



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**TRABALHO DE CONCLUSÃO DE CURSO - TCC**  
**CURSO DE GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO**

**MARIANA TEIXEIRA DE CASTRO**

**REDES NEURAIS CONVOLUCIONAIS APLICADAS A IDENTIFICAÇÃO E  
CLASSIFICAÇÃO DE GALÁXIAS**

**SOBRAL - CE**

**2025**

MARIANA TEIXEIRA DE CASTRO

REDES NEURAIS CONVOLUCIONAIS APLICADAS A IDENTIFICAÇÃO E  
CLASSIFICAÇÃO DE GALÁXIAS

Trabalho de Conclusão de Curso apresentado  
ao Curso de Graduação em Engenharia da  
Computação da Universidade Federal do  
Ceará, como requisito parcial à obtenção do  
grau de bacharel em Engenharia da Computação.

Orientador: Prof. Dr. Iális Cavalcante de  
Paula Junior

SOBRAL - CE

2025



Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

C352r Castro, Mariana Teixeira de.  
Redes Neurais Convolucionais Aplicadas a Identificação e Classificação de Galáxias / Mariana Teixeira de Castro. – 2025.  
90 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Sobral, Curso de Engenharia da Computação, Sobral, 2025.  
Orientação: Prof. Dr. Iális Cavalcante de Paula Junior.

1. Astronomia. 2. Redes Neurais Convolucionais. 3. Galáxias. I. Título.

CDD 621.39

---

MARIANA TEIXEIRA DE CASTRO

REDES NEURAIIS CONVOLUCIONAIS APLICADAS A IDENTIFICAÇÃO E  
CLASSIFICAÇÃO DE GALÁXIAS

Trabalho de Conclusão de Curso apresentado  
ao Curso de Graduação em Engenharia da  
Computação da Universidade Federal do Ceará,  
como requisito parcial à obtenção do grau de  
bacharel em Engenharia da Computação.

Aprovada em: 07/03/2025

BANCA EXAMINADORA

---

Prof. Dr. Iális Cavalcante de Paula  
Junior (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Marcelo Marques Simões de Souza  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Thiago Iachiley Araújo de Souza  
Universidade Federal do Ceará (UFC)

## **AGRADECIMENTOS**

Agradeço ao Prof. Dr. Iális Cavalcante de Paula Junior, pela paciência e dedicação durante a orientação deste trabalho. Sou grata também a todos os demais professores da Universidade Federal do Ceará que, de alguma forma, me influenciaram a desenvolver minhas habilidades em machine learning e ciência de dados.

Aos meus pais, Absalão Bandeira de Castro e Vilma Teixeira de Castro e ao meu irmão Gabriel Teixeira de Castro, que apesar da distância física, estão sempre presentes em minha vida, me motivando a sempre seguir em frente.

Ao Prof. Dr. Daniel Brito de Freitas e aos estudantes de pós-graduação Paulo Cleber Farias da Silva Filho e Lucas Pedrosa Valente, pelas contribuições no desenvolvimento desta pesquisa. Aos demais amigos e colegas do curso de Engenharia da Computação pelo apoio durante esta caminhada.

Por último, mas não menos importante, agradeço a mim mesma por nunca ter dado ouvidos aos motivos que me mandavam desistir.

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

(Ariano Suassuna)

## RESUMO

Os *Surveys* são telescópios automatizados posicionados estrategicamente na superfície terrestre ou no espaço, projetados para realizar varreduras do céu para que se construa um catálogo detalhado dos corpos celestes observados em seu campo de visão. Combinados, esses instrumentos geram dezenas de *terabytes* de dados astronômicos, o que torna desafiador o registro e a identificação destes objetos devido ao enorme volume de informações produzidas. Por essa razão, são necessários profissionais capacitados e técnicas computacionais avançadas de aprendizado profundo, como Redes Neurais Convolucionais (CNNs), para gerenciar e processar essa quantidade massiva de dados. O presente trabalho investiga a aplicabilidade de diversas CNNs na classificação de galáxias utilizando imagens provenientes da base de dados *Galaxy10 SDSS Dataset*, composta por 21.785 imagens, cada uma com dimensão de 69x69 *pixels*. As arquiteturas testadas, pautadas em estudos prévios, foram *Gharat & Dandawate*, *EfficientNetV2-M*, *DenseNet121*, *ResNet50*, *AlexNet*, *VGG16* e *VGG19*. Técnicas de pré-processamento como aumento de dados, redimensionamento e ajuste de contraste foram aplicadas de maneira seletiva em diferentes experimentos, buscando identificar a combinação mais eficiente de forma a melhorar os resultados. Os modelos foram avaliados usando métricas como Acurácia, Perda, ROC-AUC e PR-AUC. A arquitetura *VGG19* com aplicação da técnica *CutMix* apresentou o melhor desempenho, alcançando uma acurácia de 0,90 e PR-AUC de 0,94. Os resultados obtidos não apenas confirmaram a robustez das CNNs na classificação de dados astronômicos, mas também ressaltaram a importância da escolha criteriosa do modelo e das técnicas complementares para otimização do processo de classificação.

**Palavras-chave:** Astronomia. Redes Neurais Convolucionais. Galáxias.

## ABSTRACT

Surveys are automated telescopes strategically positioned on the Earth's surface or in space, designed to scan the sky to build a detailed catalog of the celestial bodies observed in their field of view. Combined, these instruments generate terabytes of astronomical data, which makes it necessary to record and identify these objects due to the enormous volume of information produced. For this reason, qualified professionals and advanced deep learning computational techniques, such as Convolutional Neural Networks (CNNs), are needed to manage and process this enormous amount of data. This work investigates the applicability of several CNNs in classifying galaxies using images from the Galaxy10 SDSS Dataset, composed of 21,785 images, each with dimensions of 69x69 pixels. The tested frameworks, based on previous studies, were Gharat & Dandawate, EfficientNetV2-M, DenseNet121, ResNet50, AlexNet, VGG16 and VGG19. Preprocessing techniques such as data augmentation, resizing and contrast adjustment were applied in an active selected manner in different experiments, seeking to identify the most efficient combination in order to improve the results. The models were evaluated using analyses such as Accuracy, Loss, ROC-AUC and PR-AUC. The VGG19 architecture with application of the CutMix technique presented the best performance, achieving an accuracy of 0.90 and PR-AUC of 0.94. The results obtained not only confirmed the robustness of CNNs in the classification of astronomical data, but also highlighted the importance of choosing the model criteria and complementary techniques to optimize the classification process.

**Keywords:** Astronomy. Convolutional Neural Networks. Galaxies.

## LISTA DE FIGURAS

Figura 1	– Primeira imagem de campo profundo processada pelo Telescópio Espacial James Webb, revelando a vasta quantidade de galáxias distantes. . . . .	20
Figura 2	– Sequência de Hubble para classificação das galáxias com base em sua morfologia: elípticas, espirais, espirais barradas e irregulares. . . . .	21
Figura 3	– Morfologia das galáxias elípticas com diferentes graus de achatamento. . . .	22
Figura 4	– Representação esquemática das subdivisões da Via Láctea, destacando a localização do nosso Sistema Solar no disco galáctico. . . . .	22
Figura 5	– Morfologia das galáxias espirais, nos seus três subtipos, com base no grau de abertura dos braços espirais e no tamanho do bojo central. . . . .	23
Figura 6	– Morfologia das galáxias espirais barradas, em seus três subtipos, com base na proeminência da barra central e na abertura dos braços espirais. . . . .	23
Figura 7	– Representação da morfologia de uma galáxia lenticular ( <i>S0</i> ) à esquerda e a galáxia <i>M104</i> , um exemplo real desse tipo de galáxia, à direita. . . . .	24
Figura 8	– Galáxias irregulares. Da esquerda para direita: <i>NGC4038</i> , <i>NGC1427</i> e <i>IC4710</i> , caracterizadas pela ausência de uma estrutura definida e pela intensa formação estelar. . . . .	25
Figura 9	– O neurônio artificial, onde as entradas são ponderadas por pesos sinápticos e somadas. A função de ativação, que determina a saída da equação, adicionando não linearidade ao neurônio. . . . .	27
Figura 10	– Funcionamento do algoritmo backpropagation em uma rede neural, destacando a propagação direta ( <i>forward</i> ) dos dados até a saída e a retropropagação ( <i>backward</i> ) do erro, que ajusta os pesos sinápticos para otimizar o aprendizado da rede. . . . .	27
Figura 11	– Estrutura de uma rede neural convolucional e suas camadas. O fluxo de processamento ocorre da entrada até a classificação final. . . . .	28
Figura 12	– Camadas convolutivas em uma rede neural no processamento de uma imagem de $32 \times 32$ <i>pixels</i> . Sucessivas camadas de convolução e subamostragem reduzem a dimensionalidade e extraem características. . . . .	29

Figura 13 – Funcionamento da camada de pooling em redes neurais convolucionais. O exemplo demonstra a aplicação do <i>Max Pooling</i> com um filtro de tamanho 2×2, onde o valor máximo de cada região filtrada é selecionado para compor a nova matriz reduzida. . . . .	30
Figura 14 – Operação de <i>flattening</i> em redes neurais convolucionais. Esse processo transforma a saída multidimensional da camada de <i>pooling</i> em um vetor unidimensional para os dados serem processados por camadas densas totalmente conectadas. . . . .	30
Figura 15 – Comparação entre Aprendizado de Máquina Tradicional e Aprendizado por Transferência. No método tradicional (a), diferentes tarefas utilizam sistemas de aprendizado independentes. Já no Aprendizado por Transferência (b), o conhecimento adquirido em uma tarefa fonte é reutilizado para aprimorar o desempenho em uma tarefa alvo, tornando o treinamento mais eficiente. . .	31
Figura 16 – Ilustração da arquitetura VGG16. . . . .	32
Figura 17 – Ilustração da arquitetura VGG19. . . . .	32
Figura 18 – Ilustração da arquitetura ResNet50. . . . .	33
Figura 19 – Ilustração da arquitetura DenseNet-121. . . . .	34
Figura 20 – Ilustração da arquitetura EfficientNetV2-M. . . . .	34
Figura 21 – Ilustração da arquitetura AlexNet. . . . .	35
Figura 22 – Matriz de confusão para um problema de classificação binária. . . . .	36
Figura 23 – Curva para avaliação de modelos de classificação. A linha tracejada representa um classificador aleatório, enquanto a curva cinza indica um classificador perfeito. . . . .	38
Figura 24 – Curva PR-AUC para avaliação do desempenho de classificadores. . . . .	39
Figura 25 – Telescópio da missão Sloan Digital Sky Survey. . . . .	41
Figura 26 – Árvore de decisão do Galaxy Zoo. Ferramenta interativa que orienta voluntários na classificação de galáxias com base em suas características visuais. . .	43
Figura 27 – Exemplo de pergunta de múltipla escolha no Galaxy Zoo para classificação de galáxias. A interface interativa apresenta imagens de galáxias e opções de resposta, permitindo que voluntários as classifiquem com base em suas características visuais, como formato e estrutura. . . . .	44
Figura 28 – Ilustração da arquitetura de Gharat & Dandawate. . . . .	45



Figura 29 – Exemplos de imagens das diferentes classes do conjunto de dados <i>Galaxy10 SDSS Dataset</i> . Podemos observar diversas morfologias galácticas, classificadas por formato, estrutura espiral e orientação do disco. . . . .	48
Figura 30 – Técnicas de Aumento de Dados: <i>MixUp</i> e <i>CutMix</i> . (a) e (b) representam as imagens originais utilizadas como base, (c) apresenta a aplicação da técnica <i>MixUp</i> , combinando elementos das imagens (a) e (b), e por último, (d) representando a aplicação da técnica <i>CutMix</i> , inserindo segmentos da imagem (b) na imagem (a). . . . .	50
Figura 31 – Imagens realçadas com a aplicação do <i>CLAHE</i> . . . . .	50
Figura 32 – Quantidade de imagens por classe de Galáxias para a Arquitetura de Gharat & Dandawate. . . . .	54
Figura 33 – Treinamento II para os modelos <i>VGG16</i> , <i>VGG19</i> e <i>DenseNet121</i> , organizadas respectivamente de cima para baixo. . . . .	59
Figura 34 – ROC-AUC e PR-AUC II para os modelos <i>VGG16</i> , <i>VGG19</i> e <i>DenseNet121</i> , organizadas respectivamente de cima para baixo. . . . .	60
Figura 35 – Treinamento III para os modelos <i>VGG16 MixUp</i> , <i>VGG19 MixUP</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	63
Figura 36 – ROC-AUC e PR-AUC III para os modelos <i>VGG16 MixUp</i> , <i>VGG19 MixUP</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	64
Figura 37 – Treinamento IV para os modelos <i>VGG16 MixUp</i> , <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	67
Figura 38 – ROC-AUC e PR-AUC IV para os modelos <i>VGG16 MixUp</i> , <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	68
Figura 39 – Treinamento V para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	71
Figura 40 – ROC-AUC e PR-AUC V para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	72
Figura 41 – Treinamento VI para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	74
Figura 42 – ROC-AUC e PR-AUC VI para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	75

Figura 43 – Treinamento VII para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	77
Figura 44 – ROC-AUC e PR-AUC VII para os modelos <i>VGG16 CutMix</i> e <i>VGG19 CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	78
Figura 45 – Treinamento VIII para os modelos <i>VGG16 CutMix</i> e <i>EfficientNetV2-M CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	80
Figura 46 – ROC-AUC e PR-AUC VIII para os modelos <i>VGG16 CutMix</i> e <i>EfficientNetV2-M CutMix</i> , organizadas respectivamente de cima para baixo. . . . .	81

## LISTA DE TABELAS

Tabela 1	– Comparação das propriedades das galáxias espirais, elípticas e irregulares com relação a massa, tamanho, luminosidade, composição estelar, presença de gás e poeira, cor e idade das estrelas. . . . .	25
Tabela 2	– Resumo das Métricas de Avaliação de Modelos de Classificação. . . . .	36
Tabela 3	– Distribuição da quantidade de imagens por classe de galáxias no conjunto de dados <i>Galaxy10 SDSS Dataset</i> , apresentando as categorias morfológicas, sua nomenclatura e o número de imagens em cada classe, totalizando 21.785 registros. . . . .	48
Tabela 4	– Descrição dos oito testes conduzidos, detalhando as arquiteturas dos modelos utilizados e os métodos de pré-processamento aplicados em cada etapa. . . .	53
Tabela 5	– Acurácia por classe para a arquitetura de Gharat & Dandawate. . . . .	55
Tabela 6	– Precisão, Revocação e F1-score por classe para a Arquitetura de Gharat & Dandawate. . . . .	55
Tabela 7	– Comparação do desempenho dos modelos para o Experimento II. . . . .	56
Tabela 8	– Desempenho II do modelo VGG16 por classe. . . . .	57
Tabela 9	– Desempenho II do modelo VGG19 por classe. . . . .	57
Tabela 10	– Desempenho II do modelo DenseNet121 por classe. . . . .	57
Tabela 11	– Comparação do desempenho dos modelos para o Experimento III. . . . .	58
Tabela 12	– Desempenho III do modelo VGG16 MixUp por classe. . . . .	61
Tabela 13	– Desempenho III do modelo VGG19 MixUp por classe. . . . .	61
Tabela 14	– Desempenho III do modelo VGG19 CutMix por classe. . . . .	62
Tabela 15	– Comparação do desempenho dos modelos para o Experimento IV. . . . .	65
Tabela 16	– Desempenho IV do modelo VGG16 MixUp por classe. . . . .	65
Tabela 17	– Desempenho IV do modelo VGG16 CutMix por classe. . . . .	66
Tabela 18	– Desempenho IV do modelo VGG19 CutMix por classe. . . . .	66
Tabela 19	– Comparação do desempenho dos modelos para o Experimento V. . . . .	69
Tabela 20	– Desempenho V do modelo VGG16 CutMix por classe. . . . .	69
Tabela 21	– Desempenho V do modelo VGG19 CutMix por classe. . . . .	70
Tabela 22	– Comparação do desempenho dos modelos para o Experimento VI. . . . .	70
Tabela 23	– Desempenho VI do modelo VGG16 CutMix por classe. . . . .	72
Tabela 24	– Desempenho VI do modelo VGG19 CutMix por classe. . . . .	73

Tabela 25 – Comparação do desempenho dos modelos para o Experimento VII. . . . .	73
Tabela 26 – Desempenho VII do modelo VGG16 CutMix por classe. . . . .	76
Tabela 27 – Desempenho VII do modelo VGG19 CutMix por classe. . . . .	76
Tabela 28 – Comparação do desempenho dos modelos para o Experimento VIII. . . . .	78
Tabela 29 – Desempenho VIII do modelo VGG16 CutMix por classe. . . . .	79
Tabela 30 – Desempenho VIII do modelo VGG19 CutMix por classe. . . . .	79
Tabela 31 – Comparação dos melhores resultados dentre os oito modelos analisados . . .	81

## LISTA DE SÍMBOLOS

$E$	Galáxias Elípticas
$FN$	Falso negativo
$FP$	Falso positivo
$n$	Grau de achatamento de uma elipse
$S$	Galáxias Espirais
$Sb$	Galáxias Espirais Barradas
$VN$	Verdadeiro negativo
$VP$	Verdadeiro positivo

## LISTA DE ABREVIATURAS E SIGLAS

CIFAR-10	<i>Canadian Institute For Advanced Research</i> (Instituto Canadense de Pesquisa Avançada)
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i> (Equalização de Histograma Adaptável Limitada por Contraste)
CNN	<i>Convolutional Neural Network</i> (Redes Neurais Convolucionais)
EHT	<i>Event Horizon Telescope</i> (Telescópio do Horizonte de Eventos)
ILSVRC	<i>ImageNet Large Scale Visual Recognition Challenge</i> (Desafio de reconhecimento visual em larga escala do ImageNet)
M	<i>Messier Catalogue</i> (Catalogo Messier)
MNIST	<i>Modified National Institute of Standards and Technology</i> (Instituto Nacional de Padrões e Tecnologia modificado)
NAS	<i>Neural Architecture Search</i> (Pesquisa de Arquitetura Neural)
NASA	<i>National Aeronautics and Space Administration</i> (Administração Nacional da Aeronáutica e Espaço)
NGC	<i>New General Catalogue</i> (Novo Catalogo Geral)
PR-AUC	<i>Precision-Recall</i> (Precisão-Revocação)
ReLU	<i>Rectified Linear Unit</i> (Unidade Linear Retificada)
ResNet	Residual Networks (Redes Residuais)
RGB	<i>Red-Green-Blue</i> (Vermelho-Verde-Azul)
RNA	<i>Artificial Neural Networks</i> (Redes Neurais Artificiais)
ROC-AUC	<i>Receiver Operating Characteristic</i> (Característica Operacional do Receptor)
SDSS	<i>Sloan Digital Sky Survey</i> (Buscador Sloan Digital do Céu)
TL	<i>Transfer Learning</i> (Aprendizado por Transferência)
UFC	Universidade Federal do Ceará

## SUMÁRIO

1	INTRODUÇÃO . . . . .	17
1.1	Objetivos . . . . .	18
1.1.1	<i>Objetivo Geral</i> . . . . .	18
1.1.2	<i>Objetivos Específicos</i> . . . . .	18
1.2	Organização do Trabalho . . . . .	18
2	FUNDAMENTAÇÃO TEÓRICA . . . . .	19
2.1	Galáxias . . . . .	19
2.1.1	<i>Classificação Morfológica das Galáxias</i> . . . . .	20
2.2	Deep Learning . . . . .	26
2.2.1	<i>Redes Neurais Artificiais</i> . . . . .	26
2.2.2	<i>Redes Neurais Convolucionais</i> . . . . .	28
2.2.3	<i>Transfer Learning</i> . . . . .	30
2.2.4	<i>Modelos de CNN Pré-Treinados</i> . . . . .	31
2.3	Métricas de Desempenho . . . . .	35
2.4	Exploração Digital do Espaço: O Sloan Digital Sky Survey e a Plataforma Galaxy Zoo . . . . .	40
2.4.1	<i>Sloan Digital Sky Survey</i> . . . . .	40
2.4.2	<i>Galaxy Zoo</i> . . . . .	42
2.5	Trabalhos Relacionados . . . . .	44
2.5.1	<i>Galaxy Classification: a deep learning approach for classifying Sloan Digital Sky Survey images</i> . . . . .	44
2.5.2	<i>Aplicação de deep learning para Classificação Morfológica de Galáxias</i> . .	46
3	METODOLOGIA . . . . .	47
3.1	Base de Dados . . . . .	47
3.2	Pré-processamento . . . . .	49
3.3	Validação Cruzada . . . . .	51
3.4	Etapa Experimental . . . . .	52
3.5	Ferramentas e técnicas utilizadas . . . . .	52
4	RESULTADOS E DISCUSSÕES . . . . .	54
4.1	Experimento I . . . . .	54

<b>4.2</b>	<b>Experimento II . . . . .</b>	<b>56</b>
<b>4.3</b>	<b>Experimento III . . . . .</b>	<b>58</b>
<b>4.4</b>	<b>Experimento IV . . . . .</b>	<b>62</b>
<b>4.5</b>	<b>Experimento V . . . . .</b>	<b>66</b>
<b>4.6</b>	<b>Experimento VI . . . . .</b>	<b>70</b>
<b>4.7</b>	<b>Experimento VII . . . . .</b>	<b>73</b>
<b>4.8</b>	<b>Experimento VIII . . . . .</b>	<b>76</b>
<b>4.9</b>	<b>Comparação dos melhores desempenhos . . . . .</b>	<b>80</b>
<b>5</b>	<b>CONCLUSÕES . . . . .</b>	<b>83</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>85</b>



# 1 INTRODUÇÃO

Na vasta extensão do nosso universo, os astrônomos utilizam técnicas de inteligência artificial para explorar e compreender corpos celestes distantes como asteróides, luas, planetas em outros sistemas solares e até mesmo as primeiras galáxias, situadas a cerca de treze bilhões de anos-luz<sup>1</sup> de distância. Telescópios localizados na superfície da Terra e no espaço produzem diariamente grandes volumes de dados, tornando particularmente difícil o registro e a categorização destes astros sem o uso de métodos computacionais avançados (SANTOS, 2024).

Um exemplo significativo foi a primeira imagem direta de um buraco negro, em 2019, capturada pelo projeto Event Horizon Telescope (EHT). Devido à impossibilidade técnica de obtenção dessa imagem por meio de um único telescópio, dados provenientes de diversos radiotelescópios foram integrados utilizando algoritmos avançados e aprendizado de máquina para reconstruir a imagem final. Esse método possibilitou a criação da famosa imagem da sombra do buraco negro supermassivo da galáxia Messier 87, localizada a aproximadamente 55 milhões de anos-luz (EHT COLLABORATION *et al.*, 2019).

O surgimento das Redes Neurais Convolucionais (CNNs) revolucionou o campo do processamento de imagens (VARGAS *et al.*, 2016). Inspiradas no sistema visual humano, as CNNs se destacam pelo reconhecimento eficiente de padrões, nos permitindo analisar imensos conjuntos de dados com uma precisão única, extraindo características cruciais para discernir assinaturas sutis de objetos celestes a partir de uma imagem de campo profundo, obtidas por meio de longa exposição. A aplicação de CNNs neste contexto não só agiliza a análise de vastos pacotes de dados, mas também aumenta a precisão da classificação, facilitando a descoberta e caracterização destes objetos (LECUN *et al.*, 2015).

Neste trabalho aplicamos as CNNs na classificação morfológica de Galáxias da base de dados *Galaxy10 SDSS Dataset* (GALAXY10 SDSS DATASET, 2024), comparando a eficácia dos modelos: *Gharat & Dandawate* (2022), *EfficientNetV2-M* (TAN; LE, 2021), *DenseNet121* (HUANG *et al.*, 2017), *ResNet50* (HE *et al.*, 2016), *AlexNet* (KRIZHEVSKY *et al.*, 2012), *VGG16* e *VGG19* (SIMONYAN; ZISSERMAN, 2015).

---

<sup>1</sup> O ano-luz é uma unidade de comprimento usada para expressar distâncias astronômicas e equivale a cerca de 9,46 trilhões de quilômetros.

## 1.1 Objetivos

### 1.1.1 *Objetivo Geral*

Investigar a aplicação de modelos de aprendizagem profunda na classificação de galáxias em imagens da base de dados *Galaxy10 SDSS Dataset*

### 1.1.2 *Objetivos Específicos*

- Pesquisar bibliografia relevante sobre aplicação de redes convolucionais na classificação de objetos astronômicos;
- Analisar classificação de imagens para diferentes grupos de galáxias da base de dados *Galaxy10 SDSS Dataset*;
- Comparar a eficácia dos modelos: Gharat & Dandawate, *EfficientNetV2-M*, *DenseNet121*, *ResNet50*, *AlexNet*, *VGG16*, e *VGG19*;
- Avaliar como as estratégias de aumento de dados *DataGen*, *MixUp* e *CutMix* e aumento de contraste *CLAHE* (*Contrast Limited Adaptive Histogram Equalization*) impactam no desempenho dos modelos do item anterior.

## 1.2 Organização do Trabalho

Os demais capítulos seguem a seguinte estrutura:

- No Capítulo 2, trazemos o embasamento teórico e a revisão das principais obras, autores e conceitos relacionados ao tema do trabalho;
- No Capítulo 3, descrevemos os métodos de pesquisa utilizados, como coleta e análise de dados, técnicas, amostragem, entre outras informações relevantes;
- No Capítulo 4, apresentamos os resultados obtidos durante o desenvolvimento da pesquisa;
- Finalmente no Capítulo 5, discutimos as principais conclusões sobre os resultados obtidos na presente pesquisa e propomos algumas perspectivas para a continuidade deste trabalho.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, abordaremos os fundamentos para a compreensão deste estudo. Na Seção 2.1, exploraremos o conceito de galáxias, suas subdivisões e classificações de acordo com a Sequência de Hubble. Na Seção 2.2, falaremos de *Deep Learning* com foco em CNNs, bem como suas características, vantagens, subdivisões para treinamento, *Transfer Learning*, principais métricas de desempenho e estratégias de balanceamento de dados utilizadas neste trabalho. Na Seção 2.4, exploraremos as ferramentas empregadas para a coleta, distribuição e disponibilização de nossos dados observacionais. Finalmente, na Seção 2.5 revisaremos alguns trabalhos relacionados que complementam e contextualizam nossa pesquisa.

### 2.1 Galáxias

As galáxias são vastas ilhas cósmicas na estrutura do universo, contendo desde bilhões a trilhões de estrelas, juntamente com gás e poeira interestelar (NASA, 2024a). Elas existem em vários formatos, tamanhos e composições, preenchendo todo o nosso cosmos.

A maioria destes astros possuem entre 10 e 13,6 bilhões de anos. Algumas são quase tão antigas como o próprio Universo, que se formou há cerca de 13,8 bilhões de anos. Acredita-se que a galáxia mais jovem conhecida se formou há apenas 500 milhões de anos (NASA, 2024a). Elas costumam organizar-se em grupos e aglomerados de cerca de 100 membros, unidos por sua gravidade mútua. Estes, também pertencem a estruturas maiores, chamadas superaglomerados, podendo conter milhares de galáxias e formando as maiores estruturas do universo, como podemos observar na Figura 1.

As galáxias não são entidades estáticas, elas evoluem ao longo do tempo através de processos como formação de estrelas, explosões de supernovas e interações com outras galáxias. Seu estudo nos fornece informações cruciais sobre a história e a dinâmica do universo, ajudando-nos a compreender sua estrutura e evolução em maiores escalas (FILHO; SARAIVA, 2014). Dentro das galáxias, também existem fenômenos que influenciam em sua evolução, como buracos negros supermassivos em seus núcleos, com massas milhões de vezes maiores que a do Sol.

Figura 1 – Primeira imagem de campo profundo processada pelo Telescópio Espacial James Webb, revelando a vasta quantidade de galáxias distantes.



Fonte: <https://www.nationalgeographicbrasil.com/>

### 2.1.1 Classificação Morfológica das Galáxias

As galáxias podem ser classificadas com base em sua morfologia em cinco principais categorias: Elípticas, Lenticulares, Espirais, Espirais Barradas e Irregulares. Com base nessas subdivisões, constrói-se o esquema evolutivo, representado nas Figuras 2, conhecido como Sequência de Hubble (FILHO; SARAIVA, 2014).

Essa classificação foi proposta pelo astrônomo Edwin Hubble, que dedicou sua vida ao estudo das galáxias e responsável por provar a existência de objetos localizados além da Via Láctea, demonstrando que o universo é formado por diversas galáxias e aglomerados galácticos. Antes dessa descoberta, acreditava-se que tudo o que era observado pertencia exclusivamente à nossa galáxia, e muitos desses objetos eram erroneamente identificados como aglomerados globulares<sup>2</sup> (FILHO; SARAIVA, 2014).

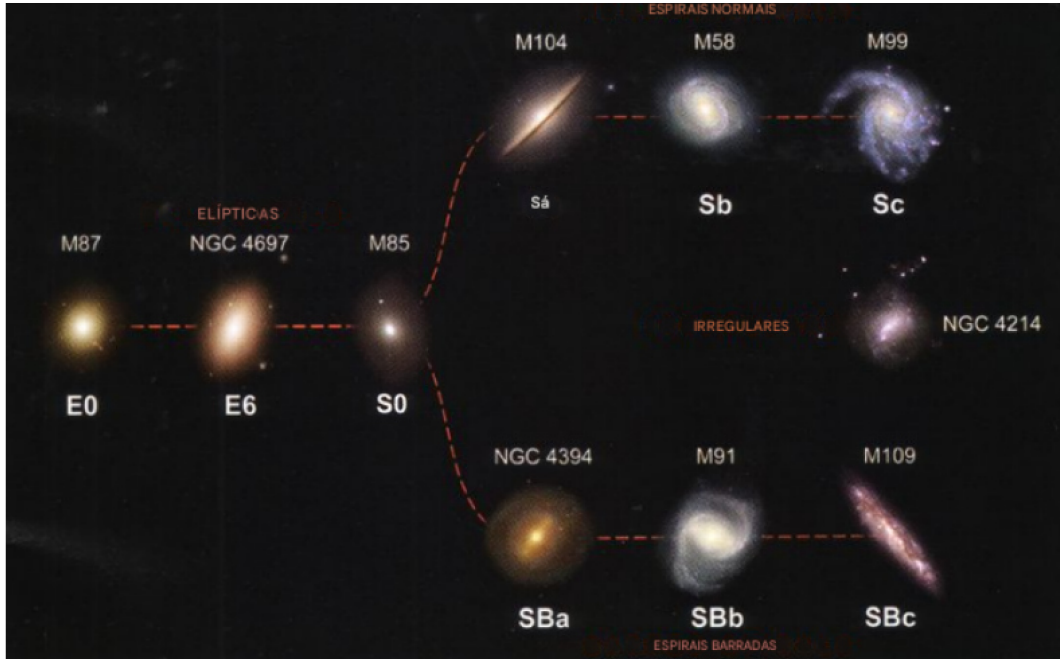
Para estabelecer essa distinção, Hubble utilizou a correlação direta entre a luminosidade e o período de pulsação de uma cefeida variável clássica<sup>3</sup> (descoberta de Henrietta Swan Leavitt em 1908) para dimensionar distâncias galácticas e extragalácticas (NASA, 2023).

