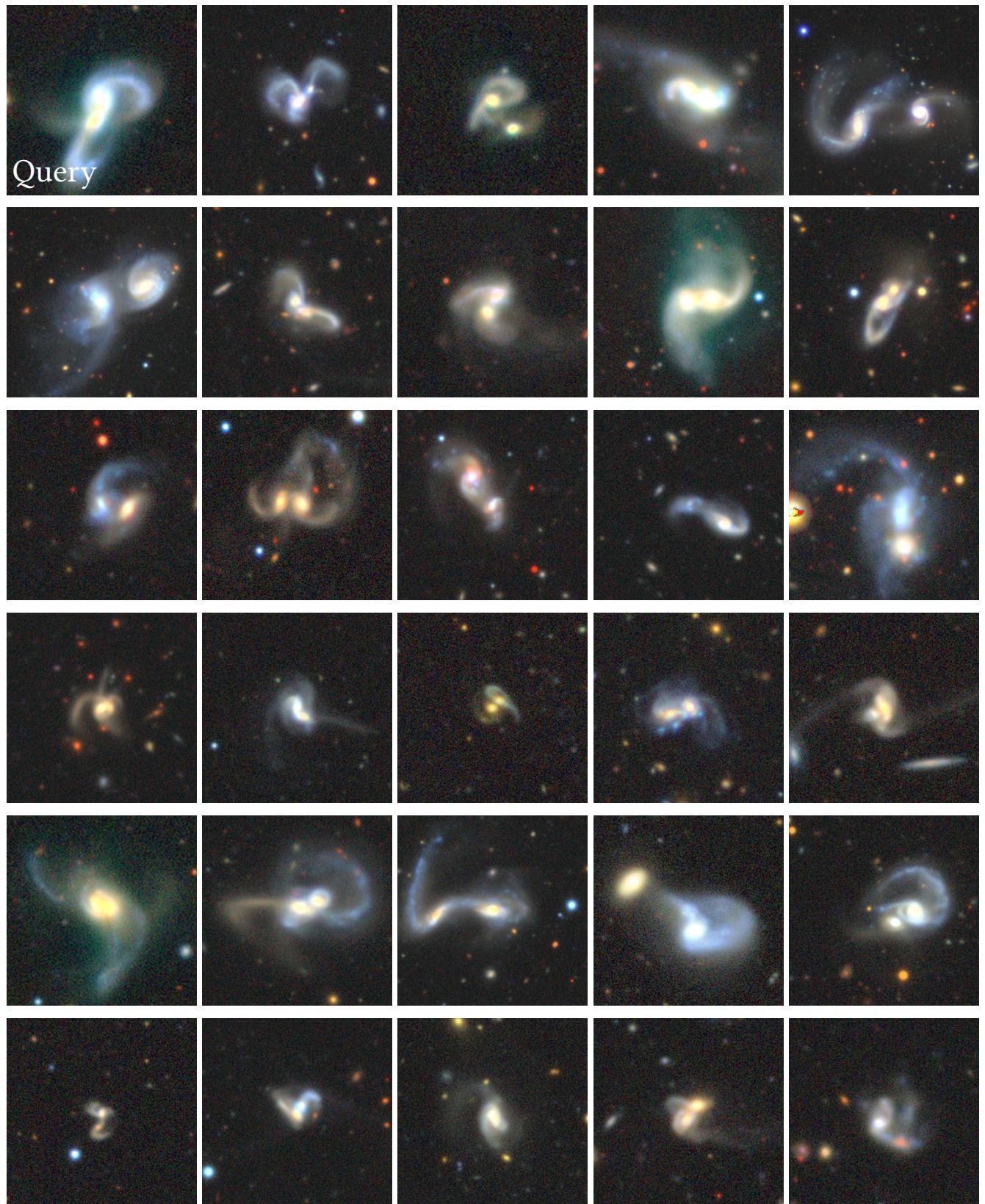


Figura 30: J151806.13+424445.2 mostra um fenômeno raro de fusão de duas galáxias. O primeiro painel (Query) mostra a imagem da galáxia J151806.13+424445.2, utilizada como referência. Os demais painéis mostram as galáxias obtidas por similaridade visual. Esta busca utilizou a distância do cosseno (Seção 5.8.1) como métrica de similaridade.

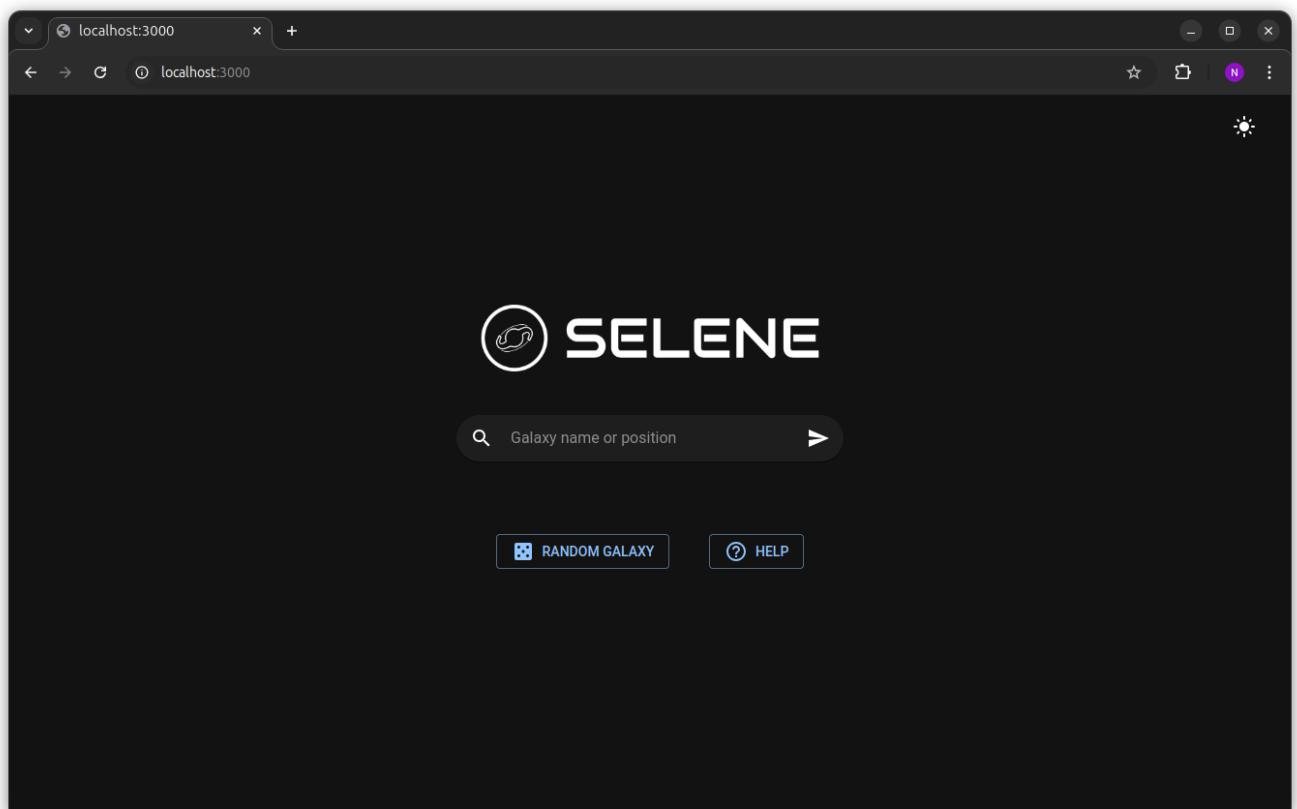
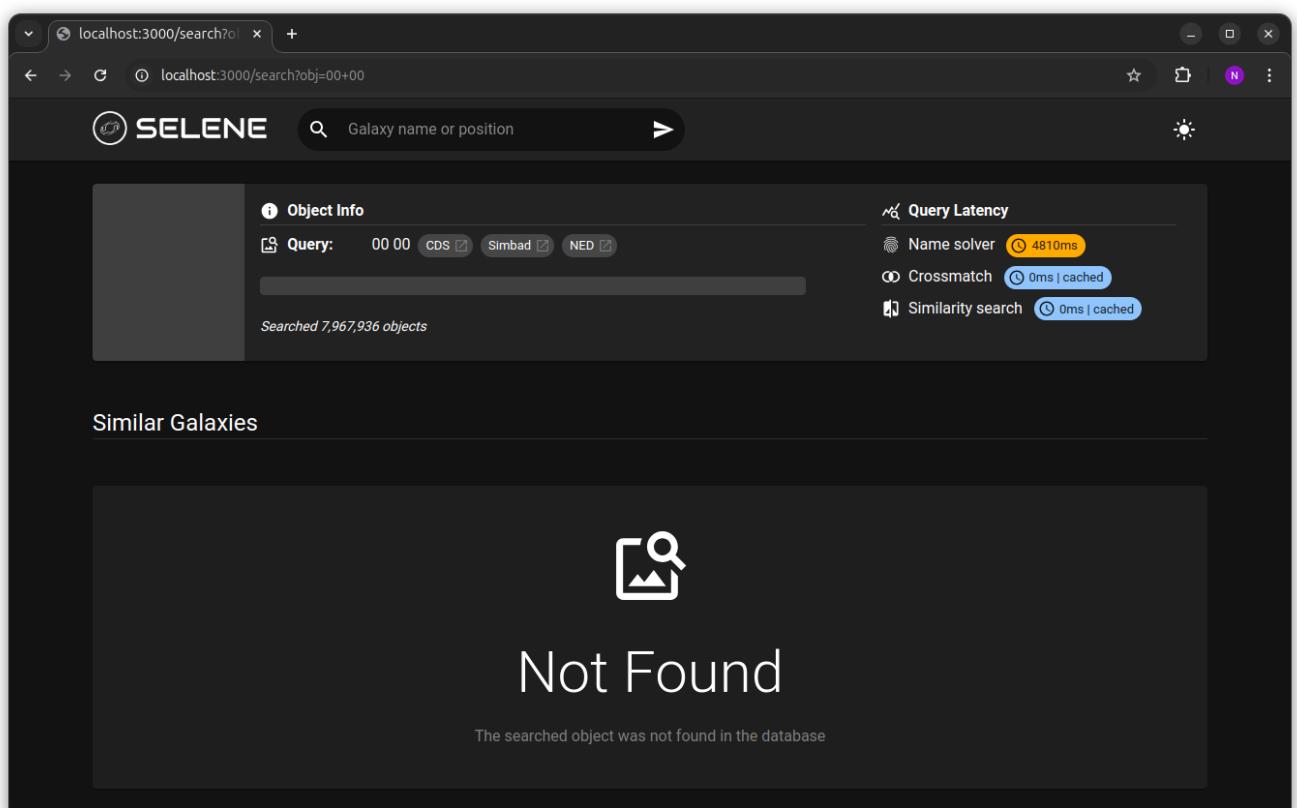


