

# Exploração de Redes Neurais para Classificação de imagens de Galáxias

Deyvisson N. Garcês<sup>1</sup>, Eduardo C. Martins<sup>1</sup>, Inayê Melo<sup>1</sup>,  
Nicoly R. da Costa<sup>1</sup>, Thamara Medeiros<sup>1</sup>

<sup>1</sup>Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP)

dgarces@usp.br, educherobin@gmail.com, inayecaroline@gmail.com,  
nicolyrodrigues1111@gmail.com, thamara.medeiros9@gmail.com.

**Abstract.** *This work investigates the use of different neural network architectures for morphological galaxy classification using the Galaxy10 DECaLS dataset. We evaluated a classical CNN, a pre-trained ResNet-50, and an implementation of the semi-supervised BYOL method. Model performance was compared using accuracy, recall, precision, and F1-score. The results, analyzed both individually and comparatively, indicate that all architectures achieve good performance on the morphological classification task for the chosen dataset.*

**Resumo.** *Este trabalho investiga o uso de diferentes arquiteturas de redes neurais para a classificação morfológica de galáxias no dataset Galaxy10 DECaLS. Foram testadas uma CNN clássica, a ResNet-50 pré-treinada e uma implementação do método BYOL semi-supervisionado. Para comparar o desempenho dos modelos, utilizamos métricas como acurácia, recall, precisão e F1-score. Os resultados foram analisados tanto individualmente quanto de forma comparativa, permitindo concluir que as arquiteturas estudadas apresentam bom desempenho na tarefa de classificação morfológica no conjunto de dados escolhido.*

## 1. Introdução

As galáxias são as unidades fundamentais que compõem o Universo, consistindo principalmente de estrelas, remanescentes estelares, gás interestelar, poeira e matéria escura [Wen et al. 2014]. Em particular, os estudos sobre a morfologia e a estrutura das galáxias são cruciais para a compreensão de sua evolução, uma vez que essas características estão intimamente relacionadas à história evolutiva das galáxias e constituem também um dos principais meios de investigação de seus parâmetros físicos [Cao et al. 2024]. A utilidade de um sistema de classificação dessas galáxias reside em sua capacidade de estabelecer ordem e simplificação [Sandage 2005].

Apesar de as pesquisas modernas revelarem detalhes morfológicos refinados, elas geram imagens de um volume de galáxias muito superior à capacidade de classificação visual manual dos cientistas. Para desimpedir essa complexa interação entre a morfologia e a evolução das galáxias, são necessárias medições morfológicas detalhadas em grandes amostras, [Slijepcevic et al. 2024]. A classificação morfológica de galáxias permanece uma das tarefas mais complexas da astrofísica. Essa dificuldade decorre tanto dos desafios inerentes à interpretação visual das estruturas galácticas quanto da presença de

ruídos e borrões nas imagens, que intensificam ainda mais a complexidade do processo classificatório [Le et al. 2025].

Para superar esse gargalo, o projeto Galaxy Zoo utiliza a ciência cidadã, solicitando a membros do público que forneçam classificações morfológicas através de uma interface web. Este esforço colaborativo permitiu a obtenção de medições visuais detalhadas para centenas de milhares de galáxias [Slijepcevic et al. 2024]. As imagens de galáxias usadas neste trabalho são geradas a partir de dados coletados por pesquisas em grande escala, como a DECals (Dark Energy Camera Legacy Survey). A DECals, por exemplo, utiliza a Câmera de Energia Escura (DECam) no telescópio Blanco de 4 metros no Observatório Interamericano de Cerro Tololo, no Chile [Slijepcevic et al. 2024].

As classificações de morfologia de galáxias baseadas em aprendizado de máquina constituem uma abordagem amplamente utilizada para lidar com os grandes volumes de dados provenientes de observações astronômicas [Yao et al. 2024]. Diante da complexidade inerente à classificação morfológica e da crescente disponibilidade de dados, este trabalho propõe a aplicação de modelos de aprendizado de máquina para a classificação automática de imagens de galáxias. Ressalta-se que as imagens empregadas no treinamento e na validação dos modelos passaram previamente por um processo de pré-classificação baseado em ciência cidadã, o que fornece rótulos confiáveis que servem como referência para o aprendizado supervisionado.

Enquanto a classificação por voluntários é fundamental para fornecer os rótulos de “verdade” iniciais, é humanamente inviável (e demorado) classificar manualmente todos os milhões de objetos observados. O aprendizado de máquina permite que, após ser treinado com o subconjunto classificado por humanos, o modelo possa extrapolar o conhecimento e classificar o restante em uma fração do tempo. A classificação rápida permite aos cientistas testar hipóteses sobre a evolução das galáxias em escalas cosmológicas de forma muito mais ágil, acelerando o ritmo da descoberta científica.

## **2. Objetivos**

O objetivo deste trabalho é explorar a aplicação de diferentes arquiteturas de redes neurais para classificação de imagens e analisar comparativamente seu desempenho. Especificamente, trabalhamos com as imagens do dataset Galaxy10 DECaLS, composto por imagens de galáxias distribuídas em 10 classes distintas, compondo um problema de classificação multiclasse. Exploramos três arquiteturas distintas para essa tarefa: uma rede convolucional (CNN) tradicional, o modelo ResNet-50 e a arquitetura auto-supervisionada BYOL. Os modelos são comparados de forma quantitativa a partir de métricas como acurácia, precisão, recall e F1-score, considerando tanto o desempenho global quanto o comportamento por classe. Também é analisado o impacto das etapas de pré-processamento e das técnicas de data augmentation no processo de treinamento e na capacidade de generalização dos modelos. Por fim, discute-se o custo computacional e a estabilidade do treinamento.

## **3. Trabalhos Relacionados**

Os avanços em astronomia tem cada vez mais permitido a exploração do universo, com uma crescente capacidade de capturar imagens e corpos celestes cada vez mais distantes e com melhores resoluções. Projetos como o [Galaxy Zoo 2007] se dedicam a

analisar e classificar tais imagens, contato com dedicação de voluntários e métodos estatísticos para agrupar galáxias.