As galáxias elípticas apresentam uma forma elíptica mais arredondada, sem braços espirais distintos (NASA, 2023). Eles variam amplamente em tamanho, desde pequenas elípticas

<sup>2</sup> Aglomerado globular é a denominação dada a um tipo de aglomerado estelar cujo formato aparente é esférico e cujo interior é muito denso e rico em estrelas antigas, mantidas juntas pela ação da gravidade. Estrutura milhares de vezes menor que uma galáxia.

<sup>3</sup> Cefeidas são estrelas instáveis muito maiores e mais brilhantes do que o Sol, que apresentam uma variação de brilho regular, uma vez medida esta variação, podemos utiliza-la para determinar distâncias astronômicas.

Figura 2 – Sequência de Hubble para classificação das galáxias com base em sua morfologia: elípticas, espirais, espirais barradas e irregulares.



Fonte: <https://spacetoday.com.br/dando-sentido-para-a-sequencia-de-hubble/>

anãs - em maior número - até enormes elípticas supergigantes encontradas nos centros de aglomerados de galáxias - mais raras. As galáxias elípticas tendem a ter populações de estrelas mais antigas e menos material interestelar para a formação de novas estrelas em comparação com as galáxias espirais (FILHO; SARAIVA, 2014).

Hubble subdividiu as galáxias elípticas em classes de  $E0$  a  $E7$ , conforme seu grau de achatamento (Figura 3), definido pela equação:

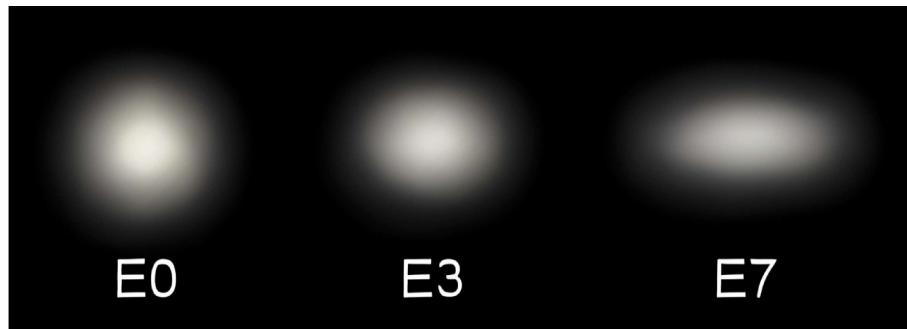
$$n = 10 \times \left( \frac{a-b}{a} \right) \quad (2.1)$$

onde  $a$  representa o eixo maior da elipse e  $b$  o eixo menor. Uma galáxia classificada como  $E0$  tem uma aparência quase redonda, podendo ser de fato esférica ou apenas parecer assim dependendo do ângulo de observação. Já uma  $E7$  é visivelmente mais alongada e achatada, tendo uma forma mais ovalada. Essa classificação é baseada na aparência da galáxia do ponto de vista da Terra, não em sua forma tridimensional real (FILHO; SARAIVA, 2014).

As galáxias espirais tem braços espirais distintos, que se estendem a partir de uma protuberância central. Esses braços são regiões ricas em gás, poeira e estrelas jovens, onde ocorrem intensos processos de formação estelar (NASA, 2023).

Estas podem ser subdividida em três componentes principais: o bojo central, o disco e o halo (Figura 4). Estas componentes são definidas:

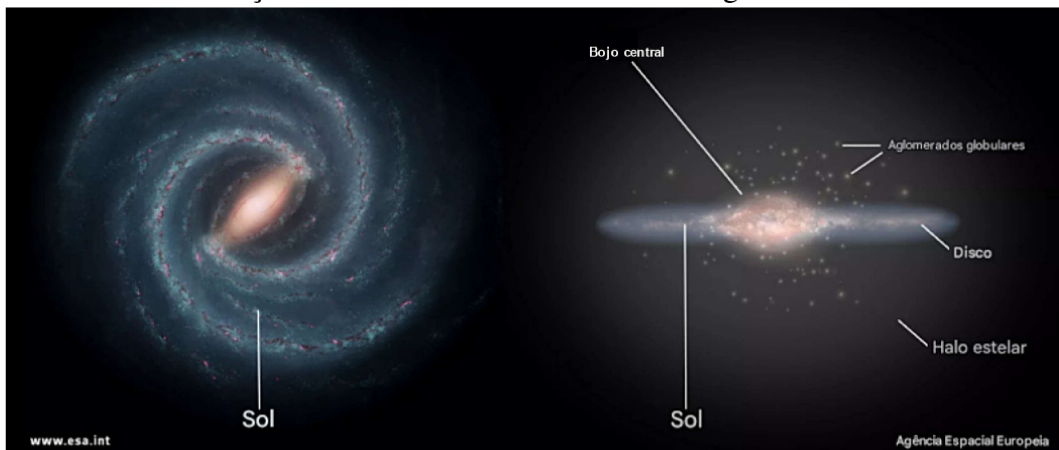
Figura 3 – Morfologia das galáxias elípticas com diferentes graus de achatamento.



Fonte: <https://www.galeriadometeorito.com/p/galaxias.html>

- **Bojo central:** Região esférica ou ovalada no centro da galáxia, composta majoritariamente por estrelas mais antigas e, possivelmente, um buraco negro supermassivo.
- **Disco:** Contém os braços espirais e abriga estrelas jovens, gás e poeira. Região da galáxia onde predominam os processos de formação estelar.
- **Halo:** Camada externa extensa e menos densa, cujas estruturas centrais são compostas por aglomerados globulares e matéria escura<sup>4</sup>, crucial para a massa e gravidade da galáxia.

Figura 4 – Representação esquemática das subdivisões da Via Láctea, destacando a localização do nosso Sistema Solar no disco galáctico.



Fonte: <https://astroPontos.org/2019/09/17/a-formacao-da-via-lactea/>

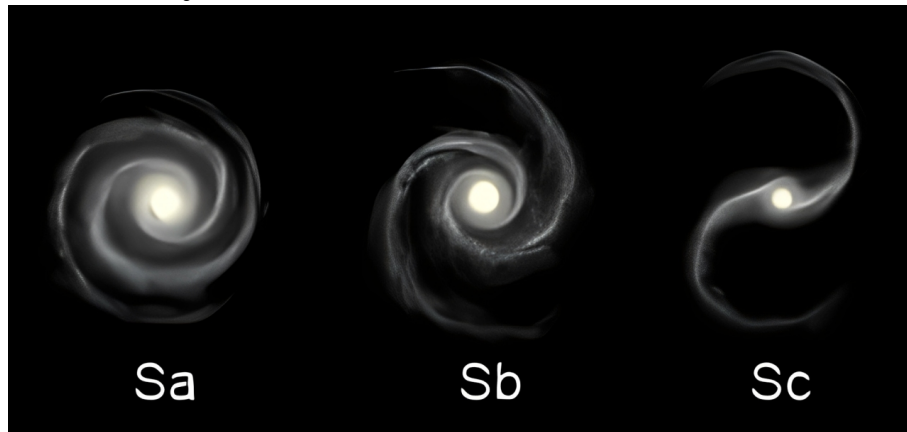
As galáxias espirais podem ainda ser subdivididas em dois tipos principais: espirais sem barra e espirais barradas. Nas espirais sem barra, os braços emergem diretamente do núcleo da galáxia, enquanto nas espirais barradas, há uma barra central de estrelas que atravessa o bojo galáctico, e os braços espirais se originam das extremidades dessa estrutura. A presença ou ausência dessa barra pode influenciar processos internos da galáxia, como a redistribuição de gás

<sup>4</sup> Matéria escura é um tipo de matéria que não emite, absorve ou reflete luz, tornando-a invisível e detectável apenas através dos efeitos gravitacionais que exerce sobre a matéria visível e a radiação cósmica.

e poeira, afetando a formação de estrelas e a evolução do núcleo galáctico (FILHO; SARAIVA, 2014).

As galáxias espirais são também classificadas em diferentes subtipos com base no grau de abertura de seus braços espirais e no tamanho do bojo central (SANTOS, 2024). Essa subdivisão segue a sequência Sa, Sb, e Sc (Figura 5).

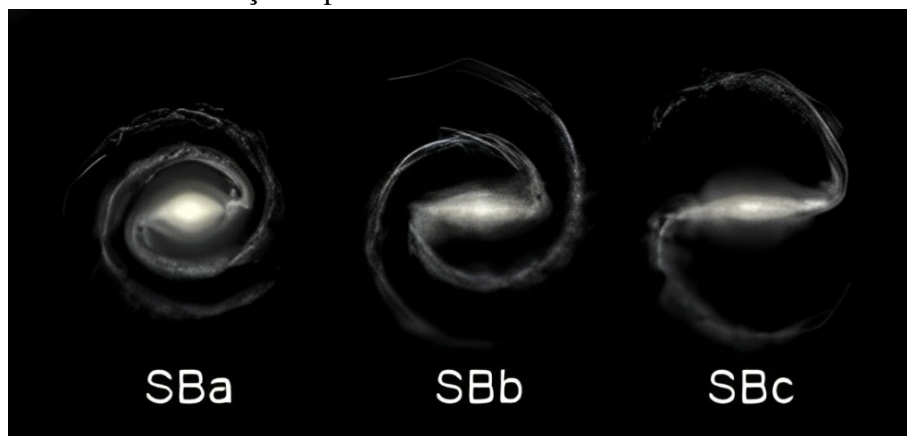
Figura 5 – Morfologia das galáxias espirais, nos seus três subtipos, com base no grau de abertura dos braços espirais e no tamanho do bojo central.



Fonte: <https://www.galeriadometeorito.com/p/galaxias.html>

As galáxias espirais barradas são classificadas de acordo com a proeminência de sua barra central e a abertura de seus braços espirais, seguindo a sequência SBa, SBb e SBc, como apresentado na Figura 6 (SANTOS, 2024). Evidências observacionais indicam que a Via Láctea é uma galáxia espiral barrada, pertencente ao tipo SBc.

Figura 6 – Morfologia das galáxias espirais barradas, em seus três subtipos, com base na proeminência da barra central e na abertura dos braços espirais.



Fonte: <https://www.galeriadometeorito.com/p/galaxias.html>



As galáxias lenticulares apresentam uma estrutura semelhante às galáxias espirais - possuindo um núcleo, um disco e um halo - porém não possuem braços espirais. Na classificação de Hubble, são representadas pela designação *S0*, onde o número 0 faz referência à sua forma, que lembra uma lente (SANTOS, 2024).

Essas galáxias são consideradas uma forma de transição entre galáxias elípticas e espirais, pois compartilham características de ambas (Figura 7). Assim como as galáxias elípticas, elas contêm pouca quantidade de gás e poeira, apresentando uma predominância de estrelas antigas. No entanto, sua estrutura em disco as aproxima das galáxias espirais. As lenticulares podem ser classificadas como barradas *SB0* ou não barradas *S0*, dependendo da presença ou ausência de uma barra central (FILHO; SARAIVA, 2014).

Figura 7 – Representação da morfologia de uma galáxia lenticular (*S0*) à esquerda e a galáxia *M104*, um exemplo real desse tipo de galáxia, à direita.



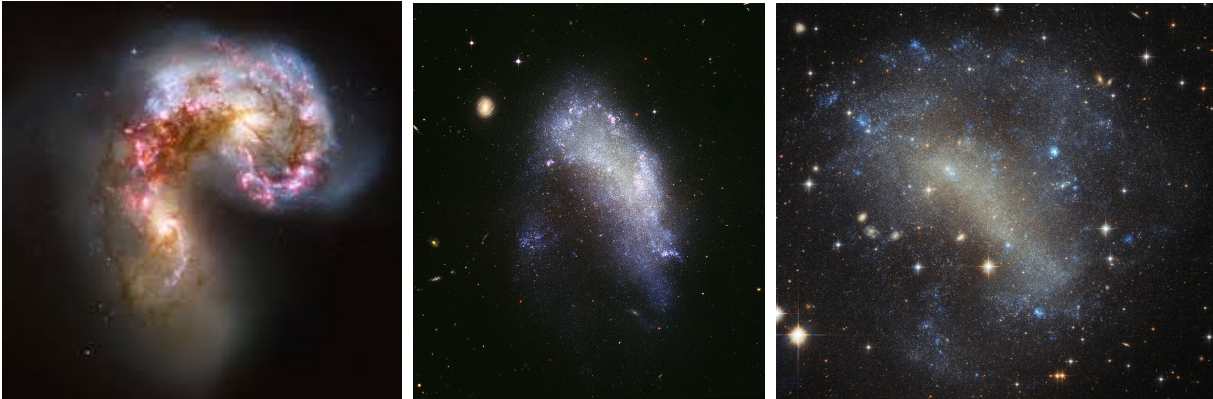
Fonte: <https://www.galeriadometeorito.com/p/galaxias.html>

As galáxias irregulares não têm uma forma definida (NASA, 2024b). Eles podem ter uma aparência caótica, sem nenhuma simetria particular (Figura 8). Galáxias irregulares geralmente resultam de interações gravitacionais ou fusões entre galáxias. Eles podem conter bolsões de intensa formação estelar, juntamente com nuvens de poeira e gás.

Além da diferenciação visual, cada classe também diz respeito a características de propriedades únicas de cada tipo de galáxia. Enquanto as espirais possuem gás e estrelas jovens, as elípticas são dominadas por estrelas antigas e pouca poeira. As irregulares exibem características variadas, com formação estelar ativa. A Tabela 1 resume as principais características das galáxias elípticas, espirais e irregulares (FILHO; SARAIVA, 2014).



Figura 8 – Galáxias irregulares. Da esquerda para direita: *NGC4038*, *NGC1427* e *IC4710*, caracterizadas pela ausência de uma estrutura definida e pela intensa formação estelar.



Fonte: <https://www.galeriadometeorito.com/p/galaxias.html>

Tabela 1 – Comparação das propriedades das galáxias espirais, elípticas e irregulares com relação a massa, tamanho, luminosidade, composição estelar, presença de gás e poeira, cor e idade das estrelas.

Propriedade	Espirais	Elípticas	Irregulares
Massa ( $M$ )	$10^9$ a $10^{12}$	$10^5$ a $10^{13}$	$10^8$ a $10^{11}$
Diâmetro ( $10^3 pc$ )	5 - 30	1 - 1000	1 - 10
Luminosidade ( $L$ )	$10^8$ a $10^{11}$	$10^6$ a $10^{12}$	$10^7$ a $2 \times 10^9$
População Estelar	Velha e jovem	Velha	Velha e jovem
Tipo Espectral	A a K	G a K	A a F
Gás	Bastante	Muito pouco	Bastante
Poeira	Bastante	Muito pouca	Varia
Cor	Azulada no disco e amarelada no bojo	Amarelada	Azulada
Estrelas mais velhas	$10^{10}$ anos	$10^{10}$ anos	$10^{10}$ anos
Estrelas mais jovens	Recentes	$10^{10}$ anos	Recentes

Fonte: Adaptado de Santos (2024)

## 2.2 Deep Learning

*Deep Learning*, ou aprendizado profundo, é uma técnica de aprendizado de máquina capaz de identificar padrões complexos em grandes volumes de dados utilizando redes neurais artificiais que processam dados em múltiplas camadas. Esta abordagem avançada aprende características automaticamente, sem necessidade de engenharia manual de atributos, beneficiando-se com o aumento dos dados disponíveis (LECUN *et al.*, 2015).

Ao contrário de outras técnicas de Machine Learning, que podem não escalar bem, o *Deep Learning* é eficaz em resolver desafios complexos, como interpretação de imagens e compreensão da fala, otimizando processos em áreas diversas como saúde, finanças, segurança, automação e astronomia (HAYKIN, 2000).

### 2.2.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no funcionamento do cérebro humano (LECUN *et al.*, 2015). Seu princípio fundamental está na plasticidade cerebral, a capacidade de adaptação e reorganização diante de mudanças ambientais. Essa propriedade permite a formação de novas conexões entre neurônios, facilitando o aprendizado (HAYKIN, 2000).

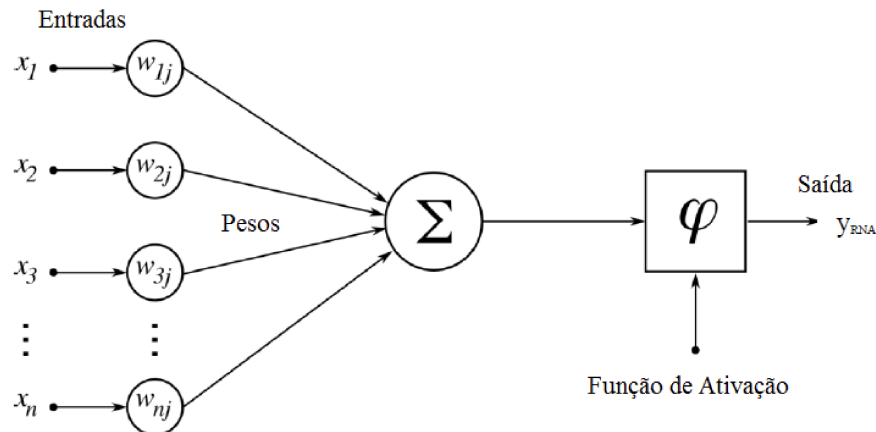
O neurônio artificial é a unidade básica das redes neurais (Figura 9). Ele recebe múltiplas entradas ponderadas, aplica uma função de ativação e gera uma saída (HAYKIN, 2000). Seu funcionamento pode ser descrito pela equação:

$$z = \sum w_i x_i + b \quad (2.2)$$

onde  $x_i$  representa as entradas,  $w_i$  os pesos associados a cada entrada,  $b$  um viés e  $z$  a soma ponderada dos valores. Uma função de ativação, como *ReLU* (*Rectified Linear Unit*) ou sigmoide, adiciona não-linearidade, permitindo que a rede aprenda padrões complexos. Assim como os neurônios biológicos transmitem sinais através das sinapses, os neurônios artificiais ajustam seus pesos para otimizar o desempenho do modelo (HAYKIN, 2000).

Para que essas redes aprendam de maneira eficiente, utiliza-se o algoritmo de *back-propagation*, responsável por ajustar os pesos das RNAs e minimizar o erro entre a saída prevista e a real (LECUN *et al.*, 2015). O aprendizado ocorre em duas etapas, como apresentado na

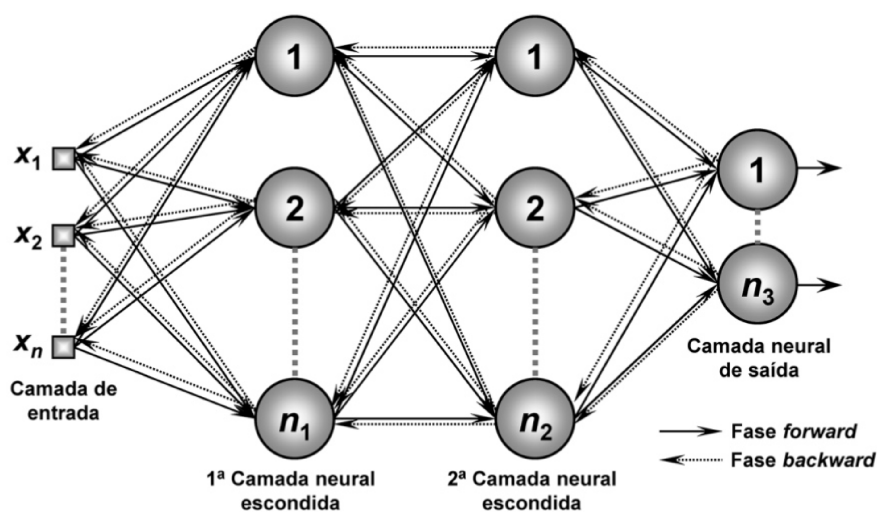
Figura 9 – O neurônio artificial, onde as entradas são ponderadas por pesos sinápticos e somadas. A função de ativação, que determina a saída da equação, adicionando não linearidade ao neurônio.



Fonte: Haykin (2000)

Figura 10. Na fase *forward* (propagação direta), os dados percorrem a rede camada por camada até gerar uma previsão. Enquanto na fase *backward* (retropropagação do erro), o erro da previsão é calculado e propagado de trás para frente, ajustando os pesos dos neurônios conforme o gradiente da função de perda (HAYKIN, 2000).

Figura 10 – Funcionamento do algoritmo backpropagation em uma rede neural, destacando a propagação direta (*forward*) dos dados até a saída e a retropropagação (*backward*) do erro, que ajusta os pesos sinápticos para otimizar o aprendizado da rede.



Fonte: Haykin (2000)

Para otimizar esse ajuste, o *backpropagation* utiliza a Regra da Cadeia do cálculo diferencial, garantindo que cada camada aprenda representações relevantes dos padrões nos

dados. O treinamento das redes é realizado com grandes volumes de dados, e a minimização do erro ocorre por meio de algoritmos de otimização baseados em gradiente descendente (FILHO, 2022). Dessa forma, o modelo ajusta seus pesos e vieses de maneira eficiente, melhorando a precisão das previsões.

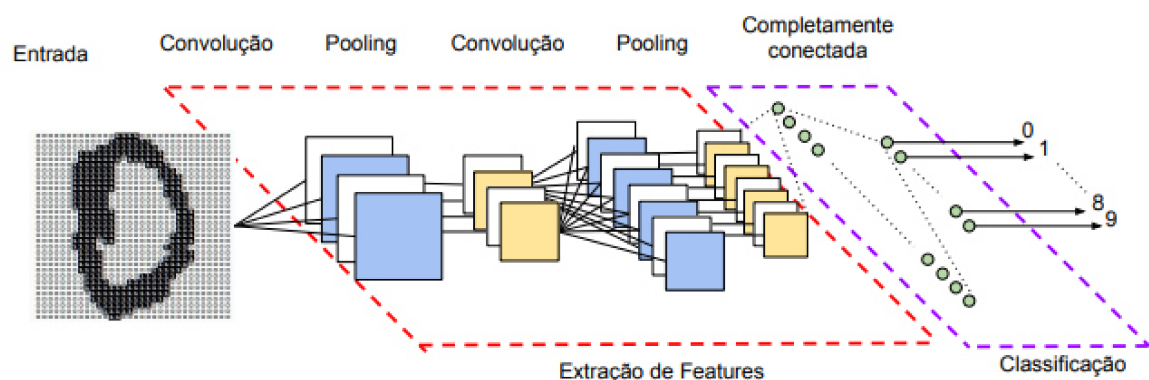
### 2.2.2 Redes Neurais Convolucionais

As CNNs são uma classe de modelos de aprendizagem profunda para processar e analisar dados visuais, como imagens e vídeos. Elas são inspirados na organização do córtex visual humano e são excepcionalmente poderosas para tarefas como detecção de objetos, reconhecimento e segmentação de imagens (LECUN *et al.*, 2015).

Uma das principais vantagens das CNNs é a sua habilidade de aprender automaticamente representações hierárquicas de recursos a partir de dados brutos, eliminando a necessidade de engenharia manual de recursos. Isso as torna extremamente eficazes em tarefas que envolvem dados visuais complexos, levando a avanços em áreas como visão computacional, análise de imagens médicas, direção autônoma e muito mais (LECUN *et al.*, 2015).

Para Filho (2024) uma CNN pode ser subdividida da seguinte forma, como apresentado na Figura 11. O processo envolve camadas de convolução e pooling, para a extração de características das imagens, seguidas por camadas totalmente conectadas, responsáveis pela classificação dos dados.

Figura 11 – Estrutura de uma rede neural convolucional e suas camadas. O fluxo de processamento ocorre da entrada até a classificação final.

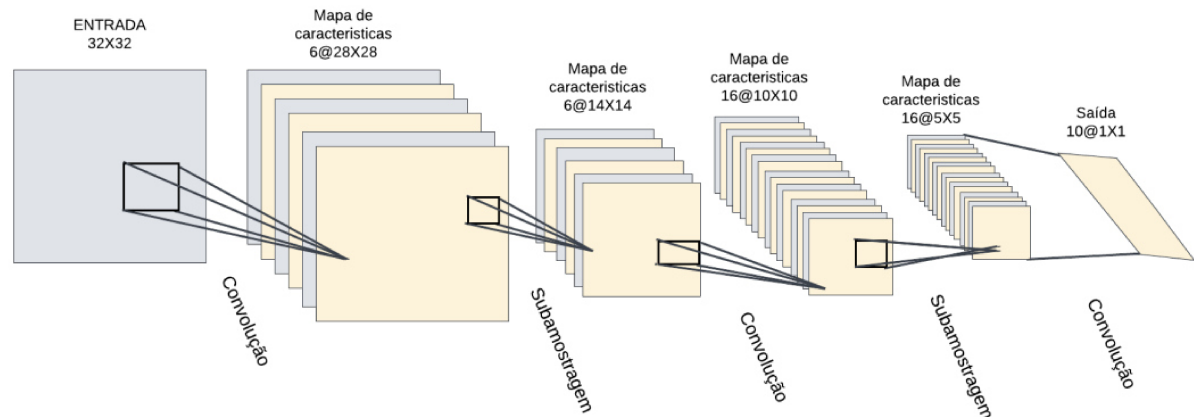


Fonte: LeCun et al. (2015)

A camada convolutiva é responsável em aplicar um conjunto de filtros (*kernels*) sob os dados de entrada (SANTOS, 2024). Cada filtro executa uma operação de convolução

deslizando pelas imagens de entrada e calculando um produto escalar entre seus pesos e os valores dos pixels que eles sobrepõem. Esta operação captura efetivamente padrões espaciais, como bordas, texturas e outros recursos em diferentes locais das figuras (FILHO, 2024). Por fim, produz-se um valor que representa uma característica local da imagem, como apresentado na Figura 12.

Figura 12 – Camadas convolutivas em uma rede neural no processamento de uma imagem de  $32 \times 32$  pixels. Sucessivas camadas de convolução e subamostragem reduzem a dimensionalidade e extraem características.



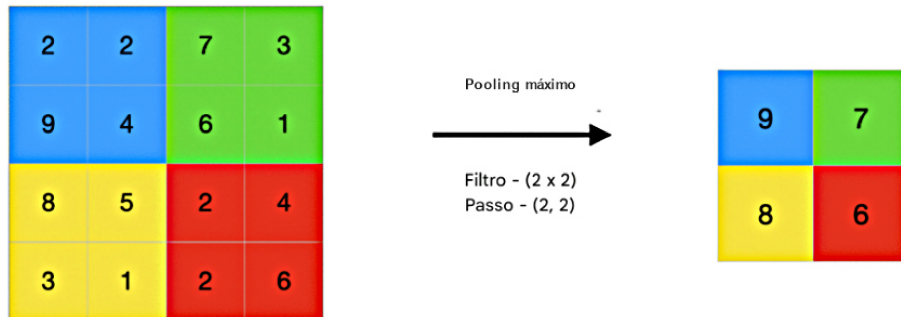
Fonte: Santos (2024)

As camadas de *pooling* são frequentemente inseridas após as camadas convolucionais para reduzir as dimensões espaciais dos mapas de recursos, ao mesmo tempo que retém informações importantes (SANTOS, 2024). O *pooling* máximo (*max pooling*), por exemplo, seleciona o valor máximo de cada região local do mapa de recursos, reduzindo efetivamente a resolução (Figura 13).

Optamos pelo uso de *pooling* em vez da análise de componentes principais (Principal Component Analysis - PCA) em nossa CNN devido à capacidade desta técnica de preservar a localidade das características espaciais, que são essenciais para tarefas de visão computacional. Em contraste, o PCA, embora eficaz na redução da dimensionalidade, não mantém essas características locais e requer mais recursos computacionais. Além disso, técnicas como o *max pooling* reduzem eficientemente a resolução enquanto preservam informações cruciais, facilitando o aprendizado de invariâncias e melhorando a eficiência computacional do modelo (SANTOS, 2024).

Após várias camadas convolucionais e de *pooling*, os recursos de alto nível são representados na forma de um vetor unidimensional e alimentados em uma ou mais camadas totalmente conectadas (*dense layers*), como apresentado na Figura 14. Essas camadas realizam

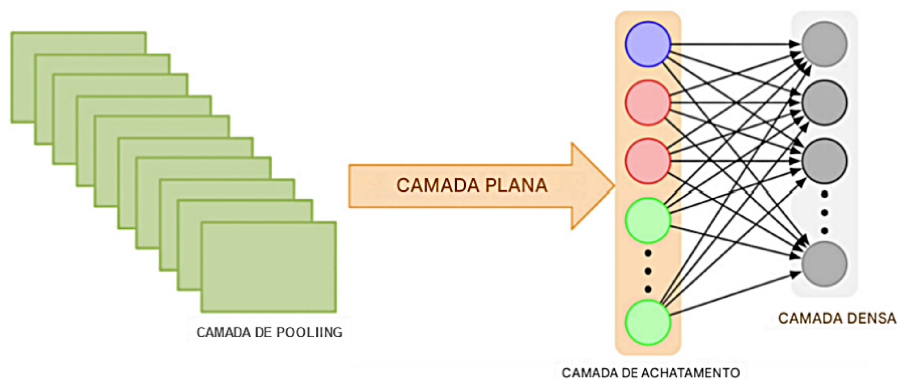
Figura 13 – Funcionamento da camada de pooling em redes neurais convolucionais. O exemplo demonstra a aplicação do *Max Pooling* com um filtro de tamanho  $2 \times 2$ , onde o valor máximo de cada região filtrada é selecionado para compor a nova matriz reduzida.



Fonte: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>

tarefas de classificação ou regressão aprendendo a mapear os recursos extraídos para os rótulos de saída desejados (SANTOS, 2024).

Figura 14 – Operação de *flattening* em redes neurais convolucionais. Esse processo transforma a saída multidimensional da camada de *pooling* em um vetor unidimensional para os dados serem processados por camadas densas totalmente conectadas.



Fonte: Santos (2024)

### 2.2.3 Transfer Learning

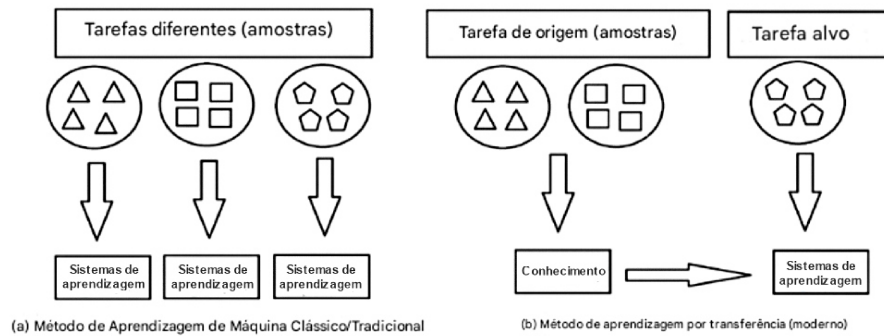
Para Hosna et al. (2022), *Transfer Learning (TL)* é uma abordagem de aprendizado de máquina que permite que um modelo treinado em um determinado conjunto de dados sejam reutilizados para resolver um problema relacionado. Diferente dos métodos tradicionais, que assumem que o treinamento e os testes ocorrem sob a mesma distribuição de dados, essa técnica possibilita a adaptação de conhecimento previamente adquirido, tornando o processo de



aprendizado mais eficiente.