Figura 31: Tela inicial da aplicação. Com uma interface simples, a página inicial possui um seu centro o seu componente principal: a barra de busca. Nela, o usuário digita o nome ou posição da galáxia

■ Figura 32: Página de erro para objetos não encontrados



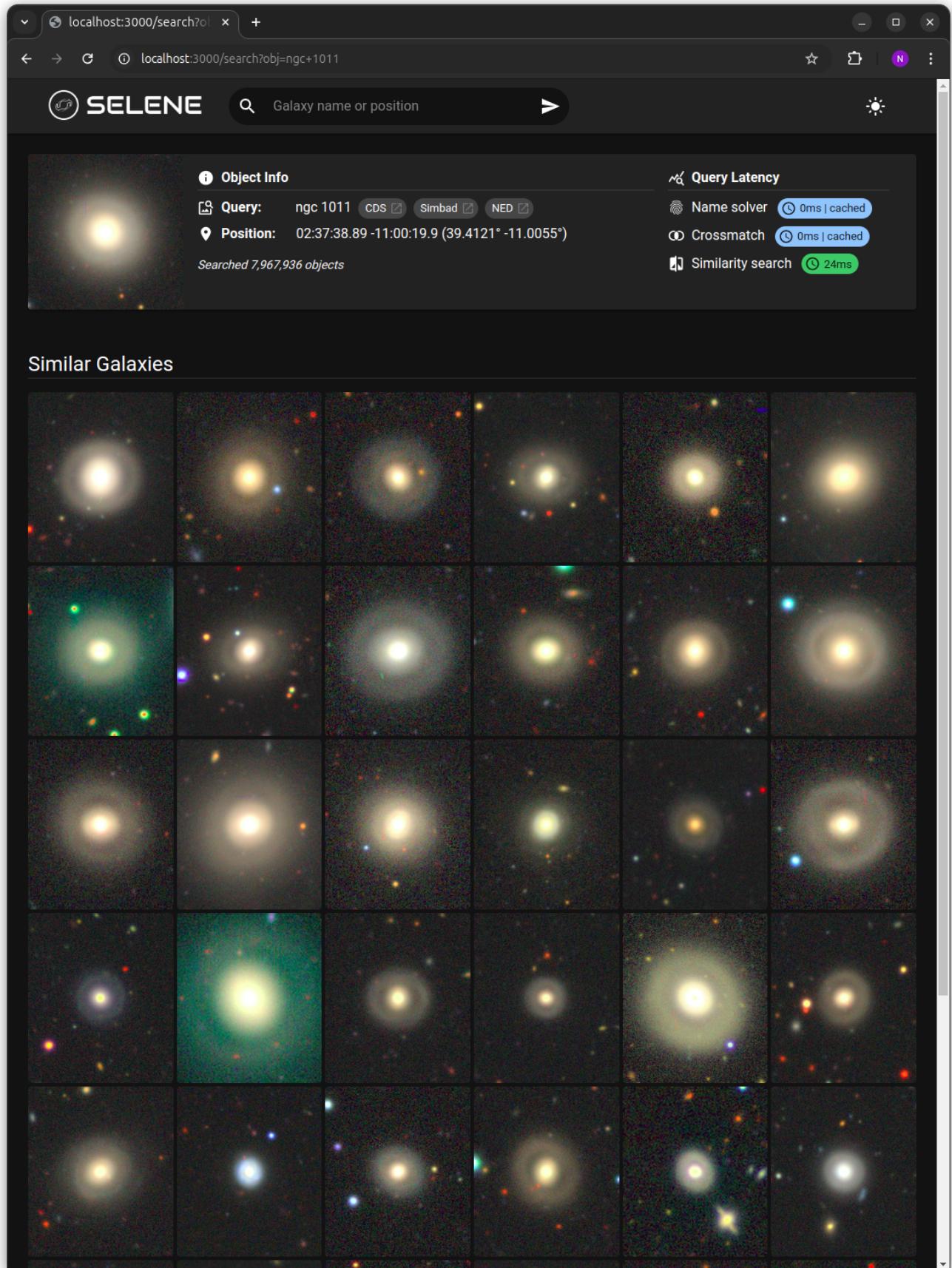


Figura 33: A tela de resultados exibe, no topo, um painel contendo informações adicionais da galáxia buscada e as medidas de tempo gasto em cada etapa da busca (resolução do nome, correlação e busca por similaridade). Abaixo, são mostradas todas as galáxias similares recuperadas.

8. CONCLUSÃO

8.1. SUMÁRIO

Este trabalho enfatiza a importância e as contribuições do desenvolvimento de um sistema inteligente para busca de objetos astronômicos por similaridade visual, que tem como objetivo principal facilitar e ampliar a análise e o estudo de dados astronômicos. Todos os procedimentos aqui documentados foram desenvolvidos neste projeto de formatura, desde a concepção (Cap. 3) até a especificação de detalhes técnicos de desenvolvimento fundamentais para garantir o funcionamento do sistema da maneira esperada, como a forma de cortar as figuras (Seção 4.2), os hiperparâmetros do modelo de aprendizagem profunda (Seção 5.4) e o tipo de indexação das colunas do banco de dados (Seção 6.1.1).

A implementação de um modelo de aprendizagem profunda, baseado em redes neurais convolucionais, permitiu a extração eficiente de características visuais (embeddings) dos objetos, possibilitando uma busca rápida e precisa de galáxias e outros corpos celestes que compartilham características semelhantes. Esse sistema não apenas facilita o processo de classificação e descoberta de novos padrões morfológicos no universo, mas também representa um avanço no campo da astronomia ao permitir análises detalhadas e o cruzamento de dados observacionais com maior precisão.

A validação do sistema evidenciou sua robustez e aplicabilidade prática. Testes realizados mostraram que o modelo é capaz de identificar objetos visualmente semelhantes com alta acurácia, mesmo em cenários desafiadores, como classes desbalanceadas e padrões visuais complexos. A interface gráfica desenvolvida não apenas assegurou a usabilidade, mas também ampliou o potencial de aplicação do sistema, permitindo que pesquisadores de diferentes níveis de especialização utilizem a ferramenta para explorar grandes volumes de dados de maneira intuitiva.

Além disso, foi reforçada a eficácia das redes neurais convolucionais em capturar padrões visuais complexos, mas também emergiu a necessidade de um design cuidadoso de arquiteturas e ajustes de hiperparâmetros. Modelos como EfficientNet mostraram-se adequados para tarefas que demandam alta eficiência e precisão. Além disso, a transfe-

rência de aprendizado, com redes previamente treinadas em datasets como ImageNet, foi uma estratégia crucial para reduzir o custo computacional e melhorar a generalização do modelo em um domínio especializado como a astronomia.