Outras bases de dados dedicadas ao problema de classificação automática de galáxias, estrelas e outros corpos celestes disponíveis. Em 2024, o The Multimodal Universe Collaboration (2024) reuniu quase 100 TB de dados científicos sobre corpos celestes, de várias fontes e formatos. Nessa colaboração estão disponíveis bases de dados de diversos tamanhos, temas, formatos e precisão. Muitas dessas bases, como a Galaxy Zoo DECaLS e a Galaxy10 DECaLS, são dedicadas a classificação de galáxias. Em geral, essa classificação depende de características morfológicas visíveis em imagens de telescópio e câmera escura, como a presença ou ausência de braços espirais, imagem de fusão de galáxias ou barras escuras no centro.

O emprego de redes neurais para classificação de imagens de galáxias se mostra como um caminho potente para estudo desses fenômenos. Esse tipo de tarefa de aprendizado de máquina já foi explorado em trabalhos como em [de Castro 2025], com a aplicação de diferentes estruturas de aprendizado supervisionado para classificação de Galáxias. O trabalho explora diferentes técnicas de Data Augmentation e processamento para comparar as arquiteturas selecionadas na base Galaxy10 DECaLS. Em [Le et al. 2025], são exploradas três arquiteturas para classificação de galáxias no dataset Galaxy10 DECaLS: uma CNN traducional customizada, a Resnet e a Efficient Net. Esses trabalhos adicionam camadas importantes na identificação de galáxias com redes neurais, mas não exploram a classificação com arquiteturas de aprendizagem contrastivas ou outras arquiteturas auto supervisionadas.

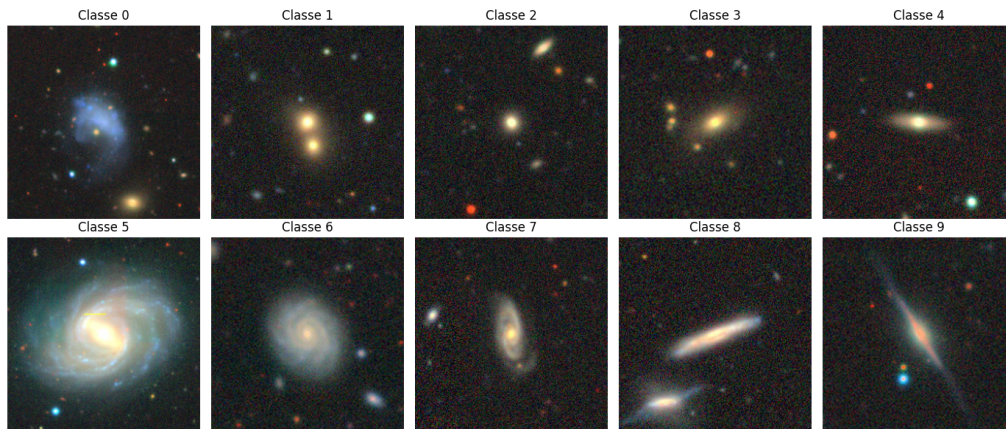
Em [Mohale and Lochner 2024] e [Slijepcevic et al. 2024], algoritmos auto supervisionados como o Bootstrap Your Own Latent [Grill et al. 2020] são aplicados em bases de dados astronômicas como Radio Galaxy Zoo e Galaxy10 DECaLS para classificação, atingindo bons resultados, comparáveis aos melhores modelos de aprendizado supervisionado para imagens, porém não comparando diretamente com redes mais simples como CNNs tradicionais customizadas.

Neste trabalho, buscamos explorar simultaneamente modelos de aprendizado supervisionados e não supervisionados, comparando os resultados entre eles e com a literatura, para compreender a aplicação desses métodos em observações e classificações astronômicas.

## **4. Materiais e Métodos**

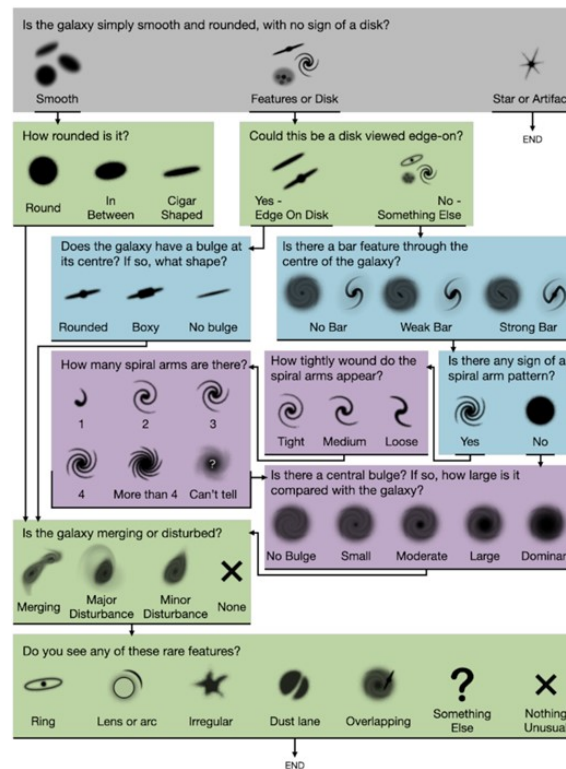
### **4.1. Dataset**

O dataset Galaxy10 DECaLS é um conjunto de dados com quase 18 mil registros com resolução de  $256 \times 256$  pixels coloridos composto por imagens rigorosamente filtradas de galáxias classificadas pelo projeto Galaxy Zoo (GZ). Tal projeto conta com voluntários para realizar a classificação de galáxias em 10 classes distintas (Figura 1) com base em suas características morfológicas, como braços espirais, presença de barras no centro, características de disco, entre outras.



**Figura 1. Exemplos de galáxias no dataset Galaxy Zoo**

Para a classificação, voluntários do GZ respondiam perguntas relativas à morfologia das galáxias, sintetizadas em uma árvore de decisão (Figura 2). As respostas a tais perguntas possibilitaram posteriormente a classificação das imagens em tipos distintos de galáxias com base em métodos estatísticos, constituindo as 10 classes disponíveis no dataset.