A TL reduz a necessidade de grandes volumes de dados rotulados e acelera o treinamento dos modelos. Essa abordagem se destaca especialmente em cenários onde há pouca disponibilidade de dados, aproveitando informações de domínios similares para obter melhores resultados e melhorar o desempenho dos algoritmos. A Figura 15, ilustra o funcionamento deste modelo de aprendizagem.

Figura 15 – Comparação entre Aprendizado de Máquina Tradicional e Aprendizado por Transferência. No método tradicional (a), diferentes tarefas utilizam sistemas de aprendizado independentes. Já no Aprendizado por Transferência (b), o conhecimento adquirido em uma tarefa fonte é reutilizado para aprimorar o desempenho em uma tarefa alvo, tornando o treinamento mais eficiente.



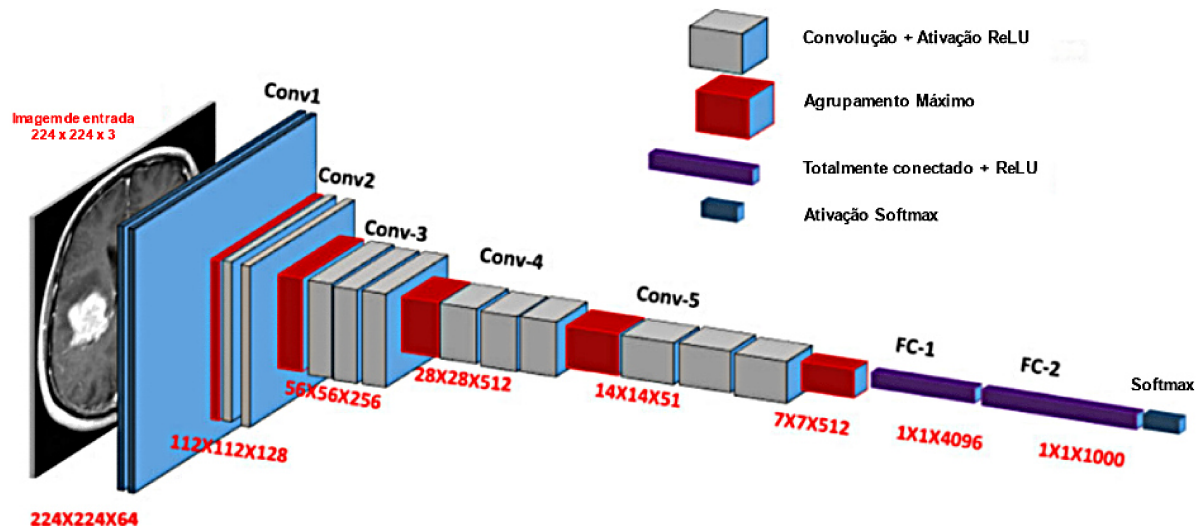
Fonte: Hosna et al. (2022)

#### 2.2.4 Modelos de CNN Pré-Treinados

Neste capítulo apresentaremos as seis arquiteturas de modelos pré-treinados que serão adotadas na etapa experimental deste trabalho, são elas: *VGG16*, *VGG19*, *ResNet50*, *DenseNet121*, *EfficientNetV2-M* e *AlexNet*.

A *VGG16* (SIMONYAN; ZISSERMAN, 2015) é uma arquitetura desenvolvida para aumentar a profundidade das redes convolucionais e melhorar o desempenho em tarefas de reconhecimento de imagens. Essa rede possui 16 camadas treináveis, incluindo 13 camadas convolucionais com filtros pequenos de  $3 \times 3$  e três camadas totalmente conectadas, como apresentado na Figura 16. A rede utiliza camadas de *max pooling* ( $2 \times 2$ ) após conjuntos de camadas convolucionais para reduzir a dimensionalidade dos mapas de características. Todas as camadas escondidas usam a função de ativação *ReLU*, e a camada final emprega *Softmax* para classificação.

Figura 16 – Ilustração da arquitetura VGG16.

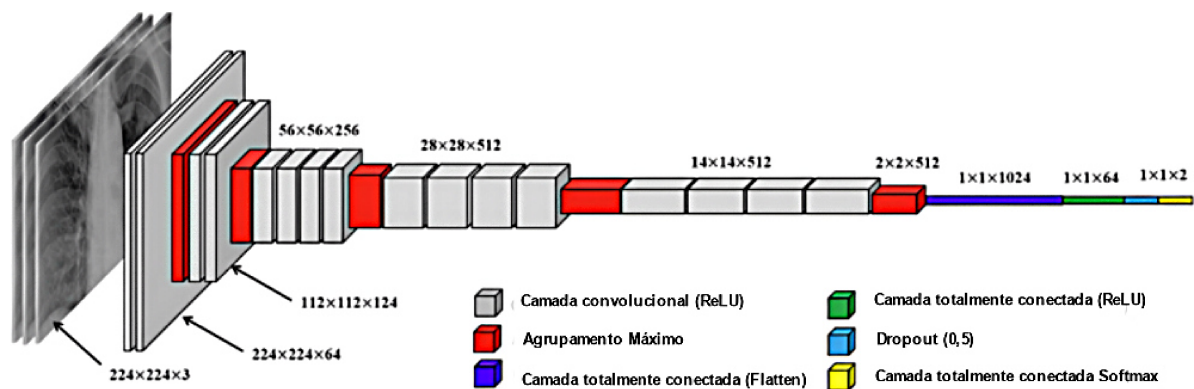


Fonte: <https://medium.com/@siddheshb008/vgg-net-architecture-explained-71179310050f>

A *VGG19* (SIMONYAN; ZISSERMAN, 2015) é uma variação mais profunda da *VGG16*, composta por 19 camadas treináveis (16 convolucionais e três totalmente conectadas), foi projetada para melhorar a precisão na classificação de imagens em larga escala.

A estrutura da *VGG19* é ilustrada na Figura 17. Essa é composta por pequenos filtros convolucionais  $3 \times 3$  em todas as camadas, mantendo um *stride* de um e preservando a resolução espacial com preenchimento de um pixel. A rede inclui cinco camadas de *max pooling* ( $2 \times 2$ , *stride* 2) para reduzir a dimensionalidade dos mapas de características. Após a extração de características convolucionais, há três camadas totalmente conectadas, onde as duas primeiras possuem 4096 neurônios e a última realiza a classificação final com *Softmax*. Todas as camadas ocultas utilizam a função de ativação *ReLU*, garantindo melhor aprendizado de características profundas (SIMONYAN; ZISSERMAN, 2015).

Figura 17 – Ilustração da arquitetura VGG19.



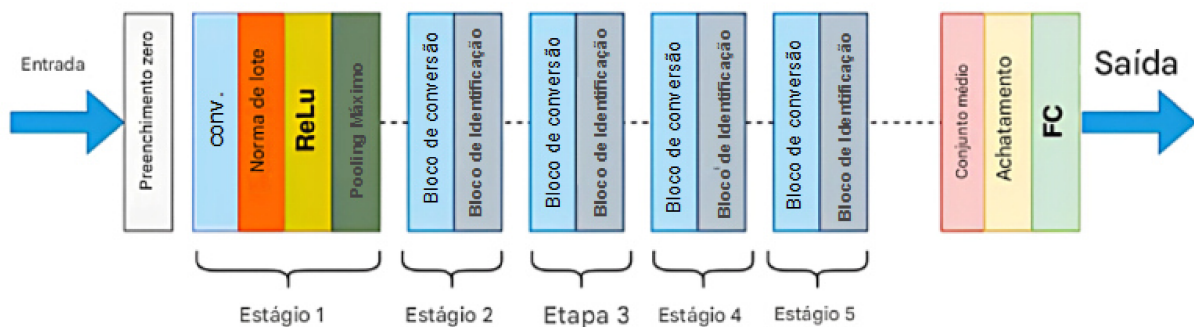
Fonte: <https://medium.com/@siddheshb008/vgg-net-architecture-explained-71179310050f>



A *ResNet50* (HE *et al.*, 2016), faz parte de uma família de redes convolucionais denominada ResNet (*Residual Network*). Essas foram projetadas para superar desafios no treinamento de redes muito profundas, introduzindo o conceito de aprendizado residual.

A arquitetura da *ResNet50*, ilustrada na Figura 18, possui 50 camadas treináveis, organizadas em blocos residuais. Esses blocos incluem atalhos (*skip connections*) que permitem a propagação direta das informações entre camadas, evitando problemas de perda de gradiente. A arquitetura é composta por uma camada convolucional inicial de  $7 \times 7$ , seguida por uma sequência de blocos residuais compostos por convoluções  $1 \times 1$ ,  $3 \times 3$  e  $1 \times 1$ , maximizando a eficiência da rede. O uso desses blocos permite que a rede aprenda transformações residuais em vez de tentar modelar diretamente a saída desejada, facilitando a otimização (HE *et al.*, 2016).

Figura 18 – Ilustração da arquitetura ResNet50.



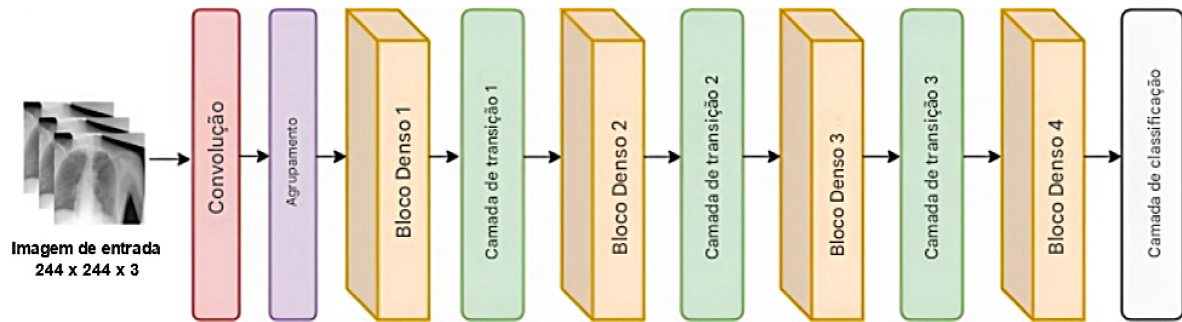
Fonte: <https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f>

A *DenseNet121* (HUANG *et al.*, 2017) introduziu uma nova abordagem de conexão entre camadas, onde cada camada recebe como entrada os mapas de características de todas as camadas anteriores, formando conexões densas que melhoram a propagação da informação e o fluxo de gradientes.

Sua arquitetura possui 121 camadas treináveis, organizadas em blocos densos, intercalados com camadas de transição que realizam convolução  $1 \times 1$  e *pooling*  $2 \times 2$  para reduzir a dimensionalidade, como apresentado na Figura 19. A principal inovação da *DenseNet* em relação a outras arquiteturas, como *ResNet*, é o uso de concatenação dos mapas de características em vez de soma residual, permitindo um melhor reaproveitamento das informações extraídas ao longo da rede. Isso resulta em um modelo mais eficiente em termos de parâmetros, melhorando o desempenho com um número menor de filtros (HUANG *et al.*, 2017).

A *EfficientNetV2-M* (TAN; LE, 2021), apresentada na Figura 20, foi projetada para melhorar a eficiência de treinamento e reduzir o número de parâmetros em comparação com

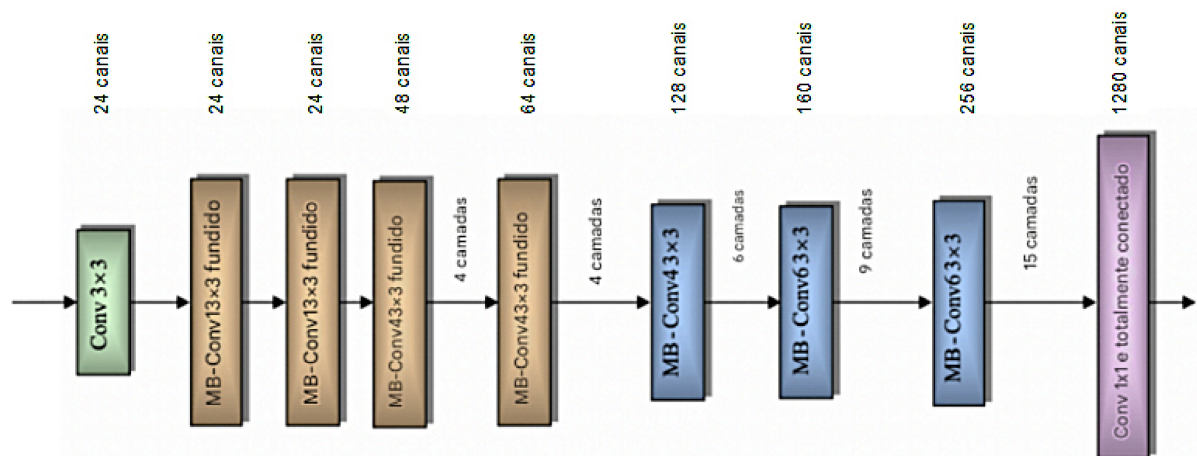
Figura 19 – Ilustração da arquitetura *DenseNet-121*.



Fonte: <https://link.springer.com/article/10.1007/s11277-024-11467-8/figures/2>

modelos anteriores, mantendo um alto desempenho em classificação de imagens.

Figura 20 – Ilustração da arquitetura *EfficientNetV2-M*.



Fonte: [https://www.researchgate.net/figure/Architecture-of-efficientNetv2\\_fig2\\_369050073](https://www.researchgate.net/figure/Architecture-of-efficientNetv2_fig2_369050073)

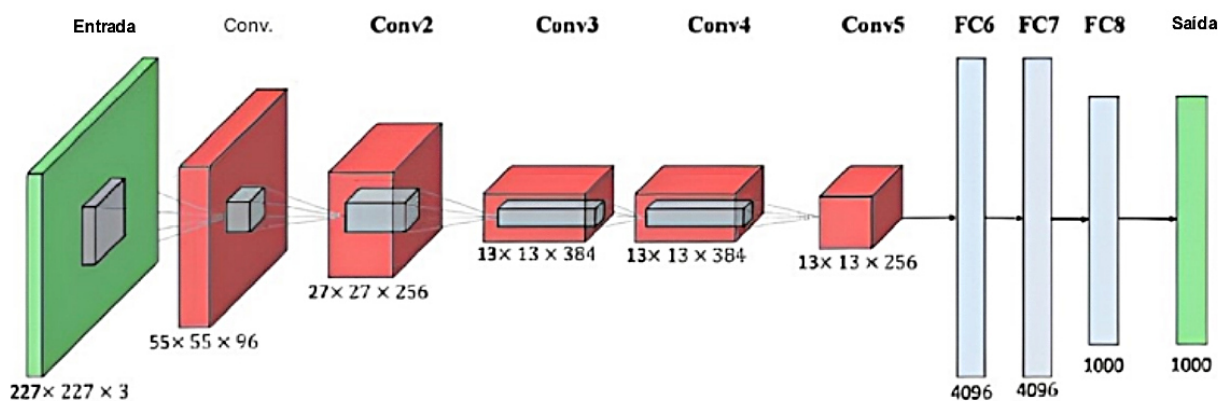
Este modelo utiliza uma combinação de *Neural Architecture Search (NAS)* e escalonamento composto para otimizar a relação entre velocidade de treinamento e eficiência de parâmetros. Diferente da versão anterior (*EfficientNetV1*), esta nova arquitetura introduz os blocos *Fused-MBConv*, que substituem convoluções nos estágios iniciais para acelerar o processamento. Além disso, incorpora um método de aprendizado progressivo, onde o tamanho da imagem e a regularização são ajustados dinamicamente ao longo do treinamento, melhorando a generalização do modelo (TAN; LE, 2021).

A *AlexNet* (KRIZHEVSKY *et al.*, 2012), ilustrada na Figura 21, possui oito camadas treináveis, sendo cinco camadas convolucionais seguidas por três camadas totalmente conectadas. Suas principais inovações incluem o uso da função de ativação *ReLU* para acelerar o treinamento, camadas convolucionais sobrepostas, normalização local para melhorar a generalização e a técnica de *dropout* para reduzir o *overfitting*. Além disso, foi uma das primeiras

arquiteturas a explorar treinamento em múltiplas GPUs, permitindo lidar com redes profundas mais complexas. Essa arquitetura foi a primeira a demonstrar o potencial do *deep learning* para classificação de imagens em larga escala, influenciando o desenvolvimento de redes neurais modernas (KRIZHEVSKY *et al.*, 2012).

Diferente de outras arquiteturas populares, a *AlexNet* não está disponível na biblioteca *Keras*, o que exigiu sua implementação, sem o uso de TL. Isso porque treinar essa rede com *Imagenet* seria inviável, pois demandaria dias ou até semanas, devido às limitações de *hardware* disponíveis.

Figura 21 – Ilustração da arquitetura *AlexNet*.



Fonte: <https://jefferson023.medium.com/alexnet-para-classificacao-de-imagens-d86c482a44b3>

## 2.3 Métricas de Desempenho

A eficácia de um algoritmo após treinado é verificada examinando sua eficiência nos dados do conjunto de teste. Para isso, esse desempenho do algoritmo é avaliado através de métricas como matrix de confusão, acurácia, precisão, revocação, entre outras (FILHO, 2024).

A matriz de confusão permite visualizar o desempenho de um modelo de classificação, ajudando a entender melhor os erros e a qualidade das previsões do modelo (JUNIOR, 2023).

Em uma matriz de confusão, as linhas representam as classes previstas, enquanto as colunas correspondem às classes reais. Cada célula contém os valores das frequências de ocorrência para as combinações possíveis entre previsões e valores reais. Existem quatro tipos de valores que podem ser encontrados nas células da matriz: Verdadeiro Positivo (VP), Falso Positivo (FP), Verdadeiro Negativo (VN) e Falso Negativo (FN). Esses valores são ilustrados na Figura 22:

Figura 22 – Matriz de confusão para um problema de classificação binária.

		----- REAL -----	
		POSITIVO	NEGATIVO
--- PREVISTO ---	POSITIVO	VP	FP
	NEGATIVO	FN	VN

Fonte: Junior (2023)

- **Verdadeiro positivo (VP):** Classificação correta na classe positiva, ou seja, pertence à classe positiva e foi classificado como pertencente à classe positiva;
- **Falso positivo (FP):** Classificação incorreta na classe positiva, ou seja, pertence à classe negativa e foi classificado como pertencente à classe positiva;
- **Verdadeiro negativo (VN):** Classificação correta na classe negativa, ou seja, pertence à classe negativa e foi classificado como pertencente à classe negativa;
- **Falso negativo (FN):** Classificação incorreta na classe negativa, ou seja, pertence à classe positiva e foi classificado como pertencente à classe negativa.

A quantidade de linhas e colunas em uma matriz de confusão varia de acordo com o número de classes de saída identificadas após o treinamento. Para uma avaliação abrangente em vários cenários, é possível recorrer a outras métricas derivadas dos resultados dessa matriz, como apresentado na Tabela 2. Cada uma dessas métricas oferece uma perspectiva única sobre o desempenho do modelo, facilitando uma análise mais completa.

Tabela 2. Resumo das Métricas de Avaliação de Modelos de Classificação.

Classe	Descrição
Acurácia	Proporção de previsões corretas sobre o total de casos.
Precisão	Proporção de previsões positivas corretas sobre o total de previsões positivas.
Revocação	Proporção de positivos reais corretamente identificados pelo modelo.
<i>F1-Score</i>	Média harmônica entre precisão e revocação.
ROC-AUC	Avalia a capacidade do modelo de distinguir entre classes.
PR-AUC	Avalia a eficácia do modelo em contextos de classes desbalanceadas.

A acurácia é a proporção de previsões corretas (tanto positivas quanto negativas) em relação ao total de previsões feitas (JUNIOR, 2023). Em outras palavras, é a fração de todas as classificações que o modelo acertou, considerando tanto as classes positivas quanto as negativas.

Fazendo uma relação com os valores presentes na matriz de confusão temos:

$$\text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN} \quad (2.3)$$

A acurácia geral é útil quando as classes estão balanceadas, ou seja, quando há aproximadamente o mesmo número de exemplos em cada classe. Para modelos desbalanceados, podemos avaliar a acurácia por classe (SANTOS, 2024).

A precisão, por outro lado, foca apenas nas previsões positivas. Ela mede a proporção de exemplos corretamente classificados como positivos em relação a todos os exemplos que o modelo classificou como positivos. Ela é especialmente útil quando se quer minimizar o número de falsos positivos, ou seja, quando o custo de classificar algo incorretamente como positivo é alto (JUNIOR, 2023). Fazendo uma relação com os valores presentes na matriz de confusão temos:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2.4)$$

A revocação é calculada como a proporção de verdadeiros positivos em relação ao total de casos que realmente são positivos, ou seja, a soma dos verdadeiros positivos e dos falsos negativos (JUNIOR, 2023). Ele responde à pergunta "Entre todos os exemplos que realmente pertencem à classe positiva, quantos o modelo conseguiu identificar corretamente?". Relacionando com os valores presentes na matriz de confusão temos:

$$\text{Revocação} = \frac{VP}{VP + FN} \quad (2.5)$$

A *F1-Score* consiste na métrica de desempenho que combina os valores das métricas precisão e revocação em um único valor. Ele é especialmente útil quando há um desequilíbrio entre as classes, e você deseja um equilíbrio entre a precisão e o revocação (JUNIOR, 2023). Podemos calcula-la pela fórmula a seguir:

$$F1\text{-score} = 2 \times \frac{\text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (2.6)$$

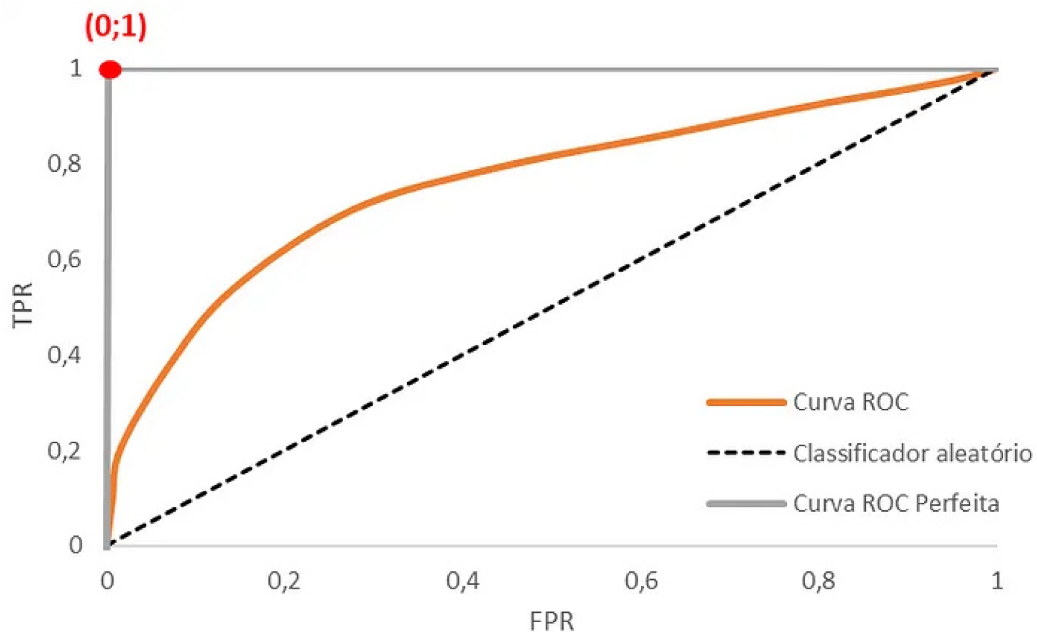
A curva *Receiver Operating Characteristic (ROC)* relaciona a Taxa de Falso Positivo (FPR) ou revocação, com a Taxa de Verdadeiro Positivo (TPR) (JUNIOR, 2023). Estes valores podem ser obtidos a partir de:

$$FPR = 1 - Revocação = \frac{FP}{FP + VN} \quad (2.7)$$

O modelo ideal seria aquele em que  $TPR = 1$ , ou seja, todas as classes realmente positivas fossem classificadas corretamente. Da mesma forma, a condição ideal para FPR seria 0, indicando que nenhuma classe realmente negativa foi erroneamente classificada como positiva. Isso resultaria no ponto (0,1) no gráfico da curva ROC. Assim, quanto mais próxima a curva ROC estiver desse ponto, melhor será o desempenho do classificador (JUNIOR, 2023). A Figura 23 apresenta um gráfico ilustrativo desta métrica.

A área sob a curva (AUC), derivada da curva ROC, equivale a métrica que mede a capacidade discriminativa do modelo. Um valor de AUC mais alto indica uma maior probabilidade do modelo distinguir corretamente entre classes positivas e negativas. Esta métrica é especialmente valiosa para avaliar a eficiência do classificador em contextos com desequilíbrio de classes ou onde as consequências de erros de classificação são consideráveis (JUNIOR, 2023).

Figura 23 – Curva para avaliação de modelos de classificação. A linha tracejada representa um classificador aleatório, enquanto a curva cinza indica um classificador perfeito.

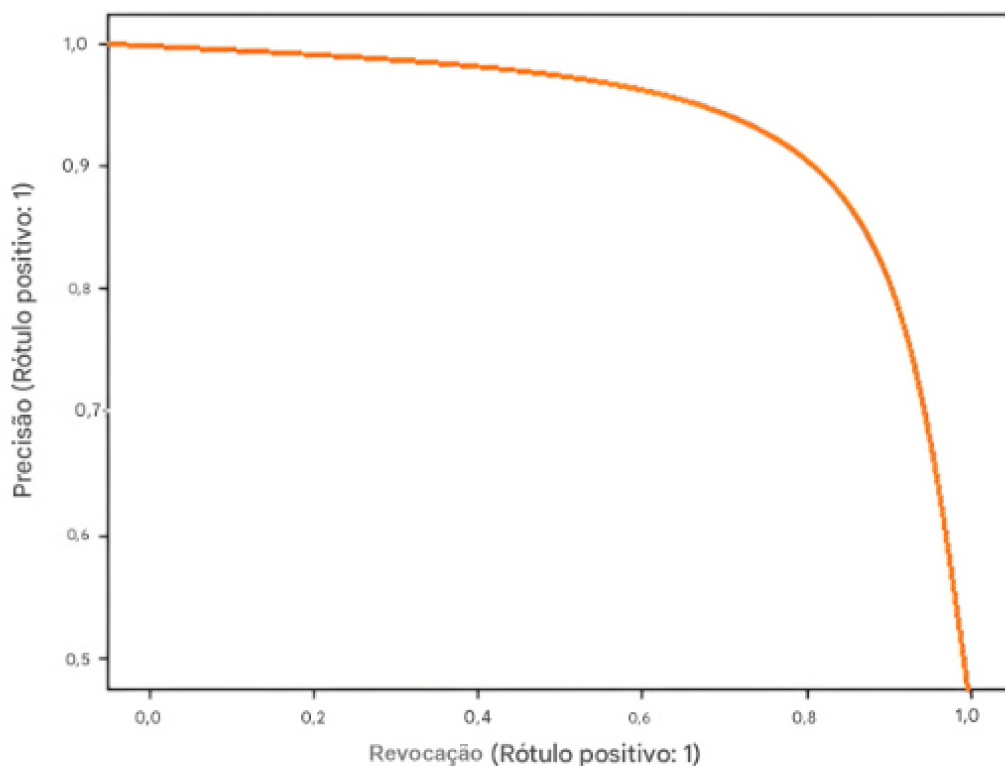


Fonte: Junior (2023)

Por fim, a curva precisão-revoação é uma métrica que relaciona o revocação com a precisão (JUNIOR, 2023). Na Figura 24 a seguir, podemos observar um exemplo.

A área sob a curva precisão-revoação (PR-AUC) indica a qualidade do modelo em distinguir classes, refletindo sua habilidade em manter alta precisão e revocação em diversos limiares (JUNIOR, 2023). Um valor alto de AUC mostra que o modelo efetivamente identifica instâncias positivas com uma taxa mínima de falsos positivos, sendo crucial em contextos com classes desequilibradas.

Figura 24 – Curva PR-AUC para avaliação do desempenho de classificadores.



Fonte: Junior (2023)



## 2.4 Exploração Digital do Espaço: O Sloan Digital Sky Survey e a Plataforma Galaxy Zoo

*Surveys* são telescópios instalados no espaço ou em regiões remotas da Terra que realizam observações sistemáticas do céu para catalogar objetos celestes, estudar suas propriedades e mapear o universo em várias escalas. Estes equipamentos cobrem diferentes comprimentos de onda, desde ondas de rádio até raios gama, fornecendo vastos volumes de dados para análise e apuração (ALMEIDA *et al.*, 2023).

Muitas pesquisas astronômicas adotam políticas de dados abertos, disponibilizando seus conjuntos de dados gratuitamente à comunidade científica e ao público. Isto incentiva a colaboração, permite a verificação independente dos resultados e promove a inovação nas técnicas de investigação. Dessas bases de dados, foram realizadas inúmeras descobertas no campo da astronomia nos últimos anos, incluindo a detecção de exoplanetas<sup>6</sup>, o mapeamento da radiação cósmica de fundo em micro-ondas<sup>7</sup> e a identificação de galáxias e quasares<sup>8</sup> distantes (SLOAN, 2024). Atualmente, os *Surveys* são essenciais para a abordagem de questões fundamentais sobre a origem, evolução e estrutura do universo, abrindo caminho para novos avanços na astronomia e na cosmologia.

### 2.4.1 Sloan Digital Sky Survey

O *Sloan Digital Sky Survey* (SDSS) é o mais ambicioso levantamento astronômico em andamento na atualidade. É um esforço colaborativo que envolve centenas de astrônomos de todo o mundo, com o objetivo de mapear bilhões de objetos celestes e sondar a estrutura em grande escala do universo. Este projeto teve início no ano 2000 e já está em sua quinta versão, o SDSS-V (SLOAN, 2024). Sua versão original consistia da observação de imagens e espectros usando um telescópio dedicado de 2,5 metros localizado em *Apache Observatory Point* no Novo México - EUA (Figura 25) (SDSS, 2022).

Suas imagens são essenciais para a identificação de alvos para a observação de espectros, nos permitindo construir um mapa tridimensional de milhões de galáxias e quasares.