Por fim, o sistema desenvolvido não apenas atende às demandas crescentes da astronomia moderna, mas também abre caminho para novas aplicações em outros domínios científicos que enfrentam desafios semelhantes relacionados ao big data e à análise visual automatizada. Este trabalho representa uma contribuição relevante para a ciência computacional e astronômica, fornecendo uma base sólida para futuras inovações e pesquisas no campo de mineração de dados multimídia.

8.2. LIÇÕES APRENDIDAS

Na manipulação de grandes volumes de dados astronômicos, uma das principais lições foi a integração de técnicas avançadas de processamento paralelo e indexação vetorial, como KD-trees, foi essencial para lidar com o volume de imagens e embeddings gerados. Esse aprendizado destacou a importância da eficiência na manipulação dos dados, desde o pré-processamento das imagens até a estruturação dos conjuntos de treinamento, validação e teste.

Além disso, a implementação de uma função de custo baseada na distribuição dos votos do GalaxyZoo representou um avanço significativo. Foi necessário projetar uma função que equilibrasse a variabilidade dos rótulos com a necessidade de estabilidade nos gradientes, resultando em um modelo que capturasse melhor as incertezas e distribuições observadas nos dados reais. Esse processo evidenciou a importância de alinhar os objetivos da modelagem com as especificidades do problema.

Por fim, a colaboração entre ciência da computação e astronomia não apenas impulsionou o desenvolvimento de soluções tecnológicas, mas também possibilitou a validação contínua dos resultados no contexto científico.

8.3. CONTRIBUIÇÃO CIENTÍFICA

Atualmente, os dados gerados por este sistema são usados em diversas pesquisas do Instituto de Astronomia Geofísica e Ciências Atmosféricas (IAG-USP), incluindo um projeto de iniciação científica, um projeto de mestrado e dois projetos de pós-doutorado.

Em agosto deste ano, este projeto foi apresentado na *19^a International Meeting of the Southern Photometric Local Universe Survey (S-PLUS) Collaboration* e, entre 7 e 10 de abril de 2025, será apresentada, na Argentina, uma palestra e dois hands-on na *XI La Plata International School* (<https://congresos.unlp.edu.ar/xilapis>) sobre as técnicas de aprendizado profundo e lições aprendidas no decorrer deste projeto.

8.4. TRABALHOS FUTUROS

Dando continuidade ao desenvolvimento deste projeto, previsto para ser aprofundado em nível de mestrado, os trabalhos futuros buscarão explorar melhorias na eficiência e na precisão do sistema proposto. Desta forma, pretende-se investigar o uso de modelos de aprendizado profundo baseados em arquiteturas de transformadores, que têm mostrado significativo em tarefas de visão computacional. Além disso, as redes neurais baseadas em grafos, têm mostrado eficácia na tarefa de recuperação de conteúdo. Pois, ao representar os objetos astronômicos como nós em um grafo, com arestas refletindo as similaridades, é possível capturar interações de alto nível que as redes convolucionais tradicionais não conseguem representar. Por fim, está planejada a publicação de um artigo sobre este projeto desenvolvido.

REFERÊNCIAS

- ABAS, MOHAMAD AQIB HAQMI et al. (out. de 2018). “VGG16 for Plant Image Classification with Transfer Learning and Data Augmentation”. Em: *International Journal of Engineering & Technology* 7.4.11, p. 90. DOI: [10.14419/ijet.v7i4.11.20781](https://doi.org/10.14419/ijet.v7i4.11.20781). URL: <https://doi.org/10.14419/ijet.v7i4.11.20781>.
- AKIBA, TAKUYA et al. (jul. de 2019). “Optuna”. Em: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, pp. 2623–2631. ISBN: 9781450362016. DOI: [10.1145/3292500.3330701](https://doi.org/10.1145/3292500.3330701). URL: <https://doi.org/10.1145/3292500.3330701>.
- ALMEIDA, ANDRÉS et al. (ago. de 2023). “The Eighteenth Data Release of the Sloan Digital Sky Surveys: Targeting and First Spectra from SDSS-V”. Em: *The Astrophysical Journal Supplement Series* 267.2, p. 44. ISSN: 1538-4365. DOI: [10.3847/1538-4365/acda98](https://doi.org/10.3847/1538-4365/acda98). URL: <https://doi.org/10.3847/1538-4365/acda98>.
- BAHCALL, NETA A. (ago. de 1995). “CLUSTERING AND LARGE-SCALE STRUCTURE WITH THE SLOAN DIGITAL SKY SURVEY”. Em: *Publications of the Astronomical Society of the Pacific* 107 (714), p. 790. ISSN: 1538-3873. DOI: [10.1086/133625](https://doi.org/10.1086/133625). URL: <https://doi.org/10.1086/133625>.
- BALEISIS, AUDRA et al. (jun. de 1998). “Searching for large-scale structure in deep radio surveys”. Em: *Monthly Notices of the Royal Astronomical Society* 297 (2), pp. 545–558. ISSN: 0035-8711. DOI: [10.1046/j.1365-8711.1998.01536.X](https://doi.org/10.1046/j.1365-8711.1998.01536.x). URL: <https://dx.doi.org/10.1046/j.1365-8711.1998.01536.x>.
- BELAID, OUIZA NAIT e MALIK LOUDINI (2020). “Classification of Brain Tumor by Combination of Pre-Trained VGG16 CNN”. Em: *Journal of Information Technology Management* 12.2, pp. 13–25. ISSN: 2008-5893. DOI: [10.22059/jitm.2020.75788](https://doi.org/10.22059/jitm.2020.75788). eprint: https://jitm.ut.ac.ir/article_75788_e36c948ee9258c82b9398f136692f3f5.pdf. URL: https://jitm.ut.ac.ir/article_75788.html.
- BLEI, DAVID M., ANDREW Y. NG e MICHAEL I. JORDAN (mar. de 2003). “Latent dirichlet allocation”. Em: *The Journal of Machine Learning Research* 3, pp. 993–1022. ISSN: 1532-4435. DOI: [10.5555/944919.944937](https://doi.org/10.5555/944919.944937). URL: <https://doi.org/10.5555/944919.944937>.

- BOM, C. R. et al. (out. de 2021). “Deep Learning assessment of galaxy morphology in S-PLUS Data Release 1”. Em: *Mon. Not. R. Astron. Soc.* 507.2, pp. 1937–1955. ISSN: 0035-8711. DOI: [10.1093/mnras/stab1981](https://doi.org/10.1093/mnras/stab1981). URL: <https://doi.org/10.1093/mnras/stab1981>.
- BONNAREL, F. et al. (abr. de 2000). “The ALADIN interactive sky atlas”. Em: *Astronomy and Astrophysics Supplement Series* 143 (1), pp. 33–40. ISSN: 0365-0138. DOI: [10.1051/aas:2000331](https://doi.org/10.1051/aas:2000331). URL: <http://doi.org/10.1051/aas:2000331>.
- BONNEY, RICK et al. (mar. de 2014). “Citizen science. Next steps for citizen science.” Em: *Science (New York, N.Y.)* 343.6178, pp. 1436–7. ISSN: 1095-9203. DOI: [10.1126/science.1251554](https://doi.org/10.1126/science.1251554). URL: <https://doi.org/10.1126/science.1251554>.
- BOVIK, ALAN C. (2009). “Basic Gray Level Image Processing”. Em: *The Essential Guide to Image Processing*. Elsevier, pp. 43–68. DOI: [10.1016/b978-0-12-374457-9.00003-2](https://doi.org/10.1016/b978-0-12-374457-9.00003-2). URL: <https://doi.org/10.1016/b978-0-12-374457-9.00003-2>.
- BRIDLE, JOHN (1990). “Training Stochastic Model Recognition Algorithms as Networks can Lead to Maximum Mutual Information Estimation of Parameters”. Em: *Advances in Neural Information Processing Systems*. Ed. por D. TOURETZKY. Vol. 2. Morgan-Kaufmann. URL: <https://proceedings.neurips.cc/paper/1989/file/0336dcbab05b9d5ad24f4333c76Paper.pdf>.
- BUTA, RONALD J. (2011). *Galaxy Morphology*. arXiv: [1102.0550 \[astro-ph.CO\]](https://arxiv.org/abs/1102.0550).
- CANINI, KEVIN, LEI SHI e THOMAS GRIFFITHS (nov. de 2009). “Online Inference of Topics with Latent Dirichlet Allocation”. Em: *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. Ed. por DAVID van DYK e MAX WELLING. Vol. 5. PMLR, pp. 65–72. URL: <https://proceedings.mlr.press/v5/canini09a.html>.
- CECCONI, BAPTISTE et al. (jan. de 2023). *UCD1+ controlled vocabulary - Updated List of Terms Version 1.5 Version 1.5*. IVOA Endorsed Note 25 January 2023. DOI: [10.5479/ADS/bib/2023ivoa.spec.0125C](https://doi.org/10.5479/ADS/bib/2023ivoa.spec.0125C). URL: <https://doi.org/10.5479/ADS/bib/2023ivoa.spec.0125C>.
- CHAMBERLIN, DONALD D e RAYMOND F BOYCE (1974). “SEQUEL: A structured English query language”. Em: *Proceedings of the 1974 ACM SIGFIDET (Now SIGMOD) Workshop on Data Description, Access and Control*. Association for Computing Machinery, pp. 249–264. ISBN: 9781450374156. DOI: [10.1145/800296.811515](https://doi.org/10.1145/800296.811515). URL: <https://doi.org/10.1145/800296.811515>.
- CHOLLET, FRANCOIS (2017). *Deep Learning with Python*. 1st. USA: Manning Publications Co. ISBN: 1617294438.