**Figura 2. Árvore de decisão utilizada na classificação das galáxias por voluntários segundo sua morfologia (WALMSLEY et al., 2022)**

A consolidação dos votos dos voluntários construiu algumas bases conhecidas para estudo de morfologia de galáxias, como o Galaxy Zoo DECaLS, com mais de 441.000 registro. A partir da filtragem dos melhores registros disponíveis nessa base, originou-se a Galaxy10 DECaLS, com uma média de 30 avaliações de voluntários por

registro. Alguns tipos de galáxia são mais raros ou mais difíceis de observar e classificar, o que faz com que as classes do dataset sejam naturalmente desbalanceadas. Uma sumarização da quantidade de registros das classes e da sua descrição pode ser encontrada na Tabela 1.

**Tabela 1. Tamanho e descrição das classes do dataset Galaxy10 DECaLS.**

Classe	Registros	Descrição
0	1081	Galáxias perturbadas
1	1853	Fusão de galáxias
2	2645	Galáxias redondas e lisas
3	2027	Galáxias redondas e lisas intermediárias
4	334	Galáxias em formato de charuto
5	2043	Galáxias espirais com barras
6	1829	Galáxias espirais compactas sem barras
7	2628	Galáxias espirais soltas sem barras
8	1423	Galáxias de perfil sem protuberância
9	1873	Galáxias de perfil com protuberância

## 4.2. Redes Neurais

Esse trabalho explora o uso de 3 arquiteturas diferentes para a classificação de galáxias. A primeira, uma CNN customizada, é tomada como baseline. A segunda, o modelo ResNet 50, bem explorado na literatura para tarefas análogas. A terceira, a arquitetura BYOL, auto supervisionada. A escolha das arquiteturas utilizadas foi feita com objetivo de explorar complexidades e diferentes formas de aprendizado, funções de custos e comparação entre modelos supervisionados e não supervisionados.

### 4.2.1. CNN

As Convolutional Neural Networks (CNNs) são modelos projetados para processar dados com estrutura espacial, especialmente imagens, explorando padrões locais por meio de operações de convolução que permitem extrair bordas, texturas e formas [Rawat and Wang 2017]. Conforme discutido no trabalho, o poder das CNNs reside na capacidade de aprender representações hierárquicas, onde camadas iniciais capturam padrões simples e camadas profundas aprendem estruturas complexas de alto nível.

A Figura 3 apresenta um modelo típico de CNN aplicado a uma tarefa simples de classificação de imagens. A entrada da rede é a própria imagem, que passa por uma sequência de camadas de convolução e pooling. As características extraídas nessas etapas são então encaminhadas para uma ou mais camadas totalmente conectadas. A última dessas camadas é responsável por indicar a classe prevista. Tal arquitetura constitui a base conceitual dos modelos contemporâneos e fundamenta a estrutura adotada neste trabalho.

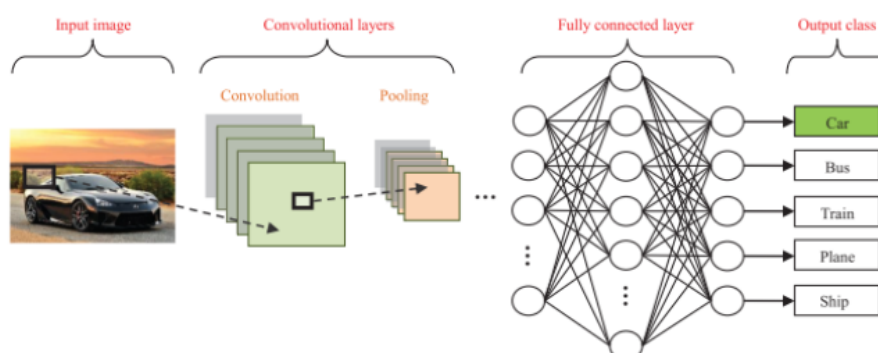


Figura 3. Exemplo de arquitetura típica de CNN. (RAWAT; WANG, 2017).

#### 4.2.2. ResNet-50

Apresentada em 2015 por Kaiming He e colaboradores [He et al. 2015], a Residual Network (ResNet-50) consolidou-se rapidamente como uma das arquiteturas mais influentes na área, destacando-se pela elevada precisão e pelo desempenho eficiente em tarefas de reconhecimento de imagens em larga escala. No trabalho publicado por He, foram apresentadas outras variações da ResNet, como a ResNet-34 e ResNet-101, indicando mudanças apenas no número de layers.

Sabemos que, em modelos muito profundos, informações importantes podem se perder à medida que se movem por muitas camadas, dificultando o aprendizado do modelo, como mostra a figura (3) evidenciado no trabalho [He et al. 2015].

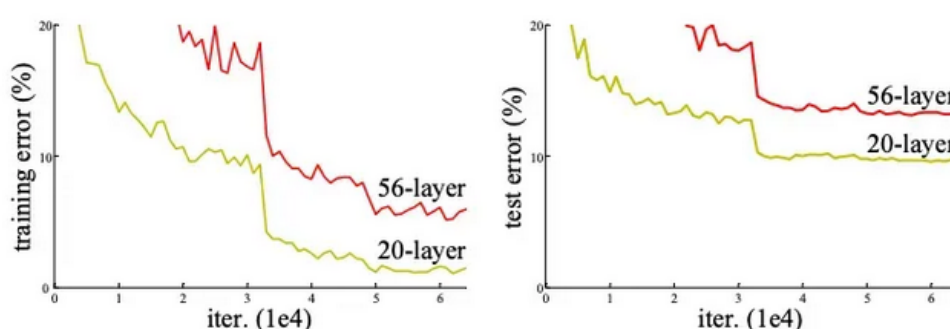
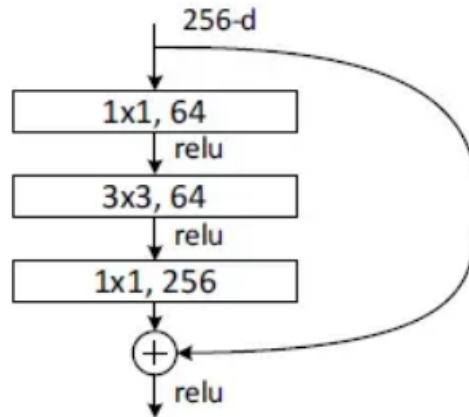


Figura 4. Erro de treinamento (esquerda) e erro de teste (direita) no CIFAR-10 com redes de 20 camadas e 56 camadas.