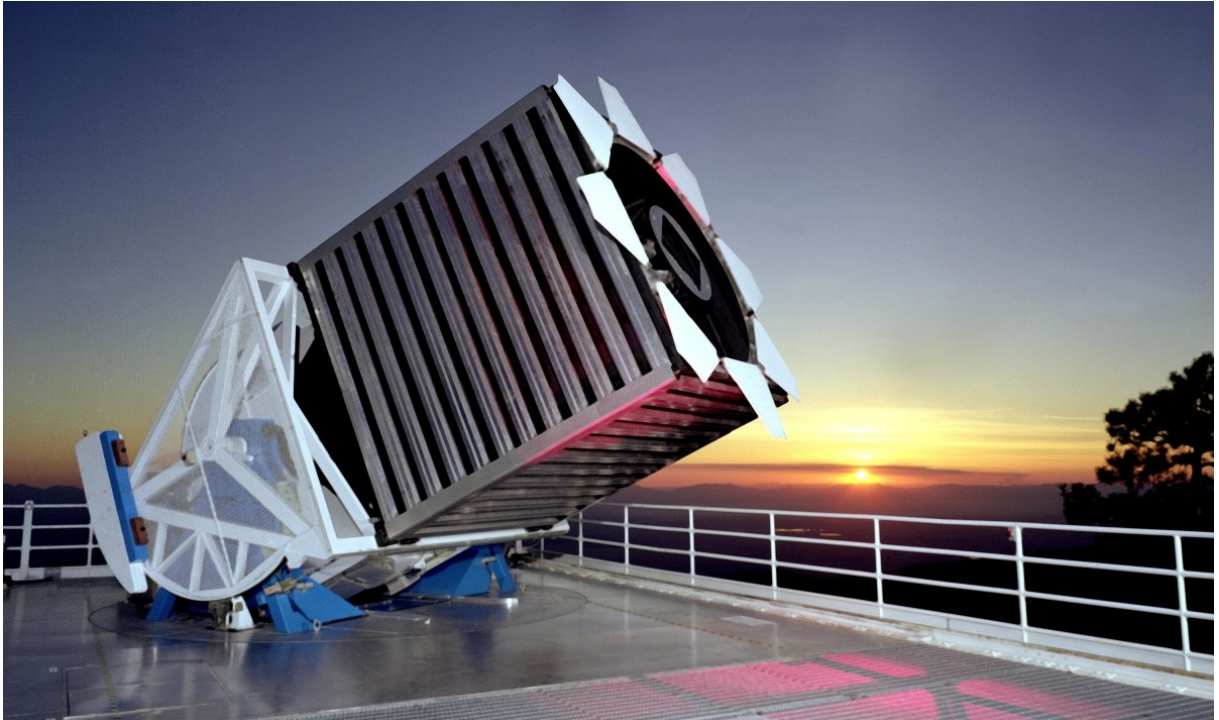
<sup>6</sup> Um exoplaneta é um planeta que se encontra fora do nosso Sistema Solar.

<sup>7</sup> A Radiação Cósmica de Fundo em microondas (RCFM) é um sinal eletromagnético, de origem cosmológica, que pode ser observado em todo o céu. Ela está associada a uma época em que o Universo ainda era muito jovem, quando a matéria era predominantemente constituída por prótons e elétrons que formavam uma espécie de "gás primordial".

<sup>8</sup> Um quasar, é um núcleo galáctico ativo, de tamanho maior que o de uma estrela, porém menor do que o tamanho mínimo para ser considerado uma galáxia.



Figura 25 – Telescópio da missão *Sloan Digital Sky Survey*.



Fonte: <https://sloan.org/programs/research/sloan-digital-sky-survey>

A medida que o levantamento progride, os dados são disponibilizados à comunidade científica (e para o público em geral) em incrementos anuais através do site *SkyServer*<sup>9</sup> (ALMEIDA *et al.*, 2023).

Dentre as conquistas do SDSS, destacam-se: criação de um mapa 3D detalhado do universo; descoberta de milhões de objetos celestes, incluindo estrelas, galáxias e quasares; medição da estrutura em grande escala do universo e impressão das oscilações acústicas bariônicas<sup>10</sup>; estudo da estrutura, dinâmica e composição química da Via Láctea; numerosas contribuições para vários campos da astronomia, incluindo cosmologia, evolução de galáxias, astrofísica estelar e muito mais (SLOAN, 2024).

O SDSS exerce um impacto profundo em nossa compreensão do universo, impulsionando uma vasta gama de pesquisas e definindo a trajetória da astronomia contemporânea (SLOAN, 2024).

<sup>9</sup> <https://www.sdss4.org/>

<sup>10</sup> Ondas de pressão geradas nos primórdios do universo causadas pela interação dos átomos com a radiação. Este fenômeno está correlacionado com a abundância de Aglomerados Galácticos em determinadas regiões.

### 2.4.2 *Galaxy Zoo*

O *Galaxy Zoo* é um projeto de ciência cidadã lançado em 2007, que utiliza classificações visuais de mais de um milhão de galáxias extraídas do SDSS, para criar um amplo catálogo morfológico desses corpos celestes. Com dados publicamente acessíveis, esse projeto facilita que pesquisadores e o público em geral explorem e empreguem esses recursos (LINTOTT *et al.*, 2008).

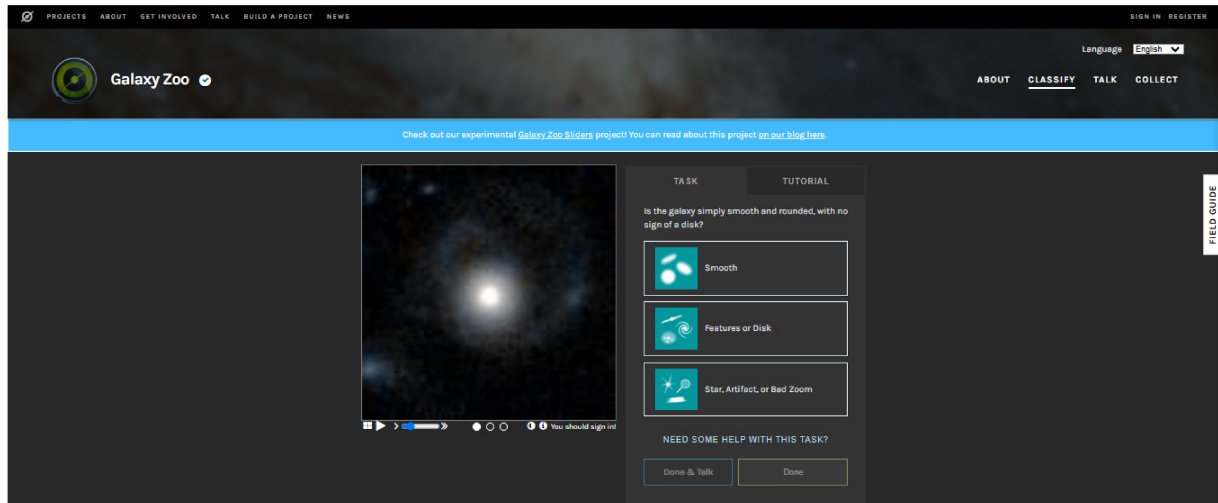
Com uma abordagem de *crowdsourcing*, no qual mais de 150 mil voluntários ao redor do mundo classificam galáxias em uma escala muito maior do que seria possível apenas por astrônomos profissionais. Este método não apenas democratiza a participação na ciência, mas também melhora a qualidade dos dados coletados (MASTERS, 2019). Os voluntários recebem treinamento e passam por um teste de categorização para garantir a precisão das classificações, que são rigorosamente verificadas e comparadas com avaliações profissionais para assegurar conformidade com padrões científicos (LINTOTT *et al.*, 2008).

A árvore de decisão do *Galaxy Zoo* é um recurso interativo que orienta os voluntários na categorização destes corpos celestes com base em suas características visuais, conforme a Figura 26. O processo inicia-se com uma pergunta geral sobre a forma básica do objeto, questionando se é liso e arredondado, possui características distintas ou é difícil de classificar devido à visibilidade ou distorção, exemplo na Figura 27. Dependendo da resposta, os participantes são direcionados para perguntas mais detalhadas, como a presença de um bojo central em objetos lisos. À medida que avançam, as perguntas tornam-se mais específicas, abordando aspectos como barras, número e orientação dos braços espirais. O processo é finalizado com a classificação precisa da categoria morfológica da galáxia, baseada nas respostas dadas (WILLETT *et al.*, 2013).

A equipe do *Galaxy Zoo* já contribuiu para mais de 60 publicações científicas, com cerca de 100 citações cada, desvendando fenômenos pouco comuns como as *Green Peas* (CARDAMONE *et al.*, 2009) e *Hanny's Voorwerp* (LINTOTT *et al.*, 2009; KEEL *et al.*, 2018). Além disso, o projeto abordou o impacto de barras nas dinâmicas galácticas e na formação estelar, investigando também como os braços espirais afetam sua estrutura e evolução (SKIBBA *et al.*, 2012; SCHAWINSKI *et al.*, 2014; SMETHURST *et al.*, 2015). Uma de suas revelações mais importantes foi esclarecer a complexa relação entre a cor e a morfologia das galáxias (SKIBBA *et al.*, 2009; MASTERS, 2019). Contrariando a suposição anterior de que esses atributos estavam diretamente relacionados, as observações demonstraram que a cor não é um indicador consis-



Figura 27 – Exemplo de pergunta de múltipla escolha no *Galaxy Zoo* para classificação de galáxias. A interface interativa apresenta imagens de galáxias e opções de resposta, permitindo que voluntários as classifiquem com base em suas características visuais, como formato e estrutura.



Fonte: Galaxy Zoo (2024)

abriga uma variedade de programas de ciência cidadã em múltiplas disciplinas como: química, biologia, história, entre outras (ZOOUNIVERSE, 2025). Isso não apenas evidencia a eficácia do modelo de *crowdsourcing* na condução de pesquisas, mas também demonstra como a participação pública pode ser integrada de maneira produtiva na investigação científica, permitindo explorar as complexidades da morfologia galáctica e da evolução do universo de forma inovadora e inclusiva (MASTERS, 2019).

## 2.5 Trabalhos Relacionados

Nesta seção, apresentamos dois estudos que exploram técnicas de aprendizado profundo aplicadas à classificação de imagens de galáxias. Ambos utilizam o mesmo conjunto de dados empregado neste trabalho e oferecem abordagens complementares, contribuindo para a compreensão e aprimoramento dos modelos utilizados.

### 2.5.1 *Galaxy Classification: a deep learning approach for classifying Sloan Digital Sky Survey images*

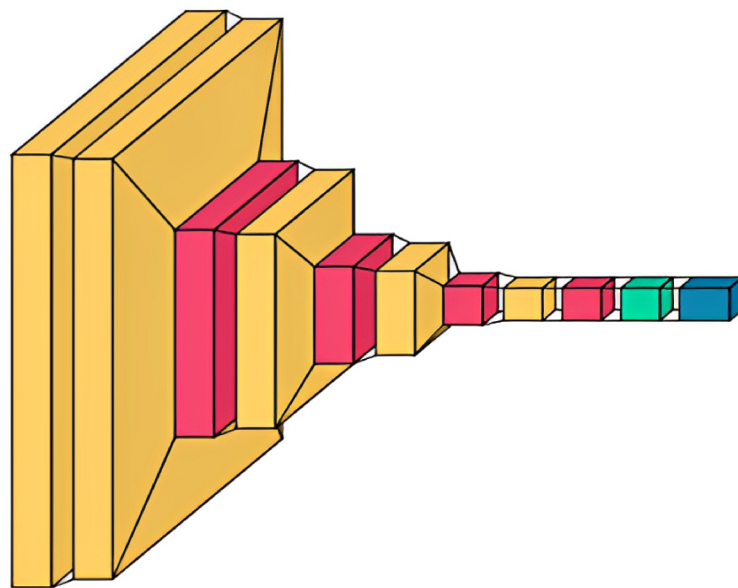
Gharat & Dandawate (2022) propuseram, neste trabalho, o uso de deep learning para classificar galáxias a partir de imagens do SDSS. No estudo utilizaram 21.785 amostras, distribuídas em 10 classes, com base no modelo estendido da classificação de Hubble.

A base de dados utilizada, *Galaxy10 Dataset* (GALAXY10 SDSS DATASET, 2024),

passou por um pré-processamento para minimizar erros de classificação no projeto *Galaxy Zoo*. Para garantir maior confiabilidade, foram selecionadas apenas imagens em que mais de 55% dos avaliadores concordaram com a mesma classe. No entanto, o conjunto de dados apresenta um desbalanceamento, que não foi tratado no estudo.

A arquitetura proposta por Gharat & Dandawate (2022) inicia com uma camada densa de 64 neurônios, seguida por outra com 32 neurônios, conforme ilustrado na Figura 28. Como o modelo classifica dez categorias, a camada de saída possui igual número de neurônios.

Figura 28 – Ilustração da arquitetura de Gharat & Dandawate.



Fonte: Gharat & Dandawate (2022)

A função de ativação *ReLU* é utilizada tanto nas camadas iniciais quanto nas cinco camadas convolutivas do modelo. Essa escolha se deve à sua capacidade de mitigar o desaparecimento do gradiente e de gerar ativações esparsas, ativando apenas um subconjunto de neurônios por vez, o que otimiza o treinamento e melhora o desempenho do modelo. Após a extração de características pelas camadas convolutivas, as saídas são convertidas em uma matriz unidimensional e processadas por duas camadas densas, também com ativação *ReLU* (GHARAT; DANDAWATE, 2022).

A classificação final é realizada pela camada que utiliza a função *Softmax*. Este arranjo permitiu que o treinamento ocorresse sem transferência de aprendizado, ou seja, sem recorrer a pesos pré-ajustados, alcançando uma acurácia de 84,78% nos testes (GHARAT; DANDAWATE, 2022).

### 2.5.2 Aplicação de deep learning para Classificação Morfológica de Galáxias

Santos (2024), avaliou em seu trabalho, seis arquiteturas de redes neurais convolucionais - *EfficientNetV2-M*, *DenseNet121*, *ResNet50*, *AlexNet*, *VGG19* e *VGG16* - além da arquitetura proposta por Gharat & Dandawate (2022). Ademais, técnicas como classificação hierárquica e *Ensemble Learning* (GANAIE *et al.*, 2022) foram aplicadas para melhorar os resultados.

A base de dados empregada foi a mesma reportada no trabalho de Gharat & Dandawate (2022). No estudo, foram conduzidos experimentos considerando o desbalanceamento dos dados aplicando técnicas de aumento de dados, tanto para equilibrar todas as classes quanto para reforçar exclusivamente a classe minoritária.

Com a aplicação de transferência de aprendizagem, a arquitetura *EfficientNetV2-M* obteve uma acurácia de 82,90%. Quando realizado balanceamento das classes através de aumento de dados, a *ResNet50* se destacou, alcançando 88,20% de acurácia. No entanto, a estratégia de *Ensemble Learning* superou esses resultados, atingindo 85,53% com dados desbalanceados e 91,75% com dados balanceados. Esses achados destacam o *Ensemble Learning* como uma abordagem promissora para aprimorar a classificação morfológica de galáxias, tornando o modelo mais robusto e eficaz.

### 3 METODOLOGIA

Neste capítulo, descrevemos detalhadamente a metodologia adotada para a condução dos experimentos que sustentam o desenvolvimento deste trabalho. A estrutura deste estudo é inspirada nas pesquisas de Gharat & Dandawate (2022), nas quais se utiliza a arquitetura de uma CNN construída pelos próprios autores. Além da CNN de Gharat & Dandawate, também trabalhamos com outros modelos de redes neurais pré-treinadas, conforme apresentado anteriormente. Na Seção 3.1, apresentamos a base de dados utilizada. A Seção 3.2 aborda as técnicas de pré-processamento aplicadas às imagens. Em seguida, na Seção 3.3, explicamos o procedimento de validação cruzada adotado. Por fim, na Seção 3.4, detalhamos a etapa experimental.

#### 3.1 Base de Dados

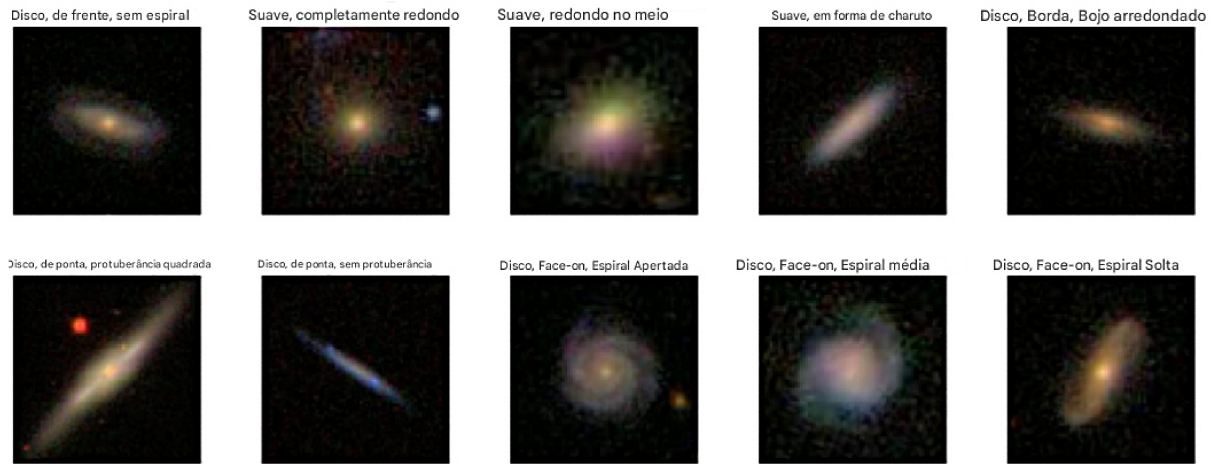
A base de dados utilizada neste trabalho deriva de pré-processamentos aplicados às imagens oriundas da *Galaxy10 SDSS Dataset* (GALAXY10 SDSS DATASET, 2024). Essa é composta por 21785 imagens coloridas de galáxias, observadas pelo SDSS e pré-classificadas por voluntários. Originalmente com resolução de 424x424 pixels, as imagens foram centralizadas e redimensionadas para 207x207 pixels, e posteriormente para 69x69 pixels através de interpolação bilinear, para torná-las gerenciáveis na maioria dos computadores e memórias de placas de vídeo (GHARAT; DANDAWATE, 2022). As imagens foram então organizadas em 10 classes distintas, conforme ilustrado na Figura 29.

Para garantir a confiabilidade das classificações no *Galaxy10 SDSS Dataset*, foi estabelecido como limiar de aceitação das imagens que mais de 55% dos votos dos voluntários concordem em classifica-las em uma única classe dentre as 10 possíveis. Este limiar foi selecionado com o objetivo futuro de posicionar o *Galaxy10 SDSS Dataset* (GALAXY10 SDSS DATASET, 2024) como uma alternativa viável aos conjuntos de dados tradicionais usados em aprendizado profundo, como MNIST (*Modified National Institute of Standards and Technology*) (DENG, 2012) e Cifar\_10 (*Canadian Institute For Advanced Research*) (DOON *et al.*, 2018), com foco para astrônomos.

Tomando o modelo *astroNN.models.Cifar10\_CNN* como referência, Gharat & Dandawate (2022) observaram que um limite de 50% resultava em baixa precisão de classificação devido ao grande número de imagens potencialmente mal classificadas, dificultando o aprendi-



Figura 29 – Exemplos de imagens das diferentes classes do conjunto de dados *Galaxy10 SDSS Dataset*. Podemos observar diversas morfologias galácticas, classificadas por formato, estrutura espiral e orientação do disco.



Fonte: Gharat & Dandawate (2022)

zado da rede. Todavia, a definição de um limite de 60% produziu resultados semelhantes aos obtidos com 55%, mas optou-se pelo limite de 55% por permitir a inclusão de mais imagens no conjunto de dados.

A Figura 3 exibe a distribuição das imagens do *dataset* produzido por classe, realçando a diversidade e amplitude do conjunto de dados. Observa-se claramente que as classes não encontram-se balanceadas, o que demandou a implementação de estratégias para assegurar um desempenho satisfatório dos modelos adotados.

Tabela 3 – Distribuição da quantidade de imagens por classe de galáxias no conjunto de dados *Galaxy10 SDSS Dataset*, apresentando as categorias morfológicas, sua nomenclatura e o número de imagens em cada classe, totalizando 21.785 registros.

Classe	Descrição	Acurácia
Classe 0	Disco, Vista Frontal, Sem Espiral	3461
Classe 1	Suave, Completamente Redonda	6997
Classe 2	Suave, Parcialmente Redonda	6992
Classe 3	Suave, Formato de Charuto	394
Classe 4	Disco, Vista Lateral, Bojo Arredondado	1534
Classe 5	Disco, Vista Lateral, Bojo Quadrado	17
Classe 6	Disco, Vista Lateral, Sem Bojo	589
Classe 7	Disco, Vista Frontal, Espiral Apertada	1121
Classe 8	Disco, Vista Frontal, Espiral Média	906
Classe 9	Disco, Vista Frontal, Espiral Solta	519
Total		21785

Fonte: Adaptado de Gharat & Dandawate (2022)



### 3.2 Pré-processamento

Antes do treinamento do modelo, realizamos o pré-processamento dos dados, com técnicas de aumento de dados (*data augmentation*) e de redução de dados (*undersampling*), a depender do experimento. Este passo nos permitiu ampliar a diversidade das amostras disponíveis e reduzir possíveis vieses, tornando o modelo mais robusto e capaz de generalizar melhor.

Avaliamos três tipos de estratégias de aumento de dados: *Datagen* (LU *et al.*, 2024), *MixUp* (ZHANG *et al.*, 2018) e *CutMix* (YUN *et al.*, 2019) (Figura 30).

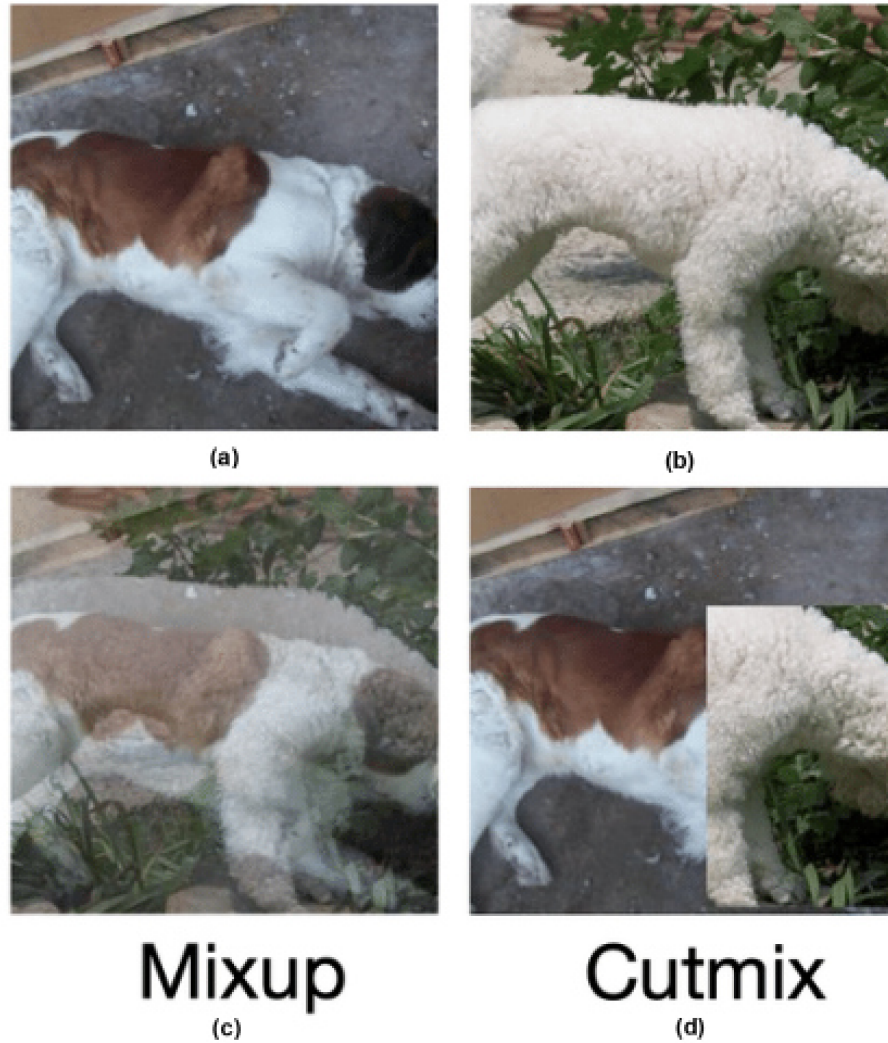
- **Datagen:** Estratégia clássica, com transformações como espelhamento horizontal e vertical aleatórios (*Random Horizontal Flip* e *Random Vertical Flip*), corte aleatório e rotação aleatória de até 30 graus (*Random Rotation*);
- **MixUp:** Técnica que cria novos exemplos sintéticos ao interpolar linearmente tanto os dados de entrada quanto seus respectivos rótulos;
- **CutMix:** Substitui aleatoriamente uma região de uma imagem por um recorte de outra, ajustando os rótulos proporcionalmente.

Para subamostragem e sobreamostragem, aplicamos técnicas de redução das classes majoritárias para evitar que o modelo desenvolvesse viés excessivo em relação às classes com mais amostras. Utilizamos o *Random Undersampling (RUS)* (HASANIN; KHOSHGOFTAAR, 2018), que seleciona aleatoriamente um subconjunto das amostras da classe majoritária, equilibrando a distribuição dos dados.

Além disso, em alguns experimentos, optamos pelo descarte de determinadas classes e ajustamos a proporção das imagens para 244x244 para favorecer o treinamento do modelo. Também exploramos o uso da técnica *CLAHE* para aumentar o contraste das imagens escurecidas. A Figura 31 representa as imagens com a aplicação de *CLAHE* no pré-processamento.

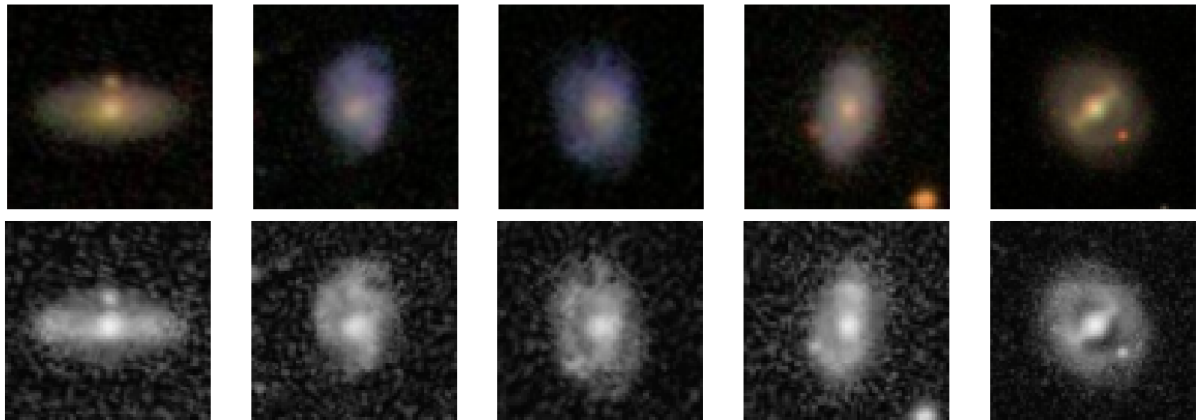
As imagens de entrada foram previamente convertidas para o formato exigido por cada modelo antes do treinamento. Por exemplo, a *ResNet50* requer imagens no formato RGB (0-255), enquanto a *EfficientNetV2-M* aceita imagens em RGB (0-1), normalizadas para o intervalo [0,1]. Além disso, foi aplicada a normalização baseada na média e desvio-padrão do conjunto *ImageNet*, conforme recomendado para modelos pré-treinados.

Figura 30 – Técnicas de Aumento de Dados: *MixUp* e *CutMix*. (a) e (b) representam as imagens originais utilizadas como base, (c) apresenta a aplicação da técnica *MixUp*, combinando elementos das imagens (a) e (b), e por último, (d) representando a aplicação da técnica *CutMix*, inserindo segmentos da imagem (b) na imagem (a).



Fonte: Gharat & Dandawate (2022)

Figura 31 – Imagens realçadas com a aplicação do *CLAHE*.



### 3.3 Validação Cruzada

Esta etapa consistiu na divisão de nossa base de dados em três grupos: treino, validação e teste.

- **Treino:** Este lote foi empregado para treinar a CNN, ajustando os pesos e parâmetros do modelo conforme as imagens e categorias foram apresentadas. Normalmente, uma proporção significativa do conjunto de dados é reservada para treinamento, uma vez que, é nesta fase que o modelo efetivamente aprende.
- **Validação:** O conjunto de validação permite avaliar o desempenho da CNN ao longo do treinamento e refinar os hiperparâmetros do modelo, visando sua otimização. O conjunto de dados de validação é usado para avaliar a precisão e o desempenho geral do modelo com exemplos não vistos durante o treinamento.
- **Teste:** Após a CNN ser aprimorada com os dados de treinamento e validação, é crucial avaliar sua eficácia final em um conjunto de dados completamente separado, conhecido como conjunto de testes.

A divisão dos dados para treino, validação e teste seguiu, respectivamente, uma proporção de 70%, 15% e 15% do conjunto total de imagens, conforme sugerem Gharat & Dandawate (2022).

Na fase de treinamento da rede neural é configurado o número de épocas, que representa uma iteração completa em que todos os dados do conjunto de treinamento são apresentados à rede (SANTOS, 2024). O tamanho do lote (batch size) é determinado para definir o número de amostras utilizadas para atualizar os pesos da rede a cada iteração. Neste trabalho, o número máximo de épocas foi definido para 20 interações e quanto ao tamanho dos lotes para processamento, optou-se por valores de 8, 16 e 32, conforme em Santos (2024).

Aplicamos *dropout* de 50% e utilizamos *callbacks* durante o treinamento para mitigar a ocorrência de *overfitting*, garantindo uma melhor generalização dos modelos. Além disso, adotamos a técnica de *fine-tuning* para os modelos pré-treinados, inicialmente congelando as primeiras camadas da rede, responsáveis por aprender características mais gerais das imagens, e refinando apenas as camadas superiores. À medida que o treinamento evolui, descongelam-se as demais camadas, permitindo que o modelo ajuste seus pesos de forma progressiva e adapte-se melhor aos padrões do novo conjunto de dados.

Por fim, de forma a avaliar o desempenho dos modelos, aplicamos as métricas matriz de confusão, acurácia, revocação, precisão, *F1-Score*, ROC AUC e PR-AUC, *Macro avg* e

*Weighted avg.* O *Macro Average (Macro avg)* é uma métrica que calcula a média simples das métricas da figura acima, atribuindo o mesmo peso a todas as classes, enquanto o *Weighted Average (Weighted avg)* pondera essas métricas de acordo com o número de amostras em cada classe, dando mais influência às classes mais frequentes.