- CODD, E. F. (jun. de 1970). “A relational model of data for large shared data banks”. Em: *Communications of the ACM* 13 (6), pp. 377–387. ISSN: 15577317. DOI: [10.1145/362384.362685](https://doi.org/10.1145/362384.362685). URL: <https://doi.org/10.1145/362384.362685>.
- (dez. de 1979). “Extending the database relational model to capture more meaning”. Em: *ACM Transactions on Database Systems (TODS)* 4 (4), pp. 397–434. ISSN: 15574644. DOI: [10.1145/320107.320109](https://doi.org/10.1145/320107.320109). URL: <https://doi.org/10.1145/320107.320109>.
- DENG, JIA et al. (2009). “ImageNet: A large-scale hierarchical image database”. Em: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. DOI: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848). URL: <https://doi.org/10.1109/CVPR.2009.5206848>.
- DEY, ARJUN et al. (abr. de 2019). “Overview of the DESI Legacy Imaging Surveys”. Em: *The Astronomical Journal* 157.5, p. 168. ISSN: 1538-3881. DOI: [10.3847/1538-3881/ab089d](https://doi.org/10.3847/1538-3881/ab089d). URL: <https://doi.org/10.3847/1538-3881/ab089d>.
- DOSOVITSKIY, ALEXEY et al. (out. de 2020). “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”. Em: *arXiv*. DOI: [10.48550/arXiv.2010.11929](https://doi.org/10.48550/arXiv.2010.11929). eprint: [2010.11929](https://doi.org/10.48550/arXiv.2010.11929). URL: <https://doi.org/10.48550/arXiv.2010.11929>.
- DOWLER, PATRICK, FRANÇOIS BONNAREL e DOUG TODY (dez. de 2015). *IVOA Simple Image Access Version 2.0*. IVOA Recommendation 23 December 2015. DOI: [10.5479/ADS/bib/2015ivoa.spec.1223D](https://doi.org/10.5479/ADS/bib/2015ivoa.spec.1223D). URL: <https://doi.org/10.5479/ADS/bib/2015ivoa.spec.1223D>.
- DOWLER, PATRICK, MARKUS DEMLEITNER et al. (mai. de 2017). *Data Access Layer Interface Version 1.1*. IVOA Recommendation 17 May 2017. DOI: [10.5479/ADS/bib/2017ivoa.spec.0517D](https://doi.org/10.5479/ADS/bib/2017ivoa.spec.0517D). URL: <https://doi.org/10.5479/ADS/bib/2017ivoa.spec.0517D>.
- DOWLER, PATRICK, JANET EVANS et al. (nov. de 2021). *IVOA Architecture Version 2.0*. IVOA Endorsed Note 01 November 2021. DOI: [10.5479/ADS/bib/2021ivoa.spec.1101D](https://doi.org/10.5479/ADS/bib/2021ivoa.spec.1101D). URL: <https://doi.org/10.5479/ADS/bib/2021ivoa.spec.1101D>.
- DOWLER, PATRICK, GUY RIXON et al. (set. de 2019). *Table Access Protocol Version 1.1*. IVOA Recommendation 27 September 2019. DOI: [10.5479/ADS/bib/2019ivoa.spec.0927D](https://doi.org/10.5479/ADS/bib/2019ivoa.spec.0927D). URL: <https://doi.org/10.5479/ADS/bib/2019ivoa.spec.0927D>.
- DOZAT, TIMOTHY (2015). “Incorporating Nesterov Momentum into”. Em.
- FANSON, JAMES L. et al. (ago. de 2022). *Overview and status of the Giant Magellan Telescope project*. Ed. por HEATHER K. MARSHALL, JASON SPYROMILIO e TOMONORI USUDA. DOI: [10.1117/12.2631694](https://doi.org/10.1117/12.2631694). URL: <https://doi.org/10.1117/12.2631694>.

- FORBES, CATHERINE et al. (nov. de 2010a). “Beta Distribution”. Em: 4^a ed. Wiley, pp. 55–61. ISBN: 9780470390634. DOI: [10.1002/9780470627242.ch8](https://doi.org/10.1002/9780470627242.ch8). URL: <https://doi.org/10.1002/9780470627242.ch8>.
- (nov. de 2010b). “Binomial Distribution”. Em: 4^a ed. Wiley, pp. 62–65. ISBN: 9780470390634. DOI: [10.1002/9780470627242.ch9](https://doi.org/10.1002/9780470627242.ch9). URL: <https://doi.org/10.1002/9780470627242.ch9>.
- (nov. de 2010c). “Dirichlet Distribution”. Em: 4^a ed. Wiley, pp. 77–78. ISBN: 9780470390634. DOI: [10.1002/9780470627242.ch13](https://doi.org/10.1002/9780470627242.ch13). URL: <https://doi.org/10.1002/9780470627242.ch13>.
- (nov. de 2010d). “Gamma Distribution”. Em: 4^a ed. Wiley, pp. 109–113. ISBN: 9780470390634. DOI: [10.1002/9780470627242.ch22](https://doi.org/10.1002/9780470627242.ch22). URL: <https://doi.org/10.1002/9780470627242.ch22>.
- (nov. de 2010e). “Multinomial Distribution”. Em: 4^a ed. Wiley, pp. 135–136. ISBN: 9780470390634. DOI: [10.1002/9780470627242.ch30](https://doi.org/10.1002/9780470627242.ch30). URL: <https://doi.org/10.1002/9780470627242.ch30>.
- FUJIOKA, TOMOYUKI, LEONA KATSUTA, KAZUNORI KUBOTA et al. (jun. de 2020). “Classification of Breast Masses on Ultrasound Shear Wave Elastography using Convolutional Neural Networks”. Em: *Ultrasonic Imaging* 42.4-5, pp. 213–220. DOI: [10.1177/0161734620932609](https://doi.org/10.1177/0161734620932609). URL: <https://doi.org/10.1177/0161734620932609>.
- FUKUGITA, M. et al. (abr. de 1996). “The Sloan Digital Sky Survey Photometric System”. Em: *The Astronomical Journal* 111, p. 1748. DOI: [10.1086/117915](https://doi.org/10.1086/117915). URL: <https://doi.org/10.1086/117915>.
- GRAEFE, GOETZ (jan. de 1993). “Query Evaluation Techniques for Large Databases”. Em: *ACM Computing Surveys (CSUR)* 25 (2), pp. 73–169. ISSN: 15577341. DOI: [10.1145/152610.152611](https://doi.org/10.1145/152610.152611).
- GRAY, N. et al. (dez. de 2023). *Units in the VO Version 1.1*. IVOA Recommendation 15 December 2023. DOI: [10.5479/ADS/bib/2023ivoa.spec.1215G](https://doi.org/10.5479/ADS/bib/2023ivoa.spec.1215G). URL: <https://doi.org/10.5479/ADS/bib/2023ivoa.spec.1215G>.
- GUNN, J. E. et al. (dez. de 1998). “The Sloan Digital Sky Survey Photometric Camera”. Em: *The Astronomical Journal* 116.6, pp. 3040–3081. ISSN: 0004-6256. DOI: [10.1086/300645](https://doi.org/10.1086/300645). URL: <https://doi.org/10.1086/300645>.
- GUNN, JAMES E. et al. (abr. de 2006). “The 2.5 m Telescope of the Sloan Digital Sky Survey”. Em: *The Astronomical Journal* 131.4, pp. 2332–2359. ISSN: 1538-3881. DOI: [10.1086/500975](https://doi.org/10.1086/500975). URL: <https://doi.org/10.1086/500975>.