Conforme a profundidade das redes aumenta, o ganho de precisão deixa de aumentar e, posteriormente, passa a piorar de forma acentuada. Dessa forma, muitas vezes a rede mais profunda tem maior erro de treinamento e, portanto, erro de teste.

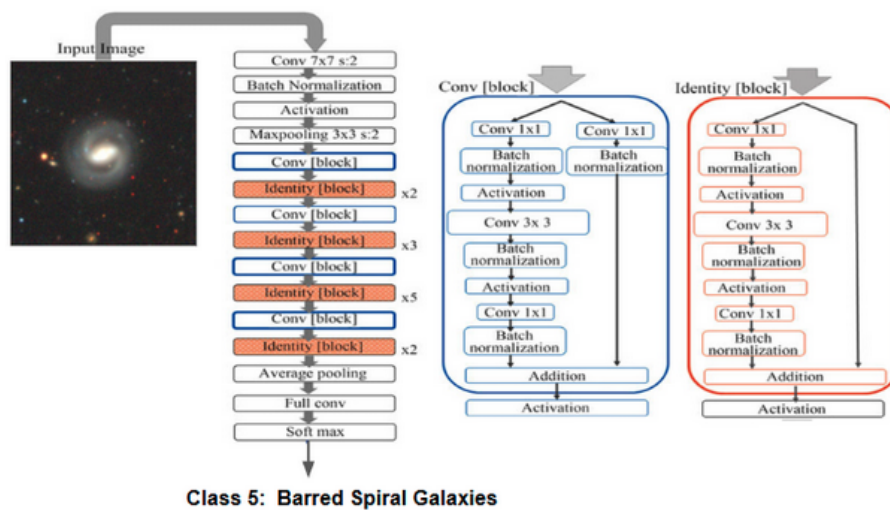
A principal característica da ResNet-50 é a adoção de blocos residuais, estruturados por meio de skip connections. Essas conexões criam caminhos alternativos nos

quais a informação pode fluir sem necessariamente atravessar todas as transformações intermediárias. O bloco residual tem a seguinte estrutura.



**Figura 5. Bloco Residual da ResNet-50.**

Em essência, em vez de obrigar o modelo a processar cada detalhe em todas as camadas, os atalhos permitem que sinais relevantes sejam preservados e transmitidos diretamente. A figura 6 ilustra a estrutura da ResNet-50 de forma mais detalhada.



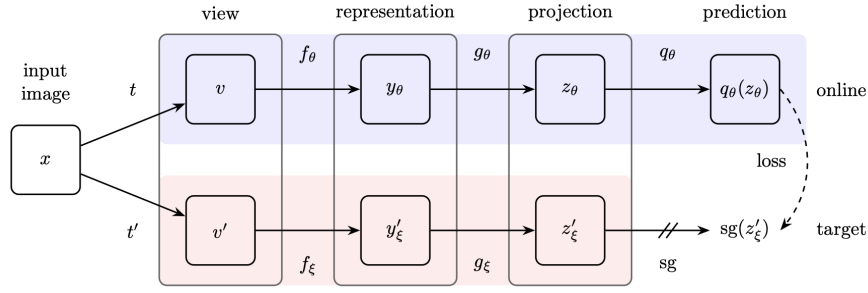
**Figura 6. Representação esquemática da ResNet-50 (2015, Kaiming He et al.)**

Tal mecanismo acelera o treinamento e aumenta a estabilidade da aprendizagem, reduzindo problemas típicos de redes profundas, como o vanishing gradient. As conexões residuais ajudam a evitar isso, mantendo o fluxo de informações claro do início ao fim.

#### 4.2.3. BYOL

Bootstrap Your Own Latent (BYOL) é um algoritmo não-supervisionado para aprendizado de imagens, publicado em 2020 [Grill et al. 2020].

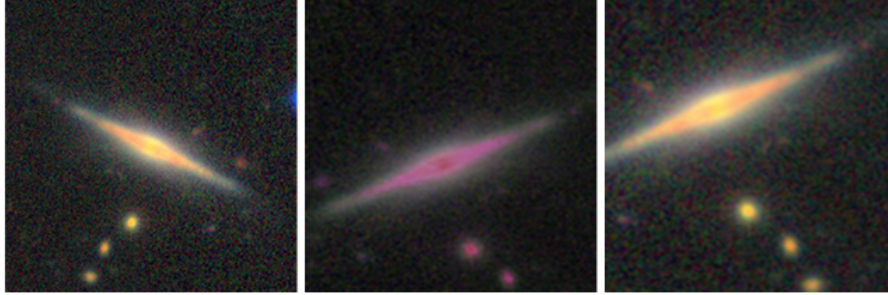




**Figura 7. Arquitetura BYOL (Grill et al., 2020).**

Nessa arquitetura, são criadas duas redes iguais e paralelas. A primeira, chamada “Rede Target”, é constituída de um classificador de imagens (no nosso caso, ResNet-50) e um projetor, que cria a representação da imagem. A segunda, chamada “Rede Online”, tenta, então prever a representação da imagem na Rede Target.

Durante o treinamento, a mesma imagem é usada para treinar ambas as redes, usando data augmentations diferentes. Além disso, a rede online é treinada com back-propagation do gradiente de erro, enquanto os pesos da rede target são atualizados com a média móvel exponencial dos pesos da rede online.



**Figura 8. Exemplo de modificações realizadas nas imagens durante o Data Augmentation do BYOL. A imagem original (da esquerda), e as duas modificações para a rede online e target.**

O aprendizado de representação da rede online do BYOL pode ser usado para outras tarefas, tais como classificação, com uma adequação em fine tuning. O BYOL é um método estável e moderno, que alcança métricas equiparáveis a métodos supervisionados.