### 3.4 Etapa Experimental

Realizamos oito experimentos a saber: Experimento I consistiu na reprodução do modelo de Gharat & Dandawate (2022). As etapas II, III e IV, empregaram a técnica de *Transfer Learning* com base no trabalho de Santos (2024). Por fim, V a VIII, também adotaram TL, mas para 20000 imagens, redimensionamento 244x244 pixels e aplicação de *CLAHE*.

A escolha dos modelos pré-treinados e as técnicas de aumento de dados nos experimentos V a VIII tomou como base os dois melhores desempenhos em II, III e IV. A Tabela 4 apresenta a descrição dos testes conduzidos.

### 3.5 Ferramentas e técnicas utilizadas

Todo o desenvolvimento foi realizado utilizando os recursos oferecidos pelo *Google Colab*, os detalhes são apresentado abaixo:

Para os experimentos I, II, III e IV:

- **Disco:** 235,7,25 GB de armazenamento;
- **GPU:** *NVIDIA Tesla T4* (15 GB de memória de vídeo);
- **RAM:** 51 GB de memória RAM.

Para os experimentos V, VI, VII e VIII:

- **Disco:** 235,7,25 GB de armazenamento;
- **GPU:** *NVIDIA Ampere A100* (40 GB de memória de vídeo);
- **RAM:** 83,5 GB de memória RAM.

Tabela 4 – Descrição dos oito testes conduzidos, detalhando as arquiteturas dos modelos utilizados e os métodos de pré-processamento aplicados em cada etapa.

<b>Exp.</b>	<b>Arquitetura</b>	<b>Pré-processamento</b>
I	<i>Gharat &amp; Dandawate</i>	-
II	<i>EfficientNetV2-M, DenseNet121, ResNet50, AlexNet, VGG16 e VGG19</i>	-
III	<i>EfficientNetV2-M, DenseNet121, ResNet50, AlexNet, VGG16 e VGG19</i>	710 imagens por classe e <i>Datagen, MixUp</i> ou <i>CutMix</i>
IV	<i>EfficientNetV2-M, DenseNet121, ResNet50, AlexNet, VGG16 e VGG19</i>	4898 imagens por classe e <i>Datagen, MixUp</i> ou <i>CutMix</i>
V	<i>VGG16, VGG19</i> e <i>EfficientNetV2-M</i>	20000 imagens por classe e <i>MixUp</i> ou <i>CutMix</i>
VI	<i>VGG16, VGG19</i> e <i>EfficientNetV2-M</i>	4898 imagens por classe, <i>MixUp</i> ou <i>CutMix</i> e 244x244 <i>pixels</i>
VII	<i>VGG16, VGG19</i> e <i>EfficientNetV2-M</i>	20000 imagens por classe, <i>MixUp</i> ou <i>CutMix</i> e <i>CLAHE</i>
VIII	<i>VGG16, VGG19</i> e <i>EfficientNetV2-M</i>	4898 imagens por classe, <i>MixUp</i> ou <i>CutMix</i> , 244x244 <i>pixels</i> e <i>CLAHE</i>

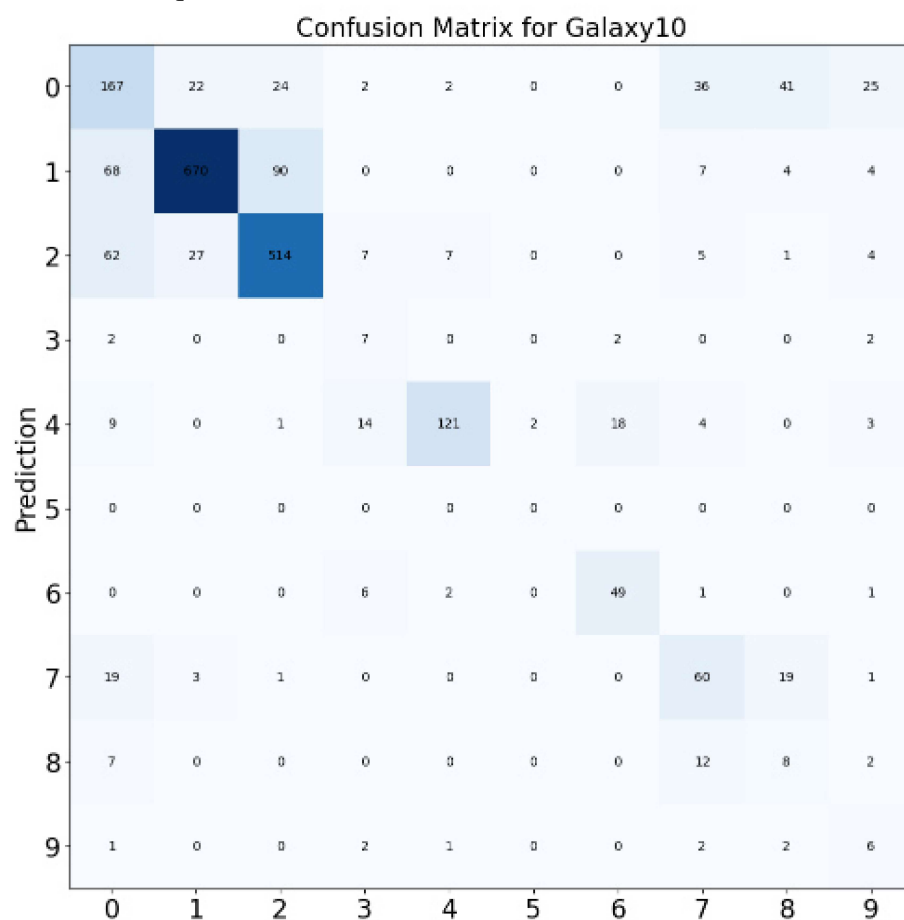
## 4 RESULTADOS E DISCUSSÕES

Neste capítulo, são apresentados os resultados obtidos ao longo deste estudo. A Seção 4.1 aplica a CNN adotada por Gharat & Dandawate (2022). Posteriormente, nas seções seguintes, são descritos os resultados alcançados utilizando a técnica de *Transfer Learning* com base nas CNNs utilizadas por Sousa (2024). Por fim, a Seção 4.9 compara os melhores resultados atingidos.

### 4.1 Experimento I

A Arquitetura de Gharat & Dandawate (2022) sem balanceamento de dados, apresentou acurácia de 0,83. Ao analisarmos a matriz de confusão deste modelo (Figura 32), observamos uma incoerência entre o desempenho do modelo para as classes individuais e o resultado geral de acurácia.

Figura 32 – Quantidade de imagens por classe de Galáxias para a Arquitetura de Gharat & Dandawate.



O modelo mostrou boa performance em classificar as galáxias das Classes 1 e 2, mas fraco desempenho para as demais classes. Especificamente, o modelo não obteve acertos para a Classe 5.

Avaliando o treinamento com base na acurácia por Classe (Tabela 5) e no relatório de desempenho (Tabela 6), observa-se que a acurácia geral não representa a qualidade do modelo em um conjunto de dados desbalanceado. Isso porque essa métrica é influenciada pela predominância de classes majoritárias, mascarando o real impacto do desbalanceamento na classificação. Outra justificativa seria a complexidade do modelo, contando apenas com cinco camadas convolucionais, e a escassez de imagens em algumas classes.

Tabela 5. Acurácia por classe para a arquitetura de Gharat & Dandawate.

Classe	Descrição	Acurácia
0	Disco, Vista Frontal, Sem Espiral	0,52
1	Suave, Completamente Redonda	0,79
2	Suave, Parcialmente Redonda	0,82
3	Suave, Formato de Charuto	0,54
4	Disco, Vista Lateral, Bojo Arredondado	0,70
5	Disco, Vista Lateral, Bojo Quadrado	0,00
6	Disco, Vista Lateral, Sem Bojo	0,83
7	Disco, Vista Frontal, Espiral Apertada	0,58
8	Disco, Vista Frontal, Espiral Média	0,28
9	Disco, Vista Frontal, Espiral Solta	0,26

Tabela 6. Precisão, Revocação e F1-score por classe para a Arquitetura de Gharat & Dandawate.

Classe	Descrição	Precisão	Revocação	F1-Score
Classe 0	Disco, Vista Frontal, Sem Espiral	0,50	0,52	0,51
Classe 1	Suave, Completamente Redonda	0,92	0,79	0,85
Classe 2	Suave, Parcialmente Redonda	0,80	0,82	0,81
Classe 3	Suave, Formato de Charuto	0,30	0,54	0,39
Classe 4	Disco, Vista Lateral, Bojo Arredondado	0,91	0,70	0,79
Classe 5	Disco, Vista Lateral, Bojo Quadrado	0,00	0,00	0,00
Classe 6	Disco, Vista Lateral, Sem Bojo	0,71	0,83	0,77
Classe 7	Disco, Vista Frontal, Espiral Apertada	0,45	0,58	0,51
Classe 8	Disco, Vista Frontal, Espiral Média	0,11	0,28	0,15
Classe 9	Disco, Vista Frontal, Espiral Solta	0,12	0,26	0,17
<b>Média Macro</b>		0,48	0,53	0,50
<b>Média Ponderada</b>		0,77	0,73	0,75

No conjunto de dados analisado, o *Macro avg* é significativamente menor que o *Weighted avg*, indicando que o modelo tem um desempenho desigual entre as classes, classificando melhor aquelas mais representadas e pior as menos frequentes.

## 4.2 Experimento II

O desbalanceamento da base de dados desse experimento impacta a interpretação dos resultados, no que diz respeito as métricas globais acurácia e *F1-Score*. Portanto, avaliações do desempenho do modelo por classe são mais informativas do que análises generalizadas. A avaliação das curvas ROC-AUC e PR-AUC possibilita entender mais precisamente o desempenho dos modelos em diferentes classes e sua capacidade de generalização. A Tabela 7 apresenta os resultados deste experimento.

Tabela 7. Comparação do desempenho dos modelos para o Experimento II.

Modelo	Acurácia	Perda	ROC-AUC	PR-AUC
<b>VGG16</b>	<b>0,8252</b>	<b>0,5391</b>	<b>0,9729</b>	<b>0,7143</b>
<b>VGG19</b>	<b>0,8323</b>	<b>0,4911</b>	<b>0,9719</b>	<b>0,7330</b>
<i>ResNet50</i>	0,7922	0,5913	0,9697	0,6979
<b>DenseNet121</b>	<b>0,8185</b>	<b>0,5667</b>	<b>0,9754</b>	<b>0,7335</b>
<i>EfficientNetV2-M</i>	0,7998	0,5731	0,9671	0,6750
<i>AlexNet</i>	0,7821	0,6225	0,9443	0,6395

Observamos que os modelos *VGG16* e *VGG19* apresentaram o melhor desempenho geral, destacando-se pelos valores das métricas acurácia, perda, ROC-AUC e PR-AUC. O *VGG19* obteve a maior acurácia 0,8323, a menor perda 0,4911 e a maior PR-AUC 0,7330, demonstrando sua superioridade na classificação. Embora o desempenho do *VGG16* seja ligeiramente inferior, este apresentou o maior ROC-AUC 0,9729, evidenciando sua capacidade de discriminar corretamente entre as classes. Além disso, o modelo *DenseNet121* se destacou como a terceira melhor opção, com uma boa acurácia 0,8185 e um ROC-AUC competitivo 0,9754, demonstrando um equilíbrio entre precisão e generalização.

A análise do desempenho por classe dos modelos, apresentados nas Tabelas 8, 9 e 10, confirmam o *VGG19* como o melhor modelo, apresentando maior acurácia global 0,83, menor perda e valores mais equilibrados entre classes, com destaque para a Classe 2 0,96 e a Classe 4 0,92. Embora o *VGG16* apresente desempenho muito próximo do *VGG19*, o primeiro mostrou maior variação da acurácia entre classes, com fraco desempenho para a Classe 8 0,48 e a Classe 9 0,56. Apesar do modelo *DenseNet121* alcançar uma acurácia global de 0,82, este obteve um *F1-score* médio inferior aos resultados dos dois melhores modelos, com uma variação ainda mais acentuada na acurácia por classe.

Importante salientar que todos os modelos tiveram dificuldades em classificar nas classes menos representadas, como a Classe 5, que obteve precisão, revocação e *F1-score* zerados



Tabela 8. Desempenho II do modelo VGG16 por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,68	0,62	0,68	0,65
Classe 1	0,93	0,91	0,93	0,92
Classe 2	0,90	0,89	0,90	0,90
Classe 3	0,62	0,68	0,62	0,65
Classe 4	0,89	0,86	0,89	0,88
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,77	0,82	0,77	0,80
Classe 7	0,59	0,71	0,59	0,64
Classe 8	0,48	0,65	0,48	0,55
Classe 9	0,56	0,80	0,56	0,66
<b>Média Macro</b>	-	0,69	0,64	0,66
<b>Média Ponderada</b>	0,83	0,82	0,83	0,82

Tabela 9. Desempenho II do modelo VGG19 por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,55	0,73	0,55	0,63
Classe 1	0,92	0,92	0,92	0,92
Classe 2	0,96	0,84	0,96	0,89
Classe 3	0,56	0,81	0,56	0,66
Classe 4	0,92	0,86	0,92	0,89
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,85	0,80	0,85	0,82
Classe 7	0,76	0,62	0,76	0,68
Classe 8	0,56	0,67	0,56	0,61
Classe 9	0,59	0,81	0,59	0,68
<b>Média Macro</b>	-	0,71	0,67	0,68
<b>Média Ponderada</b>	0,83	0,83	0,83	0,83

Tabela 10. Desempenho II do modelo DenseNet121 por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,61	0,68	0,61	0,64
Classe 1	0,93	0,88	0,93	0,91
Classe 2	0,90	0,86	0,90	0,88
Classe 3	0,62	0,73	0,62	0,67
Classe 4	0,92	0,83	0,92	0,87
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,84	0,76	0,84	0,80
Classe 7	0,64	0,71	0,64	0,67
Classe 8	0,40	0,62	0,40	0,49
Classe 9	0,63	0,73	0,63	0,68
<b>Média Macro</b>	-	0,68	0,65	0,66
<b>Média Ponderada</b>	0,82	0,81	0,82	0,81

devido à baixa quantidade de amostras. No entanto, o *VGG19* mostrou maior robustez ao lidar com classes problemáticas, o que justifica sua melhor performance nas métricas globais.

As Figuras 33 e 34 mostram os resultados de treinamento e as curvas ROC-AUC e PR-AUC para os modelos *VGG16*, *VGG19* e *DenseNet121*. Durante o treinamento, observa-se uma melhoria na acurácia para todos os modelos, mas com sinais de leve *overfitting* nos modelos *VGG*. A perda diminui consistentemente, indicando boa convergência dos modelos. Nas análises de ROC-AUC e PR-AUC, o modelo *DenseNet121* destacou-se por apresentar curvas superiores, indicando uma melhor taxa de verdadeiros positivos e maior precisão na classificação positiva, sugerindo que é o mais eficaz em termos de generalização e precisão entre os modelos avaliados.

### 4.3 Experimento III

Nesta análise, a base de dados utilizada foi balanceada para 710 imagens por classe e utilizamos também três técnicas de aumento de dados para cada modelo: *MixUp*, *Datagen* e *CutMix*. As métricas ROC-AUC e PR-AUC foram consideradas para caracterizar o desempenho dos modelos em diferentes classes e sua capacidade de generalização. A Tabela 11 apresenta os resultados obtidos.

Tabela 11. Comparação do desempenho dos modelos para o Experimento III.

<b>Modelo</b>	<b>Acurácia</b>	<b>Perda</b>	<b>ROC-AUC</b>	<b>PR-AUC</b>
<b><i>VGG16 MixUp</i></b>	<b>0,7717</b>	<b>0,6310</b>	<b>0,9601</b>	<b>0,6996</b>
<i>VGG16 Datagen</i>	0,7604	0,9217	0,9385	0,6029
<i>VGG16 CutMix</i>	0,7616	0,6739	0,9665	0,6878
<b><i>VGG19 MixUp</i></b>	<b>0,7775</b>	<b>0,6488</b>	<b>0,9274</b>	<b>0,6899</b>
<i>VGG19 Datagen</i>	0,7239	0,8949	0,9205	0,5881
<b><i>VGG19 CutMix</i></b>	<b>0,7836</b>	<b>0,6458</b>	<b>0,9630</b>	<b>0,7052</b>
<i>ResNet50 MixUp</i>	0,6364	0,9993	0,9336	0,6080
<i>ResNet50 Datagen</i>	0,6493	1,1590	0,9126	0,5219
<i>ResNet50 CutMix</i>	0,6900	0,8584	0,9498	0,6534
<i>DenseNet121 MixUp</i>	0,7279	0,8212	0,9517	0,6588
<i>DenseNet121 Datagen</i>	0,6735	1,0769	0,9301	0,5425
<i>DenseNet121 CutMix</i>	0,7322	0,8171	0,9204	0,6455
<i>EfficientNetV2-M MixUp</i>	0,7083	0,8235	0,9436	0,6089
<i>EfficientNetV2-M Datagen</i>	0,6958	0,9173	0,9155	0,5837
<i>EfficientNetV2-M CutMix</i>	0,6407	0,9806	0,9119	0,5260
<i>AlexNet MixUp</i>	0,6967	0,7894	0,9490	0,6137
<i>AlexNet Datagen</i>	0,7093	0,9010	0,9313	0,5491
<i>AlexNet CutMix</i>	0,7325	0,7405	0,9590	0,6158

Os modelos *VGG19 CutMix*, *VGG16 MixUp* e *VGG19 MixUp* obtiveram os melhores desempenhos. *VGG19 CutMix* apresentou a maior acurácia e PR-AUC, tornando-se o mais eficaz. O *VGG16 MixUp*, se destaca pela menor perda 0,6310, o que indica um treinamento

Figura 33 – Treinamento II para os modelos *VGG16*, *VGG19* e *DenseNet121*, organizadas respectivamente de cima para baixo.

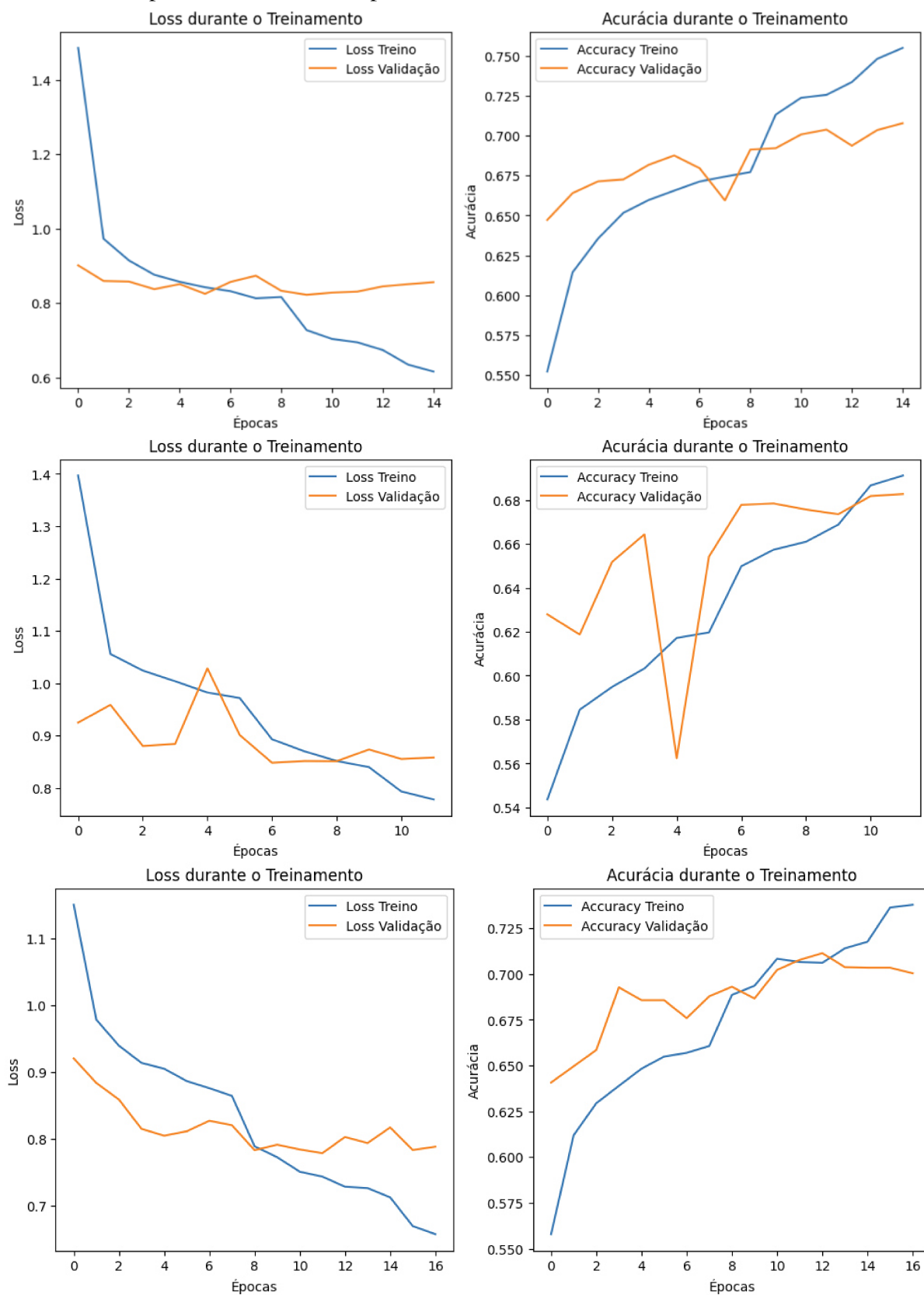
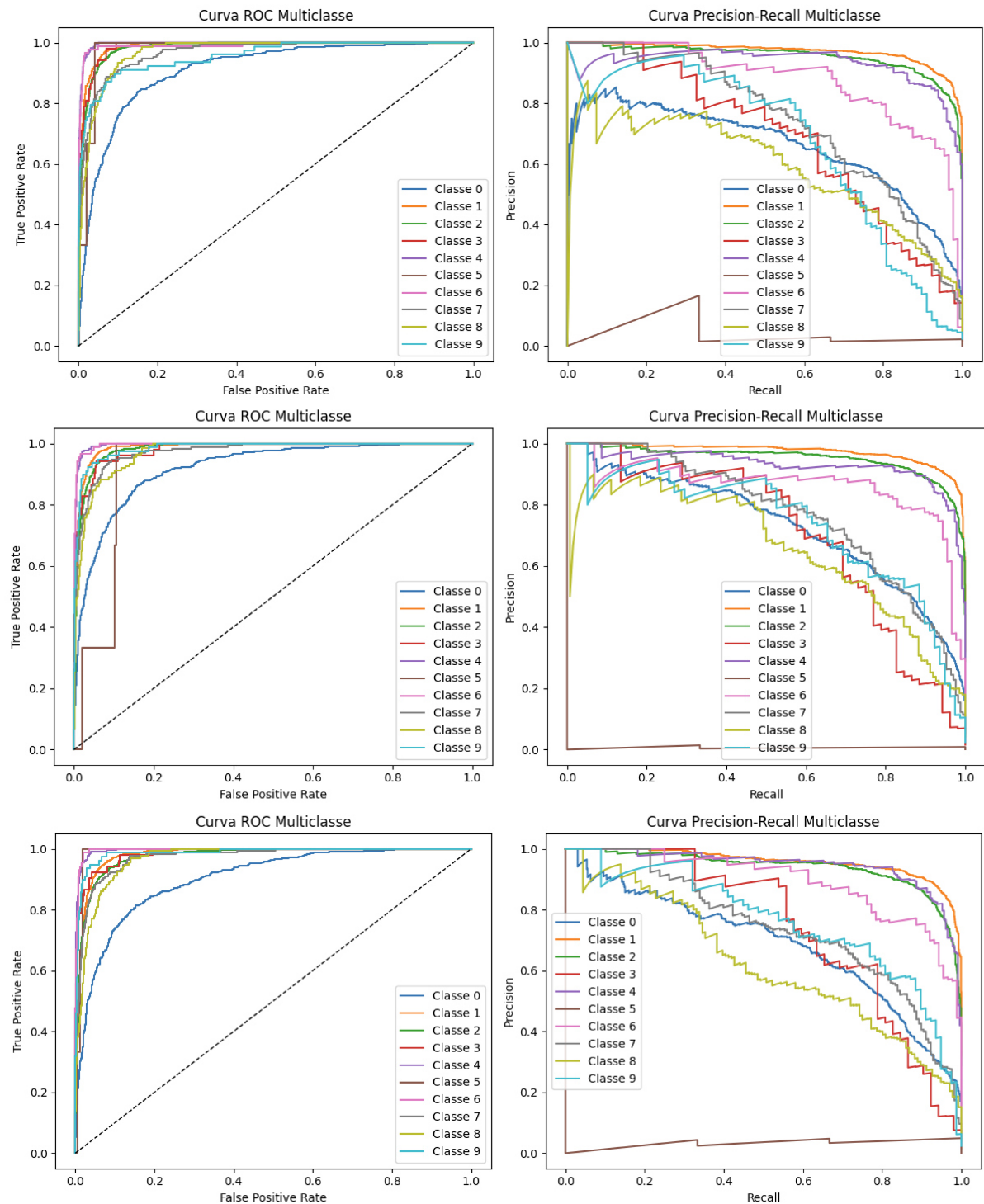


Figura 34 – ROC-AUC e PR-AUC II para os modelos *VGG16*, *VGG19* e *DenseNet121*, organizadas respectivamente de cima para baixo.

ROC AUC: 0.9729  
PR AUC: 0.7143



mais estável e melhor otimizado. Apesar de o *VGG19 CutMix* apresentar uma acurácia superior ao *VGG16 MixUp*, a métrica ROC-AUC apresentou menor valor 0,9274, o que impacta sua capacidade de distinção entre classes. Entretanto, os modelos da família *VGG* foram os mais robustos, enquanto *CutMix* e *MixUp* demonstraram ser técnicas eficazes de aumento de dados.

A análise do desempenho por classe dos modelos, apresentados nas Tabelas 12, 13 e 14, confirmam o *VGG19 CutMix* como o melhor modelo, apresentando maior acurácia global 0,78, baixa perda e valores mais equilibrados entre classes, com destaque para a Classe 1 0,91, a Classe 4 0,87 e a Classe 6 0,88. Embora o *VGG19 MixUp* apresente desempenho muito próximo do *VGG19 CutMix*, o primeiro mostrou maior variação da acurácia entre classes dentre dos dois, com fraco desempenho para Classe 0 0,42, Classe 3 0,56 e a Classe 8 0,54. Os modelos *VGG16 MixUp* e *VGG19 MixUp* também obtiveram um *F1-score* médio superior aos demais resultados, figurando entre os três melhores modelos.

Tabela 12. Desempenho III do modelo VGG16 MixUp por classe.

Classe	Acurácia	Precisão	Revocação	<i>F1-Score</i>
Classe 0	0,37	0,62	0,37	0,47
Classe 1	0,94	0,82	0,94	0,88
Classe 2	0,80	0,87	0,80	0,83
Classe 3	0,67	0,51	0,67	0,58
Classe 4	0,89	0,89	0,89	0,89
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,89	0,78	0,89	0,83
Classe 7	0,71	0,56	0,71	0,63
Classe 8	0,69	0,52	0,69	0,59
Classe 9	0,67	0,63	0,67	0,65
<b>Média Macro</b>	-	0,62	0,66	0,63
<b>Média Ponderada</b>	0,77	0,77	0,77	0,76

Tabela 13. Desempenho III do modelo VGG19 MixUp por classe.

Classe	Acurácia	Precisão	Revocação	<i>F1-Score</i>
Classe 0	0,42	0,68	0,42	0,58
Classe 1	0,93	0,85	0,93	0,88
Classe 2	0,84	0,84	0,84	0,84
Classe 3	0,56	0,56	0,56	0,56
Classe 4	0,87	0,82	0,87	0,84
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,84	0,73	0,84	0,78
Classe 7	0,71	0,57	0,71	0,63
Classe 8	0,54	0,54	0,54	0,54
Classe 9	0,82	0,60	0,82	0,70
<b>Média Macro</b>	-	0,62	0,65	0,63
<b>Média Ponderada</b>	0,77	0,77	0,78	0,77

Os gráficos de treinamento (Figura 35) indicam que os modelos *VGG16 MixUp*, *VGG19 MixUP* e *VGG19 CutMix* aprendem eficazmente, com perda decrescente e acurácia crescente durante as épocas de treino. As curvas ROC-AUC e precisão-revocação (Figura 36)

Tabela 14. Desempenho III do modelo VGG19 CutMix por classe.

Classe	Acurácia	Precisão	Revocação	F1-Score
Classe 0	0,55	0,55	0,55	0,55
Classe 1	0,91	0,89	0,91	0,90
Classe 2	0,79	0,89	0,79	0,84
Classe 3	0,56	0,63	0,56	0,59
Classe 4	0,87	0,89	0,87	0,88
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,88	0,75	0,88	0,81
Classe 7	0,70	0,61	0,70	0,66
Classe 8	0,65	0,51	0,65	0,57
Classe 9	0,74	0,67	0,74	0,70
<b>Média Macro</b>	-	0,64	0,67	0,65
<b>Média Ponderada</b>	0,78	0,79	0,78	0,79

mostram bom desempenho em discriminar e classificar as classes, especialmente notável nas classes com maior representatividade. Esses resultados sugerem que os modelos são robustos e eficientes para classificação multi-classe.

#### 4.4 Experimento IV

Neste experimento, a base de dados foi balanceada para 4898 imagens por classe e utilizamos também três técnicas de aumento de dados para cada modelo: *MixUp*, *Datagen* e *CutMix*. Dessa forma, a avaliação das curvas ROC-AUC e PR-AUC continua sendo essencial para compreender melhor o desempenho dos modelos em diferentes classes e sua capacidade de generalização. A Tabela 15 apresenta os resultados deste experimento:

Observamos que os modelos que utilizaram a técnica de aumento de dados *CutMix* apresentaram os melhores desempenhos em todas as métricas avaliadas. Os modelos *VGG16 CutMix* com acurácia 0,8298, ROC-AUC 0,9722 e PR-AUC 0,7465 e *VGG19 CutMix* com acurácia 0,8261, ROC-AUC 0,9750 e PR-AUC 0,7411 destacaram-se como os mais eficazes, seguidos pelo *VGG16 MixUp* com acurácia 0,8191, ROC-AUC 0,9667 e PR-AUC 0,7172.

Os modelos que utilizaram e *Datagen* apresentaram desempenhos inferiores, especialmente em acurácia e PR-AUC, sugerindo que essa técnica pode não ter sido tão eficaz para esta base de dados quanto o *CutMix* e *MixUp*. Modelos como *ResNet50 Datagen* com acurácia 0,5559, ROC-AUC 0,9102, PR-AUC 0,6138, obtiveram dificuldades em obter um bom equilíbrio entre acurácia e capacidade de discriminação entre classes.