- HARRISON, P. A. e G. RIXON (out. de 2016). *Universal Worker Service Pattern Version 1.1*. IVOA Recommendation 24 October 2016. doi: 10.5479/ADS/bib/2016ivoa.spec.1024H. URL: <https://doi.org/10.5479/ADS/bib/2016ivoa.spec.1024H>.
- HART, ROSS E. et al. (out. de 2016). “Galaxy Zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias”. Em: *Monthly Notices of the Royal Astronomical Society* 461 (4), pp. 3663–3682. ISSN: 0035-8711. doi: 10.1093/MNRAS/STW1588. URL: <https://dx.doi.org/10.1093/mnras/stw1588>.
- HE, KAIMING et al. (dez. de 2015). “Deep Residual Learning for Image Recognition”. Em: *arXiv*. doi: 10.48550/arXiv.1512.03385. eprint: 1512.03385. URL: <https://doi.org/10.48550/arXiv.1512.03385>.
- (2016). “Deep Residual Learning for Image Recognition”. Em: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778. doi: 10.1109/CVPR.2016.90. URL: <https://doi.org/10.1109/CVPR.2016.90>.
- HINTON, GEOFFREY, NITISH SRIVASTAVA e KEVIN SWERSKY (2012). *On the Variance of the Adaptive Learning Rate and Beyond*.
- HOGG, DAVID W. et al. (out. de 2001). “A Photometricity and Extinction Monitor at the Apache Point Observatory”. Em: *The Astronomical Journal* 122.4, pp. 2129–2138. ISSN: 0004-6256. doi: 10.1086/323103. URL: <https://doi.org/10.1086/323103>.
- HUANG, GAO et al. (2018). *Densely Connected Convolutional Networks*. arXiv: 1608.06993 [cs.CV]. URL: <https://doi.org/10.48550/arXiv.1608.06993>.
- HUBBLE, E. P. (dez. de 1926). “Extragalactic nebulae.” Em: *Astrophysical Journal* 64, pp. 321–369. doi: 10.1086/143018.
- II, DAVID WILLIAM EDWARDS e IMREN DINC (nov. de 2020). “Classification of Protein Crystallization Images using EfficientNet with Data Augmentation”. Em: *CSBio '20: Proceedings of the Eleventh International Conference on Computational Systems-Biology and Bioinformatics*. ACM. doi: 10.1145/3429210.3429220. URL: <https://doi.org/10.1145/3429210.3429220>.
- IVEZIĆ, ŽELJKO et al. (mar. de 2019). “LSST: From Science Drivers to Reference Design and Anticipated Data Products”. Em: *Astrophys. J.* 873.2, p. 111. ISSN: 0004-637X. doi: 10.3847/1538-4357/ab042c. URL: <https://doi.org/10.3847/1538-4357/ab042c>.
- JARRETT, THOMAS (2004). “Large Scale Structure in the Local Universe – The 2MASS Galaxy Catalog”. Em: *Publications of the Astronomical Society of Australia* 21 (4),

pp. 396–403. ISSN: 1323-3580. DOI: [10.1071/AS04050](https://doi.org/10.1071/AS04050). URL: <https://doi.org/10.1071/AS04050>.

JELODAR, HAMED et al. (jun. de 2019). “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey”. Em: *Multimedia Tools and Applications* 78 (11), pp. 15169–15211. ISSN: 1380-7501. DOI: [10.1007/s11042-018-6894-4](https://doi.org/10.1007/s11042-018-6894-4). URL: <https://doi.org/10.1007/s11042-018-6894-4>.

JIANG, LIANGXIAO et al. (fev. de 2016). “Structure extended multinomial naive Bayes”. Em: *Information Sciences* 329, pp. 346–356. ISSN: 0020-0255. DOI: [10.1016/J.INS.2015.09.037](https://doi.org/10.1016/J.INS.2015.09.037).

KALCHEVA, NELI, GINKA MARINOVA e MAYA TODOROVA (out. de 2023). “Comparative Analysis of the Bernoulli and Multinomial Naive Bayes Classifiers for Text Classification in Machine Learning”. Em: *2023 International Conference Automatics and Informatics (ICAI)*. IEEE, pp. 28–31. ISBN: 979-8-3503-1291-1. DOI: [10.1109/ICAI58806.2023.10339077](https://doi.org/10.1109/ICAI58806.2023.10339077). URL: <https://ieeexplore.ieee.org/document/10339077/>.

KANG, DAE-KI et al. (2005). “Multinomial Event Model Based Abstraction for Sequence and Text Classification”. Em: vol. 3607 LNAI. Springer, Berlin, Heidelberg, pp. 134–148. ISBN: 978-3-540-31882-8. DOI: [10.1007/11527862_10](https://doi.org/10.1007/11527862_10). URL: [http://doi.org/10.1007/11527862_10](https://doi.org/10.1007/11527862_10).

KIBRIYA, ASHRAF M. et al. (2004). “Multinomial Naive Bayes for Text Categorization Revisited”. Em: *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)* 3339, pp. 488–499. ISSN: 1611-3349. DOI: [10.1007/978-3-540-30549-1_43](https://doi.org/10.1007/978-3-540-30549-1_43). URL: https://doi.org/10.1007/978-3-540-30549-1_43.

KINGMA, DIEDERIK P. e JIMMY BA (2017). *Adam: A Method for Stochastic Optimization*. arXiv: [1412.6980 \[cs.LG\]](https://arxiv.org/abs/1412.6980). URL: <https://doi.org/10.48550/arXiv.1412.6980>.

KOPOSOV, S e O BARTUNOV (jul. de 2006). “Q3C, Quad Tree Cube – The new Sky-indexing Concept for Huge Astronomical Catalogues and its Realization for Main Astronomical Queries (Cone Search and Xmatch) in Open Source Database PostgreSQL”. Em: *Astronomical Data Analysis Software and Systems XV*. Ed. por C GABRIEL et al. Vol. 351, p. 735. URL: <https://ui.adsabs.harvard.edu/abs/2006ASPC..351..735K>.

KRIEGSMANN, MARK, CHRISTIAN HAAG, CLEO-ARON WEIS et al. (jun. de 2020). “Deep Learning for the Classification of Small-Cell and Non-Small-Cell Lung Cancer”. Em: *Cancers* 12.6, p. 1604. DOI: [10.3390/cancers12061604](https://doi.org/10.3390/cancers12061604). URL: <https://doi.org/10.3390/cancers12061604>.

KRIZHEVSKY, ALEX, ILYA SUTSKEVER e GEOFFREY E HINTON (2012). “ImageNet Classification with Deep Convolutional Neural Networks”. Em: *Advances in Neural Information*

Processing Systems. Ed. por F. PEREIRA et al. Vol. 25. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

LEMMENS, ROB et al. (2021). “Citizen Science in the Digital World of Apps”. Em: ed. por KATRIN VOHLAND et al. Springer International Publishing, pp. 461–474. ISBN: 978-3-030-58278-4. DOI: [10.1007/978-3-030-58278-4_23](https://doi.org/10.1007/978-3-030-58278-4_23). URL: https://doi.org/10.1007/978-3-030-58278-4_23.

LI, LISHA et al. (mar. de 2016). “Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization”. Em: *arXiv*. DOI: [10.48550/arXiv.1603.06560](https://doi.org/10.48550/arXiv.1603.06560). eprint: [1603.06560](https://doi.org/10.48550/arXiv.1603.06560). URL: <https://doi.org/10.48550/arXiv.1603.06560>.

LI, XIA et al. (2020). “Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet)”. Em: *PLoS ONE* 15. DOI: [10.1371/journal.pone.0232127](https://doi.org/10.1371/journal.pone.0232127). URL: <https://doi.org/10.1371/journal.pone.0232127>.

LI, YALIANG et al. (ago. de 2021). “AutoML: A Perspective where Industry Meets Academy”. Em: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 4048–4049. DOI: [10.1145/3447548.3470827](https://doi.org/10.1145/3447548.3470827). URL: <https://doi.org/10.1145/3447548.3470827>.

LI, ZHONG et al. (2020). “Protein Contact Map Prediction Based on ResNet and DenseNet”. Em: *BioMed Research International* 2020. DOI: [10.1155/2020/7584968](https://doi.org/10.1155/2020/7584968). URL: <https://doi.org/10.1155/2020/7584968>.

LIN, TSUNG-YI et al. (mai. de 2014). “Microsoft COCO: Common Objects in Context”. Em: *arXiv*. DOI: [10.48550/arXiv.1405.0312](https://doi.org/10.48550/arXiv.1405.0312). eprint: [1405.0312](https://doi.org/10.48550/arXiv.1405.0312). URL: <https://doi.org/10.48550/arXiv.1405.0312>.

LINTOTT, CHRIS et al. (jan. de 2011). “Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies”. Em: *Monthly Notices of the Royal Astronomical Society* 410 (1), pp. 166–178. ISSN: 0035-8711. DOI: [10.1111/j.1365-2966.2010.17432.X](https://doi.org/10.1111/j.1365-2966.2010.17432.X). URL: <https://dx.doi.org/10.1111/j.1365-2966.2010.17432.x>.