## 5. Experimentos

Nesta seção são descritos os experimentos conduzidos para avaliar o desempenho das arquiteturas de redes neurais estudadas na tarefa de classificação morfológica de galáxias. São apresentados os procedimentos adotados para o pré-processamento dos dados, a configuração dos conjuntos de treino, validação e teste, bem como as estratégias de treinamento empregadas para cada modelo. Além disso, são detalhadas as escolhas relacionadas às arquiteturas utilizadas, aos hiperparâmetros e às técnicas de data augmentation, de modo a garantir uma comparação justa e consistente entre as abordagens avaliadas. Os resultados obtidos a partir desses experimentos são posteriormente analisados e discutidos na seção seguinte.



## 5.1. Pré-processamento

A base de dados Galaxy10 DECaLS foi dividida em conjuntos de treino, validação e teste correspondentes respectivamente a 80% (14.188 registros), 5% (887 registros) e 15% (2.261 registros) do total. Esses splits foram armazenados em arquivos e mantidos de forma consistente para o treinamento supervisionado das três arquiteturas avaliadas. Abaixo, a tabela explicita a quantidade de exemplos por classe, para os conjuntos de treino, validação e teste.

**Tabela 2. Distribuição dos exemplos por classe nos conjuntos de treino, validação e teste**

Classe	Treino	Validação	Teste
0	865	54	162
1	1482	93	278
2	2116	132	397
3	1622	101	304
4	267	17	50
5	1634	102	307
6	1463	91	275
7	2102	132	394
8	1139	71	213
9	1498	94	281

Os experimentos foram conduzidos em Python utilizando as bibliotecas NumPy, Matplotlib, Torch e o pacote Scikit-learn. As execuções ocorreram no ambiente Google Colab, conectado ao serviço gratuito e limitado de GPU disponível, devido à facilidade de acesso e colaboração para os testes realizados.

Durante as etapas de preparação dos dados e experimentação do treinamento, avaliámos diferentes técnicas de Data Augmentation. Observamos que algumas delas, ao contrário do esperado, acabaram prejudicando o aprendizado do modelo.

Um exemplo foi o uso do T.RandomResizedCrop, esse tipo de Crop realiza cortes aleatórios na imagem, alterando a posição e o enquadramento do objeto principal. No entanto, em nosso conjunto de dados, as galáxias estão centralizadas. Assim, a rede recebia exemplos que não representavam a label associada a ele. Como consequência, o modelo tinha mais dificuldade para aprender, resultando em um pior desempenho.

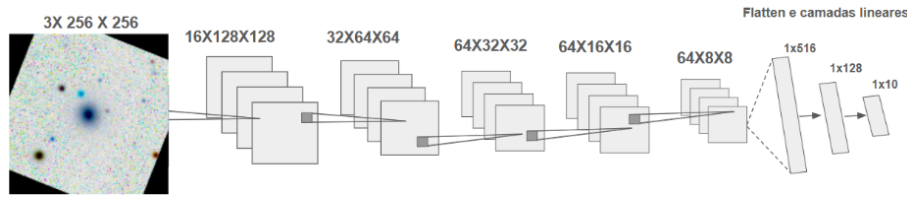
Por esse motivo, descartamos o uso desse augmentation e as transformações adotadas no conjunto de treino foram: conversão para tensor (ToTensor), pequenas variações de cor (ColorJitter), borramento gaussiano suave (GaussianBlur), espelhamentos horizontal e vertical com probabilidade de 0,5, rotações aleatórias entre  $-90^\circ$  e  $90^\circ$  (RandomRotation) e normalização dos canais (Normalize). Tais operações aumentam a variabilidade das imagens sem alterar a estrutura central das galáxias.



**Figura 9. Random Crop: O problema do corte aleatório como Data Augmentation no nosso dataset.**

## 5.2. CNN

Desenvolvemos uma arquitetura CNN personalizada composta por cinco blocos convolucionais seguidos por três camadas totalmente conectadas para a tarefa de classificação. As imagens de entrada possuem dimensão  $256 \times 256 \times 3$ , preservando o formato original do conjunto Galaxy10 DECaLS. A Figura 6 apresenta a estrutura completa do modelo.



**Figura 10. Arquitetura da CNN utilizada.**

Cada bloco convolucional emprega kernels  $3 \times 3$  com 16, 32 ou 64 filtros, acompanhados por Batch Normalization, função de ativação ReLU, dropout e MaxPool  $2 \times 2$ . Essa combinação reduz progressivamente a resolução espacial enquanto permite a extração de padrões morfológicos cada vez mais complexos. A saída da última camada convolucional ( $64 \times 8 \times 8$ ) é achatada e encaminhada para três camadas lineares com 512, 128 e 10 neurônios. As duas primeiras utilizam ReLU e dropout, enquanto a camada final gera as pontuações para cada uma das classes, que são posteriormente convertidas em probabilidades pela função softmax na etapa de inferência. O treinamento da rede foi realizado ao longo de 20 épocas, com tempo total aproximado de 35 minutos.

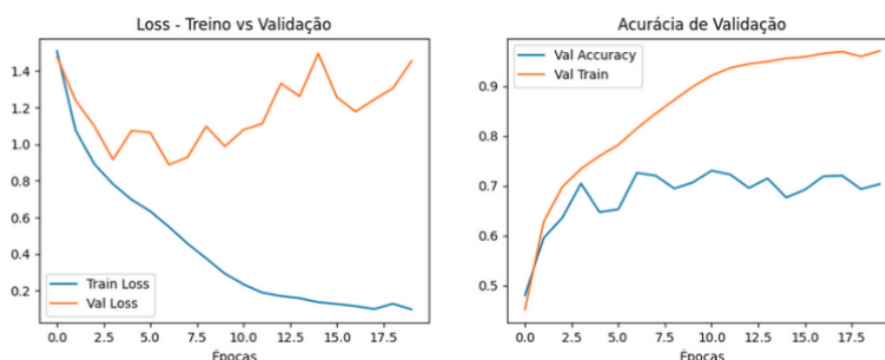
## 5.3. ResNet-50

Inicialmente, foram conduzidos diversos experimentos para buscar um desempenho superior com a ResNet-50. Implementamos a arquitetura manualmente, camada por camada, e também avaliamos a ResNet-50 pré-treinada no ImageNet, testando tanto o congelamento total da rede com ajuste apenas da camada totalmente conectada, quanto o descongelamento combinado da fully connected e do bloco correspondente à layer 4. Realizamos ainda testes com as variantes ResNet-101 e ResNet-152, mas o custo computacional tornou o treinamento inviável. Com base nesses experimentos, o melhor desempenho foi obtido com a versão da ResNet-50 implementada integralmente do zero.