A análise do desempenho por classe dos modelos, apresentados nas Tabelas 16, 17 e 18, confirmam o *VGG16 CutMix* como o melhor modelo, apresentando maior acurácia global

Figura 35 – Treinamento III para os modelos *VGG16 MixUp*, *VGG19 MixUP* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.

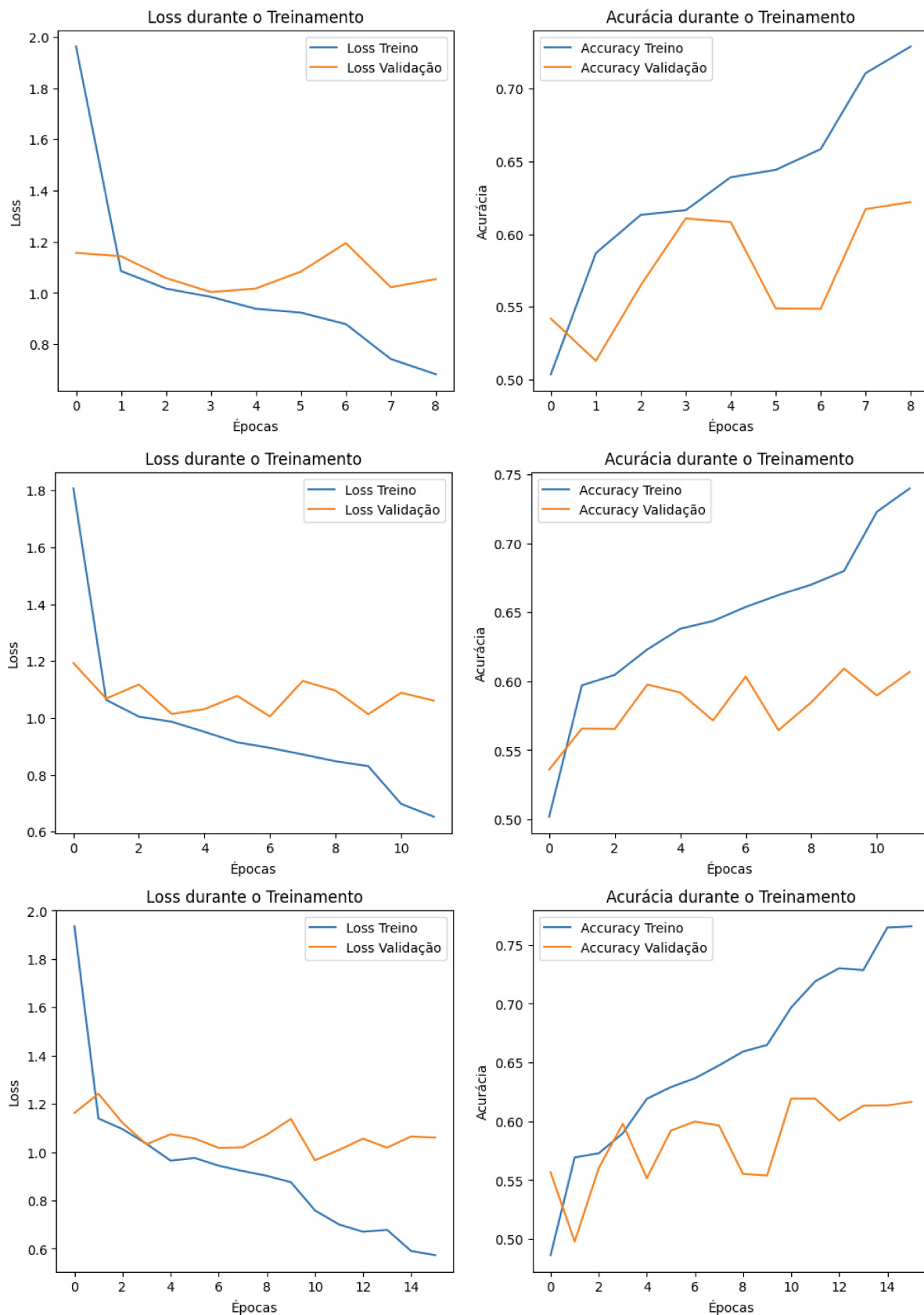
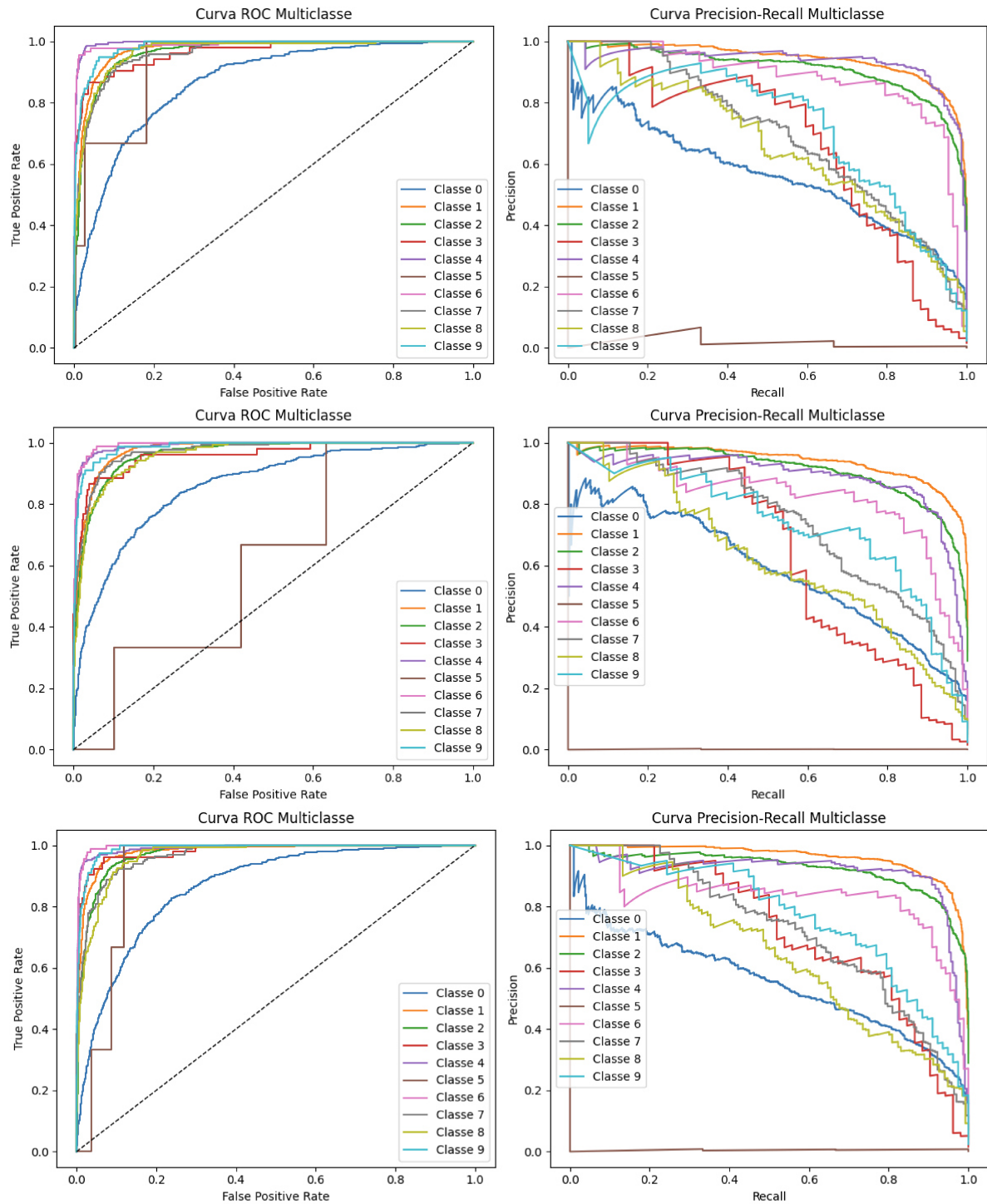


Figura 36 – ROC-AUC e PR-AUC III para os modelos *VGG16 MixUp*, *VGG19 MixUP* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



0,8298. *VGG16 CutMix* e *VGG19 CutMix* apresentaram baixa perda e valores mais equilibrados entre classes, com destaque para a Classe 1 0,96. Os modelos *VGG16 MixUp* e *VGG16 CutMix* também obtiveram *F1-score* médio superior aos demais resultados, figurando entre os três melhores modelos.

Os modelos *VGG16* e *VGG19* usando a técnica *CutMix* apresentam desempenho



Tabela 15. Comparação do desempenho dos modelos para o Experimento IV.

<b>Modelo</b>	<b>Acurácia</b>	<b>Perda</b>	<b>ROC-AUC</b>	<b>PR-AUC</b>
<b>VGG16 MixUp</b>	<b>0,8191</b>	<b>0,6290</b>	<b>0,9667</b>	<b>0,7172</b>
<i>VGG16 Datagen</i>	0,6542	1,4778	0,9302	0,6868
<b>VGG16 CutMix</b>	<b>0,8298</b>	<b>0,5649</b>	<b>0,9722</b>	<b>0,7465</b>
<i>VGG19 MixUp</i>	0,8182	0,6559	0,9592	0,7214
<i>VGG19 Datagen</i>	0,6490	1,2433	0,9510	0,6910
<b>VGG19 CutMix</b>	<b>0,8261</b>	<b>0,5912</b>	<b>0,9750</b>	<b>0,7411</b>
<i>ResNet50 MixUp</i>	0,8000	0,6728	0,9634	0,6978
<i>ResNet50 Datagen</i>	0,5559	2,9864	0,9102	0,6138
<i>ResNet50 CutMix</i>	0,8044	0,6561	0,9667	0,7099
<i>DenseNet121 MixUp</i>	0,8011	0,6978	0,9303	0,7189
<i>DenseNet121 Datagen</i>	0,5055	1,9635	0,9039	0,5366
<i>DenseNet121 CutMix</i>	0,7879	0,7115	0,9584	0,7115
<i>EfficientNetV2-M MixUp</i>	0,7714	0,6782	0,9441	0,6487
<i>EfficientNetV2-M Datagen</i>	0,6835	1,0741	0,9584	0,7062
<i>EfficientNetV2-M CutMix</i>	0,8026	0,6408	0,9380	0,6386
<i>AlexNet MixUp</i>	0,7548	0,7004	0,9356	0,6328
<i>AlexNet Datagen</i>	0,5578	0,5578	0,9395	0,5971
<i>AlexNet CutMix</i>	0,7834	0,6253	0,6958	0,6490

Tabela 16. Desempenho IV do modelo VGG16 MixUp por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,51	0,70	0,51	0,59
Classe 1	0,94	0,89	0,94	0,91
Classe 2	0,92	0,86	0,92	0,88
Classe 3	0,67	0,62	0,67	0,65
Classe 4	0,86	0,90	0,86	0,88
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,88	0,79	0,88	0,83
Classe 7	0,73	0,62	0,73	0,67
Classe 8	0,56	0,62	0,56	0,59
Classe 9	0,69	0,72	0,69	0,71
<b>Média Macro</b>	-	0,67	0,67	0,67
<b>Média Ponderada</b>	0,82	0,81	0,82	0,81

superior em comparação com o *VGG16 MixUp*, conforme observado durante o treinamento (Figura 37) e nas análises das curvas ROC-AUC e PR-AUC (Figura 38). O *VGG16 CutMix* mostra uma evolução consistente e sustentável na acurácia de validação, indicando um aproveitamento efetivo do treinamento prolongado. O *VGG19 CutMix* também melhora ao longo do tempo, mas com um leve indicativo de superajuste no final.

Tabela 17. Desempenho IV do modelo VGG16 CutMix por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,60	0,69	0,60	0,65
Classe 1	0,96	0,89	0,96	0,92
Classe 2	0,86	0,91	0,86	0,89
Classe 3	0,67	0,71	0,67	0,69
Classe 4	0,91	0,90	0,91	0,90
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,78	0,84	0,78	0,81
Classe 7	0,75	0,68	0,72	0,70
Classe 8	0,71	0,55	0,71	0,62
Classe 9	0,67	0,72	0,67	0,69
<b>Média Macro</b>	-	0,69	0,69	0,69
<b>Média Ponderada</b>	0,83	0,83	0,83	0,83

Tabela 18. Desempenho IV do modelo VGG19 CutMix por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,56	0,71	0,56	0,63
Classe 1	0,96	0,88	0,96	0,91
Classe 2	0,88	0,88	0,88	0,88
Classe 3	0,62	0,58	0,62	0,60
Classe 4	0,82	0,92	0,82	0,87
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,94	0,72	0,94	0,82
Classe 7	0,73	0,70	0,73	0,72
Classe 8	0,60	0,67	0,60	0,63
Classe 9	0,83	0,68	0,83	0,75
<b>Média Macro</b>	-	0,67	0,69	0,68
<b>Média Ponderada</b>	0,83	0,82	0,83	0,82

#### 4.5 Experimento V

Nesta seção, são apresentados os resultados do Experimento V para os dois modelos pré-treinados com melhor desempenho nos experimentos anteriores *VGG16* e *VGG19*. Também optamos por avaliar o desempenho do modelo *EfficientNetV2-M*, que apresentou resultado mediano, com o objetivo de analisar sua adequação a um cenário com mais imagens para treino. A análise considera as métricas de desempenho acurácia, perda, ROC-AUC e PR-AUC.

A base de dados utilizada foi balanceada para 20000 imagens por classe, e adotamos os dois tipos de aumento de dados que melhor se adequaram nos experimentos anteriores *MixUp* e *CutMix*. A Tabela 19 apresenta os resultados deste experimento:

A análise de desempenho revela o *VGG16 CutMix* como a melhor opção entre os avaliados, alcançando a maior accuracy 0,8307 e os melhores valores para ROC-AUC 0,9743 e

Figura 37 – Treinamento IV para os modelos *VGG16 MixUp*, *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.

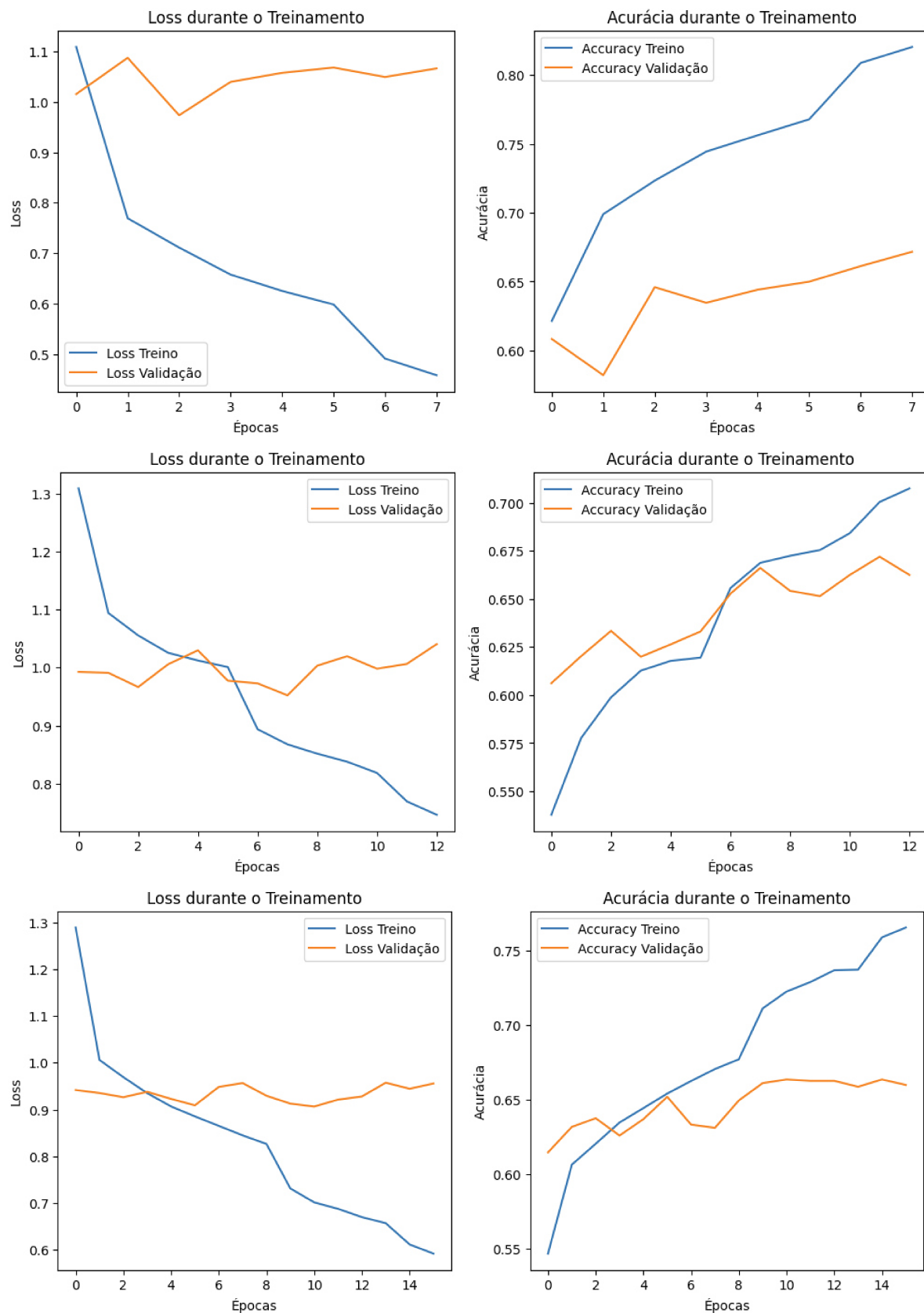
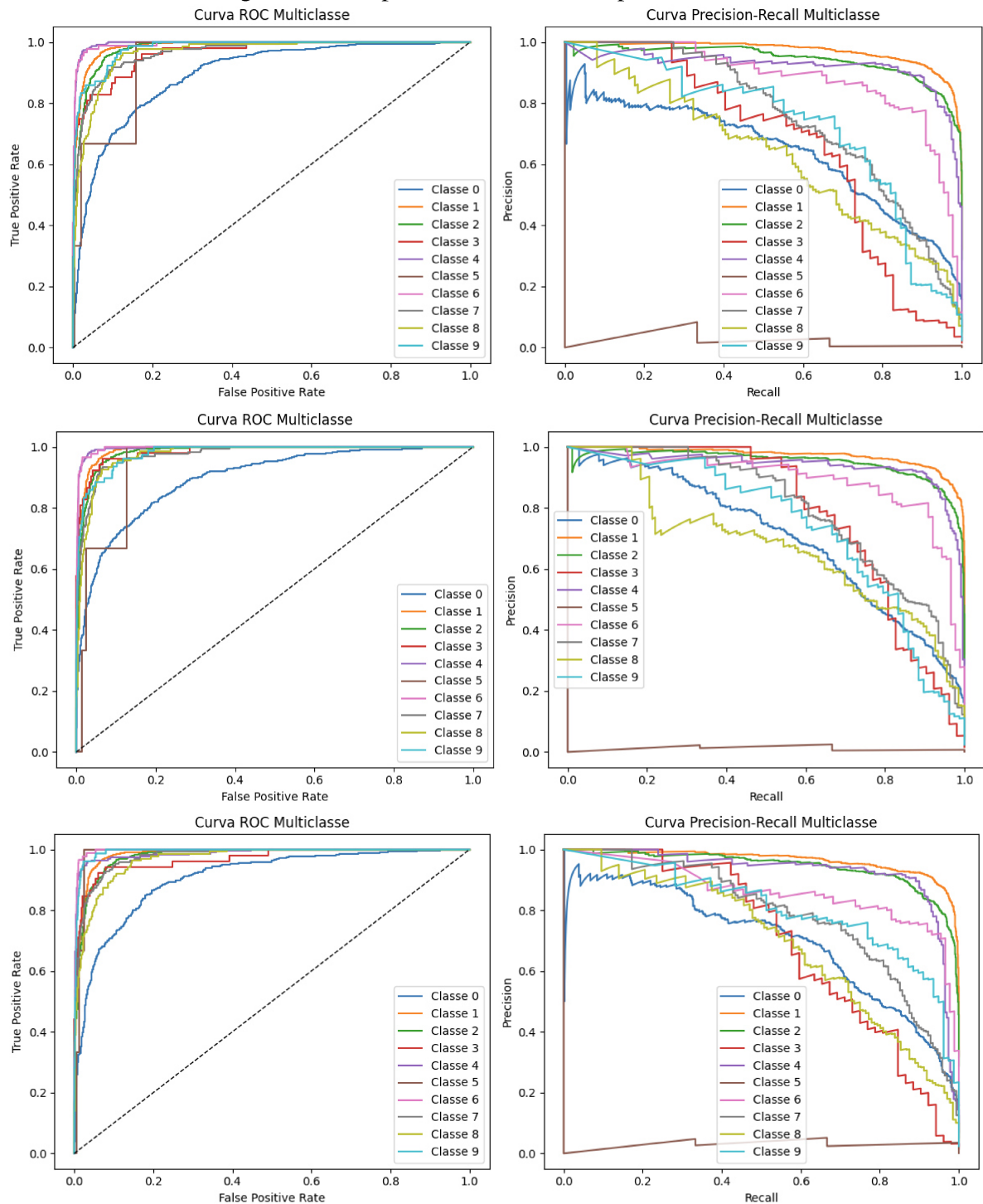


Figura 38 – ROC-AUC e PR-AUC IV para os modelos *VGG16 MixUp*, *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



PR-AUC 0,7467. Esses resultados indicam que o modelo possui uma excelente capacidade de generalização e discriminação entre classes, tornando-se a escolha mais robusta para a tarefa considerada. O *VGG19 CutMix* também apresentou um desempenho competitivo, com accuracy de 0,8255 e o menor valor de perda 0,5781, além de valores competitivos de ROC-AUC 0,9719 e PR-AUC 0,7422, sendo assim uma alternativa ao modelo anterior.

Tabela 19. Comparação do desempenho dos modelos para o Experimento V.

<b>Modelo</b>	<b>Acurácia</b>	<b>Perda</b>	<b>ROC-AUC</b>	<b>PR-AUC</b>
<i>VGG16 MixUp</i>	0,8059	0,7699	0,9570	0,7062
<b><i>VGG16 CutMix</i></b>	<b>0,8307</b>	<b>0,6030</b>	<b>0,9743</b>	<b>0,7467</b>
<i>VGG19 MixUp</i>	0,8102	0,6957	0,9401	0,7025
<b><i>VGG19 CutMix</i></b>	<b>0,8255</b>	<b>0,5781</b>	<b>0,9719</b>	<b>0,7422</b>
<i>EfficientNetV2-M MixUp</i>	0,7637	0,7950	0,9420	0,6266
<i>EfficientNetV2-M CutMix</i>	0,7836	0,6273	0,9512	0,6573

O *EfficientNetV2-M CutMix* destacou-se pelo seu equilíbrio entre métricas, atingindo um ROC-AUC de 0,9512 e um PR-AUC de 0,6573, embora com acurácia menor 0,7836. Isso sugere esse modelo como uma escolha interessante em cenários com um número maior de amostras e que necessitam de maior eficiência computacional. Em geral, os modelos que utilizaram a técnica *CutMix* obtiveram os melhores resultados, reforçando a evidência de que essa técnica de aumento de dados é adequada na melhoria da capacidade de generalização dos modelos de classificação de galáxias.

Os modelos *VGG16* e *VGG19 CutMix* mostram desempenhos diferentes em acurácia por classe (Tabelas 20 e 21), com o *VGG16 CutMix* superando o *VGG19* em várias classes, apesar de ambos falharem na Classe 5 com acurácia zero.

Tabela 20. Desempenho V do modelo VGG16 CutMix por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,70	0,64	0,70	0,67
Classe 1	0,88	0,94	0,88	0,91
Classe 2	0,91	0,86	0,91	0,88
Classe 3	0,73	0,73	0,73	0,73
Classe 4	0,91	0,91	0,91	0,91
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,84	0,83	0,84	0,84
Classe 7	0,58	0,82	0,58	0,68
Classe 8	0,69	0,60	0,69	0,64
Classe 9	0,69	0,74	0,69	0,72
<b>Média Macro</b>	-	0,71	0,69	0,70
<b>Média Ponderada</b>	0,83	0,84	0,83	0,83

Os gráficos de treinamento dos modelos *VGG16* e *VGG19 CutMix* revelam uma redução consistente na perda de treinamento (Figura 39), com estabilização na perda de validação, sugerindo sinais de *overfitting*, especialmente no *VGG19*. A acurácia de treinamento aumenta para ambos, mas o *VGG16* demonstra uma generalização ligeiramente melhor. Nas curvas ROC-AUC e precisão-revocação (Figura 40), o *VGG16* geralmente supera o *VGG19*, destacando-se

Tabela 21. Desempenho V do modelo VGG19 CutMix por classe.

Classe	Acurácia	Precisão	Revocação	F1-Score
Classe 0	0,59	0,67	0,59	0,63
Classe 1	0,95	0,89	0,95	0,92
Classe 2	0,87	0,90	0,87	0,88
Classe 3	0,67	0,73	0,67	0,70
Classe 4	0,91	0,92	0,91	0,91
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,85	0,80	0,85	0,82
Classe 7	0,66	0,72	0,66	0,69
Classe 8	0,67	0,52	0,67	0,59
Classe 9	0,76	0,65	0,76	0,70
<b>Média Macro</b>	-	0,68	0,69	0,68
<b>Média Ponderada</b>	0,83	0,83	0,83	0,82

nas curvas ROC. Contudo, ambas as curvas de precisão-revocação indicam dificuldades em prever os valores corretos para a Classe 5, em ambos os modelos.

#### 4.6 Experimento VI

Nesta seção, são apresentados os resultados do Experimento VI para os dois modelos pré-treinados que obtiveram melhor desempenho nos experimentos II, III e IV: *VGG16* e *VGG19*. Também optamos por avaliar o desempenho do modelo *EfficientNetV2-M*, um modelo que apresentou resultado mediano, com o objetivo de analisar como ele se adequa a um cenário com imagens de dimensionalidade 244x244 *pixels*. A análise inclui as principais métricas para a avaliação de desempenho, entre as quais se destacam: Acurácia, Perda, ROC-AUC e PR-AUC.

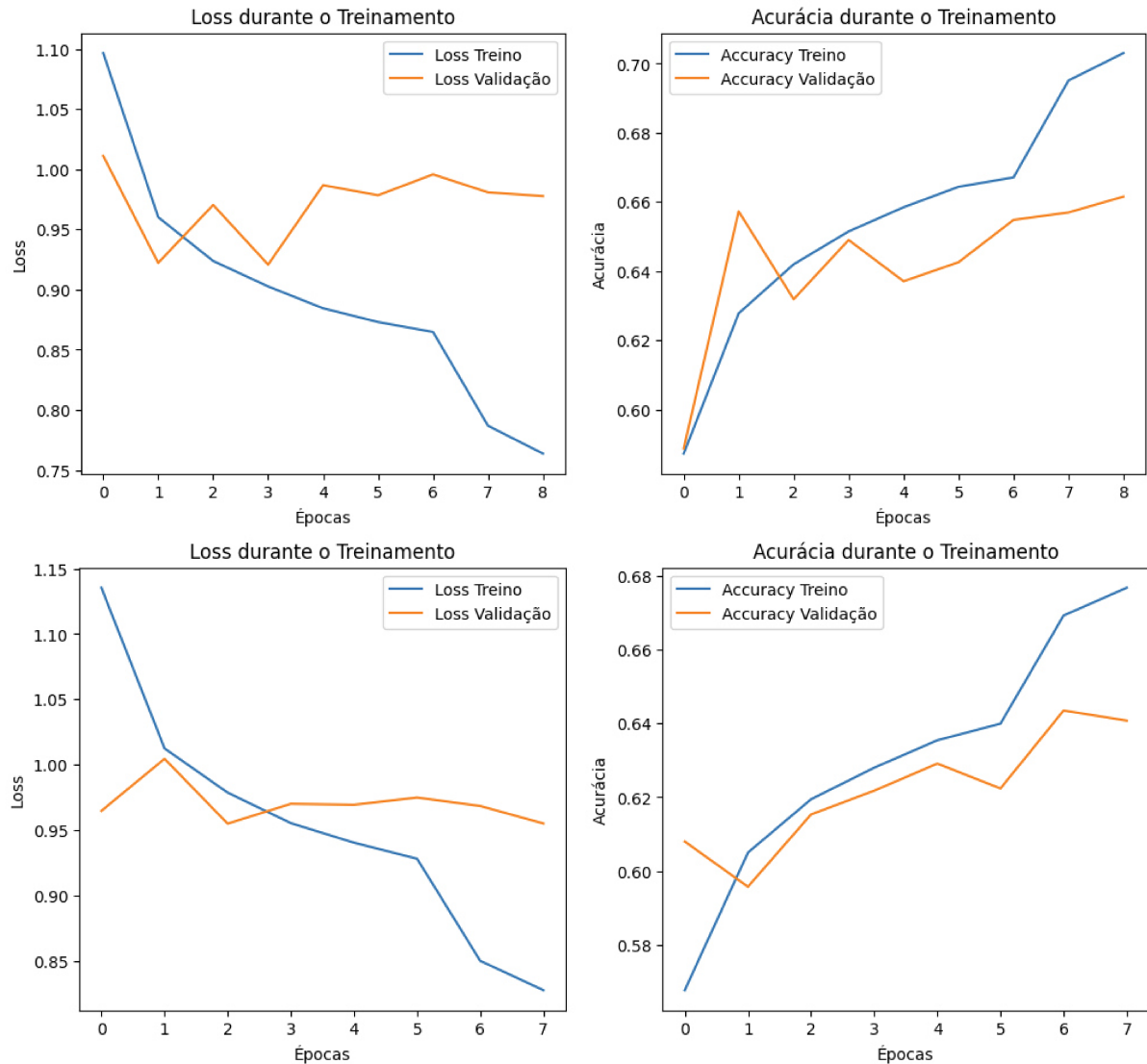
A base de dados utilizada esta balanceada para 4898 imagens por classe, mas agora com imagens redimensionadas para 244x244 *pixels*. Utilizamos novamente as duas técnicas de aumento de dados que melhor se adequaram aos experimentos anteriores *MixUp* e *CutMix*. A Tabela 22 apresenta os resultados deste experimento:

Tabela 22. Comparação do desempenho dos modelos para o Experimento VI.

Modelo	Acurácia	Perda	ROC-AUC	PR-AUC
<i>VGG16 MixUp</i>	0,8835	0,3246	0,9813	0,9356
<b><i>VGG16 CutMix</i></b>	<b>0,8886</b>	<b>0,2979</b>	<b>0,9821</b>	<b>0,9415</b>
<i>VGG19 MixUp</i>	0,8811	0,3308	0,9791	0,9344
<b><i>VGG19 CutMix</i></b>	<b>0,9010</b>	<b>0,3048</b>	<b>0,9824</b>	<b>0,9467</b>
<i>EfficientNetV2-M MixUp</i>	0,8880	0,4388	0,9779	0,9345
<i>EfficientNetV2-M CutMix</i>	0,8629	0,3647	0,9762	0,9217

A análise dos treinados com as técnicas de aumento de dados *MixUp* e *CutMix*

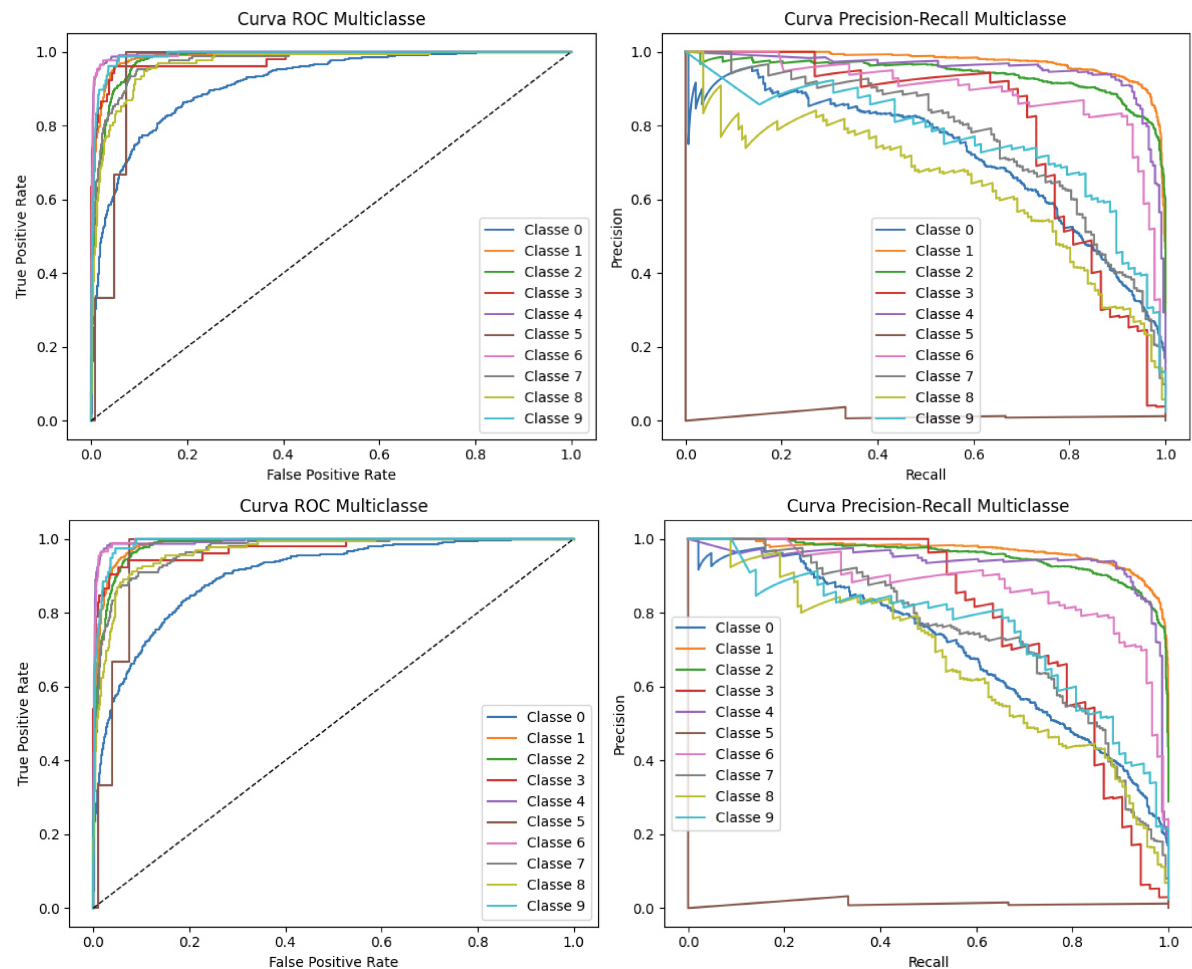
Figura 39 – Treinamento V para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



revelou que os modelos *VGG16* e *VGG19* apresentaram os melhores desempenhos. O *VGG19 CutMix* se destacou com a maior acurácia 0,9010, maior ROC-AUC 0,9824 e maior PR-AUC 0,9467, indicando uma excelente capacidade de generalização. Já o *VGG16 CutMix* obteve a menor perda 0,2979 e um dos maiores ROC-AUC 0,9821, sugerindo um bom ajuste aos dados de treinamento e uma excelente distinção entre classes. Dentre os modelos avaliados, os que utilizaram *CutMix* tiveram melhor desempenho geral do que aqueles treinados com *MixUp*, especialmente para *VGG19* e *VGG16*.

Embora os modelos baseados no *EfficientNetV2-M* tenham obtido resultados competitivos no experimento anterior, na configuração atual seu desempenho foi inferior em todas as métricas quando comparado ao *VGG* com *CutMix*. Isso sugere que o *EfficientNetV2-M* possa ter um desempenho mais eficaz em treinamentos com um maior volume de dados.

Figura 40 – ROC-AUC e PR-AUC V para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



Os modelos *VGG16 CutMix* e *VGG19 CutMix* apresentam variações na acurácia por classe (Tabelas 23 e 24). No modelo *VGG16 CutMix*, a acurácia varia de 0,67 a 0,98, com médias macro e ponderada de 0,89. No *VGG19 CutMix*, as acurácias são geralmente mais altas, variando de 0,75 a 0,99, com uma média macro de 0,87 e média ponderada de 0,90. Isso indica que o *VGG19 CutMix*, apesar de uma média macro ligeiramente inferior, oferece um desempenho geral melhor, especialmente com uma melhora significativa na Classe 0.

Tabela 23. Desempenho VI do modelo *VGG16 CutMix* por classe.

Classe	Acurácia	Precisão	Revocação	<i>F1-Score</i>
Classe 0	0,67	0,84	0,67	0,75
Classe 1	0,98	0,88	0,98	0,93
Classe 2	0,90	0,91	0,90	0,91
Classe 4	0,98	0,96	0,98	0,97
Classe 7	0,82	0,81	0,82	0,81
<b>Média Macro</b>	-	0,88	0,87	0,87
<b>Média Ponderada</b>	0,89	0,89	0,89	0,89



Tabela 24. Desempenho VI do modelo VGG19 CutMix por classe.

Classe	Acurácia	Precisão	Revocação	F1-Score
Classe 0	0,75	0,82	0,75	0,79
Classe 1	0,95	0,92	0,95	0,94
Classe 2	0,91	0,91	0,91	0,91
Classe 4	0,99	0,96	0,99	0,97
Classe 7	0,87	0,88	0,87	0,88
<b>Média Macro</b>	-	0,90	0,90	0,90
<b>Média Ponderada</b>	0,90	0,90	0,90	0,90

Analisando a Figura 41 Os modelos *VGG16* e *VGG19 CutMix* exibem um treinamento eficaz com a perda diminuindo e a acurácia aumentando ao longo das épocas, indicando boa generalização sem sobreajuste. As análises das curvas ROC-AUC e precisão-revocação (Figura 42) revelam altos desempenhos para ambos os modelos, com o *VGG19* apresentando ligeiramente melhores resultados. Isso mostra que os modelos são eficientes na classificação, com o *VGG19* possivelmente oferecendo uma melhor precisão e revocação geral.

#### 4.7 Experimento VII

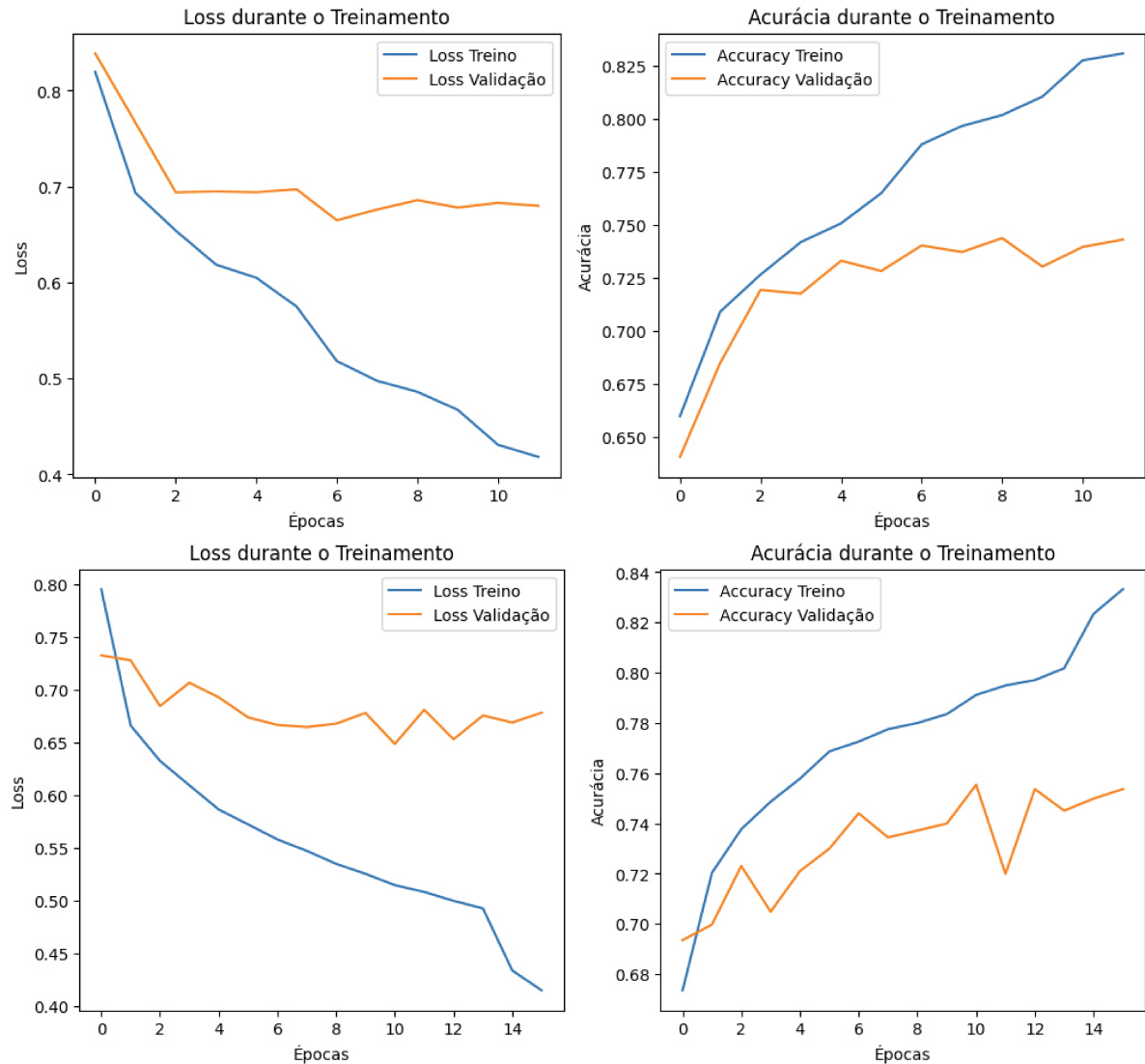
Nesta seção, são apresentados os resultados do Experimento VII para os dois modelos pré-treinados que obtiveram melhor desempenho nos experimentos II, III e IV: *VGG16* e *VGG19*. Também optamos por avaliar o desempenho do modelo *EfficientNetV2-M*, um modelo que apresentou resultado mediano, com o objetivo de analisar como ele se adequa a um cenário com mais imagens para treino. A análise inclui as principais métricas para a avaliação de desempenho, entre as quais se destacam: Acurácia, Perda, ROC-AUC e PR-AUC.

Neste caso, a base de dados utilizada estava balanceada para 20000 imagens por classe, optamos por adotar os dois tipos de aumento de dados que se adequaram melhor nos experimentos anteriores *MixUp* e *CutMix*. O diferencial deste experimento para o experimento V foi a aplicação de *CLAHE* para aumento do contraste das imagens durante a etapa de pré-processamento. A Tabela 25 apresenta os resultados deste experimento:

Tabela 25. Comparação do desempenho dos modelos para o Experimento VII.

Modelo	Acurácia	Perda	ROC-AUC	PR-AUC
<i>VGG16 MixUp</i>	0,7775	0,7749	0,9331	0,6880
<b><i>VGG16 CutMix</i></b>	<b>0,8167</b>	<b>0,5465</b>	<b>0,9695</b>	<b>0,7373</b>
<i>VGG19 MixUp</i>	0,7962	0,6672	0,9642	0,6945
<b><i>VGG19 CutMix</i></b>	<b>0,7952</b>	<b>0,5995</b>	<b>0,9607</b>	<b>0,7205</b>
<i>EfficientNetV2-M MixUp</i>	0,7567	0,6878	0,9344	0,6307
<i>EfficientNetV2-M CutMix</i>	0,7858	0,6649	0,9323	0,6765

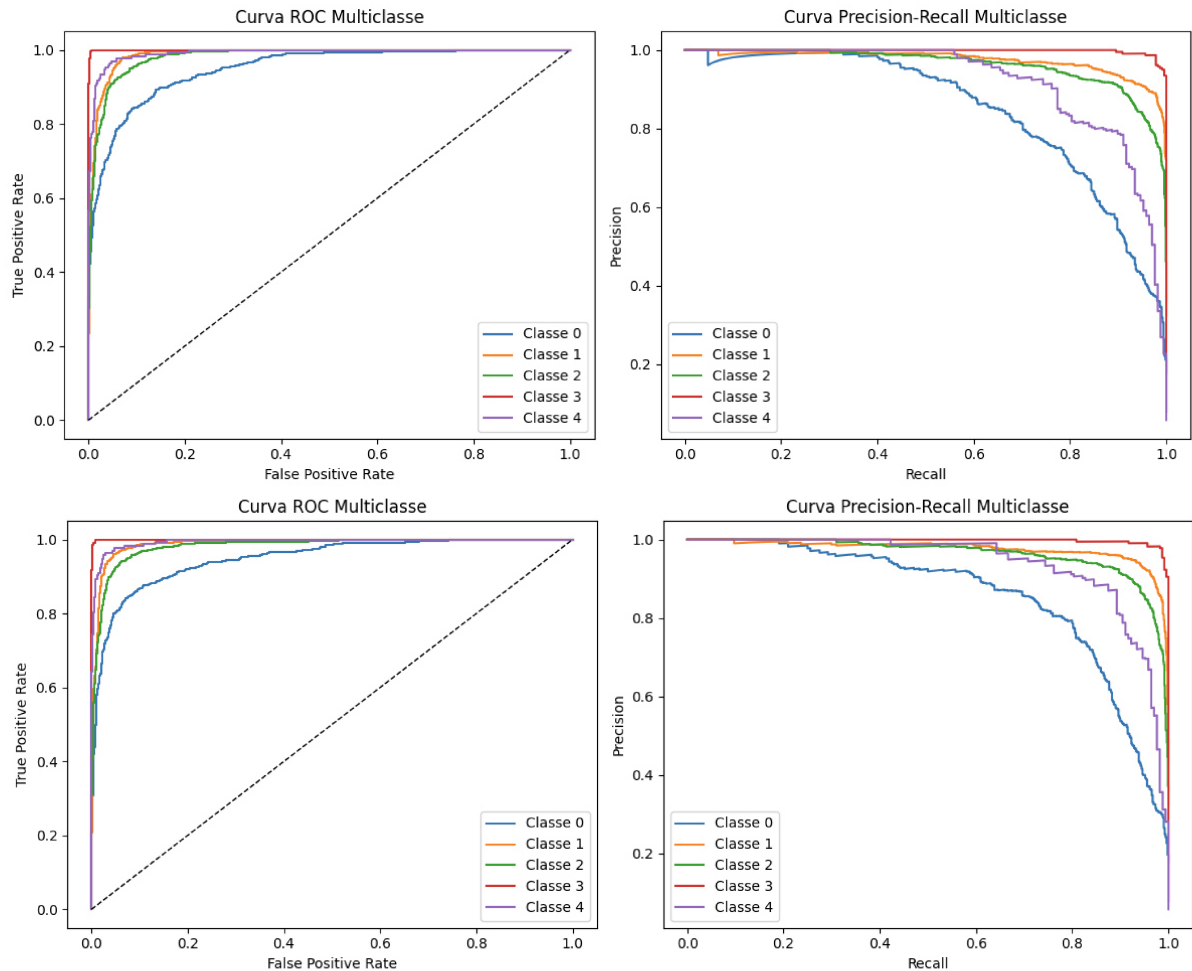
Figura 41 – Treinamento VI para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



Para o experimento VII, o modelo *VGG16 CutMix* apresentou o melhor desempenho geral, obtendo a maior acurácia 0,8167, a menor perda 0,5465 e os melhores valores de ROC-AUC 0,9695 e PR-AUC 0,7373. Esses resultados indicam que a técnica de aumento de dados *CutMix* foi particularmente eficaz quando aplicada à arquitetura *VGG16*, melhorando a capacidade do modelo de distinguir entre classes e reduzindo o erro durante a classificação.

O modelo *VGG19 CutMix* também apresentou um bom desempenho, com um ROC-AUC de 0,9607 e PR-AUC de 0,7205, mas ainda inferior ao *VGG16 CutMix*. Entre os modelos *EfficientNetV2-M*, a versão com *CutMix* superou a variante com *MixUp*, porém ambas tiveram resultados inferiores às arquiteturas *VGG16* e *VGG19*. Assim, os resultados sugerem que a combinação da arquitetura *VGG16* com a técnica *CutMix* é a mais adequada para este conjunto de dados balanceado com aplicação de *CLAHE*.

Figura 42 – ROC-AUC e PR-AUC VI para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



As Tabelas 26 e 27 destacam *VGG16 CutMix* e *VGG19 CutMix*, em termos de acurácia, precisão, revocação e *F1-Score* para diferentes classes. O modelo *VGG16 CutMix* atinge uma acurácia média ponderada de 0,81, enquanto o *VGG19 CutMix* tem uma acurácia média ponderada ligeiramente menor, de 0,80. Ambos os modelos exibem variações no desempenho entre as classes, sugerindo que alguns tipos de dados ou características são mais desafiadores de prever do que outros.

As Figuras 43 e 44 analisam o treinamento e a performance dos modelos *VGG16* e *VGG19 CutMix*. Os gráficos de treinamento mostram uma redução na perda e um aumento na acurácia ao longo das épocas, indicando melhora no aprendizado, mas com sinais de leve *overfitting*, visto que a perda de validação estagna ou aumenta. As curvas ROC-AUC e PR-AUC revelam um bom desempenho na distinção entre as classes, com algumas variações entre elas.

Comparando os resultados dos Experimentos V (Seção 4.5) e VII (Seção 4.7), observa-se que a aplicação do *CLAHE* no Experimento VII não trouxe melhorias consistentes

Tabela 26. Desempenho VII do modelo VGG16 CutMix por classe.

Classe	Acurácia	Precisão	Revocação	F1-Score
Classe 0	0,51	0,72	0,51	0,59
Classe 1	0,92	0,89	0,92	0,91
Classe 2	0,92	0,82	0,92	0,87
Classe 3	0,79	0,59	0,79	0,67
Classe 4	0,90	0,88	0,90	0,89
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,81	0,85	0,81	0,83
Classe 7	0,62	0,68	0,62	0,65
Classe 8	0,67	0,65	0,67	0,66
Classe 9	0,71	0,65	0,71	0,69
<b>Média Macro</b>	-	0,67	0,68	0,67
<b>Média Ponderada</b>	0,81	0,82	0,81	0,83

Tabela 27. Desempenho VII do modelo VGG19 CutMix por classe.

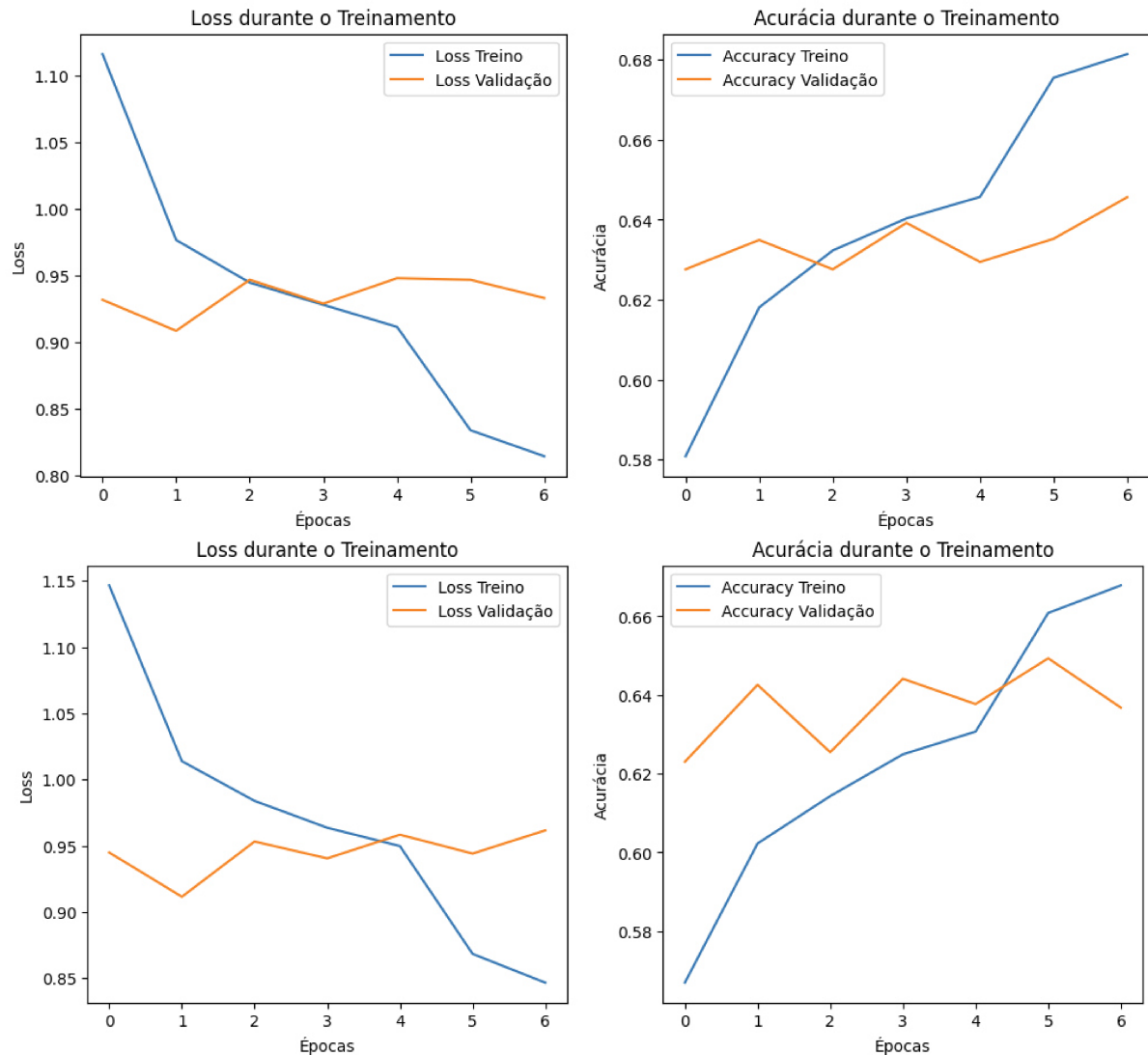
Classe	Acurácia	Precisão	Revocação	F1-Score
Classe 0	0,53	0,63	0,53	0,58
Classe 1	0,87	0,92	0,87	0,90
Classe 2	0,88	0,84	0,88	0,86
Classe 3	0,63	0,51	0,63	0,59
Classe 4	0,88	0,85	0,88	0,86
Classe 5	0,00	0,00	0,00	0,00
Classe 6	0,84	0,67	0,84	0,75
Classe 7	0,73	0,57	0,73	0,64
Classe 8	0,61	0,65	0,61	0,63
Classe 9	0,73	0,70	0,73	0,71
<b>Média Macro</b>	-	0,63	0,67	0,65
<b>Média Ponderada</b>	0,80	0,80	0,79	0,81

no desempenho dos modelos. Embora o modelo *VGG16 CutMix* tenha mantido um bom desempenho, com queda da acurácia de 0,8307 no Exp. V e 0,8167 para no Exp. VII. Houve pequenas quedas nas métricas PR-AUC de 0,7467 para 0,7373 e ROC-AUC de 0,9743 para 0,9695. O mesmo padrão foi observado nos demais modelos, com ligeira redução na acurácia e PR-AUC, sugerindo que o uso do CLAHE não impactou positivamente o treinamento e, em alguns casos, pode até ter prejudicado levemente o desempenho dos modelos.

#### 4.8 Experimento VIII

Nesta seção, são apresentados os resultados do Experimento VIII para os dois modelos pré-treinados que obtiveram melhor desempenho nos experimentos II, III e IV: *VGG16* e *VGG19*. Também optamos por avaliar o desempenho do modelo *EfficientNetV2-M*, um modelo que apresentou resultado mediano, com o objetivo de analisar como ele se adequa a um cenário

Figura 43 – Treinamento VII para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.



com imagens de dimensionalidade  $244 \times 244$  pixels. A análise inclui as principais métricas para a avaliação de desempenho, entre as quais se destacam: Acurácia, Perda, ROC-AUC e PR-AUC.

Neste experimento, a base de dados utilizada estava balanceada para 4898 imagens por classe, agora com imagens redimensionadas para  $244 \times 244$  pixels. Optamos também por utilizar os dois tipos de aumento de dados que se adequaram melhor nos experimentos anteriores *MixUp* e *CutMix*. O diferencial deste experimento para o experimento VI foi a aplicação de *CLAHE* para aumento do contraste das imagens durante a etapa de pré-processamento. A Tabela 28 apresenta os resultados deste experimento:

Os resultados da mostram que o modelo *EfficientNetV2-M CutMix* foi o melhor desempenho geral, apresentando a maior acurácia 0,8883, a menor perda 0,3295 e os melhores valores de ROC-AUC 0,9769 e PR-AUC 0,9315. Isso indica que esse modelo teve a melhor

Figura 44 – ROC-AUC e PR-AUC VII para os modelos *VGG16 CutMix* e *VGG19 CutMix*, organizadas respectivamente de cima para baixo.

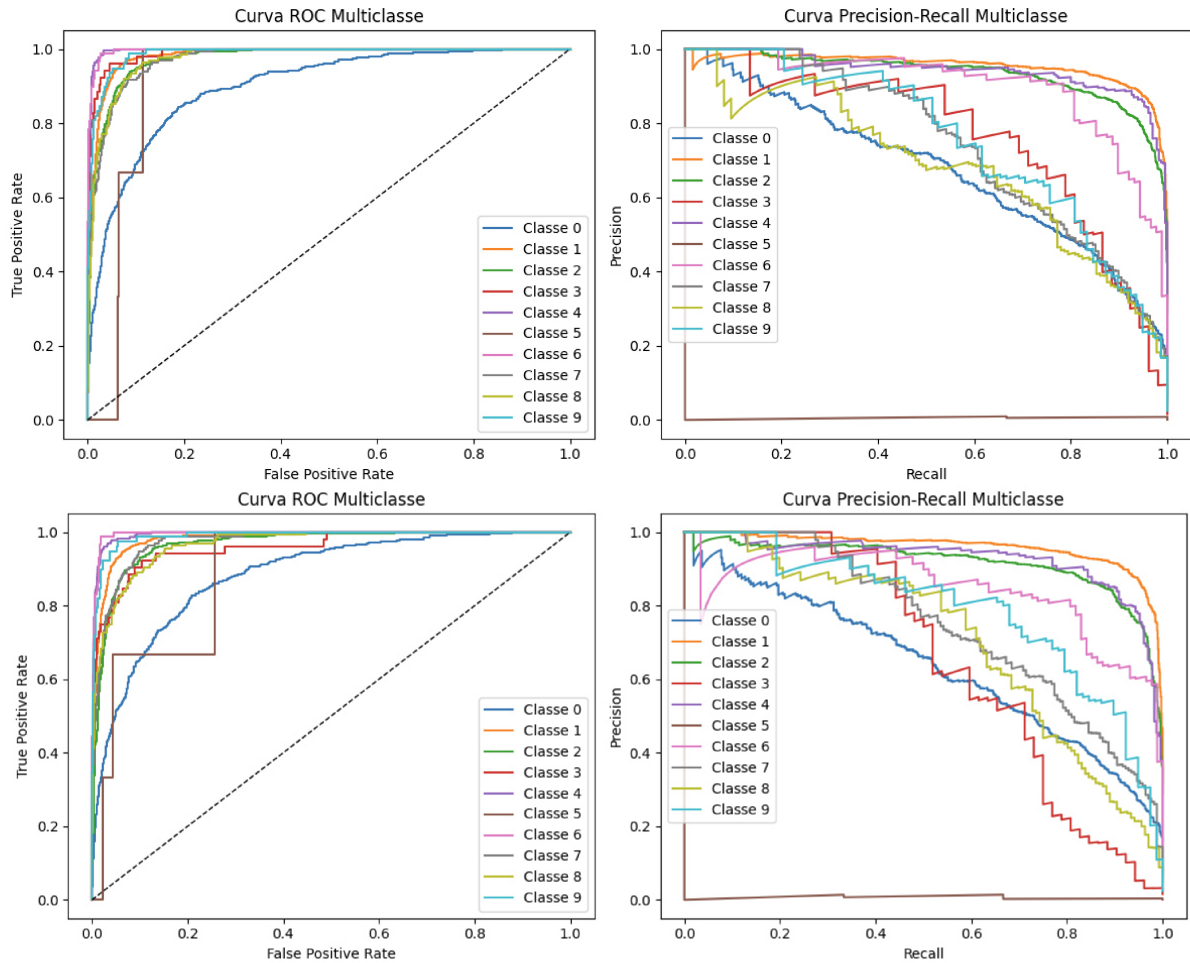


Tabela 28. Comparação do desempenho dos modelos para o Experimento VIII.

Modelo	Acurácia	Perda	ROC-AUC	PR-AUC
<i>VGG16 MixUp</i>	0,8612	0,4240	0,9690	0,9061
<b><i>VGG16 CutMix</i></b>	<b>0,8711</b>	<b>0,3876</b>	<b>0,9720</b>	<b>0,9172</b>
<i>VGG19 MixUp</i>	0,8643	0,4033	0,9702	0,9115
<i>VGG19 CutMix</i>	0,8601	0,4000	0,9691	0,9079
<i>EfficientNetV2-M MixUp</i>	0,8498	0,4222	0,9695	0,9163
<b><i>EfficientNetV2-M CutMix</i></b>	<b>0,8883</b>	<b>0,3295</b>	<b>0,9769</b>	<b>0,9315</b>

capacidade de classificar corretamente as imagens, com menor erro e alta separabilidade entre classes. O segundo melhor modelo foi o *VGG16 CutMix*, que obteve uma acurácia de 0,8711, uma perda de 0,3876 e boas pontuações de ROC-AUC 0,9720 e PR-AUC 0,9172, destacando-se como a melhor opção entre os modelos *VGG*.

Os modelos *VGG16 MixUp*, *VGG19 MixUp* e *VGG19 CutMix* também tiveram desempenhos competitivos, com acurácias de 0,8612, 0,8643 e 0,8601, respectivamente, e boas métricas de ROC-AUC e PR-AUC, embora inferiores ao *VGG16 CutMix*. De forma geral, os resultados reforçam o destaque da arquitetura *EfficientNetV2-M*, especialmente quando

combinada com a técnica de aumento de dados *CutMix*, redimensionamento 244x244 *pixels* e aplicação de *CLAHE*, enquanto o *VGG16 CutMix* se mostrou a melhor opção dentro da família *VGG*.

As Tabelas 29 e 30 mostram o desempenho de *VGG16 CutMix* e *VGG19 CutMix*, em várias classes. Observando os dados, o modelo *VGG19* geralmente supera o *VGG16* em quase todas as métricas por classe. Por exemplo, na Classe 0, o *VGG19* tem maior precisão, revocação e F1-Score. Similarmente, em métricas agregadas como média ponderada, o *VGG19* também mostra desempenho superior, com uma média ponderada de 0,89 contra 0,87 do *VGG16*. Este padrão é consistente através das outras classes, indicando que o modelo *VGG19* pode ser mais eficaz para esse conjunto específico de dados e tarefas de classificação.

Tabela 29. Desempenho VIII do modelo *VGG16 CutMix* por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,69	0,76	0,69	0,72
Classe 1	0,90	0,93	0,90	0,91
Classe 2	0,94	0,84	0,94	0,89
Classe 4	0,99	0,94	0,99	0,96
Classe 7	0,73	0,92	0,73	0,81
<b>Média Macro</b>	-	0,88	0,85	0,86
<b>Média Ponderada</b>	0,87	0,87	0,87	0,89

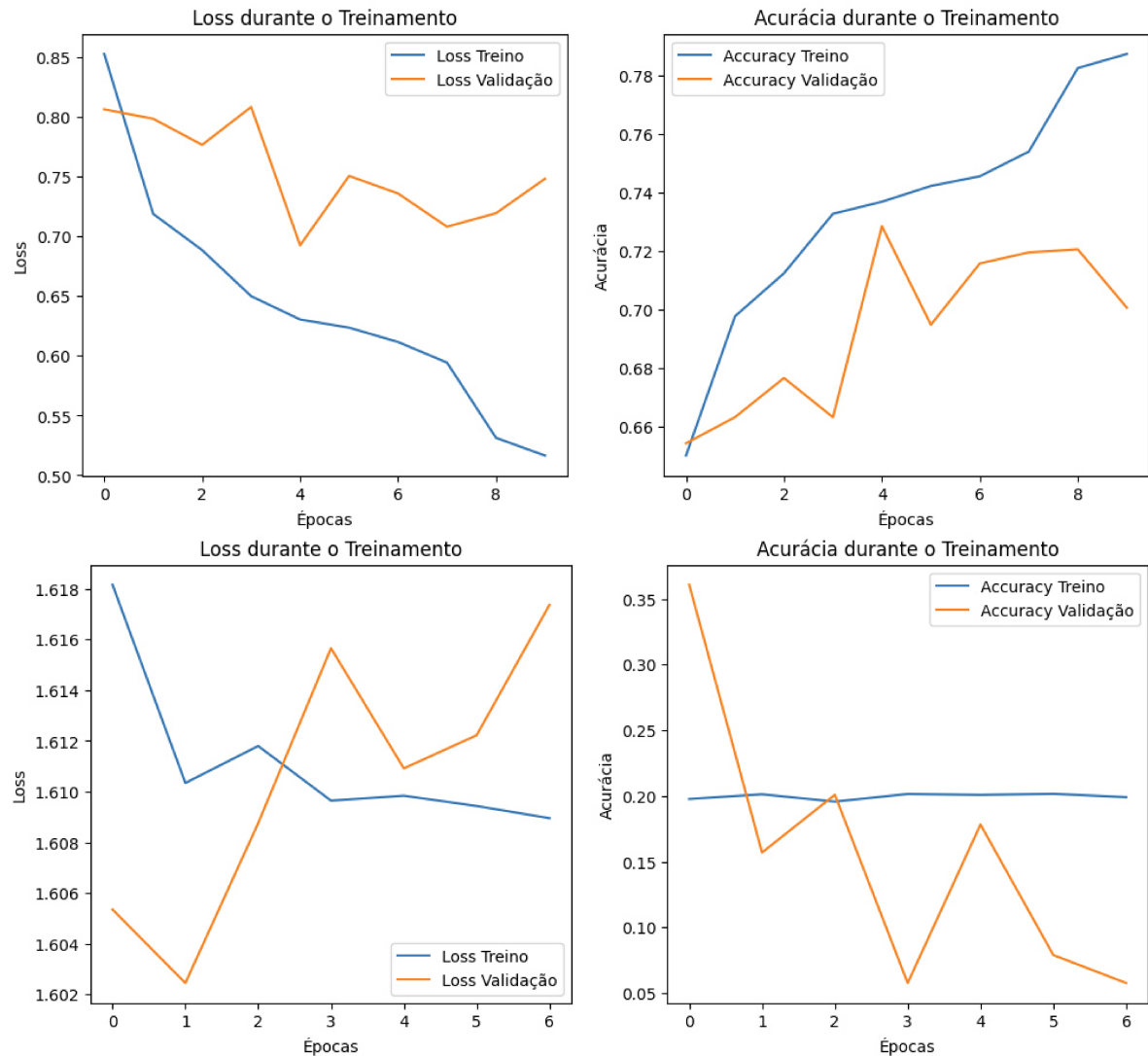
Tabela 30. Desempenho VIII do modelo *VGG19 CutMix* por classe.

<b>Classe</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Revocação</b>	<b>F1-Score</b>
Classe 0	0,68	0,82	0,68	0,75
Classe 1	0,97	0,90	0,97	0,93
Classe 2	0,90	0,89	0,90	0,89
Classe 4	0,98	0,98	0,98	0,98
Classe 7	0,83	0,86	0,83	0,85
<b>Média Macro</b>	-	0,89	0,87	0,88
<b>Média Ponderada</b>	0,89	0,89	0,89	0,90

Com base nas Tabelas 45 e 46, o modelo *VGG16 CutMix* apresentou treinamento mais estável e consistente, com melhores resultados nas curvas ROC-AUC e PR-AUC, indicando maior capacidade de generalização e melhor desempenho na distinção entre classes em comparação ao *EfficientNetV2-M CutMix*.

Comparando os Experimentos VI (Seção 4.6) e VIII (Seção 4.8), observa-se que a aplicação do *CLAHE* no Experimento VIII não melhorou o desempenho dos modelos *VGG*, que apresentaram quedas em todas as métricas, especialmente no PR-AUC e na acurácia. Por

Figura 45 – Treinamento VIII para os modelos *VGG16 CutMix* e *EfficientNetV2-M CutMix*, organizadas respectivamente de cima para baixo.



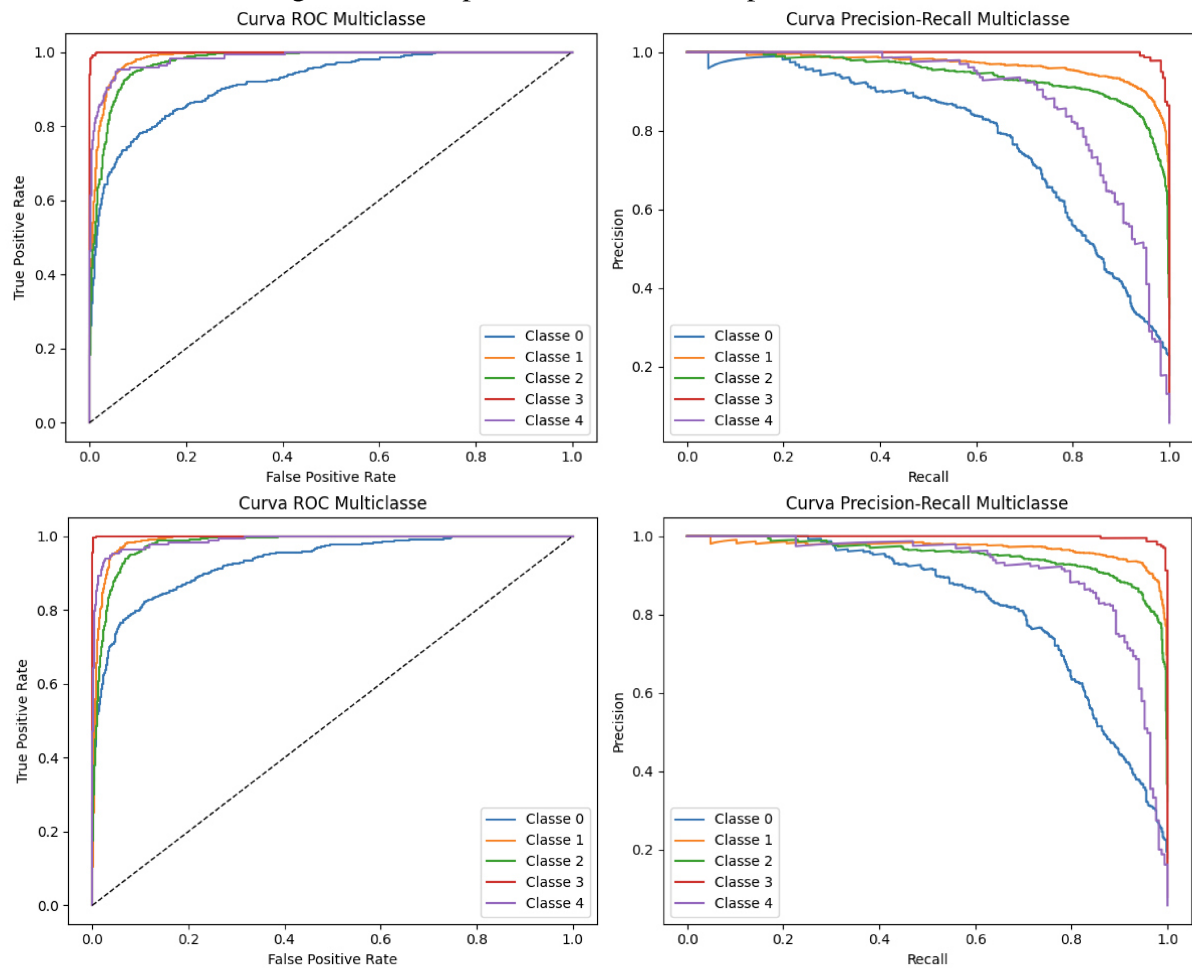
exemplo, o *VGG19 CutMix* caiu de 0,9010 para 0,8601 em acurácia e de 0,9467 para 0,9079 em PR-AUC. Por outro lado, os modelos *EfficientNetV2-M*, principalmente com *CutMix*, mostraram melhora com a aplicação do *CLAHE*: a acurácia subiu de 0,8629 para 0,8883, e o PR-AUC de 0,9217 para 0,9315. Esses resultados indicam que o *CLAHE* pode ser mais benéfico para arquiteturas como a *EfficientNetV2-M*, enquanto pode prejudicar ligeiramente o desempenho das arquiteturas *VGG*.

#### 4.9 Comparação dos melhores desempenhos

A Tabela 31 apresenta o desempenho geral de diferentes experimentos com modelos de redes neurais. O melhor desempenho global foi obtido no Exp VI por *VGG19 CutMix*, que alcançou a maior acurácia 0,9010, a menor perda 0,3048 e os maiores valores de ROC-AUC



Figura 46 – ROC-AUC e PR-AUC VIII para os modelos *VGG16 CutMix* e *EfficientNetV2-M CutMix*, organizadas respectivamente de cima para baixo.



0,9824 e PR-AUC 0,9467, indicando excelente poder de classificação e generalização.

Tabela 31. Comparação dos melhores resultados dentre os oito modelos analisados

Modelo	Acurácia	Perda	ROC-AUC	PR-AUC
Exp I - <i>Gharat &amp; Dandawate</i>	0,8300	-	-	-
Exp II - <i>VGG19</i>	0,8323	0,4911	0,9719	0,7330
Exp III - <i>VGG19 CutMix</i>	0,7836	0,6458	0,9630	0,7052
Exp IV - <i>VGG16 CutMix</i>	0,8298	0,5649	0,9722	0,7465
Exp V - <i>VGG16 CutMix</i>	0,8307	0,6030	0,9743	0,7467
<b>Exp VI - <i>VGG19 CutMix</i></b>	<b>0,9010</b>	<b>0,3048</b>	<b>0,9824</b>	<b>0,9467</b>
Exp VII - <i>VGG16 CutMix</i>	0,8167	0,5465	0,9695	0,7373
Exp VIII - <i>EfficientNetV2-M CutMix</i>	0,8883	0,3295	0,9769	0,9315

Em contraste, o *EfficientNetV2-M CutMix* (Exp VIII), embora tenha apresentado boa acurácia 0,8883 e métricas AUC elevadas, ROC-AUC 0,9769 e PR-AUC 0,9315, ainda ficou levemente abaixo do desempenho do *VGG19 CutMix*. Modelos sem *CutMix*, como o *VGG19* (Exp II), também obtiveram resultados satisfatórios, porém inferiores ao uso combinado com a

técnica *CutMix*.

Portanto, a combinação *VGG19* com *CutMix* se destacou como a mais eficaz entre todos os experimentos avaliados.

## 5 CONCLUSÕES

Diante dos objetivos propostos, este trabalho explorou a aplicação de Redes Neurais Convolucionais na classificação de objetos astronômicos, com foco na identificação e categorização de galáxias da base de dados *Galaxy10 SDSS Dataset*. Analisamos as principais abordagens utilizadas na literatura, bem como o desempenho de diferentes modelos pré-treinados, incluindo *EfficientNetV2-M*, *DenseNet121*, *ResNet50*, *AlexNet*, *VGG16* e *VGG19*. Além disso, investigamos a eficácia de estratégias de aumento de dados, como *Datagen*, *MixUp*, *CutMix* e aprimoramento de contraste por meio de *CLAHE*, visando melhorar a robustez e a generalização dos modelos testados.

A reprodução da arquitetura de Gharat & Dandawate (2022) evidenciou que uma acurácia geral elevada pode ser enganosa, uma vez que o modelo apresentou dificuldades significativas em classes menos representadas, reforçando a necessidade de métricas complementares, como ROC-AUC e PR-AUC.

As técnicas de *Transfer Learning* aprimoraram significativamente a robustez dos modelos *VGG16* e *VGG19*. Tanto aplicados isoladamente, quanto em combinação com o método de aumento de dados *CutMix* e o redimensionamento para 244x244 *pixels*, esses modelos demonstraram uma excelente capacidade de generalização. Embora o ajuste de contraste *CLAHE* não tenha tido um impacto significativo na eficácia dos modelos, sua aplicação juntamente com o redimensionamento para 244x244 *pixels* e a técnica *CutMix*, revelou um potencial competitivo para o modelo *EfficientNetV2-M*, demonstrando os melhores resultados no experimento final.

Os modelos *VGG16* e *VGG19* superaram redes mais modernas, como a *ResNet50*, possivelmente por sua arquitetura simples e profunda, que favorece a extração de padrões em imagens astronômicas, além do uso eficaz de pesos pré-treinados no *ImageNet*. Já as demais arquiteturas podem ter exigido mais ajustes ou sido mais sensíveis ao *overfitting* nesse contexto. Vale destacar, ainda, que aspectos como o consumo de memória e o tempo de inferência, que diferenciam significativamente modelos como *VGG* e *EfficientNetV2-M*, também podem ter contribuído nos resultados.

Durante o desenvolvimento deste estudo, dois desafios principais foram identificados: *overfitting* e custo computacional. O problema de *overfitting* foi observado especialmente no Experimento II, que utilizou classes desbalanceadas, e no desempenho do modelo *AlexNet*, que apresentou os piores resultados em todos os experimentos — possivelmente por não ter sido pré-treinado com o *ImageNet*, ao contrário das demais arquiteturas. Já o custo computacional

impôs limitações relevantes, especialmente em experimentos que exigiram redimensionamento das imagens, o que tornou necessário reduzir o número de classes de 10 para 5 para viabilizar os treinamentos. Apesar desses desafios, destaca-se como ponto positivo o bom desempenho dos modelos pré-treinados com imagens astronômicas coloridas, sugerindo que essas arquiteturas têm potencial para generalização em outras bases de dados com características semelhantes.

Este estudo abre diversas possibilidades de pesquisas futuras, por exemplo, a aplicação das técnicas de aumento de dados e balanceamento testadas nos experimentos aos demais modelos pré-treinados, permitindo uma avaliação mais ampla do impacto dessas abordagens em diferentes arquiteturas. Além disso, estratégias mais avançadas, como a classificação hierárquica, podem ser incorporadas ao *pipeline* de classificação. Outra abordagem relevante seria a implementação de métodos de *Ensemble Learning*, combinando múltiplos modelos para potencializar o desempenho da classificação e reduzir a variabilidade dos resultados.

Adicionalmente, a classificação de galáxias com CNNs baseada em seu espectro surge como uma ferramenta complementar essencial, possibilitando a extração de informações detalhadas, como composição química, idade estelar, taxa de formação estelar e a identificação de fenômenos ativos, incluindo buracos negros supermassivos e núcleos galácticos ativos (AGN).

Os resultados obtidos não apenas confirmam a robustez das CNNs na classificação de dados astronômicos, mas também ressaltam a importância da escolha criteriosa do modelo e das técnicas complementares para otimizar o processo de classificação. Assim, este trabalho contribui para o avanço da aplicação de *Deep Learning* na astronomia, proporcionando *insights* valiosos para futuras pesquisas e fortalecendo a base para novas abordagens na análise e categorização de objetos astronômicos.

## REFERÊNCIAS

- ALMEIDA, A.; ANDERSON, S. F.; FERNÁNDEZ, M. A.; BADENES, C. *et al.* **The Eighteenth Data Release of the Sloan Digital Sky Surveys: Targeting and First Spectra from SDSS-V.** 2023. Disponível em: <<https://skyserver.sdss.org/dr15/en/skyserver/paper/paper.aspx>>. Acesso em: 26 mar. 2024.
- CARDAMONE, C.; SCHAWINSKI, K.; SARZI, M.; BAMFORD, S. P.; BENNERT, N.; URRY, C. M.; LINTOTT, C.; KEEL, W. C.; PAREJKO, J.; NICHOL, R. C.; THOMAS, D.; ANDREESCU, D.; MURRAY, P.; RADDICK, M. J.; SLOSAR, A.; SZALAY, A.; VANDENBERG, J. Galaxy zoo green peas: discovery of a class of compact extremely star-forming galaxies\*. **Monthly Notices of the Royal Astronomical Society**, v. 399, n. 3, p. 1191–1205, 10 2009. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1111/j.1365-2966.2009.15383.x>>.
- DENG, L. The mnist database of handwritten digit images for machine learning research [best of the web]. **IEEE Signal Processing Magazine**, v. 29, n. 6, p. 141–142, 2012.
- DOON, R.; RAWAT, T. K.; GAUTAM, S. *et al.* Cifar-10 classification using deep convolutional neural network. In: **2018 IEEE Punecon**. [S.l.: s.n.], 2018. p. 1–5.
- EHT COLLABORATION *et al.* First m87 event horizon telescope results. i. the shadow of the supermassive black hole. **ApJL**, v. 875, p. 1, 2019. Disponível em: <<https://iopscience.iop.org/article/10.3847/2041-8213/ab0ec7>>.
- FILHO, K. S. O.; SARAIVA, M. F. **Astronomia e Astrofísica**. 4. ed. Porto Alegre, RS: Departamento de Astronomia - Instituto de Física, Universidade Federal do Rio Grande do Sul, 2014. Disponível em: <<http://astro.if.ufrgs.br/#gsc.tab=0>>.
- FILHO, P. C. F. da S. **Um estudo sobre a detecção de Exoplanetas com Inteligência Artificial**. 84 f. Monografia (Trabalho de Conclusão de Curso) — Universidade Federal do Ceará, Fortaleza, 2022.
- FILHO, P. C. F. da S. **Light Curve Imaging for Exoplanet detection with Deep Learning: A Conceptual Trial**. 126 f. Monografia (Dissertação de Mestrado) — Universidade Federal do Ceará, Fortaleza, 2024.
- GALAXY ZOO. **Galaxy Zoo: People's Science**. 2024. <<https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/>>. Acesso em: 20 janeiro 2025.
- GALAXY10 SDSS DATASET. **Galaxy10 SDSS Dataset**. 2024. <<https://astronn.readthedocs.io/en/latest/galaxy10sdss.html>>. Acesso em: 20 agosto 2024.
- GANAI, M.; HU, M.; MALIK, A.; TANVEER, M.; SUGANTHAN, P. Ensemble deep learning: A review. **Engineering Applications of Artificial Intelligence**, Elsevier BV, v. 115, p. 105151, out. 2022. ISSN 0952-1976. Disponível em: <<http://dx.doi.org/10.1016/j.engappai.2022.105151>>.
- GHARAT, S.; DANDAWATE, Y. Galaxy classification: a deep learning approach for classifying sloan digital sky survey images. **Monthly Notices of the Royal Astronomical Society**, Oxford University Press (OUP), v. 511, n. 4, p. 5120–5124, fev. 2022. ISSN 1365-2966. Disponível em: <<http://dx.doi.org/10.1093/mnras/stac457>>.

HASANIN, T.; KHOSHGOFTAAR, T. The effects of random undersampling with simulated class imbalance for big data. In: **2018 IEEE International Conference on Information Reuse and Integration (IRI)**. [S.l.: s.n.], 2018. p. 70–79.

HAYKIN, S. **Redes Neurais - 2ed.** Bookman, 2000. ISBN 9788573077186. Disponível em: <<https://books.google.com.br/books?id=IBp0X5qfyjUC>>.

HE, K.; ZHANG, X.; REN, S.; SUN, J. *et al.* Deep residual learning for image recognition. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [s.n.], 2016. p. 770–778. Disponível em: <<https://arxiv.org/abs/1512.03385>>.

HOSNA, A.; MERRY, E.; GYALMO, J.; ALOM, Z.; AUNG, Z.; AZIM, M. A. *et al.* Transfer learning: a friendly introduction. **Journal of Big Data**, Springer, v. 9, n. 1, p. 102, 2022.

HUANG, G.; LIU, Z.; MAATEN, L. van der; WEINBERGER, K. Q. *et al.* Densely connected convolutional networks. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [s.n.], 2017. p. 4700–4708. Disponível em: <<https://arxiv.org/abs/1608.06993>>.

JUNIOR, E. **Principais Métricas de Classificação de Modelos em Machine Learning**. 2023. Acesso em: 24 fev. 2025. Disponível em: <<https://medium.com/data-hackers/principais-modelos-em-machine-learning>>.

KEEL, W. C.; BENNERT, V. N.; PANCOAST, A.; HARRIS, C. E.; NIERENBERG, A.; CHOJNOWSKI, S. D.; MOISEEV, A. V.; OPARIN, D. V.; LINTOTT, C. J.; SCHAWINSKI, K.; MITCHELL, G.; CORNEN, C. Agn photoionization of gas in companion galaxies as a probe of agn radiation in time and direction. **Monthly Notices of the Royal Astronomical Society**, v. 483, n. 4, p. 4847–4865, 12 2018. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1093/mnras/sty3332>>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. *et al.* Imagenet classification with deep convolutional neural networks. In: **Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)**. [s.n.], 2012. p. 1097–1105. Disponível em: <<https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>>.

LECUN, Y.; Bengio, Y.; Hinton, G. *et al.* Deep learning. *Nature*, 2015.

LINTOTT, C. J.; SCHAWINSKI, K.; KEEL, W.; ARKEL, H. V.; BENNERT, N.; EDMONDSON, E.; THOMAS, D.; SMITH, D. J. B.; HERBERT, P. D.; JARVIS, M. J.; VIRANI, S.; ANDREESCU, D.; BAMFORD, S. P.; LAND, K.; MURRAY, P.; NICHOL, R. C.; RADDICK, M. J.; SLOSAR, A.; SZALAY, A.; VANDENBERG, J. Galaxy zoo: ‘hanny’s voorwerp’, a quasar light echo?\*. **Monthly Notices of the Royal Astronomical Society**, v. 399, n. 1, p. 129–140, 10 2009. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1111/j.1365-2966.2009.15299.x>>.

LINTOTT, C. J.; SCHAWINSKI, K.; SLOSAR, A.; LAND, K.; BAMFORD, S.; THOMAS, D.; RADDICK, M. J.; NICHOL, R. C.; SZALAY, A.; ANDREESCU, D.; MURRAY, P.; VANDENBERG, J. *et al.* Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. **Monthly Notices of the Royal Astronomical Society**, Oxford University Press (OUP), v. 389, n. 3, p. 1179–1189, set. 2008. ISSN 1365-2966. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2966.2008.13689.x>>.

LU, Y.; SHEN, M.; WANG, H.; WANG, X.; RECHEM, C. van; FU, T.; WEI, W. **Machine Learning for Synthetic Data Generation: A Review**. 2024. Disponível em: <<https://arxiv.org/abs/2302.04062>>.

MASTERS, K. L. Twelve years of galaxy zoo. **Proceedings of the International Astronomical Union**, Cambridge University Press (CUP), v. 14, n. S353, p. 205–212, jun. 2019. ISSN 1743-9221. Disponível em: <<http://dx.doi.org/10.1017/S1743921319008615>>.

NASA. **Edwin Hubble**. 2023. Disponível em: <<https://science.nasa.gov/people/edwin-hubble/>>.

NASA. **Galaxy Basics**. 2024. Disponível em: <<https://science.nasa.gov/universe/galaxies/>>.

NASA. **Galaxy Types**. 2024. Disponível em: <<https://science.nasa.gov/universe/galaxies/types/>>.

SANTOS, E. L. G. dos. **Aplicação de Deep Learning para classificação morfológica de Galáxias**. 84 f.80. Monografia (Trabalho de Conclusão de Curso) — Universidade Federal de Ouro Preto, Minas Gerais, 2024.

SCHAWINSKI, K.; URRY, C. M.; SIMMONS, B. D.; FORTSON, L.; KAVIRAJ, S.; KEEL, W. C.; LINTOTT, C. J.; MASTERS, K. L.; NICHOL, R. C.; SARZI, M.; SKIBBA, R.; TREISTER, E.; WILLETT, K. W.; WONG, O. I.; YI, S. K. The green valley is a red herring: Galaxy zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies. **Monthly Notices of the Royal Astronomical Society**, Oxford University Press (OUP), v. 440, n. 1, p. 889–907, mar. 2014. ISSN 0035-8711. Disponível em: <<http://dx.doi.org/10.1093/mnras/stu327>>.

SDSS. **SDSS Instruments**. 2022. Disponível em: <<https://www.sdss.org/instruments/>>.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: **International Conference on Learning Representations (ICLR)**. [s.n.], 2015. Disponível em: <<https://arxiv.org/abs/1409.1556>>.

SKIBBA, R. A.; BAMFORD, S. P.; NICHOL, R. C.; LINTOTT, C. J.; ANDREESCU, D.; EDMONDSON, E. M.; MURRAY, P.; RADDICK, M. J.; SCHAWINSKI, K.; SLOSAR, A.; SZALAY, A. S.; THOMAS, D.; VANDENBERG, J. Galaxy zoo: disentangling the environmental dependence of morphology and colour\*. **Monthly Notices of the Royal Astronomical Society**, v. 399, n. 2, p. 966–982, 10 2009. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1111/j.1365-2966.2009.15334.x>>.

SKIBBA, R. A.; MASTERS, K. L.; NICHOL, R. C.; ZEHAVID, I.; HOYLE, B.; EDMONDSON, E. M.; BAMFORD, S. P.; CARDAMONE, C. N.; KEEL, W. C.; LINTOTT, C.; SCHAWINSKI, K. Galaxy zoo: the environmental dependence of bars and bulges in disc galaxies. **Monthly Notices of the Royal Astronomical Society**, v. 423, n. 2, p. 1485–1502, 06 2012. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1111/j.1365-2966.2012.20972.x>>.

SLOAN, A. P. **Sloan Digital Sky Survey**. 2024. Disponível em: <<https://sloan.org/programs/research/sloan-digital-sky-survey#about>>.

SMETHURST, R. J.; LINTOTT, C. J.; SIMMONS, B. D.; SCHAWINSKI, K.; MARSHALL, P. J.; BAMFORD, S.; FORTSON, L.; KAVIRAJ, S.; MASTERS, K. L.; MELVIN, T.; NICHOL, R. C.; SKIBBA, R. A.; WILLETT, K. W. Galaxy zoo: evidence for

diverse star formation histories through the green valley. **Monthly Notices of the Royal Astronomical Society**, v. 450, n. 1, p. 435–453, 04 2015. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1093/mnras/stv161>>.

TAN, M.; LE, Q. V. Efficientnetv2: Smaller models and faster training. In: **Proceedings of the 38th International Conference on Machine Learning (ICML)**. [s.n.], 2021. Disponível em: <<https://arxiv.org/abs/2104.00298>>.

VARGAS, A. C. G.; Paes, A.; Vasconcelos, C. N. *et al.* Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. XXXV Conference on Graphics, Patterns and Images, 2016.

WILLETT, K. W.; LINTOTT, C. J.; BAMFORD, S. P.; MASTERS, K. L.; SIMMONS, B. D.; CASTEELS, K. R. V.; EDMONDSON, E. M.; FORTSON, L. F.; KAVIRAJ, S.; KEEL, W. C.; MELVIN, T.; NICHOL, R. C.; RADDICK, M. J.; SCHAWINSKI, K.; SIMPSON, R. J.; SKIBBA, R. A.; SMITH, A. M.; THOMAS, D. *et al.* Galaxy zoo 2: detailed morphological classifications for 304122 galaxies from the sloan digital sky survey. **Monthly Notices of the Royal Astronomical Society**, v. 435, n. 4, p. 2835–2860, 09 2013. ISSN 0035-8711. Disponível em: <<https://doi.org/10.1093/mnras/stt1458>>.

YUN, S.; HAN, D.; OH, S. J.; CHUN, S.; CHOE, J.; YOO, Y. **CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features**. 2019. Disponível em: <<https://arxiv.org/abs/1905.04899>>.

ZHANG, H.; CISSE, M.; DAUPHIN, Y. N.; LOPEZ-PAZ, D. **mixup: Beyond Empirical Risk Minimization**. 2018. Disponível em: <<https://arxiv.org/abs/1710.09412>>.

ZOOUNIVERSE. **Zooniverse Projects**. 2025. <<https://www.zooniverse.org/projects>>. Acesso em: 9 de fevereiro de 2025.