LINTOTT, CHRIS J. et al. (set. de 2008). “Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey”. Em: *Monthly Notices of the Royal Astronomical Society* 389 (3), pp. 1179–1189. ISSN: 13652966. DOI: [10.1111/j.1365-2966.2008.13689.X/2/MNRAS0389-1179-F12.JPG](https://doi.org/10.1111/j.1365-2966.2008.13689.X/2/MNRAS0389-1179-F12.JPG). URL: <https://dx.doi.org/10.1111/j.1365-2966.2008.13689.x>.

- LIU, LIYUAN et al. (2020). *On the Variance of the Adaptive Learning Rate and Beyond*. arXiv: 1908.03265 [cs.LG]. URL: <https://doi.org/10.48550/arXiv.1908.03265>.
- LIU, ZE et al. (mar. de 2021). “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows”. Em: *arXiv*. doi: [10.48550/arXiv.2103.14030](https://doi.org/10.48550/arXiv.2103.14030). eprint: 2103.14030. URL: <https://doi.org/10.48550/arXiv.2103.14030>.
- LUO, JIAHUA, CHI-MAN WONG e CHI-MAN VONG (jan. de 2021). “Multinomial Bayesian extreme learning machine for sparse and accurate classification model”. Em: *Neurocomputing* 423, pp. 24–33. ISSN: 09252312. doi: [10.1016/j.neucom.2020.09.061](https://doi.org/10.1016/j.neucom.2020.09.061). URL: <https://doi.org/10.1016/j.neucom.2020.09.061>.
- MAHDIANPARI, MASOUD et al. (jul. de 2018). “Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery”. Em: *Remote Sensing* 10.7, p. 1119. doi: [10.3390/rs10071119](https://doi.org/10.3390/rs10071119). URL: <https://doi.org/10.3390/rs10071119>.
- MALKOV, YU A. e D. A. YASHUNIN (abr. de 2020). “Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs”. Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (4), pp. 824–836. ISSN: 19393539. doi: [10.1109/TPAMI.2018.2889473](https://doi.org/10.1109/TPAMI.2018.2889473). URL: <https://doi.org/10.1109/TPAMI.2018.2889473>.
- MANTELET, GRÉGORY et al. (dez. de 2023). *Astronomical Data Query Language Version 2.1*. IVOA Recommendation 15 December 2023. doi: [10.5479/ADS/bib/2023ivoa.spec.1215M](https://doi.org/10.5479/ADS/bib/2023ivoa.spec.1215M). URL: <https://doi.org/10.5479/ADS/bib/2023ivoa.spec.1215M>.
- McCULLOCH, WARREN S. e WALTER PITTS (dez. de 1943). “A logical calculus of the ideas immanent in nervous activity”. Em: *Bull. Math. Biophys.* 5.4, pp. 115–133. ISSN: 1522-9602. doi: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259). URL: <https://doi.org/10.1007/BF02478259>.
- McREYNOLDS, TOM e DAVID BLYTHE (2005). “Antialiasing”. Em: *Advanced Graphics Programming Using OpenGL*. Elsevier, pp. 169–184. doi: [10.1016/b978-155860659-3.50012-3](https://doi.org/10.1016/b978-155860659-3.50012-3). URL: <https://doi.org/10.1016/b978-155860659-3.50012-3>.
- MEALY, GEORGE H. (set. de 1955). “A method for synthesizing sequential circuits”. Em: *The Bell System Technical Journal* 34 (5), pp. 1045–1079. ISSN: 0005-8580. doi: [10.1002/j.1538-7305.1955.tb03788.x](https://doi.org/10.1002/j.1538-7305.1955.tb03788.x). URL: <https://doi.org/10.1002/j.1538-7305.1955.tb03788.x>.
- MENDES DE OLIVEIRA, C et al. (ago. de 2019). “The Southern Photometric Local Universe Survey (S-PLUS): improved SEDs, morphologies, and redshifts with 12 optical filters”.

Em: *Monthly Notices of the Royal Astronomical Society* 489.1, pp. 241–267. ISSN: 0035-8711. doi: [10.1093/mnras/stz1985](https://doi.org/10.1093/mnras/stz1985). URL: <https://doi.org/10.1093/mnras/stz1985>.

MICHEL, LAURENT et al. (jun. de 2023). *Model Instances in VOTables Version 1.0*. IVOA Recommendation 20 June 2023. doi: [10.5479/ADS/bib/2023ivoa.spec.0620M](https://doi.org/10.5479/ADS/bib/2023ivoa.spec.0620M). URL: <https://doi.org/10.5479/ADS/bib/2023ivoa.spec.0620M>.

MICKAELIAN, AREG M. (mar. de 2016). “Astronomical Surveys and Big Data”. Em: *Open Astronomy* 25 (1), pp. 75–88. doi: [10.1515/ASTRO-2017-0113](https://doi.org/10.1515/ASTRO-2017-0113). MACHINEREADABLECITATION/RIS. URL: <https://doi.org/10.1515/astro-2017-0113.html>.

MITTAL, ANSH et al. (dez. de 2019). “Data augmentation based morphological classification of galaxies using deep convolutional neural network”. Em: *Earth Science Informatics* 13.3, pp. 601–617. doi: [10.1007/s12145-019-00434-8](https://doi.org/10.1007/s12145-019-00434-8). URL: <https://doi.org/10.1007/s12145-019-00434-8>.

NEVASALMI, LAURI (nov. de 2020). “Forecasting multinomial stock returns using machine learning methods”. Em: *The Journal of Finance and Data Science* 6, pp. 86–106. ISSN: 24059188. doi: [10.1016/j.jfds.2020.09.001](https://doi.org/10.1016/j.jfds.2020.09.001). URL: <https://doi.org/10.1016/j.jfds.2020.09.001>.

NONAKA, NAOKI e JUN SEITA (dez. de 2020). “Electrocardiogram Classification by Modified EfficientNet with Data Augmentation”. Em: *2020 Computing in Cardiology Conference (CinC)*. Computing in Cardiology. doi: [10.22489/cinc.2020.063](https://doi.org/10.22489/cinc.2020.063). URL: <https://doi.org/10.22489/cinc.2020.063>.

OCHSENBEIN, FRANCOIS et al. (out. de 2019). *VOTable Format Definition Version 1.4*. IVOA Recommendation 21 October 2019. doi: [10.5479/ADS/bib/2019ivoa.spec.1021O](https://doi.org/10.5479/ADS/bib/2019ivoa.spec.1021O). URL: <https://doi.org/10.5479/ADS/bib/2019ivoa.spec.1021O>.

ODEH, AYMAN HUSSIEN, MUNTHER ODEH e NADA ODEH (2022). “Using Multinomial Naive Bayes Machine Learning Method to Classify, Detect, and Recognize Programming Language Source Code”. Em: *Proceedings - 2022 23rd International Arab Conference on Information Technology, ACIT 2022*. doi: [10.1109/ACIT57182.2022.9994117](https://doi.org/10.1109/ACIT57182.2022.9994117). URL: <https://doi.org/10.1109/ACIT57182.2022.9994117>.

OLIVEIRA TORRES, GABRIEL de, MARCELO XAVIER GUTERRES e VICTOR RAFAEL REZENDE CELESTINO (abr. de 2023). “Legal actions in Brazilian air transport: A machine learning and multinomial logistic regression analysis”. Em: *Frontiers in Future Transportation* 4, p. 1070533. ISSN: 2673-5210. doi: [10.3389/ffutr.2023.1070533](https://doi.org/10.3389/ffutr.2023.1070533). URL: <https://doi.org/10.3389/ffutr.2023.1070533>.