**Tabela 3. Configurações utilizadas no treinamento da ResNet-50**

Épocas	Loss	Tempo de treino	Learning Rate	Otimizador
20	CrossEntropy	2 horas e 10 minutos	1e-4	Adam

Outro experimento relevante consistiu em treinar o modelo sem utilizar data augmentation. Considerando que o dataset Galaxy10 DECaLS é desbalanceado, buscou-se avaliar o impacto que essa técnica exercia sobre o processo de treinamento.



**Figura 11. Gráficos de Loss e Acurácia dos conjuntos de treino e validação da ResNet-50 sem data augmentation.**

Os gráficos mostram que, sem data augmentation, a ResNet-50 rapidamente entra em overfitting. Enquanto o loss de treino cai continuamente e a acurácia de treino cresce até valores muito altos, o desempenho de validação estagna e o loss de validação aumenta ao longo das épocas. Isso indica que o modelo memoriza o conjunto de treino, mas não generaliza bem para novos dados, evidenciando a importância do data augmentation em nosso contexto.

#### 5.4. BYOL

O algoritmo padrão do BYOL empregado fornecido no artigo original de Grill et al., 2020, está disponível em repositório público e foi utilizado nos experimentos que seguem. Foi utilizada a estrutura ResNet-50 pré-treinada no Imagenet e disponível na biblioteca PyTorch. O algoritmo BYOL utilizado neste trabalho foi treinado em três etapas.

A primeira etapa consiste em um longo treino auto supervisionado em 40 épocas, utilizando todos os registros disponíveis na base. O objetivo é desenvolver uma boa representação das imagens da base, ainda que não estejamos trabalhando com as classes. Nesse estágio, foi utilizado o otimizador Adam W com Learning Rate de  $10^{-4}$  e decaimento de  $10^{-6}$ .

A segunda e terceira etapas são relacionadas ao Fine Tuning do modelo online para a tarefa de classificação das Galáxias, com um treinamento supervisionado seguindo a divisão de conjuntos de dados para treino, validação e teste.

Inicialmente, fizemos o Linear Probing com uma nova camada classificadora, sem back propagation nas demais camadas, por 10 épocas. Aqui, o otimizador Adam W continuava com Learning Rate de  $10^{-4}$  e decaimento  $10^{-6}$ .

Para a última etapa, o otimizador foi ajustado para Learning Rate de  $10^{-5}$  e decaimento maior que nas etapas anteriores, de  $10^{-4}$ . Realizamos o treinamento por mais 10 épocas para toda a rede, finalizando o fine tuning.

Durante a implementação, outros otimizadores (Adam, SGD) e Learning Rates foram testados, e a convergência para a configuração final aconteceu para evitar erros de treino como o overfitting.

Os tempos de treinamento do BYOL superaram os demais experimentos devido a sua complexidade e a nossa limitação de processamento, que foi o principal ditador de número de épocas nos experimentos. Pelo mesmo motivo, escolhemos diminuir a definição das imagens, inicialmente de 256 x 256 pixels para 128x128 pixels.

**Tabela 4. Configurações utilizadas no treinamento do BYOL**

Etapa	Épocas	Loss	Tempo	LR	Otim.	Decay
Auto Sup.	40	Cos. Loss	4h30min	1e-4	AdamW	1e-6
Lin. Probing	10	CrossEntropy	45min	1e-4	AdamW	1e-6
Fine Tuning	10	CrossEntropy	45min	1e-5	AdamW	1e-4

## 6. Resultados

Nesta seção são apresentados e discutidos os resultados obtidos a partir dos experimentos realizados com as diferentes arquiteturas de redes neurais avaliadas. A análise considera o desempenho das CNNs, da ResNet-50 e do modelo BYOL na classificação morfológica de galáxias do dataset Galaxy10 DECaLS, utilizando métricas como acurácia, recall, precisão e F1-score. Além da comparação quantitativa entre os modelos, são discutidos aspectos como estabilidade durante o treinamento, capacidade de generalização e influência das características do conjunto de dados nos resultados observados.

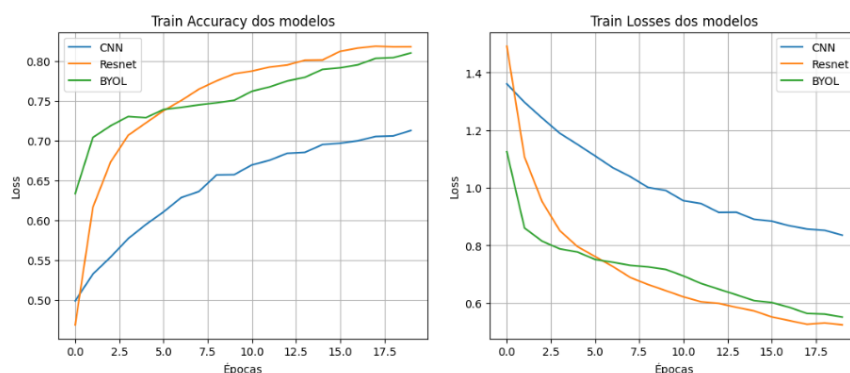
### 6.1. Fase de treino

A partir das curvas de acurácia e loss, pode-se inferir diferenças significativas entre o desempenho dos modelos 12, tendo o Resnet50 e o BYOL apresentado menores valores de loss e maior acurácia no conjunto de validação durante o treinamento. Além disso, a ResNet apresenta a convergência mais estável entre as curvas.

O comportamento de não estabilização das curvas de acurácia e loss do BYOL pode indicar que haveriam melhorias na sua convergência caso fosse treinada em mais épocas. Na literatura, trabalhos como [Mohale et al. 2024] e [Slijepcevic et al. 2024] alcançaram acurácias próximas a 90% com essa arquitetura em datasets similares, treinando em mais épocas com configurações também comparáveis.

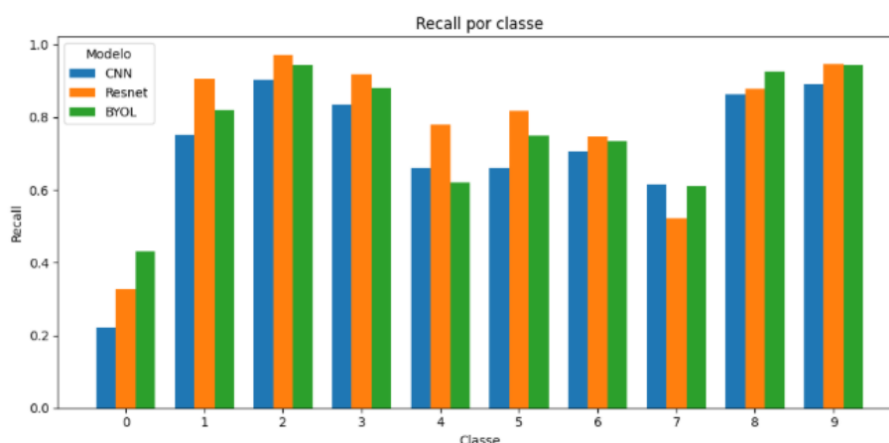
### 6.2. Análise dos resultados

No conjunto de teste, a superioridade do ResNet e BYOL também é destacada em relação às CNN. Os índices de recall ficam superiores em quase todas as classes para os dois modelos mais avançados (Figura 12), confirmando a tendência do conjunto de treino e validação. A análise do recall dos modelos revelou que o desempenho interclasses não foi uniforme, mas há coerência no desempenho dentro das classes 3 arquiteturas (Figura



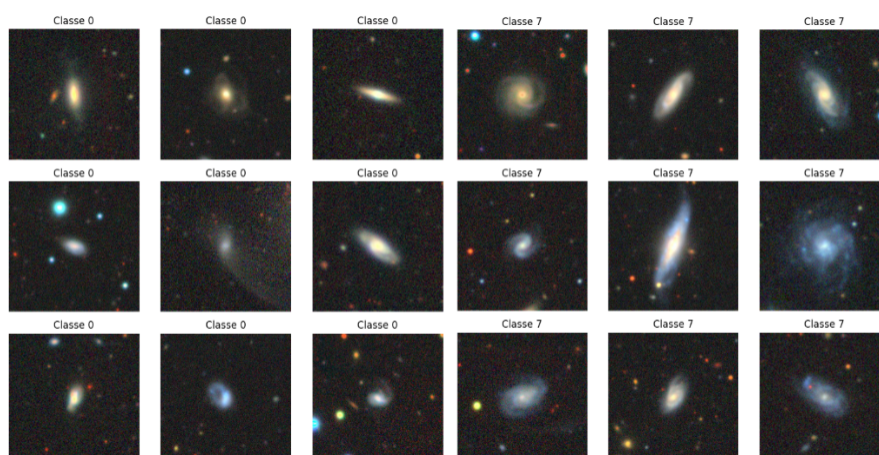
**Figura 12. Desempenho dos classificadores nas etapas de treinamento.**

13), indicando algumas morfologias com desafios de generalização variável, em especial nas classes 0 e 7.



**Figura 13. Recall da classificação dos modelos por classe**

As classes 0 e 7 apresentaram os menores recalls, que podem ser justificados por serem as classes com maiores variabilidades morfológicas em relação às demais. Vale lembrar que os grupos morfológicos do Galaxy10 DECaLS não representam a totalidade dos conjuntos morfológicos existentes entre as galáxias, portanto espera-se que haja um agrupamento de morfologias distintas em torno das classes dessa base, sendo a classe 0 denominada “Galáxias perturbadas” e a classe 7 “Galáxias espirais soltas sem barra” os grupos com maior variabilidade morfológica.



**Figura 14. Exemplos de variabilidade morfológica nas imagens das Classes 0 (Galáxias perturbadas) e 7 (Galáxias espirais soltas sem barra).**

O modelo Resnet-50 atingiu os recalls mais elevados na maioria das classes, todavia, o BYOL se sai melhor em relação às classes mais desafiadoras nessa métrica. Por trabalharmos em um dataset desbalanceado e com classes muito semelhantes entre si, foi realizada a análise do F1-score, para determinar a performance geral dos modelos intra-classe. Nessa análise, o BYOL se equipara ao ResNet, com uma pequena vantagem na média geral das classes, indicando bom poder de generalização, apesar do desafio (Tabela X).

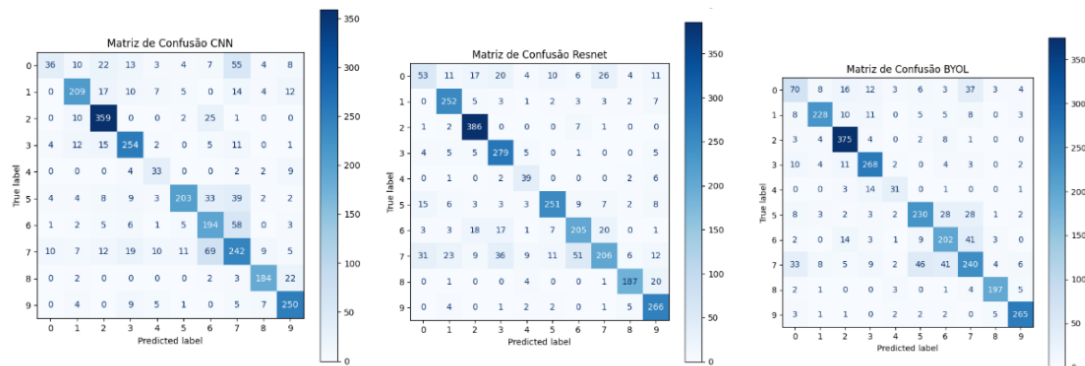
**Tabela 5. Recall da classificação dos modelos por classe**

Classe	CNN			Resnet			BYOL		
	Prec.	Recall	F1	Prec.	Recall	F1	Prec.	Recall	F1
0	0,655	0,222	0,332	0,495	0,327	0,394	0,504	0,432	0,465
1	0,804	0,752	0,777	0,818	0,906	0,860	0,887	0,820	0,852
2	0,820	0,904	0,860	0,871	0,972	0,919	0,858	0,945	0,899
3	0,784	0,836	0,809	0,773	0,918	0,839	0,827	0,882	0,854
4	0,516	0,660	0,579	0,574	0,780	0,661	0,674	0,620	0,646
5	0,879	0,661	0,755	0,887	0,818	0,851	0,767	0,749	0,758
6	0,579	0,705	0,636	0,727	0,745	0,736	0,685	0,735	0,709
7	0,563	0,614	0,587	0,777	0,523	0,625	0,663	0,609	0,635
8	0,868	0,864	0,866	0,899	0,878	0,888	0,925	0,925	0,925
9	0,801	0,890	0,843	0,792	0,947	0,862	0,920	0,943	0,931
<b>Macro avg</b>	0,727	0,711	0,704	0,761	0,781	0,764	0,771	0,766	0,767
<b>Weighted avg</b>	0,744	0,738	0,731	0,793	0,798	0,788	0,787	0,791	0,788

A partir das matrizes de confusão (Figura 15) é possível avaliar as confusões entre as classes pelo modelos. Destacam-se as classes 6 e 7 com mais exemplos classificados de forma trocada. Sendo a classe 6 “Galáxias espirais compactadas sem barra” e 7 “Galáxias espirais soltas sem barra”, apresentando parte significativa de exemplos com classificação trocada e morfologias semelhantes, possuindo também representantes em estágios intermediários entre uma morfologia e outra.

Como esperado, a classe zero também apresenta parte significativa dos exemplos

classificados como classe 7, indicando que parte dos exemplares daquela classe possuem características morfológicas comparáveis aos dessa classe.



**Figura 15. Matrizes de confusão dos modelos.**

O tempo de treinamento para os modelos foi crescente em relação à sua complexidade 6. O BYOL apresentou o maior tempo total somando os dois estágios de treinamento (supervisionado e auto supervisionado), quatro vezes maior que o tempo total da ResNet. Com o resultado final comparável e a grande diferença de gasto computacional entre as duas redes, torna-se relevante avaliar a aplicabilidade da BYOL nesse problema.

**Tabela 6. Tempo de treino dos modelos**

Modelo	Tempo de treino
CNN	35 minutos
Resnet	90 minutos
BYOL (auto supervisionado)	270 minutos
BYOL (supervisionado)	90 minutos

## 7. Conclusão

O trabalho explorou a classificação de galáxias a partir de suas características morfológicas com o uso de Redes Neurais, avaliando três diferentes arquiteturas, suas características e desempenhos.

Os modelos explorados neste trabalho apresentaram boa performance em acurácia, com recall superior a 0.75 na maioria das classes, com exceção das classes com morfologia de alta variabilidade. Os resultados alcançados foram comparáveis a estudos similares na literatura.

Dentre as arquiteturas estudadas, o modelo CNN atingiu bons resultados, mas ficou atrás dos outros dois modelos. A Resnet-50 desempenhou melhor na classificação, muito próxima e comparável com a BYOL. Esses resultados evidenciam o grande potencial de utilização de arquiteturas auto supervisionadas em comparação com seus pares supervisionados bem estabelecidos, para classificação de imagens de galáxias.

Notamos também que a rede auto supervisionada ainda não tinha alcançado a total estabilização com o treinamento realizado, que foi limitado pela capacidade de processamento restrita. Com um treinamento mais longo, essa arquitetura poderia alcançar um



resultado até melhores do que os atingidos com a ResNet, como foi constatado em outros trabalhos semelhantes na literatura.

Por outro lado, o custo computacional para o treinamento da rede auto supervisionada é muito maior do que o das outras arquiteturas estudadas, com maior tempo de treino e complexidade em etapas. Por esse motivo, com a disponibilidade de hardware que temos atualmente, a ResNet ainda se destaca como a melhor escolha para esta tarefa dentre os métodos estudados. Para outras configurações e demandas de acurácia, será necessário realizar uma avaliação ponderando essas características.

## Referências

- Cao, J. et al. (2024). Galaxy morphology classification based on convolutional vision transformer (cvt). *Astronomy & Astrophysics*, 683:A42.
- de Castro, M. T. (2025). *Redes Neurais Convolucionais Aplicadas a Identificação e Classificação de Galáxias*. Universidade Federal do Ceará.
- Galaxy Zoo (2007). Galaxy zoo data. [data.galaxyzoo.org](https://data.galaxyzoo.org). Disponível em: <https://data.galaxyzoo.org/>.
- Grill, J.-B., Strub, F., Altché, F., Tallec, C., Richemond, P. H., Buchatskaya, E., et al. (2020). Bootstrap your own latent (byol): A new approach to self-supervised learning. In *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. *arXiv preprint arXiv:1512.03385*.
- Le, T. et al. (2025). Enhanced comparative analysis of pretrained and custom deep convolutional neural networks for galaxy morphology classification. *Engineering Proceedings*, page 36.
- Mohale, K. et al. (2024). Enabling unsupervised discovery in astronomical images. *Monthly Notices of the Royal Astronomical Society*, page 1274.
- Mohale, K. and Lochner, M. (2024). Enabling unsupervised discovery in astronomical images through self-supervised representations. *Monthly Notices of the Royal Astronomical Society*, 530(1):1274–1295.
- Rawat, W. and Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29:2352–2449.
- Sandage, A. (2005). The classification of galaxies: early history and ongoing developments. *Annual Review of Astronomy and Astrophysics*, 43:581–624.
- Slijepcevic, I. V., Scaife, A. M. M., Walmsley, M., Bowles, M., Wong, O. I., Shabala, S. S., and White, S. V. (2024). Radio galaxy zoo: towards building the first multi-purpose foundation model for radio astronomy with self-supervised learning. *RAS Techniques and Instruments*, pages 19–32.
- Wen, Z. Z., Zheng, X. Z., and An, F. X. (2014). Probing asymmetric structures in the outskirts of galaxies. *The Astrophysical Journal*, 787(2):130.

Yao, Y. et al. (2024). Um método de aumento de imagem galaxy baseado em aprendizado de poucos disparos e redes adversárias gerativas. *Pesquisa em Astronomia e Astrofísica*, 24(3).