- OZAKI, YOSHIHIKO et al. (jun. de 2020). “Multiobjective tree-structured parzen estimator for computationally expensive optimization problems”. Em: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. ACM, pp. 533–541. ISBN: 9781450371285. DOI: [10.1145/3377930.3389817](https://doi.org/10.1145/3377930.3389817). URL: <https://doi.org/10.1145/3377930.3389817>.
- PAN, JAMES JIE, JIANGUO WANG e GUOLIANG LI (set. de 2024). “Survey of vector database management systems”. Em: *VLDB Journal* 33 (5), pp. 1591–1615. ISSN: 0949877X. DOI: [10.1007/S00778-024-00864-X/FIGURES/10](https://doi.org/10.1007/S00778-024-00864-X/FIGURES/10). URL: <https://doi.org/10.1007/s00778-024-00864-x>.
- PLANTE, RAYMOND et al. (fev. de 2008). *Simple Cone Search Version 1.03*. IVOA Recommendation 22 February 2008. DOI: [10.5479/ADS/bib/2008ivoa.specQ0222P](https://doi.org/10.5479/ADS/bib/2008ivoa.specQ0222P). arXiv: 1110.0498 [astro-ph.IM]. URL: <https://doi.org/10.5479/ADS/bib/2008ivoa.specQ0222P>.
- QUINN, PETER J. et al. (set. de 2004). “The International Virtual Observatory Alliance: recent technical developments and the road ahead”. Em: *Optimizing Scientific Return for Astronomy through Information Technologies*. Ed. por PETER J. QUINN e ALAN BRIDGER. Vol. 5493. SPIE, pp. 137–145. DOI: [10.1117/12.551247](https://doi.org/10.1117/12.551247). URL: <https://doi.org/10.1117/12.551247>.
- REZENDE, EDMAR et al. (2018). “Malicious Software Classification Using VGG16 Deep Neural Network’s Bottleneck Features”. Em: *Advances in Intelligent Systems and Computing*. Springer International Publishing, pp. 51–59. DOI: [10.1007/978-3-319-77028-4_9](https://doi.org/10.1007/978-3-319-77028-4_9). URL: https://doi.org/10.1007/978-3-319-77028-4_9.
- ROSENBLATT, FRANK (1960). “Perceptron Simulation Experiments”. Em: *Proceedings of the IRE* 48.3, pp. 301–309. DOI: [10.1109/JRPROC.1960.287598](https://doi.org/10.1109/JRPROC.1960.287598). URL: <https://doi.org/10.1109/JRPROC.1960.287598>.
- RUMELHART, DAVID E., GEOFFREY E. HINTON e RONALD J. WILLIAMS (out. de 1986). “Learning representations by back-propagating errors”. Em: *Nature* 323, pp. 533–536. ISSN: 1476-4687. DOI: [10.1038/323533a0](https://doi.org/10.1038/323533a0). URL: <https://doi.org/10.1038/323533a0>.
- RUSSAKOVSKY, OLGA et al. (2015). “ImageNet Large Scale Visual Recognition Challenge”. Em: *International Journal of Computer Vision (IJCV)* 115.3, pp. 211–252. DOI: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y). URL: <https://doi.org/10.1007/s11263-015-0816-y>.
- SARAVANAN, U. K.BALAJI et al. (2023). “Multinomial Naive Bayes Based Machine Learning Analysis of Twitter Sentiment”. Em: *Proceedings of the 2nd International Conference on Edge Computing and Applications, ICECAA 2023*, pp. 429–434. DOI: [10.1109/ICECAA55507.2023.9750003](https://doi.org/10.1109/ICECAA55507.2023.9750003).

ICECAA58104.2023.10212150. URL: <https://doi.org/10.1109/ICECAA58104.2023.10212150>.

SCHNEIDER, PETER (out. de 2014). *Extragalactic astronomy and cosmology*. 2^a ed. Springer. ISBN: 978-3-642-54083-7. DOI: [10.1007/978-3-642-54083-7](https://doi.org/10.1007/978-3-642-54083-7). URL: <https://doi.org/10.1007/978-3-642-54083-7>.

SEBER, GEORGE A. F. (2015). “Multinomial Distribution”. Em: Springer International Publishing, pp. 181–188. ISBN: 978-3-319-21930-1. DOI: [10.1007/978-3-319-21930-1_12](https://doi.org/10.1007/978-3-319-21930-1_12). URL: https://doi.org/10.1007/978-3-319-21930-1_12.

SILVERTOWN, JONATHAN (set. de 2009). “A new dawn for citizen science.” Em: *Trends in ecology & evolution* 24.9, pp. 467–71. ISSN: 0169-5347. DOI: [10.1016/j.tree.2009.03.017](https://doi.org/10.1016/j.tree.2009.03.017). URL: <https://doi.org/10.1016/j.tree.2009.03.017>.

SIMARD, PATRICE Y., DAVE STEINKRAUS e JOHN C. PLATT (2003). “Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis”. Em: *Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2*. ICDAR ’03. USA: IEEE Computer Society, p. 958. ISBN: 0769519601. DOI: [10.5555/938980.939477](https://doi.org/10.5555/938980.939477). URL: <https://doi.org/10.5555/938980.939477>.

SIMMONS, B. D. et al. (fev. de 2017). “Galaxy Zoo: Quantitative visual morphological classifications for 48 000 galaxies from CANDELS”. Em: *Monthly Notices of the Royal Astronomical Society* 464 (4), pp. 4420–4447. ISSN: 13652966. DOI: [10.1093/mnras/stw2587](https://doi.org/10.1093/mnras/stw2587). URL: <https://doi.org/10.1093/mnras/stw2587>.

SIMONYAN, KAREN e ANDREW ZISSERMAN (set. de 2014). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Em: *arXiv*. DOI: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556). eprint: [1409.1556](https://doi.org/10.48550/arXiv.1409.1556). URL: <https://doi.org/10.48550/arXiv.1409.1556>.

— (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv: [1409.1556 \[cs.CV\]](https://doi.org/10.48550/arXiv.1409.1556). URL: <https://doi.org/10.48550/arXiv.1409.1556>.

SMEE, STEPHEN A. et al. (jul. de 2013). “THE MULTI-OBJECT, FIBER-FED SPECTROGRAPHS FOR THE SLOAN DIGITAL SKY SURVEY AND THE BARYON OSCILLATION SPECTROSCOPIC SURVEY”. Em: *The Astronomical Journal* 146.2, p. 32. DOI: [10.1088/0004-6256/146/2/32](https://doi.org/10.1088/0004-6256/146/2/32). URL: <https://doi.org/10.1088/0004-6256/146/2/32>.

SOLANO-ROJAS, BRAULIO, RICARDO VILLALÓN-FONSECA e GABRIELA MARÍN-RAVENTÓS (2020). “Alzheimer’s Disease Early Detection Using a Low Cost Three-Dimensional Densenet-121 Architecture”. Em: *Lecture Notes in Computer Science*. Springer Inter-

- national Publishing, pp. 3–15. DOI: [10.1007/978-3-030-51517-1_1](https://doi.org/10.1007/978-3-030-51517-1_1). URL: https://doi.org/10.1007/978-3-030-51517-1_1.
- SRIVASTAVA, NITISH et al. (2014). “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. Em: *Journal of Machine Learning Research* 15.56, pp. 1929–1958. URL: <http://jmlr.org/papers/v15/srivastava14a.html>.
- STEINMETZ, MATTHIAS e JULIO F NAVARRO (2002). “The hierarchical origin of galaxy morphologies”. Em: *New Astronomy* 7.4, pp. 155–160.
- SZALAY, ALEXANDER S et al. (2002). “The SDSS skyserver: public access to the sloan digital sky server data”. Em: *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data*. Association for Computing Machinery, pp. 570–581. ISBN: 1581134975. DOI: [10.1145/564691.564758](https://doi.org/10.1145/564691.564758). URL: <https://doi.org/10.1145/564691.564758>.
- SZALAY, ALEXANDER S. et al. (mai. de 2000). “Designing and mining multi-terabyte astronomy archives”. Em: *ACM SIGMOD Record* 29 (2), pp. 451–462. ISSN: 01635808. DOI: [10.1145/335191.335439](https://doi.org/10.1145/335191.335439). URL: <https://dl.acm.org/doi/10.1145/335191.335439>.
- SZEGEDY, CHRISTIAN, SERGEY IOFFE et al. (2016). *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. arXiv: [1602.07261 \[cs.CV\]](https://arxiv.org/abs/1602.07261). URL: <https://doi.org/10.48550/arXiv.1602.07261>.
- SZEGEDY, CHRISTIAN, WEI LIU et al. (set. de 2014). “Going Deeper with Convolutions”. Em: *arXiv*. DOI: [10.48550/arXiv.1409.4842](https://doi.org/10.48550/arXiv.1409.4842). eprint: [1409.4842](https://arxiv.org/abs/1409.4842). URL: <https://doi.org/10.48550/arXiv.1409.4842>.
- (2015). “Going deeper with convolutions”. Em: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9. DOI: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594). URL: <https://doi.org/10.1109/CVPR.2015.7298594>.
- TAGHIZADEH-POPP, M. et al. (out. de 2020). “SciServer: A science platform for astronomy and beyond”. Em: *Astronomy and Computing* 33. ISSN: 22131337. DOI: [10.1016/J.JASCOM.2020.100412](https://doi.org/10.1016/J.JASCOM.2020.100412).
- TAIPALUS, TONI (jun. de 2024). “Vector database management systems: Fundamental concepts, use-cases, and current challenges”. Em: *Cognitive Systems Research* 85, p. 101216. ISSN: 1389-0417. DOI: [10.1016/J.COGSYS.2024.101216](https://doi.org/10.1016/J.COGSYS.2024.101216). URL: <https://doi.org/10.1016/J.COGSYS.2024.101216>.
- TAN, MINGXING e QUOC V. LE (mai. de 2019). “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”. Em: *arXiv*. DOI: [10.48550/arXiv.1905.11946](https://doi.org/10.48550/arXiv.1905.11946). eprint: [1905.11946](https://arxiv.org/abs/1905.11946). URL: <https://doi.org/10.48550/arXiv.1905.11946>.

- TAN, MINGXING e QUOC V. LE (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. arXiv: 1905.11946 [cs.LG]. URL: <https://doi.org/10.48550/arXiv.1905.11946>.
- TODY, DOUG et al. (fev. de 2012). *Simple Spectral Access Protocol Version 1.1*. IVOA Recommendation 10 February 2012. doi: 10.5479/ADS/bib/2012ivoa.spec.0210T. arXiv: 1203.5725 [astro-ph.IM]. URL: <https://doi.org/10.5479/ADS/bib/2012ivoa.spec.0210T>.
- TU, ZHENGZHONG et al. (abr. de 2022). “MaxViT: Multi-Axis Vision Transformer”. Em: *arXiv*. doi: 10.48550/arXiv.2204.01697. eprint: 2204.01697. URL: <https://doi.org/10.48550/arXiv.2204.01697>.
- VAN DEN BERGH, SIDNEY (1998). *Galaxy morphology and classification*. Cambridge University Press.
- VASU, PAVAN KUMAR ANASOSALU et al. (mar. de 2023). “FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization”. Em: *arXiv*. doi: 10.48550/arXiv.2303.14189. eprint: 2303.14189. URL: <https://doi.org/10.48550/arXiv.2303.14189>.
- VASWANI, ASHISH et al. (jun. de 2017). “Attention Is All You Need”. Em: *arXiv*. doi: 10.48550/arXiv.1706.03762. eprint: 1706.03762. URL: <https://doi.org/10.48550/arXiv.1706.03762>.
- VOHLAND, KATRIN et al. (2021). “Editorial: The Science of Citizen Science Evolves”. Em: ed. por KATRIN VOHLAND et al. Springer International Publishing, pp. 1–12. ISBN: 978-3-030-58278-4. doi: 10.1007/978-3-030-58278-4_1. URL: https://doi.org/10.1007/978-3-030-58278-4_1.
- WALMSLEY, MIKE, TOBIAS GÉRON et al. (out. de 2023). “Galaxy Zoo DESI: Detailed morphology measurements for 8.7M galaxies in the DESI Legacy Imaging Surveys”. Em: *Monthly Notices of the Royal Astronomical Society* 526 (3), pp. 4768–4786. ISSN: 0035-8711. doi: 10.1093/MNRAS/STAD2919. URL: <https://dx.doi.org/10.1093/mnras/stad2919>.
- WALMSLEY, MIKE, CHRIS LINTOTT, TOBIAS GERON et al. (jan. de 2022). “Galaxy Zoo DECaLS: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies”. Em: *Mon. Not. R. Astron. Soc.* 509.3, pp. 3966–3988. ISSN: 0035-8711. doi: 10.1093/mnras/stab2093. URL: <https://doi.org/10.1093/mnras/stab2093>.
- WALMSLEY, MIKE, CHRIS LINTOTT, TOBIAS GÉRON et al. (dez. de 2021). “Galaxy Zoo DECaLS: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies”. Em: *Monthly Notices of the Royal Astronomical Society* 509

(3), pp. 3966–3988. ISSN: 0035-8711. DOI: [10.1093/MNRAS/STAB2093](https://doi.org/10.1093/mnras/stab2093). URL: <https://dx.doi.org/10.1093/mnras/stab2093>.

WANG, JIACUN (jun. de 2019). *Formal methods in computer science*. Chapman & Hall/CRC. ISBN: 978-1-4987-7532-8.

WANG, LENGYANG e MINGKE ZHANG (mar. de 2025). “Statistical modeling of Dengue transmission dynamics with environmental factors”. Em: *Computational Statistics & Data Analysis* 203, p. 108080. ISSN: 01679473. DOI: [10.1016/j.csda.2024.108080](https://doi.org/10.1016/j.csda.2024.108080). URL: <https://doi.org/10.1016/j.csda.2024.108080>.

WANG, SHUI-HUA e YU-DONG ZHANG (jun. de 2020). “DenseNet-201-Based Deep Neural Network with Composite Learning Factor and Precomputation for Multiple Sclerosis Classification”. Em: *ACM Trans. Multimedia Comput. Commun. Appl.* 16.2s. ISSN: 1551-6857. DOI: [10.1145/3341095](https://doi.org/10.1145/3341095). URL: <https://doi.org/10.1145/3341095>.

WELLS, D. C. e E. W. GREISEN (jan. de 1979). “FITS - a Flexible Image Transport System”. Em: *Image Processing in Astronomy*. Ed. por G. SEDMAK, M. CAPACCIOLI e R. J. ALLEN, p. 445. URL: <https://doi.org/10.5479/ADS/bib/1979ipia.coll..445W>.

WILLETT, KYLE W., MELANIE A. GALLOWAY et al. (fev. de 2017). “Galaxy Zoo: morphological classifications for 120 000 galaxies in HST legacy imaging”. Em: *Monthly Notices of the Royal Astronomical Society* 464 (4), pp. 4176–4203. ISSN: 0035-8711. DOI: [10.1093/MNRAS/STW2568](https://doi.org/10.1093/MNRAS/STW2568). URL: <https://dx.doi.org/10.1093/mnras/stw2568>.

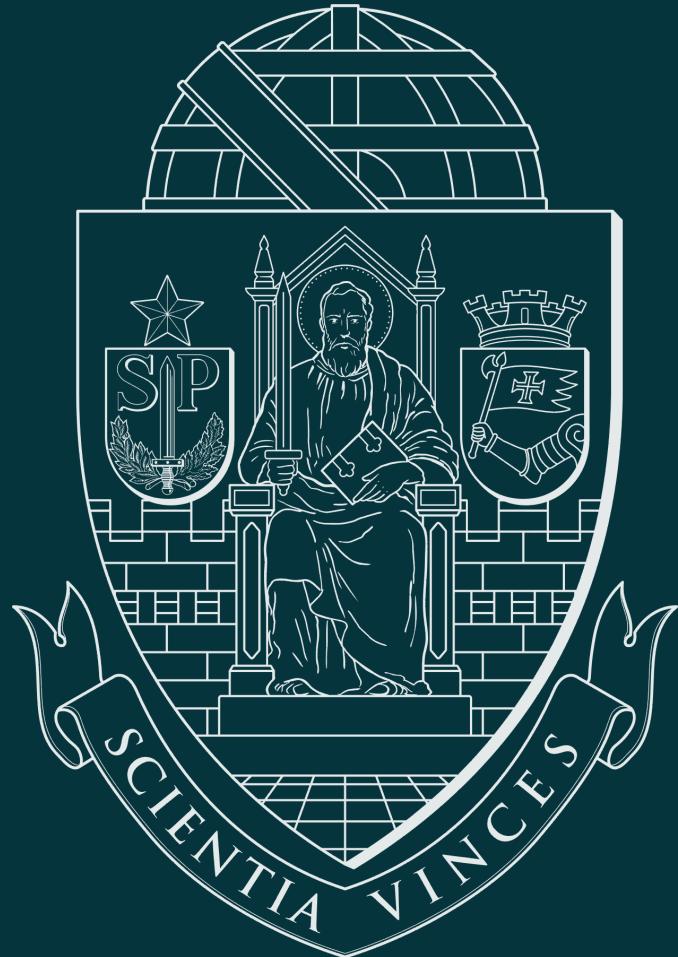
WILLETT, KYLE W., CHRIS J. LINTOTT et al. (nov. de 2013). “Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey”. Em: *Monthly Notices of the Royal Astronomical Society* 435 (4), pp. 2835–2860. ISSN: 0035-8711. DOI: [10.1093/MNRAS/STT1458](https://doi.org/10.1093/MNRAS/STT1458). URL: <https://dx.doi.org/10.1093/mnras/stt1458>.

YAEGER, LARRY, RICHARD LYON e BRANDYN WEBB (1996). “Effective Training of a Neural Network Character Classifier for Word Recognition”. Em: *Proceedings of the 9th International Conference on Neural Information Processing Systems*. NIPS’96. Denver, Colorado: MIT Press, pp. 807–813.

YORK, DONALD G. et al. (set. de 2000). “The Sloan Digital Sky Survey: Technical summary”. English (US). Em: *Astronomical Journal* 120.3, pp. 1579–1587. ISSN: 0004-6256. DOI: [10.1086/301513](https://doi.org/10.1086/301513).

ZHANG, PAN, LING YANG e DAOLIANG LI (set. de 2020). “EfficientNet-B4-Ranger: A novel method for greenhouse cucumber disease recognition under natural complex envi-

ronment”. Em: *Computers and Electronics in Agriculture* 176, p. 105652. doi: [10.1016/j.compag.2020.105652](https://doi.org/10.1016/j.compag.2020.105652). URL: <https://doi.org/10.1016/j.compag.2020.105652>.



Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia