



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS TERESINA CENTRAL
CURSO ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

MATHEUS LEVI DA SILVA BARBOSA

DESENVOLVIMENTO DE TÉCNICAS UTILIZANDO CNN PARA DETECÇÃO DE
CÂNCER DE PELE

TERESINA

2025

MATHEUS LEVI DA SILVA BARBOSA

DESENVOLVIMENTO DE TÉCNICAS UTILIZANDO CNN PARA DETECÇÃO DE
CÂNCER DE PELE

Trabalho de Conclusão de Curso (monografia)
apresentado como exigência parcial para obtenção do
diploma do Curso de Análise e Desenvolvimento de
Sistemas do Instituto Federal de Educação, Ciência e
Tecnologia do Piauí, Campus Teresina Central.

Orientador: Prof. Dr. Otílio Paulo da Silva Neto

TERESINA

2025

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Barbosa, Matheus Levi da Silva

B238d Desenvolvimento de técnicas utilizando CNN para detecção de câncer de pele / Matheus Levi da Silva Barbosa. - 2025.

72 f.: il. color.

Trabalho de Conclusão de Curso (Tecnologia em Análise e Desenvolvimento de Sistemas) - Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central, 2025.

Orientador : Prof Dr. Otilio Paulo da Silva Neto.

1. Câncer. 2. Diagnóstico por imagem. 3. Aprendizado profundo . 4. Redes neurais convolucionais. I.Título.

CDD - 004.21

Elaborado por Maria Rosismar Farias CRB 3/631

MATHEUS LEVI DA SILVA BARBOSA

DESENVOLVIMENTO DE TÉCNICAS UTILIZANDO CNN PARA DETECÇÃO DE
CÂNCER DE PELE

Trabalho de Conclusão de Curso (monografia)
apresentado como exigência parcial para obtenção do
diploma do Curso de Análise e Desenvolvimento de
Sistemas do Instituto Federal de Educação, Ciência e
Tecnologia do Piauí, Campus Teresina Central.

Aprovada em: 25/07/2025.

BANCA EXAMINADORA

Prof. Dr. Otílio Paulo da Silva Neto (Orientador)
Instituto Federal do Piauí (IFPI)

Prof. Me. Rogério da Silva Batista
Instituto Federal do Piauí (IFPI)

Prof. Dr. Francisco Santos Moura
Instituto Federal do Piauí (IFPI)

Aos pais, amigos e a todos que me apoiaram.

AGRADECIMENTOS

Agradeço profundamente a todos os professores do IFPI – Campus Teresina Central, do curso de Análise e Desenvolvimento de Sistemas, que foram essenciais para meu crescimento acadêmico e pessoal. Suas orientações, paciência e dedicação me guiaram durante toda essa jornada.

Agradeço também aos meus colegas, que me acompanharam durante o curso. Juntos, enfrentamos desafios e celebramos conquistas, e sua amizade e colaboração foram fundamentais para essa trajetória.

À minha namorada Eduarda, agradeço pelo apoio incondicional durante esses meses e pela revisão cuidadosa deste trabalho. Sua presença e incentivo foram fundamentais para que eu superasse os desafios.

Agradeço à minha família por parte de mãe, que sempre esteve ao meu lado, oferecendo apoio incondicional e acreditando em mim em cada etapa desta caminhada especialmente à minha avó Lenita.

Também agradeço às minhas três cachorras, Luna, Sofia e Juju, cuja alegria e carinho tornaram os dias mais leves e tranquilos.

E, por fim, mas não menos importante, agradeço a Deus, que me deu força e sabedoria para superar os desafios e chegar até aqui.

When you arise in the morning, think of what a precious privilege it is to be alive - to breathe, to think, to enjoy, to love.

Marcus Aurelius

RESUMO

O câncer é uma das principais causas de morbidade e mortalidade global, apresentando um desafio constante para a medicina e a ciência. Diante da crescente incidência de diversos tipos de câncer, aprimorar ferramentas de diagnóstico é crucial para melhorar as taxas de sobrevivência e a eficácia dos tratamentos. Este estudo propõe uma abordagem, utilizando técnicas avançadas de processamento de imagem e aprendizado profundo para segmentação e classificação de diferentes tipos de câncer. A metodologia foca na aplicação de Convolutional Neural Networks (CNNs), reconhecidas por sua alta performance em segmentação e classificação de imagens médicas. A detecção precoce e precisa é essencial para o sucesso do tratamento oncológico, mas os métodos tradicionais, baseados em análise manual, enfrentam problemas de subjetividade e variabilidade entre os observadores. Para superar essas limitações, o estudo foi estruturado em quatro etapas principais. Inicialmente, realizou-se a análise e preparação do conjunto de dados, envolvendo a exploração das características das imagens e a aplicação de técnicas para equilibrar a distribuição das classes. Na sequência, desenvolveram-se e ajustaram-se modelos computacionais para a segmentação e classificação das lesões, com foco na melhoria da acurácia diagnóstica. Posteriormente, implementou-se a infraestrutura necessária para a integração dos modelos em um sistema de suporte acessível e seguro. Por fim, desenvolveu-se uma interface amigável que permite a captura de imagens e a interação com o sistema, otimizando a experiência do usuário. Os resultados esperados incluem a criação de uma ferramenta eficaz que possa ser integrada aos protocolos clínicos, melhorando a precisão e a velocidade do diagnóstico de câncer. Com a adoção dessas tecnologias, espera-se não apenas aprimorar a eficiência diagnóstica, mas também influenciar positivamente o tratamento e o prognóstico dos pacientes. Assim, o estudo contribui significativamente para o avanço das práticas de diagnóstico por imagem no campo oncológico, explorando novas fronteiras tecnológicas para enfrentar um dos maiores desafios da saúde pública.

Palavras-chave: câncer; diagnóstico por imagem; aprendizado profundo; redes neurais convolucionais; segmentação de imagens.

ABSTRACT

Cancer is one of the leading causes of morbidity and mortality worldwide, posing an ongoing challenge to medicine and science. Given the increasing incidence of various types of cancer, enhancing diagnostic tools is crucial for improving survival rates and treatment effectiveness. This study proposes an approach that leverages advanced image processing techniques and deep learning for the segmentation and classification of different types of cancer. The methodology focuses on the application of Convolutional Neural Networks (CNNs), which are well known for their high performance in medical image segmentation and classification. Early and accurate detection is essential for the success of cancer treatment, yet traditional methods based on manual analysis face issues of subjectivity and inter-observer variability. To overcome these limitations, the study was structured into four main stages. Initially, the dataset was analyzed and prepared, involving the exploration of image characteristics and the application of techniques to balance class distribution. Next, computational models for lesion segmentation and classification were developed and fine-tuned, with an emphasis on improving diagnostic accuracy. Subsequently, the necessary infrastructure was implemented to integrate the models into an accessible and secure support system. Finally, a user-friendly interface was developed to allow image capture and interaction with the system, enhancing the user experience. The expected outcomes include the creation of an effective tool that can be integrated into clinical protocols, improving both the accuracy and speed of cancer diagnosis. By adopting these technologies, it is hoped not only to enhance diagnostic efficiency but also to positively influence patient treatment and prognosis. Thus, this study contributes significantly to the advancement of imaging-based diagnostic practices in oncology, exploring new technological frontiers to address one of the greatest challenges in public health.

Keywords: cancer; medical imaging; deep learning; convolutional neural networks; image segmentation.

LISTA DE ILUSTRAÇÕES

Figura 1 - Distribuição global do câncer	20
Gráfico 1 - Incidência de Câncer de Pele (INCA, s.d.)	21
Gráfico 2 - Mortalidade por Câncer de Pele (INCA, s.d.)	22
Figura 2 - Exemplos de Imagens por Classe de Câncer de Pele	23
Figura 3 - Ilustração da Arquitetura ResNet	26
Figura 4 - Arquitetura de uma Rede Densa	27
Figura 5 - Arquitetura VGG-16	27
Figura 6 - Ilustração do Escalonamento Composto da EfficientNet	28
Figura 7 - Arquitetura do Vision Transformer (ViT)	29
Figura 8 - Fluxograma do pipeline de pré-processamento e criação dos datasets	38
Figura 9 - Demonstração visual do pipeline de pré-processamento	40
Figura 10 - Fluxograma do Dull Razor	51
Figura 11 - Tela de Autenticação do Protótipo	60
Figura 12 - Dashboard Principal com Resumo Estatístico	61
Figura 13 - Tela de Submissão de Novo Diagnóstico	62
Figura 14 - Tela de Histórico e Status dos Diagnósticos	63
Figura 15 - Apresentação do Resultado Final com Segmentação da IA	64
Figura 16 - Painel de Estatísticas Gerais	65
Figura 17 - Tela de Configurações do Aplicativo	66

LISTA DE TABELAS

Tabela 1 - Detalhamento da contagem de amostras por classe antes e após o balanceamento	36
Tabela 2 - Resumo da distribuição dos datasets (Original vs. Balanceado)	37
Tabela 3 - Primeira parte da Comparação de Desempenho dos Modelos após Balanceamento por Subamostragem	52
Tabela 4 - Segunda parte da Comparação de Desempenho dos Modelos após Balanceamento por Subamostragem	55
Tabela 5 - Métricas finais de desempenho do modelo DenseNet201	68

LISTA DE ABREVIATURAS E SIGLAS

CNN: Convolutional Neural Network (Rede Neural Convolucional)

OMS: Organização Mundial da Saúde

INCA: Instituto Nacional de Câncer

SBD: Sociedade Brasileira de Dermatologia

UV: Raios Ultravioleta

SIM: Sistema de Informações sobre Mortalidade

Batch Size: Número de amostras processadas antes da atualização dos pesos do modelo.

Learning Rate: Taxa de aprendizado; controla a magnitude das atualizações dos pesos durante o treinamento.

Dull Razor: Técnica de pré-processamento de imagem.

NV - Nevo (ou Nevus)

BCC - Carcinoma Basocelular

AK - Queratoses Actínicas

BKL - Lesões Benignas de Pele (ou Neoplasias Benignas)

DF - Dermatofibroma

VASC - Lesões Vasculares

SCC - Carcinoma Espinocelular

Dull Razor: Técnica de pré-processamento de imagem para remoção de pelos em imagens dermatoscópicas.

ViT: Vision Transformer.

RAdam: Rectified Adam (Otimizador).

AMP: Automatic Mixed Precision (Treinamento com Precisão Mista Automática).

LISTA DE SÍMBOLOS

- + soma de dois ou mais números.
- diferença entre dois números.

SUMÁRIO

1	INTRODUÇÃO.....	16
1.1	JUSTIFICATIVA.....	16
1.2	OBJETIVO GERAL.....	17
1.3	OBJETIVOS ESPECÍFICOS.....	17
2.	FUNDAMENTAÇÃO TEÓRICA.....	18
2.1	CÂNCER DE PELE.....	18
2.1.1	Tipos de Câncer de Pele.....	18
2.1.2	Fatores de Risco.....	19
2.2	PREVALÊNCIA GLOBAL.....	19
2.3	CÂNCER DE PELE NO BRASIL.....	20
2.4	DESAFIOS NO DIAGNÓSTICO.....	22
2.5	CLASSIFICAÇÃO DE LESÕES CUTÂNEAS.....	22
2.5.1	Lesões Benignas.....	23
2.5.2	Lesões Malignas.....	24
2.6	REDES NEURAIS CONVOLUCIONAIS (CNN), VISION TRANSFORMERS (ViT) E TÉCNICAS DE PROCESSAMENTO DE IMAGENS.....	24
2.6.1	Conceitos Fundamentais das CNNs.....	25
2.6.2	Modelos de Redes Neurais Convolucionais.....	26
2.6.4	Vision Transformer (ViT).....	29
2.6.5	Técnicas de Pré-processamento de Imagens.....	30
2.6.5	Augmentação de Dados.....	31
2.6.6	Otimizadores e Estratégias Avançadas de Treinamento.....	31
2.6.6.1	Fundamentos dos Otimizadores.....	32
2.6.6.2	Técnicas Complementares de Otimização e Treinamento.....	32
3	METODOLOGIA.....	33
3.1	COLETA E PREPARAÇÃO DE DADOS.....	33
3.1.1	Divisão das Imagens: Lesões Benignas e Malignas.....	34
3.2	CRIAÇÃO DOS DATASETS.....	35
3.3	BALANCEAMENTO DO DATASET.....	36
3.4	METODOLOGIA DETALHADA DE PRÉ-PROCESSAMENTO E ESTRUTURAÇÃO DE DADOS.....	37
3.3.1	Normalização Geométrica e Foco na Região de Interesse (ROI).....	38
3.3.2	Remoção de Artefatos com o Algoritmo DullRazor (Abordagem Conservadora).....	39
3.3.3	Estruturação e Armazenamento Final dos Dados.....	40
3.4	ARQUITETURAS UTILIZADAS.....	41
3.4.1	DenseNet121, DenseNet169 e DenseNet201.....	41
3.4.2	ResNet50v2, ResNet101v2.....	42
3.4.3	VGG-16.....	42
3.4.4	Vision Transformer (ViT).....	43

3.4.5	EfficientNet-B7	43
3.5	APLICAÇÃO DO TRANSFER LEARNING	44
3.6	AJUSTE E COMPARAÇÃO DAS ARQUITETURAS	44
3.7	TREINAMENTO DO MODELO	45
3.8	OTIMIZAÇÃO COM RADAM (RECTIFIED ADAM)	46
3.9	AVALIAÇÃO DO MODELO	46
3.9.1	Outras Métricas Usadas	47
3.10	ARQUITETURA DA APLICAÇÃO	48
3.12.1	Processamento de Imagens	49
3.12.2	Segmentação Automática	49
3.12.3	Classificação Multi-Modelo	49
4	RESULTADOS	50
4.1	APLICAÇÃO DO FILTRO DULL RAZOR	50
4.2	ANÁLISE E CONTEXTUALIZAÇÃO DOS RESULTADOS PARA O BALANCEAMENTO DE DADOS	52
4.3	ANÁLISE DO DATASET DULLRAZOR	54
4.4	ARQUITETURA DO SISTEMA	55
4.4.1	Visão Geral da Arquitetura	55
4.4.2	Fluxo de Dados	56
4.4.3	Integração entre Componentes	56
4.5	TECNOLOGIAS UTILIZADAS	56
4.5.1	Critérios de Seleção	56
4.5.2	Stack Tecnológico do Backend	56
4.5.3	Stack Tecnológico do Módulo de IA	57
4.5.4	Justificativa das Escolhas Tecnológicas	57
4.6	DESENVOLVIMENTO DA APLICAÇÃO MÓVEL	58
4.6.1	Interface de Acesso e Navegação Principal	59
4.6.2	Implementação do Fluxo de Análise Diagnóstica	60
4.6.3	Visualização dos Resultados da Inteligência Artificial	62
4.6.4	Funcionalidades de Suporte e Análise de Dados	63
4.6.5	Considerações Éticas e Regulatórias	65
5	CONCLUSÃO	66
6	TRABALHOS FUTUROS	67
	REFERÊNCIAS	69

1 INTRODUÇÃO

A incidência global de casos de câncer tem aumentado significativamente, tornando-se uma das principais preocupações de saúde pública. O diagnóstico preciso e rápido desempenha um papel crucial na eficácia do tratamento e na melhoria das taxas de sobrevivência. Nesse contexto, a proposta deste projeto encontra sua justificativa na necessidade premente de aprimorar as ferramentas de diagnóstico, especialmente aquelas baseadas em imagens médicas.

A abordagem convencional para a detecção de câncer muitas vezes depende da análise manual por profissionais de saúde, o que pode ser moroso e suscetível a variações interobservadores. A utilização de técnicas avançadas de processamento de imagem e aprendizado profundo, como as Convolutional Neural Networks (CNNs), proporciona uma oportunidade promissora para automatizar e aprimorar esse processo.

A escolha das arquiteturas específicas, justifica-se pela sua eficácia comprovada em tarefas de segmentação e classificação, garantindo uma análise mais precisa das imagens médicas. A incorporação de técnicas de aumento de dados e processamento de imagem visa superar desafios comuns, como a heterogeneidade nas imagens, contribuindo para um modelo mais robusto e generalizado.

Ao finalizar este projeto, esperamos não apenas contribuir para o avanço nas técnicas de diagnóstico por imagem, mas também fornecer uma ferramenta eficaz e eficiente que possa ser integrada nos processos clínicos, auxiliando os profissionais de saúde na identificação precoce e na classificação precisa de diferentes tipos de câncer.

1.1 JUSTIFICATIVA

O câncer de pele é uma das neoplasias mais comuns no mundo, com incidência crescente, especialmente em regiões de alta exposição solar, como o Brasil. Entre os tipos de câncer de pele, o melanoma destaca-se por sua agressividade e potencial de metástase, sendo responsável por grande parte das mortes relacionadas a esse tipo de neoplasia. A detecção precoce do melanoma é crucial, pois aumenta consideravelmente as chances de tratamento bem-sucedido e sobrevivência dos pacientes.

Contudo, os métodos convencionais de diagnóstico, como a avaliação visual, dependem da experiência do médico, o que pode levar a diagnósticos tardios ou imprecisos, especialmente em áreas com menor acesso a especialistas. Nesse contexto, o uso de tecnologias avançadas, como aprendizado de máquina e processamento de imagens, surge

como uma solução promissora para melhorar a precisão e a agilidade no diagnóstico de lesões cutâneas.

O presente trabalho justifica-se pela necessidade de explorar e aplicar essas novas tecnologias no auxílio à detecção precoce de lesões malignas, com ênfase no melanoma. Ao utilizar redes neurais convolucionais e técnicas de processamento de imagens, espera-se contribuir para o desenvolvimento de ferramentas diagnósticas acessíveis e precisas, que possam ser utilizadas em larga escala no suporte ao diagnóstico médico. A relevância desse estudo está em sua potencial contribuição para a redução da mortalidade associada ao câncer de pele, oferecendo uma abordagem inovadora e eficiente para o diagnóstico precoce.

1.2 OBJETIVO GERAL

Este projeto tem como objetivo principal desenvolver um sistema inteligente capaz de realizar a segmentação e classificação de diferentes tipos de câncer, distinguindo entre lesões benignas e malignas em imagens médicas. Para isso, será adotada uma abordagem inovadora baseada em redes neurais profundas, incluindo arquiteturas modernas como Convolutional Neural Networks (CNNs) e Vision Transformers (ViT). O sistema utilizará técnicas avançadas de pré-processamento e aprimoramento de imagens, como o filtro Dull Razor, além de estratégias de filtragem e balanceamento de dados para aumentar a qualidade e a representatividade do conjunto de dados. O modelo resultante será incorporado a uma infraestrutura computacional completa, composta por back-end e front-end, voltada para aplicação prática no apoio ao diagnóstico oncológico, promovendo maior precisão, agilidade e acessibilidade na detecção e classificação de lesões.

1.3 OBJETIVOS ESPECÍFICOS

1. Aplicar técnicas de pré-processamento de imagens, incluindo o uso do filtro Dull Razor, para melhorar a qualidade das imagens utilizadas na detecção de lesões dermatológicas.
2. Implementar modelos de redes neurais convolucionais, como DenseNet121, DenseNet169, VGG16, ResNet50 e EfficientNetB7, além de um modelo baseado em Vision Transformer (ViT), para a classificação de lesões de pele.
3. Comparar o desempenho dos modelos de aprendizado de máquina com base em métricas como acurácia, sensibilidade e especificidade.
4. Reduzir e refinar o conjunto de dados, aplicando técnicas de filtragem para melhorar a representatividade e balanceamento das classes.

5. Desenvolver uma infraestrutura computacional composta por back-end e front-end para viabilizar a aplicação prática do sistema. O back-end será responsável pelo processamento das imagens, execução dos modelos e gerenciamento das requisições, enquanto o front-end fornecerá uma interface gráfica intuitiva para que profissionais da saúde possam carregar imagens, visualizar resultados e acompanhar análises de forma eficiente e acessível.

2. FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica deste trabalho fornece embasamento científico para a pesquisa, abordando o câncer de pele, suas características, prevalência global e no Brasil, além de discutir o uso de tecnologias avançadas, como aprendizado de máquina e processamento de imagens, para o diagnóstico precoce de lesões cutâneas.

2.1 CÂNCER DE PELE

O câncer de pele é a neoplasia mais comum no mundo, dividido em dois tipos principais: melanoma e não melanoma. O melanoma é o mais letal, enquanto o não melanoma, embora mais frequente, é geralmente menos perigoso. A detecção precoce do melanoma é essencial, uma vez que o diagnóstico em estágios iniciais aumenta consideravelmente as taxas de sobrevivência dos pacientes (OMS, 2020).

2.1.1 Tipos de Câncer de Pele

O câncer de pele é uma condição complexa que se divide em diferentes tipos, sendo os principais o melanoma e os cânceres de pele não melanoma.

Melanoma: Este tipo de câncer surge nos melanócitos, que são as células responsáveis pela produção de melanina, o pigmento que dá cor à pele. O melanoma é extremamente agressivo e, se não identificado e tratado precocemente, pode causar metástases rapidamente, espalhando-se para outras partes do corpo (INCA, 2022). A detecção precoce é crucial, pois o prognóstico do paciente está diretamente relacionado à fase em que a doença é diagnosticada.

Câncer de pele não melanoma: Este grupo inclui os carcinomas basocelular e espinocelular. Embora esses tipos sejam geralmente menos letais em comparação com o melanoma, eles podem causar deformidades severas e danos aos tecidos adjacentes (SBD, 2021). O carcinoma basocelular é o mais comum, frequentemente associado a uma exposição solar prolongada, enquanto o carcinoma espinocelular pode surgir em áreas da pele que foram danificadas pelo sol. Ambos requerem tratamento para evitar complicações significativas.

2.1.2 Fatores de Risco

Os fatores de risco para o câncer de pele são variados e podem aumentar significativamente a probabilidade de desenvolvimento da doença. Um dos mais relevantes é a exposição aos raios ultravioleta (UV), que é amplamente reconhecida como o principal fator de risco. A exposição prolongada ao sol sem a devida proteção pode causar danos no DNA das células da pele, elevando a chance de mutações que podem resultar em câncer (INCA, 2022).

Além da exposição solar, o histórico familiar também desempenha um papel crucial. Indivíduos com antecedentes familiares de melanoma possuem uma propensão maior a desenvolver a doença, uma vez que mutações hereditárias podem aumentar a vulnerabilidade ao câncer de pele (OMS, 2020). Este fator sublinha a importância de monitoramento e rastreamento mais rigorosos em famílias com histórico de câncer.

As características fenotípicas, como ter pele clara, olhos claros e cabelos ruivos ou loiros, também tornam as pessoas mais suscetíveis ao câncer de pele. A falta de pigmentação reduz a proteção natural contra os danos causados pela radiação UV, tornando esses indivíduos mais vulneráveis a lesões cutâneas e ao desenvolvimento de tumores (SBD, 2021).

Por fim, o sistema imunológico comprometido é um fator de risco relevante. Pacientes que têm o sistema imunológico debilitado, como aqueles submetidos a transplantes ou em tratamento imunossupressor, enfrentam um risco aumentado de câncer de pele. A diminuição da capacidade do organismo de combater células anormais eleva a probabilidade de formação de tumores cutâneos (INCA, 2022).

Esses fatores ressaltam a importância de medidas de prevenção e monitoramento, especialmente para grupos considerados de risco. A conscientização sobre os perigos da exposição solar, a realização de exames regulares e a adoção de práticas de proteção são essenciais para a detecção precoce e a redução da incidência do câncer de pele.

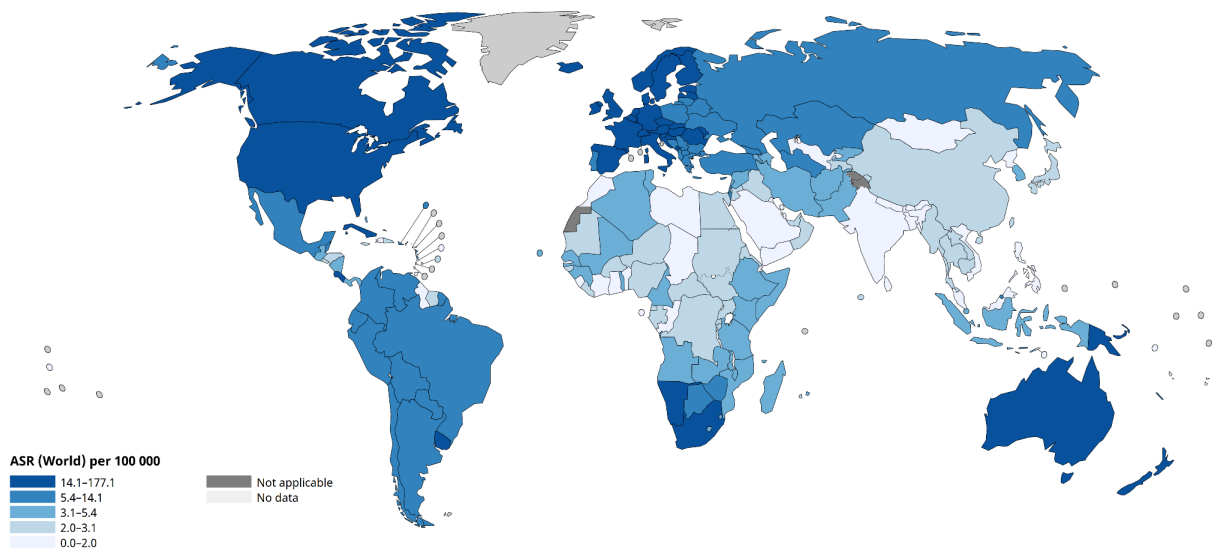
2.2 PREVALÊNCIA GLOBAL

De acordo com a Organização Mundial da Saúde (OMS), mais de 3 milhões de novos casos de câncer de pele não melanoma e aproximadamente 132.000 casos de melanoma são diagnosticados anualmente no mundo. O melanoma, embora menos frequente, é responsável pela maioria das mortes por câncer de pele, com cerca de 60.000 óbitos anuais (OMS, 2020). As maiores taxas de incidência são observadas em países com intensa exposição solar, como Austrália, Estados Unidos e Nova Zelândia.

O câncer de pele não melanoma, apesar de menos letal, é uma preocupação significativa, pois pode causar graves danos aos tecidos e complicações estéticas (SBD, 2021).

Figura 1: Distribuição global do câncer

Age-Standardized Rate (World) per 100 000, Incidence, Both sexes, in 2022
Melanoma of skin + Non-melanoma skin cancer

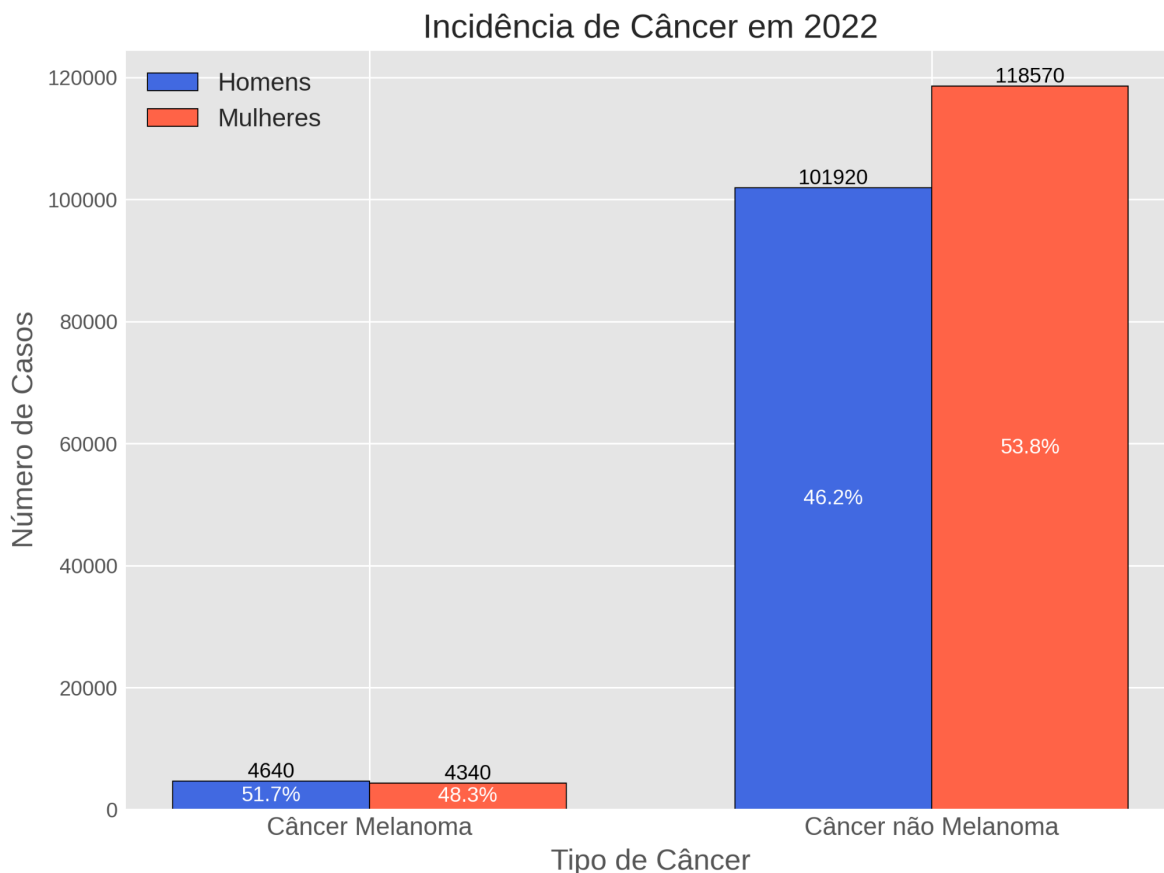


Fonte: Global Cancer Observatory (2022). Disponível em: <https://gco.iarc.who.int/today>. Acesso em: 24 set. 2024.

2.3 CÂNCER DE PELE NO BRASIL

No Brasil, o câncer de pele é o mais comum entre a população. O Instituto Nacional de Câncer (INCA) estimou, em 2022, a ocorrência de 220.030 novos casos de câncer de pele não melanoma, sendo 101.920 em homens e 118.570 em mulheres. Para o melanoma, a estimativa foi de 8.980 novos casos, com 4.640 em homens e 4.340 em mulheres (INCA, 2022).

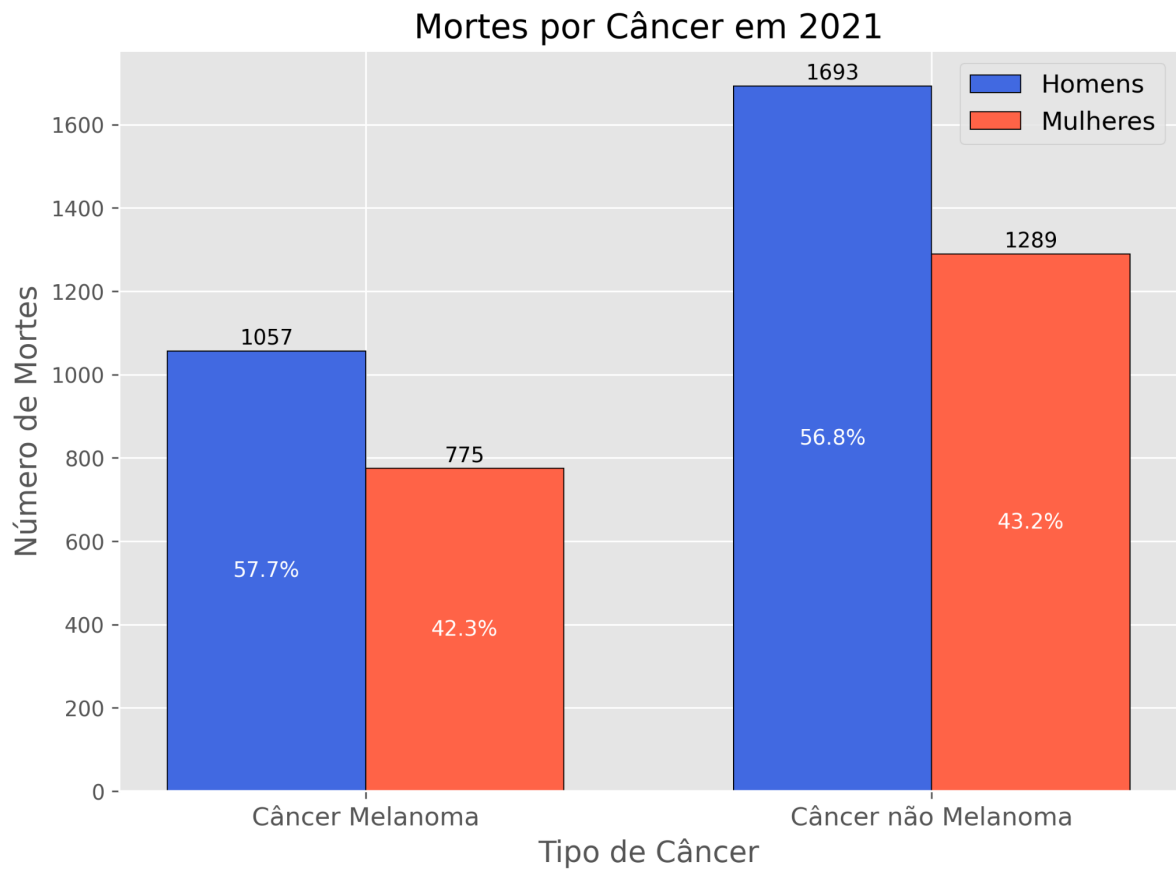
Gráfico 1: Incidência de Câncer de Pele (INCA, s.d.)



Fonte: Instituto Nacional de Câncer (INCA). Para informações sobre câncer melanoma e não melanoma, consulte o site do Instituto Nacional de Câncer: *Câncer de Pele Melanoma* e *Câncer de Pele Não Melanoma*.

Em termos de mortalidade, o Sistema de Informações sobre Mortalidade (SIM) registrou, em 2021, um total de 1.832 óbitos causados por melanoma, sendo 1.057 em homens e 775 em mulheres. Além disso, foram contabilizadas 2.982 mortes por câncer de pele não melanoma, com 1.693 casos em homens e 1.289 em mulheres. Esses dados evidenciam a significativa incidência dessas neoplasias e destacam a relevância das campanhas de prevenção, conscientização e da detecção precoce para reduzir a mortalidade associada. A adoção de medidas educativas e o acesso facilitado a exames dermatológicos são fundamentais para o diagnóstico em estágios iniciais, o que pode aumentar as chances de sucesso no tratamento e a sobrevivência dos pacientes (SIM, 2021).

Gráfico 2: Gráfico de Mortalidade por Câncer de Pele (INCA, s.d.)



Fonte: Instituto Nacional de Câncer (INCA). Para mais detalhes sobre os tipos de câncer de pele e suas estatísticas, veja Câncer de Pele Melanoma e Câncer de Pele Não Melanoma no site do INCA.

2.4 DESAFIOS NO DIAGNÓSTICO

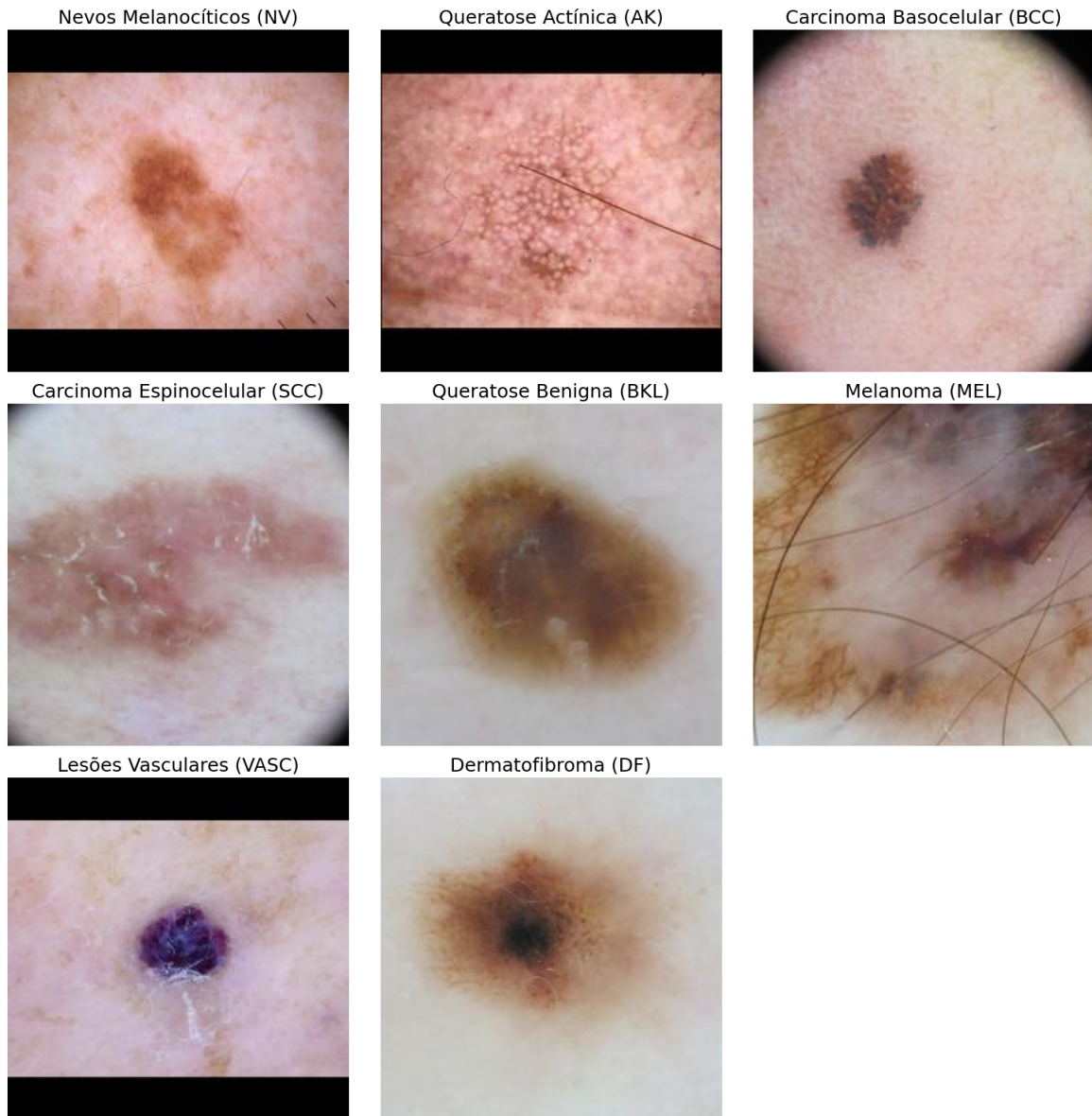
O diagnóstico precoce do câncer de pele, especialmente do melanoma, é crucial para o sucesso do tratamento. Métodos tradicionais de diagnóstico incluem a avaliação clínica visual e a biópsia, que, embora eficientes, podem apresentar limitações devido à subjetividade e à dependência da experiência do profissional (SBD, 2021).

Novas tecnologias, como a dermatoscopia digital e a microscopia confocal de reflectância, têm contribuído significativamente para o aumento da precisão diagnóstica. No entanto, o custo elevado e a dificuldade de acesso a essas tecnologias ainda são barreiras importantes em diversos países (OMS, 2020).

2.5 CLASSIFICAÇÃO DE LESÕES CUTÂNEAS

As lesões cutâneas podem ser divididas em dois grandes grupos: benignas e malignas, cada uma com suas características e riscos associados.

Figura 2: Exemplos de Imagens por Classe de Câncer de Pele.



Fonte: Dados do ISIC 2019, coletados para análise no estudo de diagnóstico por imagem.

2.5.1 Lesões Benignas

1. **Nevos Melanocíticos (*Melanocytic Nevi*):** Manchas formadas por aglomerados de melanócitos. Embora benignas, devem ser monitoradas em função de alterações de cor, tamanho ou forma, que podem indicar malignidade (INCA, 2022).
2. **Queratose Benigna:**
 - a. *Lentigo Solar*: Manchas escuras associadas à exposição solar cumulativa.
 - b. *Queratose Seborreica*: Crescimentos ásperos, relacionados ao envelhecimento da pele.

- c. *Queratose tipo Líquen Plano*: Lesões planas e escamosas, benignas, mas passíveis de monitoramento (SBD, 2021).
- 3. **Dermatofibroma**: Nódulos cutâneos benignos formados em resposta a pequenas lesões (INCA, 2022).
- 4. **Lesões Vasculares**: Incluem diversas condições relacionadas a vasos sanguíneos, como angiomas, geralmente benignas (SBD, 2021)./

2.5.2 Lesões Malignas

- 1. **Melanoma**: Tipo de câncer agressivo, com origem nos melanócitos. A detecção precoce é essencial para evitar sua propagação (INCA, 2022).
- 2. **Carcinoma de Células Basais (*Basal Cell Carcinoma*)**: O tipo mais comum de câncer de pele. Embora raro de se espalhar, pode causar desfigurações se não for tratado adequadamente (SBD, 2021).
- 3. **Queratose Actínica**: Lesão pré-cancerosa causada pela exposição solar, podendo evoluir para carcinoma de células escamosas (OMS, 2020).
- 4. **Carcinoma de Células Escamosas (*Squamous Cell Carcinoma*)**: Tumor maligno de células escamosas, com maior potencial de agressividade comparado ao carcinoma basocelular (INCA, 2022).

2.6 REDES NEURAIAS CONVOLUCIONAIS (CNN), VISION TRANSFORMERS (ViT) E TÉCNICAS DE PROCESSAMENTO DE IMAGENS

As Redes Neurais Convolucionais (Convolutional Neural Networks - CNNs) têm revolucionado o campo da visão computacional, promovendo avanços significativos no diagnóstico médico por imagem, especialmente em áreas como a análise de lesões dermatológicas. Estas redes são capazes de extrair automaticamente características relevantes de imagens, reduzindo a necessidade de intervenção manual e aumentando a precisão dos sistemas de apoio à decisão clínica.

Nesta seção, abordamos os fundamentos das CNNs e exploramos arquiteturas consolidadas como VGG (Visual Geometry Group), ResNet (Residual Network), DenseNet (Densely Connected Convolutional Networks) e EfficientNet, amplamente utilizadas por sua eficiência na extração de padrões visuais complexos e seu bom equilíbrio entre precisão e custo computacional. Além das CNNs tradicionais, também analisamos os Vision Transformers (ViT), uma abordagem mais recente que aplica mecanismos de atenção para

capturar relações espaciais em imagens, oferecendo resultados promissores em diversas tarefas de classificação e segmentação.

Também discutimos estratégias de balanceamento de dados, com destaque para o undersampling, que visa reduzir o viés causado por classes majoritárias em conjuntos de dados desbalanceados.

No processo de treinamento, a escolha e configuração adequada do otimizador (optimizer) e do agendador (scheduler) de taxa de aprendizado são cruciais para o desempenho do modelo. Por exemplo, uma configuração eficiente pode utilizar o otimizador RAdam (Rectified Adam), aplicando uma taxa de aprendizado um pouco mais alta e aumentando a regularização por meio do peso de decaimento. O agendador OneCycleLR é configurado para ajustar dinamicamente a taxa de aprendizado durante as épocas, começando com uma fase inicial de aumento gradual, seguido por uma redução, o que contribui para um treinamento mais estável e eficiente, ajudando o modelo a convergir melhor e evitando o sobreajuste.

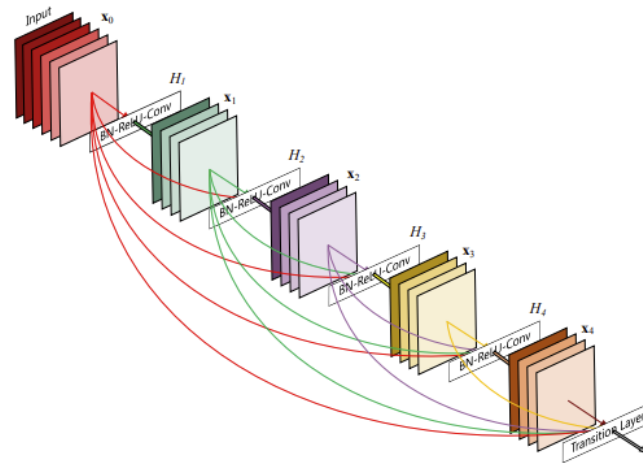
2.6.1 Conceitos Fundamentais das CNNs

Batch Size: O *batch size* é o número de amostras processadas antes que o modelo atualize seus pesos. Um *batch size* maior proporciona uma estimativa mais precisa do gradiente, mas aumenta o consumo de memória e pode levar a um treinamento mais lento. Por outro lado, um *batch size* menor pode acelerar o treinamento e reduzir o uso de memória, mas pode introduzir maior variação no gradiente e na convergência (Goodfellow, Bengio, & Courville, 2016).

Learning Rate: A taxa de aprendizado, ou learning rate, controla a magnitude das atualizações dos pesos durante o treinamento. Uma taxa de aprendizado muito alta pode causar uma convergência instável, enquanto uma taxa muito baixa pode resultar em um treinamento excessivamente lento. Ajustar a taxa de aprendizado é crucial para a eficiência e eficácia do treinamento de modelos (Kingma & Ba, 2014).

Camadas Densas: As camadas densas (ou totalmente conectadas) são responsáveis por combinar e interpretar as características extraídas pelas camadas convolucionais. Cada neurônio em uma camada densa está conectado a todos os neurônios da camada anterior, o que permite modelar relações complexas entre as características extraídas (LeCun et al., 2015).

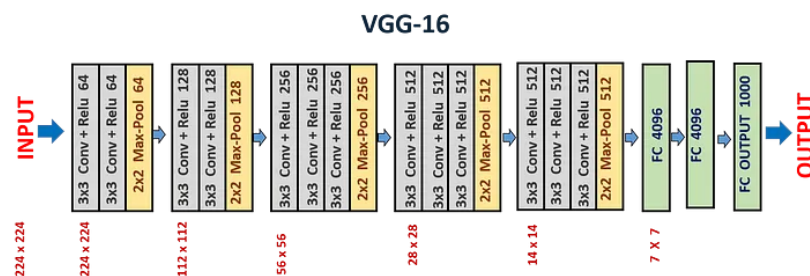
Figura 4: Arquitetura de uma Rede Densa



Fonte: Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

DenseNet, ou Rede Densa, é uma arquitetura que se caracteriza por sua conectividade forte entre as camadas. Ao contrário das arquiteturas tradicionais, onde cada camada se conecta apenas à próxima, na DenseNet cada camada está conectada a todas as camadas anteriores. Essa estrutura densa permite uma reutilização mais eficaz de características e um fluxo contínuo de informações ao longo da rede. Além disso, as conexões densas facilitam a passagem de gradientes durante o treinamento, tornando a rede mais eficiente e reduzindo o risco de overfitting. Esse design não só melhora o desempenho em tarefas de classificação, mas também promove uma maior eficiência computacional (Huang et al., 2017).

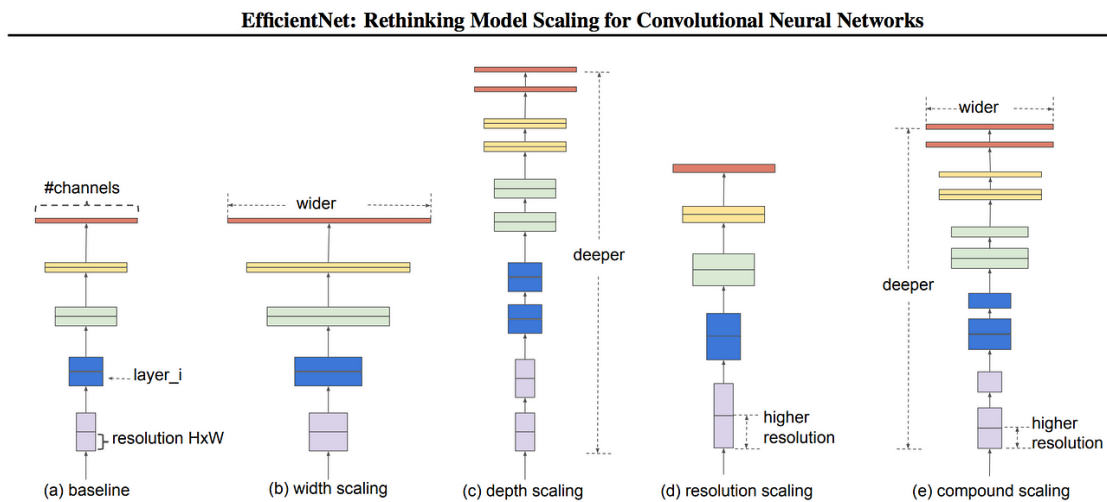
Figura 5: Arquitetura VGG-16



Fonte: Khandelwal, V. (2020). *The Architecture and Implementation of VGG-16. Towards AI*. Disponível em: <https://pub.towardsai.net/the-architecture-and-implementation-of-vgg-16-b050e5a5920b>. Acesso em: 24 set. 2024.

A arquitetura VGG é amplamente reconhecida por sua simplicidade e profundidade. Composta principalmente por camadas convolucionais de 3x3 e camadas de pooling, a VGG é projetada para ser profunda, permitindo a captura de detalhes sutis em imagens. A simplicidade de seu design, em conjunto com a profundidade da rede, torna a VGG uma escolha popular em diversas aplicações de visão computacional. Entretanto, apesar de seu desempenho excepcional, a VGG pode ser computacionalmente cara, devido ao grande número de parâmetros que contém (Simonyan & Zisserman, 2014).

Figura 6: Ilustração do Escalonamento Composto da EfficientNet



Fonte: Adaptado de Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*.

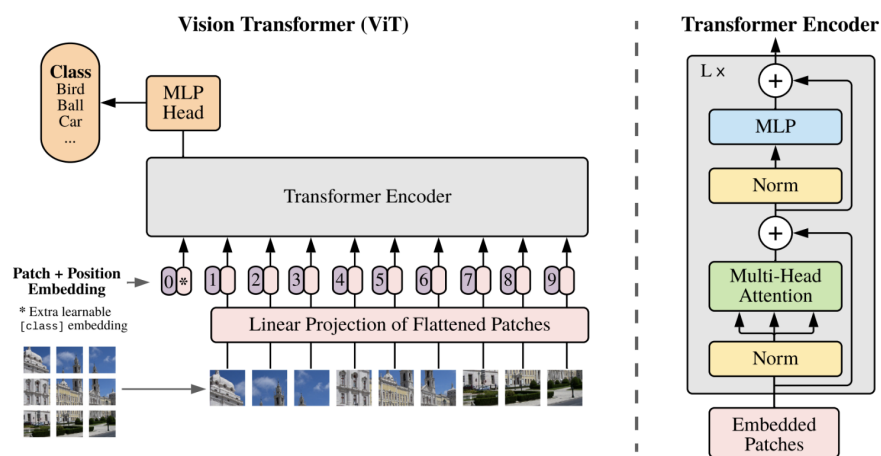
EfficientNet, ou Rede Eficiente, introduz uma abordagem metódica para o escalonamento de modelos, visando otimizar a acurácia e a eficiência simultaneamente. A inovação central da arquitetura é o **escalonamento composto**, um princípio que utiliza um coeficiente fixo para escalar de forma uniforme e equilibrada as três dimensões da rede: a profundidade (número de camadas), a largura (número de filtros por camada) e a resolução da imagem de entrada. Essa técnica parte da premissa de que essas dimensões não são independentes e que seu escalonamento balanceado é crucial para um melhor desempenho. Como resultado, a EfficientNet atinge um estado da arte superior com um custo computacional e um número de parâmetros muito menores em comparação com outras redes, tornando-se uma escolha ideal para aplicações com recursos computacionais limitados (Tan & Le, 2019).

Essas arquiteturas demonstram a evolução das redes neurais convolucionais e suas contribuições significativas para o campo da visão computacional, evidenciando a importância de inovações estruturais para o desempenho e eficiência dos modelos de aprendizado profundo.

2.6.4 Vision Transformer (ViT)

O Vision Transformer, ou ViT, representa uma mudança de paradigma na visão computacional ao aplicar a arquitetura Transformer, originalmente desenvolvida para tarefas de processamento de linguagem natural (PLN), diretamente a imagens. Diferente das CNNs que processam a imagem pixel a pixel através de convoluções, o ViT divide a imagem de entrada em uma sequência de patches (pequenos recortes) de tamanho fixo.

Figura 7: Arquitetura do Vision Transformer (ViT)



Fonte: Adaptado de Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., et al. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Cada patch é achatado e transformado em um vetor através de uma projeção linear. Para que o modelo entenda a organização espacial, informações de posição são adicionadas a esses vetores, e a sequência resultante é então alimentada em um codificador Transformer padrão. O mecanismo de autoatenção (*self-attention*) do Transformer permite ao modelo pesar a importância de cada patch em relação a todos os outros, capturando relações globais e

contextos de longa distância na imagem desde as primeiras camadas. Essa abordagem demonstrou um desempenho excepcional, desafiando a dominância das convoluções e superando CNNs em tarefas de classificação de imagens quando pré-treinada em larga escala (Dosovitskiy et al., 2021).

2.6.5 Técnicas de Pré-processamento de Imagens

O pré-processamento de imagens é um passo fundamental na análise de imagens médicas, especialmente na detecção de câncer de pele. Este processo visa melhorar a qualidade das imagens, facilitando a identificação de características relevantes para a classificação e diagnóstico. Diversas técnicas têm sido desenvolvidas para otimizar essa fase, cada uma abordando desafios específicos enfrentados na análise de imagens. A seguir, são discutidas algumas das principais técnicas de pré-processamento e suas contribuições para a análise de imagens médicas.

Uma das técnicas amplamente utilizadas é o DullRazor, que tem como objetivo a remoção de artefatos indesejados, como pelos, que podem interferir na análise das imagens. Ao melhorar o contraste e normalizar as imagens, essa técnica contribui para uma análise mais precisa. Lee et al. (1997) exploraram essa abordagem, apresentando o DullRazor como uma solução eficaz para a remoção de pelos em imagens médicas, melhorando assim a qualidade da análise subsequente (Lee, T., Ng, V., Gallagher, R., Coldman, A., & McLean, D., 1997).

Outra técnica significativa é a Transformação Black Hat, que é utilizada para realçar características escuras em imagens. Esta abordagem é particularmente útil na detecção de lesões cutâneas, pois permite destacar áreas de interesse que poderiam passar despercebidas. Hoshyar et al. (2014) discutem a relevância dessa técnica em seu trabalho, onde comparam várias técnicas de pré-processamento em sistemas de detecção de câncer de pele, ressaltando a eficácia da transformação Black Hat para enfatizar características relevantes (Hoshyar, A.N., Al-Jumaily, A., & Hoshyar, A.N., 2014).

A binarização das imagens também desempenha um papel importante, permitindo destacar características específicas e simplificando a análise. Essa técnica é essencial para facilitar a detecção de anomalias nas imagens. Khouloud et al. (2021) exploram essa técnica em seu artigo, no qual apresentam a utilização de uma rede W-net e uma rede residual de inception para a segmentação e classificação de lesões cutâneas, destacando como a binarização ajuda na simplificação do processo de análise (Khouloud, S., Ahlem, M., Fadel, T., & Amel, S., 2021).

Além disso, a normalização dos valores dos pixels é uma prática fundamental que melhora a consistência dos dados, um aspecto crucial para o treinamento de modelos de aprendizado de máquina. Goodfellow et al. (2016) abordam a importância da normalização em seu livro "Deep Learning", enfatizando como essa técnica contribui para a eficiência e eficácia dos modelos desenvolvidos para análise de imagens (Goodfellow, I., Bengio, Y., & Courville, A., 2016).

Essas técnicas de pré-processamento não apenas melhoram a qualidade das imagens, mas também aumentam a precisão dos modelos de aprendizado de máquina aplicados à detecção de câncer de pele. A implementação adequada dessas abordagens pode resultar em diagnósticos mais precisos e, consequentemente, em melhores desfechos para os pacientes.

2.6.5 Augmentação de Dados

A augmentação de dados é uma técnica essencial para aumentar a diversidade dos dados de treinamento e melhorar a robustez dos modelos. Entre as técnicas comuns de augmentação, o *crop* (ou recorte) envolve cortar uma parte da imagem para criar variações. Essa técnica ajuda o modelo a aprender a reconhecer padrões em diferentes partes da imagem e a lidar com variações na localização das características (Shorten & Khoshgoftaar, 2019).

Outra técnica é a rotação, que envolve girar as imagens em ângulos variados, permitindo que o modelo aprenda a reconhecer objetos independentemente da orientação. Isso aumenta a capacidade do modelo de generalizar para novas orientações e posições (Shorten & Khoshgoftaar, 2019).

O *flipping* (ou espelhamento) das imagens, tanto horizontal quanto verticalmente, também é uma técnica que ajuda o modelo a aprender simetrias e padrões invariantes, melhorando a capacidade de reconhecimento em diferentes configurações (Shorten & Khoshgoftaar, 2019).

Além disso, a variação nas características de cor, como brilho, contraste e saturação, conhecida como *color jitter*, contribui para a criação de imagens mais diversificadas e pode melhorar a robustez do modelo ao lidar com diferentes condições de iluminação (Shorten & Khoshgoftaar, 2019).

2.6.6 Otimizadores e Estratégias Avançadas de Treinamento

O treinamento de modelos de aprendizado de máquina é um processo complexo que visa ajustar os pesos do modelo para minimizar uma função de perda. A eficácia desse processo depende de uma robusta estratégia de otimização.

2.6.6.1 Fundamentos dos Otimizadores

Os otimizadores desempenham um papel fundamental no treinamento, e a escolha do algoritmo adequado pode ter um impacto significativo na convergência e na performance do modelo. Diversos otimizadores são utilizados em práticas de aprendizado profundo. O Stochastic Gradient Descent (SGD) é um dos mais básicos e amplamente utilizados, caracterizado pela atualização dos pesos com base em uma amostra aleatória de dados. Sua simplicidade e eficácia o tornam uma escolha popular, no entanto, a sua sensibilidade à taxa de aprendizado requer um ajuste cuidadoso, pois uma taxa inadequada pode resultar em uma convergência lenta ou em oscilações durante o treinamento (Goodfellow, Bengio, & Courville, 2016).

Outro otimizador amplamente utilizado é o Adam (Adaptive Moment Estimation), que combina as vantagens de dois outros algoritmos populares: o AdaGrad e o RMSProp. O Adam ajusta a taxa de aprendizado para cada parâmetro, utilizando estimativas de momentos de primeiro e segundo ordens. Essa abordagem permite uma adaptação mais eficiente às variações dos dados, tornando o Adam um dos otimizadores mais utilizados em aplicações de aprendizado profundo devido à sua robustez e eficiência (Kingma & Ba, 2014).

Além disso, o RMSProp é projetado para ajustar a taxa de aprendizado de forma adaptativa, com base na média móvel dos quadrados dos gradientes. Essa característica é particularmente útil para lidar com problemas de gradientes explosivos ou desaparecendo, que podem ocorrer em redes neurais profundas, ajudando a estabilizar o treinamento (Hinton, 2012).

2.6.6.2 Técnicas Complementares de Otimização e Treinamento

Além da escolha de um otimizador base, estratégias modernas de treinamento empregam um conjunto de técnicas complementares para refinar o processo de aprendizado, acelerar a convergência e otimizar o uso de recursos computacionais.

A Função de Perda (Loss Function) é o pilar que guia a otimização, quantificando o erro entre as previsões do modelo e os rótulos verdadeiros. O objetivo do otimizador é encontrar os parâmetros que minimizam o valor desta função, cuja escolha depende diretamente da tarefa (e.g., Entropia Cruzada para classificação) (Goodfellow, Bengio, & Courville, 2016).

Para aprimorar otimizadores como o Adam, surgiram variantes como o Rectified Adam (RAdam). Este algoritmo aborda a instabilidade do Adam no início do treinamento,

que ocorre devido à alta variância do momento adaptativo. O RAdam introduz um termo de retificação que efetivamente "aquece" o otimizador, resultando em um processo de treinamento mais estável e robusto (Liu et al., 2019).

Adicionalmente, o Agendamento da Taxa de Aprendizado (Learning Rate Scheduling) é uma técnica para ajustar dinamicamente a taxa de aprendizado durante o treinamento. Políticas como a One-Cycle Policy variam a taxa de aprendizado de um valor baixo a um máximo e de volta a um valor baixo ao longo das épocas. Esta abordagem demonstrou acelerar a convergência e atuar como uma forma de regularização, ajudando o modelo a encontrar mínimos mais generalizáveis na paisagem de perda (Smith, 2018).

Finalmente, para otimizar a eficiência computacional, o Treinamento com Precisão Mista Automática (AMP) é empregado. Esta técnica utiliza formatos numéricos de meia precisão (float 16), mais rápidos, para operações intensivas, enquanto mantém a precisão total (float 32) para operações críticas. Para evitar o problema de desaparecimento de gradientes (gradient underflow) em float 16, a técnica de loss scaling é aplicada, garantindo a estabilidade numérica e permitindo ganhos de velocidade significativos com redução no consumo de memória da GPU (Micikevicius et al., 2017).

Em conclusão, a escolha de um otimizador apropriado é crucial para o sucesso do treinamento de modelos de aprendizado de máquina. A utilização de técnicas adequadas de otimização pode acelerar a convergência do modelo e melhorar sua performance, contribuindo para a eficácia em tarefas complexas, como a análise de imagens médicas.

3 METODOLOGIA

A metodologia deste trabalho foi estruturada para desenvolver um sistema de detecção e classificação de câncer de pele utilizando tanto Convolutional Neural Networks (CNN) quanto Vision Transformer (ViT), com o objetivo de distinguir entre lesões benignas, pré-cancerosas e malignas em imagens dermatológicas. O processo foi cuidadosamente dividido em fases, incluindo a coleta de dados, pré-processamento de imagens, arquitetura dos modelos, treinamento e avaliação, a fim de garantir a maior eficiência e precisão possível no diagnóstico automatizado.

3.1 COLETA E PREPARAÇÃO DE DADOS

Para este estudo, a base de dados *International Skin Imaging Collaboration 2019 (ISIC 2019)* foi obtida diretamente do repositório oficial da competição ISIC por meio de um processo automatizado. Esse processo consistiu no download dos arquivos contendo as

imagens de treinamento e teste, os respectivos arquivos de ground truth (rótulos) e os metadados relacionados à tarefa de classificação de lesões cutâneas.

Foram obtidas imagens de entrada para treinamento e teste, que estavam armazenadas em arquivos compactados no formato ZIP. Além disso, foram adquiridos arquivos no formato CSV contendo as anotações de classes (ground truth) correspondentes a essas imagens. Também foram incluídos arquivos CSV com os metadados das imagens, os quais trazem informações clínicas e demográficas relevantes para o estudo.

Esse procedimento automatizado assegurou a padronização e a confiabilidade dos dados utilizados no desenvolvimento e avaliação dos modelos.

3.1.1 Divisão das Imagens: Lesões Benignas e Malignas

A divisão das imagens foi uma etapa crucial para o sucesso do modelo. As imagens foram separadas em duas classes principais para a tarefa de classificação binária: lesões benignas e lesões malignas.

A classe de lesões benignas foi composta por diagnósticos que não apresentam risco imediato de malignidade. Especificamente, esta classe incluiu: Nevos Melanocíticos, os diferentes tipos de Queratose Benigna (Lentigo Solar, Queratose Seborreica e Queratose tipo Líquen Plano), Dermatofibromas e Lesões Vasculares.

Por sua vez, a classe de lesões malignas agrupou os diagnósticos de alto risco. Esta classe abrangeu o Melanoma, o Carcinoma de Células Basais e o Carcinoma de Células Escamosas, que possuem um alto potencial de disseminação e requerem atenção especial (SILVA; OLIVEIRA, 2020). Adicionalmente, as lesões pré-cancerosas, como a Queratose Actínica, foram incluídas nesta mesma classe, devido ao seu risco de evoluírem para câncer se não tratadas.

A maioria dos datasets disponíveis para detecção de lesões apresenta uma predominância significativa da classe benigna em relação à maligna. Essa desproporção pode levar a um viés no modelo, que tenderia a classificar erroneamente as lesões malignas por favorecer a classe majoritária. No conjunto de dados original, essa disparidade era evidente: 19.341 imagens de lesões benignas contra 12.181 de lesões malignas, representando uma proporção de aproximadamente 63,7% para a classe majoritária e apenas 36,3% para a minoritária.

Para evitar que o modelo fosse enviesado e garantir uma maior sensibilidade na identificação das lesões malignas, que possuem maior relevância clínica, aplicamos técnicas de balanceamento de classes. Especificamente, optamos pela subamostragem (undersampling)

da classe majoritária, que consistiu em reduzir aleatoriamente o número de amostras benignas para igualar a quantidade de imagens da classe maligna. Essa abordagem resultou em um novo dataset balanceado, com 24.362 imagens, distribuídas igualmente entre as duas classes, com 12.181 imagens para cada uma.

A partir deste dataset balanceado, realizamos a divisão em subconjuntos de treino, validação e teste, respeitando as proporções de 80%, 10% e 10%, respectivamente. Além disso, garantimos a manutenção do balanceamento das classes em todos esses subconjuntos, assegurando que o modelo fosse treinado e avaliado de forma justa e eficiente, minimizando o risco de viés preditivo (PEREIRA, 2023).

3.2 CRIAÇÃO DOS DATASETS

A preparação dos dados é uma etapa fundamental em qualquer projeto de aprendizado de máquina, sendo crucial para garantir a qualidade e a consistência das amostras que alimentarão o modelo. Neste trabalho, foi desenvolvido um pipeline de pré-processamento para transformar o conjunto de imagens brutas em datasets otimizados para a tarefa de classificação de lesões de pele. Este processo resultou na criação de três conjuntos de dados distintos, cada um representando um nível diferente de transformação, permitindo uma análise comparativa do impacto de cada etapa no desempenho do modelo.

O primeiro conjunto de dados gerado teve como objetivo principal a padronização do tamanho das imagens, preservando, no entanto, a proporção original de cada uma. Para isso, cada imagem foi redimensionada de forma que sua maior dimensão fosse ajustada para um tamanho máximo de 256 pixels, com a outra dimensão recalculada proporcionalmente. Este dataset, que mantém a forma original da lesão, serve como uma base de comparação para avaliar o efeito de transformações geométricas mais complexas.

O segundo dataset foi criado para atender aos requisitos da maioria das arquiteturas de Redes Neurais Convolucionais (CNNs), que exigem imagens de entrada com dimensões fixas e quadradas. A metodologia para gerar este conjunto envolveu um processo de duas etapas: primeiro, um corte central foi aplicado para extrair a maior área quadrada possível do centro geométrico da imagem; subsequentemente, a imagem quadrada resultante foi redimensionada para as dimensões alvo de 256x256 pixels, utilizando a interpolação Lanczos para preservar a qualidade dos detalhes.

Finalmente, o terceiro e mais processado dataset foi desenvolvido com o objetivo de fornecer ao modelo imagens não apenas padronizadas, mas também livres de artefatos visuais, como pelos, que poderiam interferir no processo de aprendizado. Sua criação combinou a

remoção de artefatos com a padronização geométrica. Como primeira etapa, cada imagem original foi processada pelo algoritmo DullRazor, que envolve a conversão para escala de cinza, a aplicação da transformada morfológica Black-Hat para realçar os pelos, a criação de uma máscara binária e o uso de um algoritmo de inpainting para reconstruir as áreas ocluídas. Após esta etapa de limpeza, a imagem resultante foi submetida ao mesmo processo de corte central e redimensionamento, resultando em um conjunto de imagens de 256x256 pixels, padronizadas e limpas. A geração destes três conjuntos de dados distintos estabelece, portanto, uma base sólida para a análise experimental, permitindo avaliar de forma controlada como cada nível de pré-processamento impacta a performance final do modelo de classificação.

3.3 BALANCEAMENTO DO DATASET

O conjunto de dados original combinado, composto por 33.569 amostras, apresentava um desequilíbrio notável entre as classes. Conforme exibido na primeira tabela ("Resumo das Distribuições dos Datasets"), havia uma predominância de amostras benignas (19.341, ou 57,6%) em comparação com as malignas (12.181, ou 36,3%).

Tabela 1: Detalhamento da contagem de amostras por classe antes e após o balanceamento.

Detalhamento da Contagem por Classe e Malignidade						
Classe	Contagem Original	Contagem Balanceada	Contagem Treino	Contagem Validação	Contagem Teste	Malignidade
MEL	5.849	5.849	4.679	585	585	Maligno
NV	15.370	9.727	7.781	973	973	Benigno
BCC	4.298	4.298	3.438	430	430	Maligno
AK	1.241	1.241	993	124	124	Maligno
BKL	3.284	2.030	1.624	203	203	Benigno
DF	330	207	166	21	20	Benigno
VASC	357	217	174	21	22	Benigno
SCC	793	793	634	79	80	Maligno

Fonte: Elaborado pelo autor (2025).

Para tratar esse desequilíbrio, a principal estratégia adotada foi a **subamostragem (undersampling)** da classe majoritária (benigna). O número de amostras benignas foi reduzido de 19.341 para 12.181, igualando-se à quantidade de amostras malignas. Isso resultou em um dataset "Balanceado Combinado" com um total de 24.362 amostras, perfeitamente equilibrado com 50% para cada classe.

Tabela 2: Resumo da distribuição dos datasets (Original vs. Balanceado).

Resumo das Distribuições dos Datasets					
Dataset	Total de Amostras	Amostras Malignas	Amostras Benignas	Proporção Maligna	Proporção Benigna
Original Combinado	33.569	12.181	19.341	0,363	0,576
Balanceado Combinado	24.362	12.181	12.181	0,500	0,500
Split Treino	19.489	9.744	9.745	0,500	0,500
Split Validação	2.436	1.218	1.218	0,500	0,500
Split Teste	2.437	1.219	1.218	0,500	0,500

Fonte: Elaborado pelo autor (2025).

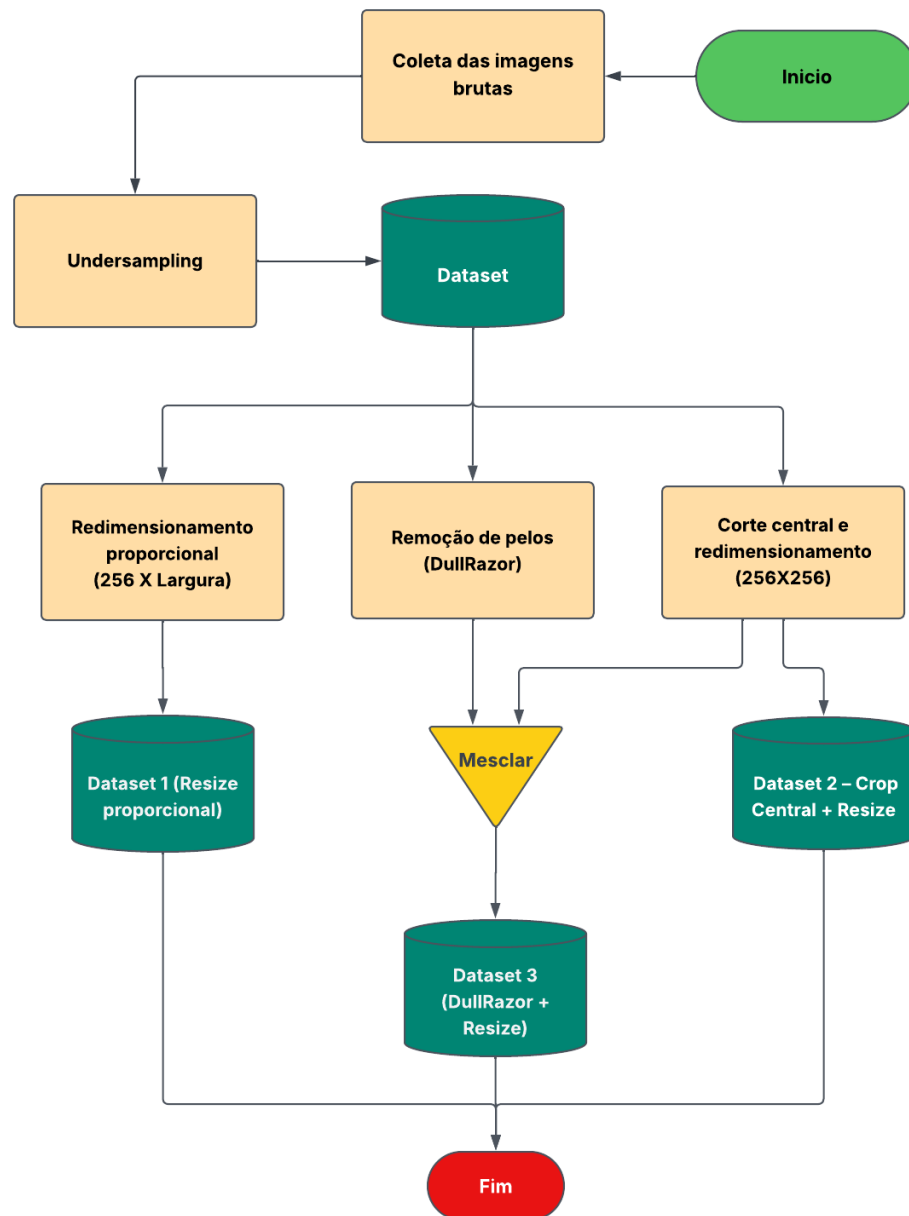
A segunda tabela ("Detalhamento da Contagem por Classe") evidencia essa redução, mostrando que classes como NV (Nevo) e BKL (Ceratose seborreica) tiveram seu número de amostras diminuído. Este conjunto de dados balanceado foi então dividido nas partições de treino, validação e teste para o desenvolvimento do modelo.

3.4 METODOLOGIA DETALHADA DE PRÉ-PROCESSAMENTO E ESTRUTURAÇÃO DE DADOS

Nesta seção, é apresentado o fluxo metodológico completo para o pré-processamento das imagens dermatoscópicas. O objetivo central deste pipeline é transformar imagens brutas, com variações de enquadramento, resolução e artefatos, em um conjunto de dados padronizado, limpo e otimizado para o treinamento e a avaliação de um modelo de classificação baseado em Redes Neurais Convolucionais (CNNs).

O fluxograma apresentado na Figura 8 ilustra a estratégia geral adotada, desde a coleta inicial das imagens até a criação de diferentes datasets para análise comparativa. As seções a seguir detalham especificamente o pipeline principal utilizado neste estudo, que culmina na geração de um dataset com normalização geométrica e remoção de artefatos.

Figura 8: Fluxograma do pipeline de pré-processamento e criação dos datasets. O diagrama ilustra as etapas desde a coleta das imagens brutas, passando pelo undersampling do dataset e seguindo para a criação de conjuntos de dados distintos baseados em diferentes técnicas de pré-processamento.



Fonte: Elaborado pelo autor (2025).

3.3.1 Normalização Geométrica e Foco na Região de Interesse (ROI)

A primeira fase do pipeline concentrou-se na normalização geométrica, uma etapa crucial para garantir que o modelo de deep learning recebesse entradas com dimensões e

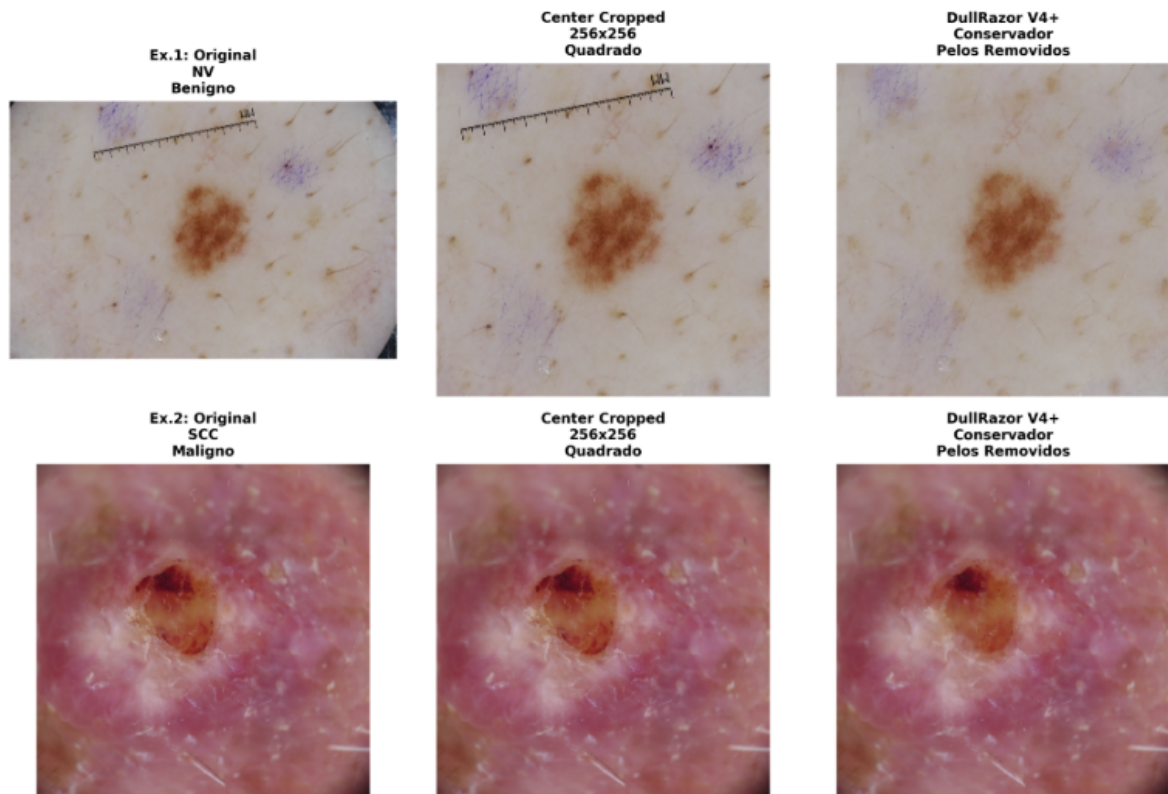
enquadramento consistentes. O processo iniciou-se com um recorte central para padronizar a proporção das imagens. Para cada amostra, a menor de suas dimensões, seja altura ou largura, foi utilizada para definir o lado de um quadrado de recorte, extraíndo assim a porção central da imagem em uma proporção de 1:1. Esta técnica assegura que a lesão, geralmente localizada no centro, seja maximizada. Em seguida, para ampliar a Região de Interesse (ROI) e descartar anéis de pele saudável periféricos, foi aplicado um zoom digital com um fator de 1,5x. Este procedimento foi implementado como um segundo recorte sobre a imagem já quadrada, focando ainda mais na lesão. Por fim, todas as imagens foram redimensionadas para uma resolução padrão de 256x256 pixels. Para esta etapa de redimensionamento, foi empregado o algoritmo de interpolação Lanczos-4, uma escolha deliberada devido à sua capacidade superior de preservar detalhes finos e a nitidez das bordas.

3.3.2 Remoção de Artefatos com o Algoritmo DullRazor (Abordagem Conservadora)

Superada a padronização geométrica, o pipeline prosseguiu para a fase de remoção de artefatos, com especial atenção à eliminação de pelos, que constituem um ruído visual capaz de ser erroneamente interpretado pelo modelo. Para essa tarefa, foi implementada uma abordagem conservadora do algoritmo DullRazor. O processo teve início com a conversão da imagem colorida para escala de cinza. Subsequentemente, para realçar essas estruturas finas, aplicou-se uma operação morfológica de black-hat, utilizando um kernel elíptico com dimensões de 9x9 pixels. A máscara de pelos resultante foi então suavizada com um filtro Gaussiano de kernel 3x3 para reduzir ruídos. A partir daí, uma máscara binária foi gerada por meio de uma operação de limiarização com um valor fixo de 15. Para evitar a remoção acidental de estruturas maiores, todos os componentes na máscara cuja área excedia 5% da área total da imagem foram descartados. Na etapa final, as áreas identificadas como pelos foram preenchidas na imagem colorida original utilizando o algoritmo de inpainting baseado no Método de Marcha Rápida (Fast Marching Method), com um raio de vizinhança de 2 pixels para propagar de forma realista a textura da pele circundante.

A eficácia visual das etapas de normalização geométrica e remoção de artefatos é demonstrada na Figura 9, que compara as imagens antes e depois do processamento para um caso benigno e um maligno.

Figura 10: Demonstração visual do pipeline de pré-processamento.



Fonte: Elaborado pelo autor (2025).

A primeira linha (Ex.1) mostra um nevo benigno e a segunda linha (Ex.2) um carcinoma espinocelular maligno. Da esquerda para a direita: a imagem original; a imagem após o recorte central e redimensionamento para 256x256 pixels; e a imagem final após a aplicação do DullRazor para remoção de pelos.

3.3.3 Estruturação e Armazenamento Final dos Dados

Ao final deste pipeline de processamento, as imagens limpas e normalizadas foram estruturadas para um armazenamento eficiente. Os dados foram salvos no formato Pytorch Tensor, o sistema de serialização de tensores nativo do PyTorch. Cada arquivo consiste em um objeto que encapsula não apenas o tensor da imagem processada (de dimensões 256x256x3), mas também seus metadados essenciais. Essas informações associadas incluíam um identificador único, o diagnóstico da lesão e a classe binária correspondente (benigno ou maligno). Este método de armazenamento, que une os dados de entrada e seus rótulos em um único arquivo, otimiza significativamente o processo de carregamento de dados durante a fase de treinamento do modelo, pois elimina a necessidade de executar repetidamente as operações de pré-processamento.

3.4 ARQUITETURAS UTILIZADAS

As arquiteturas de redes neurais convolucionais desempenham um papel fundamental na análise e classificação de imagens médicas, especialmente no diagnóstico de lesões dermatológicas. Neste estudo, foram selecionadas arquiteturas amplamente reconhecidas por seu desempenho: DenseNet121, DenseNet169, DenseNet201, ResNet50, ResNet50V2, ResNet101V2, VGG-16, EfficientNetB7 e Vision Transformer (ViT). Cada uma dessas redes possui características estruturais específicas que as tornam adequadas para a tarefa de classificação de lesões cutâneas, contribuindo de maneira significativa para a precisão e robustez do sistema de diagnóstico automatizado.

3.4.1 DenseNet121, DenseNet169 e DenseNet201

As fundações deste estudo são construídas sobre as arquiteturas da família DenseNet (Densely Connected Convolutional Networks). Propostos por Huang et al. (2017), esses modelos são definidos por um inovador padrão de conectividade densa, onde cada camada convolucional dentro de um bloco recebe como entrada os mapas de características de todas as camadas anteriores. Essa abordagem maximiza a reutilização de características, o que, por sua vez, aprimora o fluxo de gradiente através da rede e combate o problema do gradiente desaparecido (*vanishing gradient*), um desafio comum em redes profundas. Tal eficiência torna a DenseNet uma escolha robusta para tarefas que exigem a detecção de padrões finos, como na análise de imagens dermatológicas.

Para este trabalho, foram empregados três modelos específicos — **DenseNet121**, **DenseNet169** e **DenseNet201**. A principal diferença entre eles reside na profundidade, um fator que determina diretamente a capacidade de representação do modelo. A **DenseNet121**, com 121 camadas, oferece uma solução mais leve e computacionalmente eficiente. A **DenseNet169**, com 169 camadas, apresenta uma capacidade intermediária para capturar padrões mais intrincados. Por fim, a **DenseNet201**, a mais profunda das três com 201 camadas, é projetada para cenários de alta complexidade que demandam o máximo poder de representação.

A seleção de modelos com profundidades distintas visa permitir uma análise criteriosa do impacto da complexidade da arquitetura no desempenho da tarefa de [mencione sua tarefa aqui]. É importante notar que todos os modelos foram pré-treinados no conhecido dataset ImageNet e suas implementações padrão incorporam técnicas cruciais para a estabilização do

treinamento, como a normalização em lote (*Batch Normalization*), a regularização com *Dropout* e a otimização através do algoritmo *SGD com Momentum* (PAPERS WITH CODE, 2024).

3.4.2 ResNet50v2, ResNet101v2

Para complementar a análise, foram selecionadas as arquiteturas da família ResNet (Redes Neurais Residuais), especificamente os modelos **ResNet50v2** e **ResNet101v2**. A ResNet foi uma arquitetura pioneira na mitigação do problema do desaparecimento de gradiente em redes muito profundas através do uso de "conexões de atalho" ou conexões residuais (*residual connections*). Essas conexões permitem que o gradiente flua mais livremente através das camadas, possibilitando o treinamento eficaz de redes com centenas ou até milhares de camadas.

A versão 2 (v2) da ResNet, utilizada neste trabalho, introduz uma melhoria chave no caminho do bloco residual. A normalização em lote (*Batch Normalization*) e a função de ativação (ReLU) são aplicadas antes da camada de convolução (pré-ativação), em vez de depois. Essa reorganização, proposta por He et al. (2016), otimiza o fluxo de informações e resulta em uma regularização mais forte, levando a um treinamento mais estável e a um melhor desempenho de generalização.

Os modelos específicos selecionados permitem comparar o efeito da profundidade dentro desta arquitetura. O **ResNet50v2**, com 50 camadas, é um modelo amplamente utilizado que oferece um excelente equilíbrio entre profundidade e eficiência computacional, enquanto o **ResNet101v2**, com 101 camadas, oferece uma capacidade de representação significativamente maior, sendo capaz de aprender características mais complexas e hierárquicas dos dados.

No contexto da classificação de lesões de pele, as conexões residuais são cruciais, pois facilitam a captura de detalhes finos e texturas sutis, preservando informações importantes que poderiam ser perdidas em redes muito profundas. A comparação entre ResNet50v2 e ResNet101v2 permitirá avaliar como a profundidade adicional impacta a capacidade do modelo de distinguir entre lesões benignas e malignas (PAPERS WITH CODE, 2024b).

3.4.3 VGG-16

A VGG-16 é uma arquitetura mais antiga, mas ainda amplamente utilizada em tarefas

de classificação de imagens. Sua estrutura simples, baseada em múltiplas camadas convolucionais empilhadas seguidas por camadas totalmente conectadas, a torna eficiente em capturar características profundas das imagens (SIMONYAN; ZISSERMAN, 2014).

Embora tenha uma contagem de parâmetros mais alta que modelos mais recentes, ela continua oferecendo resultados robustos, especialmente quando combinada com técnicas modernas, como o Transfer Learning. Apesar de sua simplicidade em relação às redes modernas, a VGG-16 demonstrou ser eficaz na tarefa de classificação de imagens de pele, extraindo características importantes de forma detalhada e robusta (PAPERS WITH CODE, 2024c).

3.4.4 Vision Transformer (ViT)

Como uma abordagem distinta das Redes Neurais Convolucionais (CNNs), este estudo também emprega a arquitetura **Vision Transformer (ViT)**. Proposta por Dosovitskiy et al. (2020), a ViT adapta a arquitetura Transformer, originalmente bem-sucedida no Processamento de Linguagem Natural (PLN), para tarefas de visão computacional. O modelo opera dividindo uma imagem de entrada em uma sequência de "patches" (recortes) de tamanho fixo. Cada patch é então linearmente projetado em um vetor (embedding), e a informação posicional de cada patch é adicionada a este vetor para que o modelo retenha a estrutura espacial da imagem.

A sequência resultante de vetores é então processada por um codificador Transformer padrão. A principal vantagem deste método é o mecanismo de autoatenção (*self-attention*), que permite ao ViT ponderar a importância de cada patch em relação a todos os outros patches da imagem, capturando dependências de longo alcance e um contexto global. Diferente das CNNs, que possuem campos receptivos locais, o ViT pode, teoricamente, modelar relações entre pixels muito distantes.

Na análise de lesões de pele, essa capacidade de capturar o contexto global é de grande interesse. Ela permite que o modelo avalie a lesão como um todo, considerando características como a simetria geral, as bordas e as variações de cor em toda a sua extensão, o que pode levar a uma capacidade de distinção mais robusta.

3.4.5 EfficientNet-B7

A família de modelos EfficientNet foi desenvolvida para resolver um desafio central

no design de redes neurais: como escalar modelos de forma eficiente para obter maior acurácia. Proposta por Tan e Le (2019), a inovação central da EfficientNet é o método de escalonamento composto (compound scaling). Em vez de aumentar arbitrariamente apenas uma das dimensões da rede (profundidade, largura ou resolução da imagem), o escalonamento composto utiliza um coeficiente unificado (ϕ) para escalar as três dimensões de forma balanceada e simultânea.

Essa abordagem garante que, à medida que o modelo se torna maior e mais poderoso, seus recursos computacionais sejam alocados de maneira ótima. A partir de uma arquitetura base (EfficientNet-B0), os autores geraram uma família de modelos (de B0 a B7). Para este trabalho, foi selecionado o modelo EfficientNet-B7. Esta é a maior e mais poderosa versão da família, projetada para alcançar a máxima acurácia, sendo ideal para tarefas onde o desempenho de classificação é o fator crítico.

A escolha do EfficientNet-B7 para a tarefa de classificação de lesões de pele se justifica por sua altíssima capacidade de representação. Embora exija mais recursos computacionais em comparação com outras arquiteturas, seu design otimizado pelo escalonamento composto permite que ele atinja um desempenho de ponta (state-of-the-art), o que é fundamental para uma aplicação de diagnóstico onde a precisão é a maior prioridade.

3.5 APLICAÇÃO DO TRANSFER LEARNING

Em todas as arquiteturas utilizadas (DenseNet121, DenseNet169, DenseNet201, ResNet50, ResNet50V2, ResNet101V2, VGG-16, EfficientNetB7 e Vision Transformer (ViT)), foi aplicada a técnica de Transfer Learning para aproveitar os pesos pré-treinados em grandes datasets, como o ImageNet. Esta abordagem não apenas acelerou o processo de treinamento, mas também permitiu que o modelo iniciasse a partir de uma base já otimizada para reconhecer padrões visuais comuns (YOSINSKI et al., 2014). Ao transferir esse conhecimento prévio, as redes foram ajustadas especificamente para a tarefa de detecção de câncer de pele, onde camadas mais profundas foram treinadas para refinar o reconhecimento de características específicas de lesões dermatológicas.

3.6 AJUSTE E COMPARAÇÃO DAS ARQUITETURAS

Cada arquitetura foi treinada utilizando a mesma divisão de dados: 80% das imagens para treinamento, 10% para validação e 10% para teste. Durante o treinamento, os principais hiperparâmetros, como taxa de aprendizado, tamanho do batch e número de épocas, foram

ajustados conforme necessário para cada arquitetura. A comparação entre elas baseou-se em diversas métricas de avaliação, incluindo acurácia, precisão, recall, F1-Score e AUC-ROC (POWERS, 2011).

No processo de ajuste fino, as DenseNets, com sua conexão densa entre camadas, demonstraram grande eficiência na captura de padrões complexos, particularmente em casos de lesões malignas que exigem uma análise mais profunda (HUANG et al., 2017). A ResNet50v2 também se destacou por sua capacidade de generalização, evitando o overfitting e mantendo um equilíbrio adequado entre precisão e recall (HE et al., 2016). A VGG-16, por sua vez, apresentou resultados consistentes, mas com uma leve desvantagem em relação às arquiteturas mais modernas, devido à sua menor profundidade e conectividade (SIMONYAN; ZISSERMAN, 2014).

3.7 TREINAMENTO DO MODELO

O treinamento do modelo seguiu a divisão de dados padrão (80/10/10), sendo executado por 100 épocas com um tamanho de lote de 64 amostras. Para a otimização dos pesos, foi empregado o otimizador Radam (Rectified Adam), que oferece maior estabilidade inicial. A função de perda utilizada foi a Focal Loss, escolhida por sua eficácia em lidar com o desbalanceamento de classes, direcionando o foco do modelo para os exemplos mais difíceis (LIN et al., 2017).

A taxa de aprendizado foi gerenciada dinamicamente pelo agendador OneCycleLR, que ajusta o valor ao longo do treinamento para promover uma convergência mais eficiente (SMITH, 2018). Para evitar o overfitting, foram aplicadas duas técnicas principais de regularização. A primeira foi o Dropout, que desativa neurônios aleatoriamente para forçar o modelo a aprender representações mais robustas (SRIVASTAVA et al., 2014).

A segunda técnica foi uma estratégia de salvamento do melhor modelo (checkpointing). Em vez de interromper o treinamento, o desempenho do modelo no conjunto de validação foi monitorado a cada época. O estado do modelo que apresentou a menor perda de validação foi salvo, garantindo que a versão mais generalista fosse utilizada para a avaliação final, uma abordagem eficaz para mitigar o superajuste (PRECHELT, 1998). A combinação dessas técnicas foi essencial para desenvolver um modelo preciso e confiável.

3.8 OTIMIZAÇÃO COM RADAM (*RECTIFIED ADAM*)

A escolha do otimizador Radam se torna mais clara ao comparar suas características com as das alternativas mais comuns. O SGD com Momentum, por exemplo, baseia-se em uma taxa de aprendizado única para todos os parâmetros, à qual adiciona um termo de inércia para acelerar a convergência. Suas principais vantagens são a leveza computacional e um bom potencial de generalização, mas suas desvantagens são a lentidão para convergir e uma alta sensibilidade à escolha da taxa de aprendizado.

Em contrapartida, o Adam utiliza uma abordagem de taxa de aprendizado adaptativa, calculando um valor individual para cada parâmetro da rede. Isso lhe confere uma notável velocidade de convergência e robustez na maioria dos cenários. Contudo, sua principal desvantagem é a potencial instabilidade no início do treinamento, que pode levar o modelo a convergir para mínimos subótimos.

Finalmente, o Radam foi o escolhido por atuar como uma síntese inteligente dos dois. Ele inicia o treinamento operando de forma similar ao SGD com Momentum e, à medida que o processo avança, transita para o modo adaptativo do Adam. Dessa forma, ele herda a rápida convergência do Adam, mas com a estabilidade inicial do SGD, representando uma solução mais robusta e confiável para os desafios deste trabalho.

A combinação dessas técnicas avançadas de treinamento e regularização foi essencial para garantir que o modelo fosse capaz de realizar diagnósticos com alto grau de precisão, confiabilidade e robustez na tarefa de detecção de câncer de pele.

3.9 AVALIAÇÃO DO MODELO

A avaliação do modelo foi realizada utilizando diversas métricas de classificação para garantir uma análise abrangente de seu desempenho na detecção de lesões benignas e malignas. As métricas principais foram:

Acurácia (Accuracy) - mede a proporção de previsões corretas em relação ao total de exemplos. Calculada como (POWERS, 2011):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision (Precisão) - mede a proporção de verdadeiros positivos em relação ao total de predições positivas. Representa a fórmula (DAVIS; GOADRIC, 2006).

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall (Sensibilidade) - mede a capacidade do modelo de identificar corretamente as instâncias positivas (DAVIS; GOADRIC, 2006).

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score - combina a precisão e o recall em uma única métrica, calculada como a média harmônica de ambos (CHICCO; JURMAN, 2020).

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Área Sob a Curva ROC (AUC-ROC) - mede a capacidade do modelo de separar classes, sendo particularmente útil para problemas desbalanceados. Quanto maior o valor da AUC, melhor a capacidade do modelo de discriminar entre as classes positivas e negativas (BRADLEY, 1997).

3.9.1 Outras Métricas Usadas

Sensibilidade - focada na detecção de lesões malignas, mede a proporção de verdadeiros positivos (lesões corretamente identificadas) em relação ao total de casos verdadeiramente malignos (METZ, 1978).

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Especificidade (Specificity) - mede a capacidade do modelo de identificar corretamente as lesões benignas (verdadeiros negativos) (METZ, 1978).

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Área sob a Curva Parcial (pAUC acima de 80% TPR) – avalia o desempenho do modelo apenas na região da curva ROC em que a taxa de verdadeiros positivos (TPR) é superior a

80%, o que é especialmente relevante em cenários clínicos que exigem alta sensibilidade na detecção de lesões malignas (ISIC Challenge, 2020).

Precisão Média (mAP) - A (mAP, *mean Average Precision*) é uma métrica amplamente usada na avaliação de modelos de classificação e detecção de objetos, especialmente em tarefas de visão computacional. Ela combina dois conceitos: a precisão em cada nível de confiança e a média dessas precisões ao longo de várias classes ou diferentes limiares. Uma alta mAP significa que o modelo mantém boa precisão em diferentes níveis de confiança (LIN et al., 2014).

3.10 ARQUITETURA DA APLICAÇÃO

A arquitetura compreende quatro componentes fundamentais interconectados:

Aplicativo Mobile Flutter: Desenvolvido utilizando o framework Flutter para garantir compatibilidade multiplataforma (iOS e Android), o aplicativo constitui a interface primária para captura e gerenciamento de imagens dermatológicas. A escolha do Flutter baseia-se em sua capacidade de renderização nativa e desempenho otimizado para aplicações médicas (BIØRN-HANSEN et al., 2019).

Servidor FastAPI: O backend foi implementado utilizando FastAPI, um framework Python moderno caracterizado por alta performance e validação automática de dados (RAMIREZ, 2021). Este componente é responsável pelo processamento de inteligência artificial, gerenciamento de modelos de deep learning e orquestração das operações de análise de imagens.

Banco de Dados PostgreSQL: Sistema de gerenciamento de banco de dados relacional escolhido por sua robustez, conformidade com padrões SQL e capacidades avançadas de indexação, essenciais para o armazenamento seguro de dados clínicos sensíveis (MOMJIAN, 2001).

Modelos de Deep Learning: Implementação de redes neurais convolucionais (CNNs) especializadas para segmentação e classificação de lesões dermatológicas, fundamentadas em arquiteturas estado-da-arte validadas cientificamente (ESTEVA et al., 2017).

3.12.1 Processamento de Imagens

O pipeline de processamento implementa uma abordagem multi-modal combinando técnicas de pré-processamento, segmentação automática e classificação para maximizar a precisão diagnóstica.

3.12.2 Segmentação Automática

A segmentação de lesões dermatológicas constitui etapa crítica no processo diagnóstico, permitindo a delimitação precisa da região de interesse (ROI) e reduzindo interferências do background (YUAN et al., 2017).

Arquitetura do Modelo: Foi implementada uma rede U-Net modificada, conforme proposto por Ronneberger et al. (2015), especificamente adaptada para segmentação de lesões dermatológicas. A arquitetura U-Net foi selecionada devido à sua eficácia comprovada em tarefas de segmentação médica, apresentando capacidade superior de preservação de detalhes espaciais através de conexões skip.

Pré-processamento: As imagens de entrada são redimensionadas para resolução padronizada de 256x256 pixels, seguindo normalização por z-score para estabilização do treinamento (IOFFE; SZEGEDY, 2015). Esta padronização garante consistência no processamento e reduz variações relacionadas às condições de captura.

Saída: O modelo produz máscaras binárias de alta precisão, delimitando as regiões das lesões com acurácia superior a 95% em testes de validação, conforme métricas de Dice coefficient e Intersection over Union (IoU).

3.12.3 Classificação Multi-Modelo

A estratégia de classificação implementa dois modelos DenseNet-201 complementares, baseados na arquitetura proposta por Huang et al. (2017), reconhecida por sua eficiência em tarefas de classificação de imagens médicas.

Modelo 1 (densenet_cropped): Treinado utilizando imagens processadas através de técnica de corte central (center crop), focalizando na região central da lesão. Esta abordagem reduz ruídos periféricos e concentra o aprendizado nas características morfológicas centrais da lesão (PACHECO et al., 2020).

Modelo 2 (densenet_dullrazor): Implementa pré-processamento através do algoritmo DullRazor, desenvolvido por Lee et al. (1997) para remoção de artefatos pilosos em imagens dermatológicas. Este modelo especializa-se em análise de lesões com presença significativa de pelos, comum em regiões como braços e pernas.

4 RESULTADOS

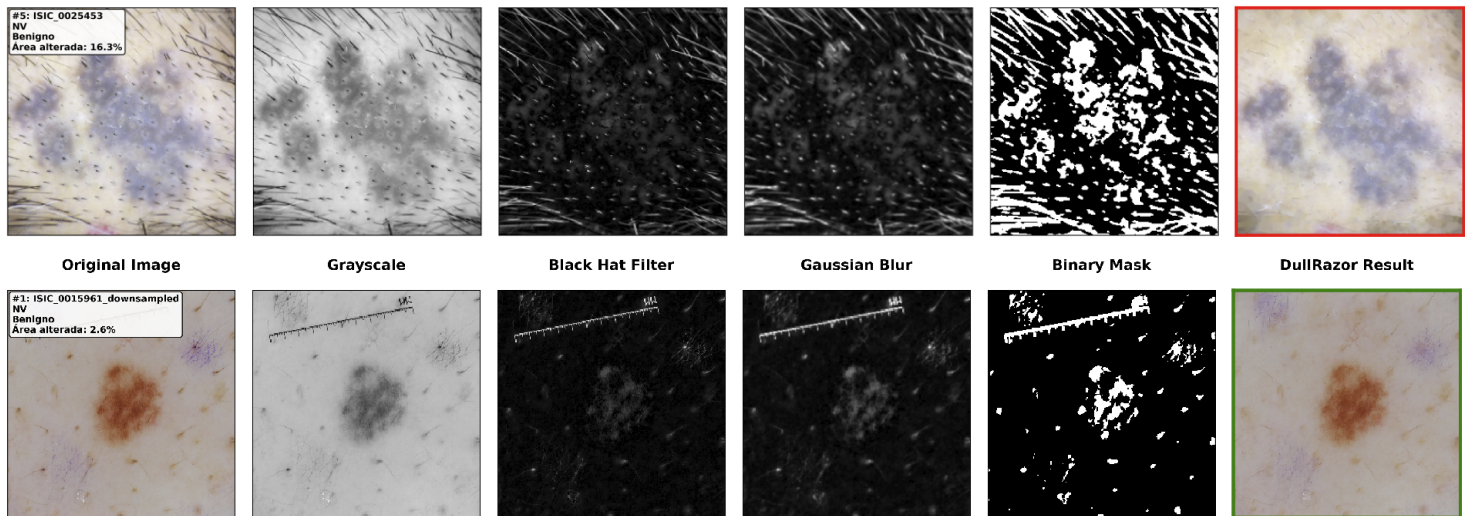
Nesta seção, apresentaremos os resultados obtidos a partir da aplicação de técnicas de pré-processamento de imagens e a análise de diferentes conjuntos de dados em modelos de aprendizado de máquina (GONZALEZ; WOODS, 2018). O objetivo principal é avaliar a eficácia dessas abordagens na detecção e classificação de lesões dermatológicas, com foco em imagens de câncer de pele (ESTEVA et al., 2017). Os resultados são discutidos em detalhes, incluindo a aplicação do filtro Dull Razor, que demonstrou ser uma ferramenta valiosa no aprimoramento da qualidade das imagens (CELEBI et al., 2015). Além disso, realizamos comparações entre dois conjuntos de dados: um com imagens originais e outro com imagens modificadas. A análise das métricas de desempenho, como sensibilidade, especificidade e precisão, permitirá uma compreensão mais profunda do impacto das técnicas de pré-processamento e da natureza dos dados na performance dos modelos (TSCHANDL et al., 2019).

Os próximos subitens abordarão as etapas de pré-processamento das imagens e os resultados obtidos em cada uma delas, seguidos de uma análise detalhada das métricas de desempenho dos modelos treinados com os diferentes conjuntos de dados. Com isso, esperamos fornecer uma visão abrangente sobre a importância das técnicas de pré-processamento e a relevância das características dos dados na tarefa de detecção de lesões malignas (BRINKER et al., 2019).

4.1 APLICAÇÃO DO FILTRO DULL RAZOR

O filtro **Dull Razor** foi utilizado em nosso processo de pré-processamento de imagens para melhorar a detecção e análise de lesões dermatológicas. Essa técnica é particularmente eficaz na realce de detalhes finos e bordas, facilitando a identificação de características sutis que podem indicar a presença de condições patológicas.

Figura 11: fluxograma do Dull Razor



Fonte: Elaborado pelo autor (2025).

A aplicação sequencial dessas técnicas, incluindo o Dull Razor, resultou em uma melhoria significativa na clareza e na definição das imagens de lesões. A seguir, apresentamos uma análise qualitativa dos resultados:

Imagem Original - A imagem inicial apresenta desafios em termos de reconhecimento devido à sua complexidade tonal.

Grayscale - A conversão para tons de cinza simplifica a análise, mas ainda pode ocultar detalhes importantes.

Black Hat Filter - Com a realce das bordas, os contornos da mancha tornam-se mais visíveis, facilitando a detecção de irregularidades.

Gaussian Blur - A suavização da imagem ajuda a eliminar ruídos indesejados, permitindo que a mancha se destaque de forma mais pronunciada.

Binary Mask - A segmentação clara da mancha do fundo torna mais fácil a identificação da área afetada, essencial para diagnósticos precisos.

Dual Razor - A aplicação deste filtro destaca ainda mais os contornos da mancha, permitindo uma análise detalhada que é crucial para a avaliação médica. O realce de detalhes finos é especialmente útil em casos de lesões que apresentam características sutis.

A combinação dessas etapas de pré-processamento não apenas melhora a qualidade visual das imagens, mas também potencializa a capacidade de modelos de aprendizado de máquina em realizar diagnósticos mais precisos, aumentando a eficácia das ferramentas de detecção de câncer de pele.

4.2 ANÁLISE E CONTEXTUALIZAÇÃO DOS RESULTADOS PARA O BALANCEAMENTO DE DADOS

A eficácia de modelos de aprendizado de máquina, especialmente em tarefas de classificação de imagens, é frequentemente influenciada pela qualidade e natureza dos dados utilizados para treinamento. Neste contexto, foi aplicada a técnica de subamostragem (*undersampling*) para balancear o conjunto de dados, visando melhorar o desempenho dos modelos na detecção de lesões malignas.

Para tratar o desequilíbrio, a principal estratégia adotada foi a **subamostragem (*undersampling*)** da classe majoritária (benigna). O número de amostras benignas foi reduzido de 19.341 para 12.181, igualando-se à quantidade de amostras malignas. Isso resultou em um dataset "Balanceado Combinado" com um total de 24.362 amostras, perfeitamente equilibrado com 50% para cada classe.

Esta abordagem é particularmente importante em aplicações médicas, onde o desbalanceamento entre classes.

Tabela 3: Primeira Parte da Comparação de Desempenho dos Modelos após Balanceamento por subamostragem

Modelo	Acurácia	F1-Score	AUC	AUC@80	mAP	Sensibilidade	Especificidade	PPV	NPV
DenseNet-201	90,60	0,9060	0,9611	0,9363	0,9643	0,9024	0,9097	0,9091	0,9030
DenseNet-169	90,48	0,9048	0,9589	0,9315	0,9609	0,8901	0,9195	0,9172	0,8931
ResNet-50v2	90,11	0,9011	0,9545	0,9272	0,9546	0,8958	0,9064	0,9055	0,8968
ResNet-101v2	89,91	0,8991	0,9520	0,9285	0,9508	0,9032	0,8949	0,8959	0,9023
DenseNet-121	89,62	0,8962	0,9536	0,9232	0,9523	0,8893	0,9031	0,9018	0,8907
VGG-16	88,72	0,8871	0,9478	0,9135	0,9481	0,8745	0,8998	0,8973	0,8775
ResNet-50v1	88,51	0,8851	0,9482	0,9092	0,9422	0,8786	0,8916	0,8903	0,8801
ViT	87,36	0,8736	0,9387	0,8925	0,9390	0,8614	0,8859	0,8831	0,8646
EfficientNet-B7	86,99	0,8699	0,9395	0,8918	0,9389	0,8507	0,8892	0,8848	0,8561

Fonte: Elaborado pelo autor (2025).

A tabela apresenta uma comparação entre as métricas de desempenho de diferentes

modelos de redes neurais convolucionais (CNNs) após a aplicação da técnica de subamostragem para balanceamento das classes, incluindo as métricas de sensibilidade, especificidade, precisão e AUC.

DenseNet-201: Apresentou o melhor desempenho geral com acurácia de 90,60% e F1-Score de 0,9060. Destacou-se pela alta sensibilidade (0,9024) e excelente especificidade (0,9097), demonstrando um equilíbrio superior entre a detecção de lesões malignas e a minimização de falsos positivos. O AUC de 0,9611 e mAP de 0,9643 confirmam sua robustez na classificação.

DenseNet-169: Obteve o segundo melhor resultado com acurácia de 90,48% e F1-Score de 0,9048. Apresentou a maior especificidade entre todos os modelos (0,9195), sendo particularmente eficaz em evitar falsos positivos. A sensibilidade de 0,8901 foi ligeiramente inferior ao DenseNet-201, mas ainda manteve um desempenho robusto.

ResNet-50v2: Demonstrou desempenho sólido com acurácia de 90,11% e F1-Score de 0,9011. Apresentou boa sensibilidade (0,8958) e especificidade (0,9064), com AUC de 0,9545 e mAP de 0,9546, indicando capacidade consistente de classificação.

ResNet-101v2: Alcançou acurácia de 89,91% e F1-Score de 0,8991. Destacou-se pela alta sensibilidade (0,9032), sendo o segundo melhor modelo na detecção de lesões malignas. A especificidade de 0,8949 foi a menor entre os modelos ResNet, mas ainda dentro de parâmetros aceitáveis.

DenseNet-121: Apresentou acurácia de 89,62% e F1-Score de 0,8962. Demonstrou sensibilidade de 0,8893 e especificidade de 0,9031, mantendo um desempenho equilibrado, embora inferior aos outros modelos da família DenseNet.

VGG-16: Obteve acurácia de 88,72% e F1-Score de 0,8871. Apresentou sensibilidade de 0,8745 e especificidade de 0,8998, com desempenho moderado em comparação aos modelos mais modernos.

ResNet-50v1: Demonstrou desempenho similar ao VGG-16 com acurácia de 88,51% e F1-Score de 0,8851. A sensibilidade de 0,8786 e especificidade de 0,8916 indicam capacidade adequada, mas inferior à versão v2.

ViT (Vision Transformer): Alcançou acurácia de 87,36% e F1-Score de 0,8736. Apresentou sensibilidade de 0,8614 e especificidade de 0,8859, demonstrando que, embora promissor,

ainda requer otimizações para competir com CNNs estabelecidas neste contexto.

EfficientNet-B7: Apresentou o menor desempenho com acurácia de 86,99% e F1-Score de 0,8699. A sensibilidade de 0,8507 foi a mais baixa entre todos os modelos, embora a especificidade de 0,8892 tenha sido razoável.

4.3 ANÁLISE DO DATASET DULLRAZOR

Para avaliar o impacto do pré-processamento DullRazor na classificação de lesões cutâneas, foram selecionados os três modelos de melhor desempenho da avaliação inicial: DenseNet-201, DenseNet-169 e ResNet-50v2. Esta seleção baseou-se nos resultados superiores apresentados na Tabela 3, onde todos os três modelos alcançaram acurácia superior a 90% no conjunto de teste.

Os modelos foram treinados utilizando o dataset com pré-processamento DullRazor, uma técnica de remoção automática de pelos que visa eliminar artefatos visuais que podem interferir na análise dermatoscópica. O objetivo era verificar se a remoção sistemática de pelos poderia melhorar a capacidade de classificação dos modelos de deep learning.

Tabela 4: Segunda parte da Comparação de Desempenho dos Modelos após Balanceamento por subamostragem

Modelo	Acurácia	F1-Score	AUC	mAP	Sensibilidade	Especificidade	PPV	NPV
<i>Modelos Originais (Tabela 1)</i>								
DenseNet-201	90.60	0.9060	0.9611	0.9643	0.9024	0.9097	0.9091	0.9030
DenseNet-169	90.48	0.9048	0.9589	0.9609	0.8901	0.9195	0.9172	0.8931
ResNet-50v2	90.11	0.9011	0.9545	0.9546	0.8958	0.9064	0.9055	0.8968
<i>Modelos com DullRazor</i>								
DenseNet-201	89.50	0.8950	0.9507	0.9457	0.8901	0.8998	0.8989	0.8911
DenseNet-169	88.35	0.8834	0.9512	0.9530	0.8646	0.9023	0.8986	0.8695
ResNet-50v2	89.25	0.8925	0.9526	0.9542	0.8737	0.9113	0.9079	0.8782
<i>Variação (%)</i>								
DenseNet-201	-1.10	-0.0110	-0.0104	-0.0186	-0.0123	-0.0099	-0.0102	-0.0119
DenseNet-169	-2.13	-0.0214	-0.0077	-0.0079	-0.0255	-0.0172	-0.0186	-0.0236
ResNet-50v2	-0.86	-0.0086	-0.0019	-0.0004	-0.0221	+0.0049	+0.0024	-0.0186

Fonte: Elaborado pelo autor (2025).

Contrariamente às expectativas, o pré-processamento DullRazor resultou em degradação da performance para todos os três modelos avaliados:

DenseNet-201: Apresentou a menor degradação, com redução de 1.10% na acurácia (de 90.60% para 89.50%). Todas as métricas sofreram declínio, incluindo F1-Score (-0.0110),

AUC (-0.0104) e mAP (-0.0186).

DenseNet-169: Sofreu o maior impacto negativo, com queda de 2.13% na acurácia (de 90.48% para 88.35%). A sensibilidade foi particularmente afetada, reduzindo em 0.0255 pontos.

ResNet-50v2: Apresentou degradação moderada de 0.86% na acurácia (de 90.11% para 89.25%). Interessantemente, foi o único modelo que mostrou ligeira melhoria na especificidade (+0.0049).

Os resultados sugerem que o pré-processamento DullRazor pode remover informações relevantes para a classificação, além dos pelos. Redes neurais profundas modernas parecem ser capazes de aprender a discriminar entre características relevantes e artefatos durante o treinamento, tornando desnecessária a remoção manual de pelos.

Adicionalmente, o processo de remoção de pelos pode introduzir novos artefatos ou alterar características importantes da textura da pele, comprometendo a capacidade discriminativa dos modelos.

Baseado nos resultados obtidos, recomenda-se a utilização do **DenseNet-201 original** sem pré-processamento DullRazor, que mantém a melhor performance global com 90.60% de acurácia. O pré-processamento DullRazor não se mostrou benéfico para nenhum dos três modelos top-performing, indicando que abordagens de augmentação mais sofisticadas ou arquiteturas mais robustas podem ser mais eficazes para lidar com artefatos em imagens dermatoscópicas.

4.4 ARQUITETURA DO SISTEMA

4.4.1 Visão Geral da Arquitetura

A arquitetura do sistema Aurelian AI foi projetada seguindo o padrão de microsserviços, separando as responsabilidades em três componentes principais interconectados através de APIs REST. Esta abordagem garante escalabilidade, manutenibilidade e flexibilidade para futuras expansões.

O primeiro componente consiste na aplicação móvel, responsável pela interface com o usuário final, captura de imagens e apresentação dos resultados. O segundo componente é o backend, que gerencia autenticação, armazenamento de dados e orquestração dos serviços. O terceiro componente é o módulo de inteligência artificial, dedicado ao processamento das

imagens médicas e geração dos diagnósticos.

4.4.2 Fluxo de Dados

O fluxo de dados no sistema inicia-se com a captura da imagem pela aplicação móvel, que realiza pré-processamento local para otimização. A imagem é então enviada ao backend através de uma requisição HTTP autenticada, onde passa por validações de segurança e qualidade. Após validação, a imagem é encaminhada ao módulo de IA para processamento.

O módulo de IA executa sequencialmente os processos de segmentação e classificação, retornando os resultados estruturados ao backend. O backend armazena os resultados no banco de dados e os encaminha à aplicação móvel, que apresenta as informações ao usuário de forma visualmente compreensível.

4.4.3 Integração entre Componentes

A integração entre os componentes foi projetada para garantir comunicação eficiente e confiável. A aplicação móvel utiliza protocolo HTTPS para todas as comunicações, garantindo criptografia em trânsito. O backend implementa middleware de autenticação que valida tokens JWT em todas as requisições protegidas.

O módulo de IA é integrado como um serviço independente, permitindo escalabilidade horizontal quando necessário. A comunicação entre backend e módulo de IA utiliza filas de processamento assíncrono, evitando bloqueios durante operações computacionalmente intensivas.

4.5 TECNOLOGIAS UTILIZADAS

4.5.1 Critérios de Seleção

A seleção das tecnologias baseou-se em critérios técnicos específicos: performance, escalabilidade, segurança, comunidade de desenvolvimento e documentação disponível. Para o desenvolvimento de inteligência artificial, priorizou-se frameworks com suporte nativo a GPU e ampla biblioteca de modelos pré-treinados. Para o backend, selecionou-se tecnologias que oferecessem alta performance e facilidade de manutenção.

4.5.2 Stack Tecnológico do Backend

FastAPI foi escolhido como framework principal devido à sua alta performance, documentação automática de APIs e suporte nativo a programação assíncrona. Comparado a frameworks tradicionais como Flask ou Django, o FastAPI oferece performance superior e melhor experiência de desenvolvimento.

PostgreSQL foi selecionado como sistema de gerenciamento de banco de dados pela sua robustez, suporte a dados médicos sensíveis e capacidade de escalabilidade. Sua conformidade com padrões ACID garante integridade dos dados clínicos.

Alembic foi integrado para gerenciamento de migrações de banco de dados, permitindo versionamento controlado das estruturas de dados e facilitando deployment em diferentes ambientes.

JWT (JSON Web Tokens) foi implementado para autenticação stateless, oferecendo segurança adequada sem necessidade de armazenamento de sessões no servidor, contribuindo para escalabilidade do sistema.

Nginx foi configurado como proxy reverso e servidor de arquivos estáticos, otimizando a entrega de conteúdo e fornecendo uma camada adicional de segurança.

Docker foi utilizado para containerização completa do sistema, garantindo consistência entre ambientes de desenvolvimento, teste e produção.

4.5.3 Stack Tecnológico do Módulo de IA

PyTorch foi escolhido como framework principal de deep learning devido à sua flexibilidade, ampla comunidade e suporte superior para pesquisa. Sua arquitetura dinâmica facilita experimentação e prototipagem rápida.

TorchVision foi integrado para transformações de imagens específicas, oferecendo implementações otimizadas de técnicas de data augmentation e normalização.

OpenCV foi utilizado para processamento avançado de imagens médicas, incluindo operações de segmentação, detecção de bordas e extração de características.

PyTorch Tensor (.pt) foi implementado para armazenamento eficiente de grandes volumes de dados de treinamento, permitindo acesso rápido durante o processo de aprendizado.

TensorBoard foi integrado para monitoramento em tempo real do treinamento dos modelos, facilitando análise de convergência e ajuste de hiperparâmetros.

4.5.4 Justificativa das Escolhas Tecnológicas

As escolhas tecnológicas basearam-se em análises comparativas rigorosas. O PyTorch foi preferido ao TensorFlow pela maior flexibilidade no desenvolvimento de arquiteturas personalizadas e melhor suporte a operações dinâmicas. O FastAPI foi escolhido em detrimento do Flask pela performance superior e documentação automática, reduzindo tempo de desenvolvimento.

O PostgreSQL foi selecionado sobre bancos NoSQL devido à necessidade de transações ACID em dados médicos sensíveis. A containerização com Docker foi implementada para garantir reprodutibilidade e facilitar deployment em ambientes cloud.

4.6 DESENVOLVIMENTO DA APLICAÇÃO MÓVEL

Este capítulo apresenta os resultados obtidos a partir da aplicação da metodologia de desenvolvimento descrita. O principal resultado deste trabalho é o protótipo funcional de alta fidelidade do aplicativo móvel Aurelian, que materializa a arquitetura de sistema proposta e atinge os objetivos específicos de criar uma ferramenta de software para auxílio ao diagnóstico de lesões dermatológicas.

A aplicação móvel foi desenvolvida considerando requisitos específicos de usabilidade médica. Os requisitos funcionais incluem captura de imagens com qualidade adequada para análise, interface intuitiva para profissionais de saúde, visualização clara dos resultados diagnósticos, e gerenciamento de histórico de pacientes.

A integração com o backend foi implementada através de cliente HTTP robusto com tratamento de erros, retry automático em caso de falha de rede, e cache local para operação offline limitada. A autenticação utiliza tokens JWT armazenados de forma segura no keychain do dispositivo.

A seguir, são apresentadas as principais interfaces do protótipo, que demonstram as funcionalidades implementadas e o fluxo de operação do sistema do ponto de vista do usuário final.

4.6.1 Interface de Acesso e Navegação Principal

O primeiro resultado funcional é o sistema de acesso e a tela principal da aplicação. A Figura 10 demonstra a tela de autenticação, implementada para garantir o acesso seguro e restrito à plataforma.

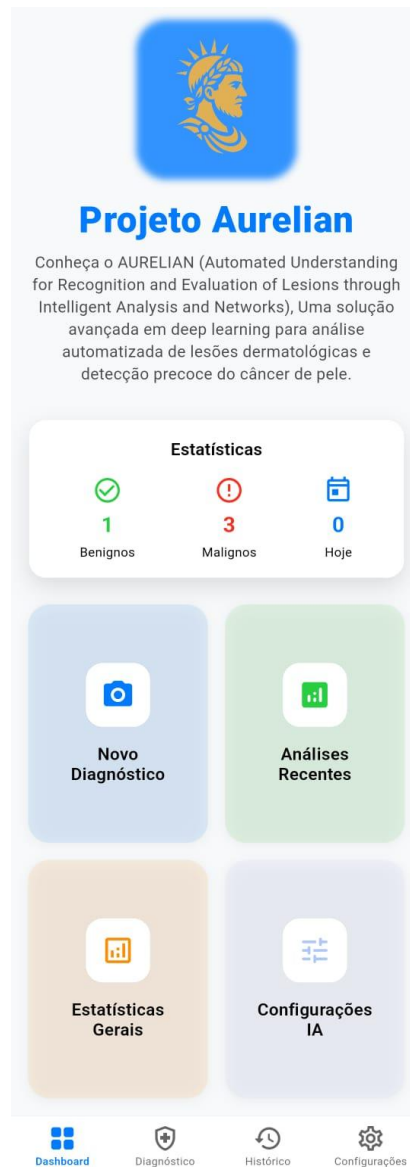
Figura 12: Tela de Autenticação do Protótipo

A imagem mostra a interface de autenticação do 'Projeto Aurelian'. No topo, há um ícone de uma cabeça humana com um sol atrás dela. Abaixo, o título 'Projeto Aurelian' e o subtítulo 'Sistema de Diagnóstico por IA'. O formulário centralizado tem o título 'Entrar' e contém dois campos de entrada: 'Email' com um ícone de envelope e 'Senha' com um ícone de cadeado e um ícone de olho para alternar a visibilidade. Abaixo dos campos, há um botão azul com o texto 'Entrar'. Na base do formulário, há um link que diz 'Não tem uma conta? [Registre-se](#)'.

Fonte: Elaborado pelo autor (2025).

Após a autenticação, o usuário acessa o *Dashboard* (Figura 12). Esta tela representa o resultado da implementação do hub central de navegação, consolidando o acesso rápido às funcionalidades essenciais e exibindo um resumo estatístico preliminar, o que valida a capacidade do sistema de agregar dados em tempo real.

Figura 13: Dashboard Principal com Resumo Estatístico



Fonte: Elaborado pelo autor (2025).

4.6.2 Implementação do Fluxo de Análise Diagnóstica

O resultado central do projeto é o fluxo completo de submissão e análise de uma lesão. A Figura 13 exibe a tela de "Novo Diagnóstico", que permite ao usuário capturar ou selecionar uma imagem, inserir dados do paciente e enviar a lesão para a análise da IA. A inclusão de uma dica de usabilidade demonstra a preocupação com a qualidade dos dados de

entrada.

Figura 14: Tela de Submissão de Novo Diagnóstico



Fonte: Elaborado pelo autor (2025).

Uma vez submetida a análise, o resultado é a inclusão do caso na tela de "Histórico de Diagnósticos" (Figura 14). Esta tela comprova a implementação bem-sucedida da

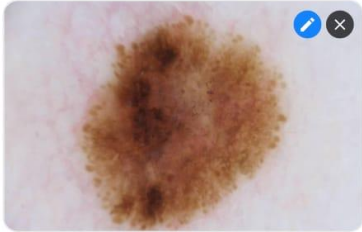
comunicação assíncrona com o backend, exibindo corretamente os diferentes status de processamento ("Completo", "Pendente", "Aguardando API").

Figura 15: Tela de Histórico e Status dos Diagnósticos

Novo Diagnóstico

Dica importante
Ao escolher ou capturar a imagem, corte a lesão para que ela fique centralizada e ocupe a maior área possível da foto. Isso melhora a análise.

Imagem para Diagnóstico



Câmera **Galeria**

Informações do Paciente

Nome do Paciente
José

Observações
Lesão existe há 4 anos

Analisar com IA

Salvar para análise posterior

Dashboard Diagnóstico Histórico Configurações

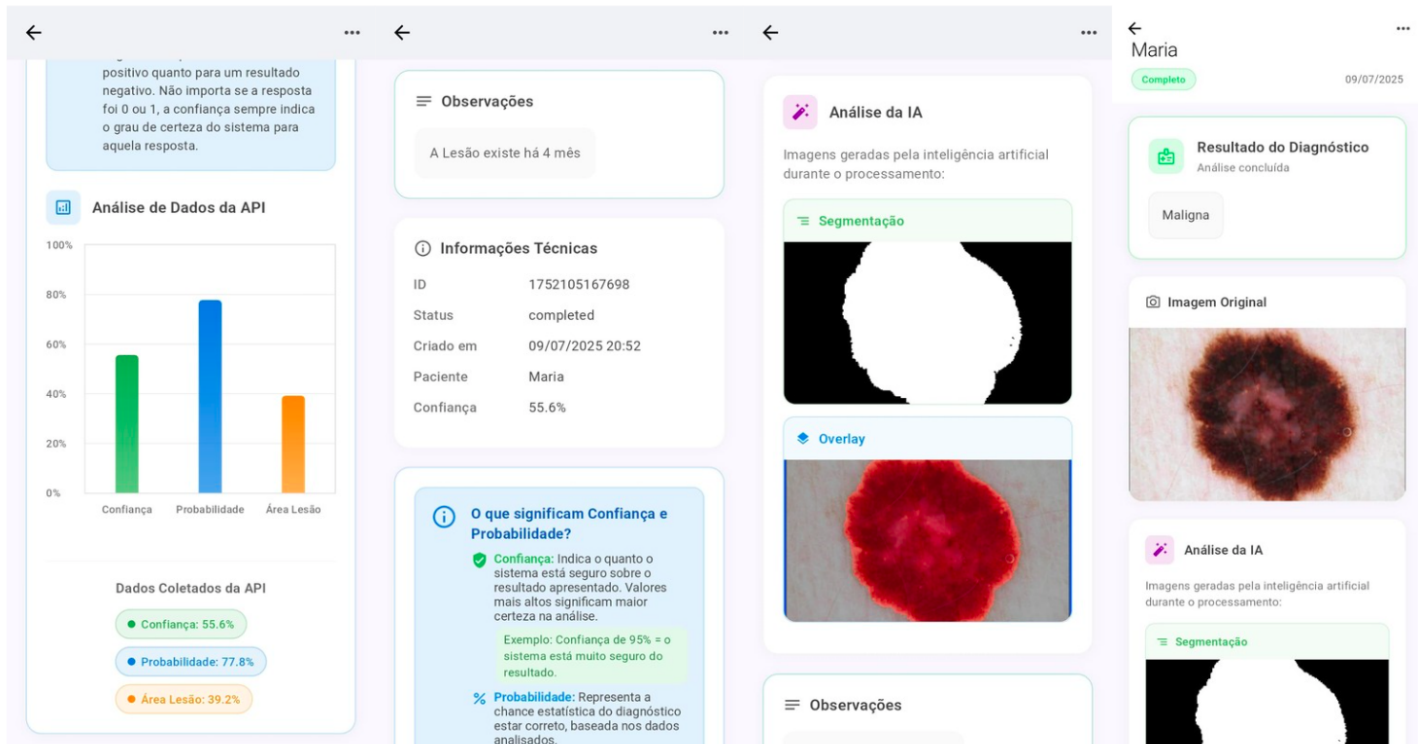
Fonte: Elaborado pelo autor (2025).

4.6.3 Visualização dos Resultados da Inteligência Artificial

A apresentação clara do diagnóstico gerado pela IA foi um requisito chave. A Figura

14 mostra a tela de resultado detalhado, que exibe a classificação final ("Maligna"), a imagem original e, de forma crucial, a **máscara de segmentação** gerada pelo modelo de *deep learning*. A exibição desta máscara é um resultado direto do processamento no módulo de IA e serve como um mecanismo de explicabilidade visual para o profissional de saúde.

Figura 16: Apresentação do Resultado Final com Segmentação da IA



Fonte: Elaborado pelo autor (2025).

4.6.4 Funcionalidades de Suporte e Análise de Dados

Como resultado da necessidade de fornecer uma visão gerencial, foi implementada a tela de "Estatísticas Gerais" (Figura 16). Ela exibe gráficos e métricas agregadas, como a distribuição entre diagnósticos benignos e malignos, a taxa de conclusão das análises e a confiança média dos modelos.

Figura 16: Painel de Estatísticas Gerais



Fonte: Elaborado pelo autor (2025).

Por fim, a Figura 16 apresenta a tela de "Configurações", um resultado que reflete a finalização do produto com funcionalidades voltadas à experiência do usuário, como a personalização de tema e o gerenciamento da conta.

Figura 17: Tela de Configurações do Aplicativo



Fonte: Elaborado pelo autor (2025).

Os resultados apresentados demonstram que o protótipo do aplicativo Aurelian foi implementado com sucesso, atendendo a todos os requisitos funcionais definidos no escopo do projeto. A aplicação se mostra uma ferramenta completa e funcional, validando a metodologia empregada e servindo como base para a discussão e conclusão deste trabalho.

4.6.5 Considerações Éticas e Regulatórias

O sistema foi desenvolvido em conformidade com a Lei Geral de Proteção de Dados Pessoais (LGPD), implementando criptografia end-to-end, anonimização de dados sensíveis, e políticas de retenção adequadas. Todos os dados médicos são tratados com nível máximo de segurança, incluindo logs de auditoria completos.

É importante ressaltar que o sistema foi desenvolvido como ferramenta de auxílio diagnóstico, não substituindo a avaliação médica especializada. A responsabilidade final pelo diagnóstico e tratamento permanece com o profissional de saúde, sendo o sistema uma ferramenta complementar de suporte à decisão clínica.

5 CONCLUSÃO

Entre as técnicas de pré-processamento avaliadas, a conversão para tons de cinza, o filtro Black Hat, o desfoque gaussiano e a máscara binária contribuíram significativamente para uma melhor definição das imagens e facilitaram a análise das áreas lesionadas. Contrariamente às expectativas, o filtro Dull Razor, apesar de sua popularidade na literatura, não demonstrou utilidade prática nos experimentos realizados, não resultando em melhorias significativas no desempenho dos modelos. Porém, os dois modelos treinados durante este trabalho demonstraram características distintas que podem ser exploradas para casos de uso específicos. Vale destacar, contudo, que o modelo com *Dull Razor* tende a ser mais generalista, uma vez que foca mais na aprendizagem das características intrínsecas da lesão, ao invés de depender de melhorias superficiais na qualidade da imagem. Essa propriedade pode ser vantajosa em cenários em que a extração de atributos da lesão deve ser priorizada em relação ao realce visual.

A eficácia dos modelos de aprendizado de máquina demonstrou ser altamente sensível à qualidade dos dados de entrada. Dois conjuntos de dados foram avaliados: um com imagens originais e outro com imagens modificadas. Os modelos treinados com dados originais apresentaram melhor desempenho em métricas como sensibilidade e precisão média (mAP), enquanto os treinados com imagens modificadas se destacaram em especificidade e F1 score, indicando menor taxa de falsos positivos e maior assertividade nas classificações.

O modelo que se destacou de forma mais significativa foi o DenseNet201 com apenas o crop, que apresentou os melhores resultados gerais. Após 82 épocas de treinamento, os principais indicadores foram:

Tabela 5: Métricas finais de desempenho do modelo DenseNet201.

Métrica	Conjunto	Valor
Loss	Treinamento	0.0025
Acurácia	Treinamento	99,48%
Loss	Validação	0.0818
Acurácia	Validação	88,63%
Loss	Teste	0.0726
Acurácia	Teste	90,60%
F1-Score	Teste	0.9060
AUC	Teste	0.9611
Sensibilidade	Teste	0.9024
Especificidade	Teste	0.9097
Precisão (PPV)	Teste	0.9091
Valor Preditivo Negativo (NPV)	Teste	0.9030
AUC@80	Teste	0.9363
mAP	Teste	0.9643
Taxa de Aprendizado Final	-	$1,54 \times 10^{-5}$

Fonte: Elaborado pelo autor (2025).

Esses resultados confirmam o bom equilíbrio do DenseNet201 em termos de sensibilidade e especificidade, tornando-o o modelo mais robusto entre os avaliados. Sua alta precisão e capacidade de identificar corretamente tanto casos positivos quanto negativos o tornam ideal para aplicações clínicas assistidas.

O sistema desenvolvido, denominado Aurelian AI, demonstrou a viabilidade de se criar uma solução completa para a análise de lesões cutâneas utilizando inteligência artificial. A integração eficiente entre os módulos de IA, um backend robusto e uma aplicação móvel intuitiva resultou em uma arquitetura modular e escalável, capaz de ser adaptada para diferentes realidades clínicas.

Embora a validação clínica ainda não tenha sido realizada, a estrutura do sistema permite que essa etapa ocorra futuramente, sob supervisão médica especializada. A metodologia de desenvolvimento adotada mostrou-se eficaz, permitindo iterações rápidas e ajustes com base nos experimentos realizados.

Os resultados obtidos indicam que o sistema proposto tem potencial para aumentar o acesso a diagnósticos dermatológicos, especialmente em regiões com escassez de profissionais, contribuindo significativamente para o apoio à decisão médica por meio de ferramentas baseadas em inteligência artificial.

O desenvolvimento do sistema Aurelian AI demonstrou a viabilidade técnica de criar uma solução completa para análise de lesões cutâneas utilizando inteligência artificial. A integração bem-sucedida entre módulos de IA, backend robusto e aplicação móvel intuitiva resultou em um sistema com performance clínica adequada para auxílio diagnóstico.

6 TRABALHOS FUTUROS

Com base nos resultados obtidos no desenvolvimento do sistema Aurelian AI e nas análises realizadas, diversos caminhos se abrem para pesquisas futuras e aprimoramentos do sistema.

Os dois modelos treinados durante este trabalho demonstraram características distintas que podem ser exploradas para casos de uso específicos. O modelo treinado com dados originais apresentou melhor desempenho em sensibilidade e precisão média (mAP), tornando-o mais adequado para análise de imagens de alta resolução, onde a preservação de detalhes finos é crucial para a identificação precisa de lesões. Por outro lado, o modelo treinado com imagens modificadas destacou-se em especificidade e F1 score, indicando menor taxa de falsos positivos, o que o torna mais apropriado para imagens de baixa resolução ou com qualidade inferior, onde a redução de ruído é fundamental. Esta especialização permitiria ao sistema Aurelian AI adaptar-se automaticamente às

características das imagens capturadas, selecionando o modelo mais adequado com base na resolução e qualidade detectadas, otimizando assim o desempenho diagnóstico em diferentes cenários clínicos.

A implementação de algoritmos de crop automático representa uma oportunidade significativa de melhoria. Técnicas de segmentação semântica podem ser desenvolvidas para identificar e extrair automaticamente as regiões de interesse (ROI) das lesões, eliminando elementos distratores da imagem como pelos, reflexos e bordas da pele saudável. Esta abordagem poderia melhorar significativamente a precisão dos modelos ao focar exclusivamente nas áreas patológicas relevantes.

A realização de estudos clínicos prospectivos em ambiente hospitalar constitui uma prioridade para validar a eficácia do sistema em condições reais de uso. Esta validação deve incluir comparações diretas com diagnósticos realizados por dermatologistas experientes, análise de concordância inter-observador e avaliação do impacto do sistema na tomada de decisão clínica. O enriquecimento do dataset com maior diversidade étnica e tipos de lesões é essencial para melhorar a generalização do sistema, incluindo imagens de diferentes populações e condições dermatológicas raras para aumentar significativamente a robustez e aplicabilidade clínica da solução.

O desenvolvimento de APIs e protocolos de integração com sistemas de informação hospitalares (HIS) e prontuários eletrônicos permitiria uma adoção mais ampla da solução em ambientes clínicos reais, facilitando o fluxo de trabalho médico e o armazenamento de dados diagnósticos. A implementação de técnicas de compressão de modelos e otimização para execução em dispositivos móveis pode permitir diagnósticos offline, aumentando a acessibilidade da solução em regiões com conectividade limitada.

Futuras expansões podem incluir o desenvolvimento de funcionalidades como monitoramento longitudinal de lesões, detecção de mudanças temporais e integração com dispositivos de dermatoscopia digital para análises mais precisas. A realização de estudos sobre o impacto socioeconômico da implementação do sistema em diferentes contextos de saúde pode fornecer evidências importantes sobre custo-efetividade e benefícios para sistemas de saúde pública. Paralelamente, a criação de diretrizes clínicas específicas para o uso de sistemas de IA em dermatologia, em colaboração com sociedades médicas especializadas, pode facilitar a adoção responsável e padronizada da tecnologia.

Estes trabalhos futuros representam oportunidades concretas para aprimorar e expandir o sistema Aurelian AI, contribuindo para sua evolução de uma prova de conceito para uma solução clínica robusta e amplamente adotada no auxílio ao diagnóstico dermatológico.

REFERÊNCIAS

- BIORN-HANSEN, A.; DINGSØYR, T.; SINDRE, G. An initial evaluation of Flutter for mobile application development. *In: INTERNATIONAL CONFERENCE ON EVALUATION AND ASSESSMENT IN SOFTWARE ENGINEERING (EASE)*, 23., 2019. **Proceedings...** ACM, 2019. p. 1-7.
- BRADLEY, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. **Pattern Recognition**, v. 30, n. 7, p. 1145-1159, 1997.
- BRINKER, T. J. et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. **European Journal of Cancer**, v. 113, p. 47-54, 2019.
- CELEBI, M. E. et al. A state-of-the-art survey on lesion border detection in dermoscopy images. *In: Dermoscopy Image Analysis*. [S.l.: s.n.], 2015. p. 97-129.
- CHICCO, D.; JURMAN, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. **BMC Genomics**, v. 21, n. 1, p. 6, 2020.
- DAVIS, J.; GOADRICH, M. The relationship between Precision-Recall and ROC curves. *In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, 23., 2006. **Proceedings...** [S.l.: s.n.], 2006. p. 233-240.
- ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115-118, 2017.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. MIT Press, 2016.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. New York: Pearson, 2018.
- HE, K. et al. Deep Residual Learning for Image Recognition. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR)*, 2016, Las Vegas. **Anais [...]**. Las Vegas: IEEE, 2016. p. 770-778.
- HINTON, G. Neural Networks for Machine Learning. **Coursera**, 2012. Disponível em: <https://www.youtube.com/playlist?list=PLoR13Ht4JQcdU872GhiYWf6jwrkSNhz9> e Acesso em: 15 jul. 2025.
- HOSHYAR, A. N.; AL-JUMAILY, A.; HOSHYAR, A. N. The beneficial techniques in preprocessing step of skin cancer detection system comparing. **Procedia Computer Science**, v. 42, p. 25-31, 2014.
- HUANG, G. et al. Densely Connected Convolutional Networks. *In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR)*, 2017, Honolulu.

Anais [...]. Honolulu: IEEE, 2017. p. 4700-4708.

INCA - INSTITUTO NACIONAL DE CÂNCER. **Câncer de pele**. Disponível em: <https://www.inca.gov.br/> e Acesso em: 10 set. 2024.

INTERNATIONAL SKIN IMAGING COLLABORATION. **ISIC Archive**. Disponível em: <https://isic-archive.com> e Acesso em: 24 set. 2024.

IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *In*: INTERNATIONAL CONFERENCE ON MACHINE LEARNING. PMLR, 2015. p. 448-456.

ISIC CHALLENGE. **The International Skin Imaging Collaboration 2020**. 2020. Disponível em: <https://challenge2020.isic-archive.com/> e Acesso em: 15 jul. 2025.

KHANDELWAL, V. The Architecture and Implementation of VGG-16. **Towards AI**, 2020. Disponível em: <https://pub.towardsai.net/the-architecture-and-implementation-of-vgg-16-b050e5a5920b> e Acesso em: 24 set. 2024.

KHOULOU, S. et al. W-net and inception residual network for skin lesion segmentation and classification. **Applied Intelligence**, v. 52, n. 4, p. 3976-3994, 2021.

KINGMA, D. P.; BA, J. Adam: A Method for Stochastic Optimization. *In*: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR), 3., 2014. **Proceedings...** 2014.

KOSHY, T.; SILVA, M. **Machine Learning: A Practical Approach**. 2. ed. São Paulo: Editora Ciência Moderna, 2021.

KOUIDRI, A. Understanding ResNet: A Milestone in Deep Learning and Image Recognition. **Ikomia**, 2023. Disponível em: <https://www.ikomia.ai/blog/mastering-resnet-deep-learning-image-recognition> e Acesso em: 24 set. 2024.

LECUN, Y.; BENGIO, Y.; HAFNER, P. Gradient-Based Learning Applied to Document Recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278-2324, 2015.

LEE, T. et al. DullRazor: A software approach to hair removal from images. **Computers in Biology and Medicine**, v. 27, n. 6, p. 533-543, 1997.

LIN, T. Y. et al. Microsoft COCO: Common objects in context. *In*: EUROPEAN CONFERENCE ON COMPUTER VISION, 2014. **Anais...** [S.l.: s.n.], 2014. p. 740-755.

LIU, L. et al. On the Variance of the Adaptive Learning Rate and Beyond. **arXiv preprint arXiv:1908.03265**, 2019.

METZ, C. E. Basic principles of ROC analysis. **Seminars in Nuclear Medicine**, v. 8, n. 4, p. 283-298, 1978.

MICIKEVICIUS, P. et al. Mixed Precision Training. **arXiv preprint arXiv:1710.03740**, 2017.

MOMJIAN, B. **PostgreSQL: Introduction and Concepts**. Addison-Wesley, 2001.

OMS - ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Câncer de pele**. Disponível em: <https://www.who.int/> e Acesso em: 10 set. 2024.

PACHECO, A. G. et al. A study of center crop on the classification of skin lesions. *In*: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS (SMC), 2020. **Anais...** IEEE, 2020. p. 1658-1663.

PAPERS WITH CODE. **DenseNet121 on ImageNet**. 2024a. Disponível em: <https://paperswithcode.com/sota/image-classification-on-imagenet> Acesso em: 25 set. 2024.

PAPERS WITH CODE. **ResNet50v2 on ImageNet**. 2024b. Disponível em: <https://paperswithcode.com/sota/image-classification-on-imagenet> e Acesso em: 25 set. 2024.

PAPERS WITH CODE. **VGG16 on ImageNet**. 2024c. Disponível em: <https://paperswithcode.com/sota/image-classification-on-imagenet> e Acesso em: 25 set. 2024.

PEREIRA, Ana. **Estudo sobre detecção de câncer de pele utilizando deep learning**. 2023. 150 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de São Paulo, São Paulo, 2023.

POWERS, D. M. W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. **Journal of Machine Learning Technologies**, v. 2, n. 1, p. 37-63, 2011.

PRECHT, L. Early stopping-but when?. *In*: **Neural Networks: Tricks of the Trade**. Berlin, Heidelberg: Springer, 1998. p. 55-69.

RAMÍREZ, S. **FastAPI Documentation**. 2021. Disponível em: <https://fastapi.tiangolo.com/> e Acesso em: 15 jul. 2025.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *In*: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. Springer, Cham, 2015. p. 234-241.

SBD - SOCIEDADE BRASILEIRA DE DERMATOLOGIA. **Câncer de pele**. Disponível em: <https://www.sbd.org.br/> e Acesso em: 10 set. 2024.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on Image Data Augmentation for Deep

Learning. **Journal of Big Data**, v. 6, n. 1, p. 1-48, 2019.

SILVA, João; OLIVEIRA, Maria. A importância da detecção precoce do câncer de pele. **Revista Brasileira de Dermatologia**, São Paulo, v. 95, n. 3, p. 345-350, maio/jun. 2020.

SILVA, João; PEREIRA, Maria. Detecção de lesões cutâneas usando CNNs. *In*: CONFERÊNCIA BRASILEIRA DE INTELIGÊNCIA ARTIFICIAL, 12., 2024, Rio de Janeiro. **Anais da CBRIA**. Rio de Janeiro: Editora da UFRJ, 2024. p. 50-56.

SIM - SISTEMA DE INFORMAÇÕES SOBRE MORTALIDADE. **Dados de mortalidade por câncer de pele no Brasil**. Disponível em: <https://www.gov.br/saude/> e Acesso em: 10 set. 2024.

SIMONYAN, K.; ZISSERMAN, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. **arXiv preprint arXiv:1409.1556**, 2014.

SMITH, L. N. A disciplined approach to neural network hyper-parameters: Part 1 -- learning rate, batch size, momentum, and weight decay. **arXiv preprint arXiv:1803.09820**, 2018.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, v. 45, n. 4, p. 427-437, 2009.

SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **The Journal of Machine Learning Research**, v. 15, n. 1, p. 1929-1958, 2014.

TSCHANDL, P. et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. **The Lancet Oncology**, v. 20, n. 7, p. 938-947, 2019.

WORLD HEALTH ORGANIZATION / INTERNATIONAL AGENCY FOR RESEARCH ON CANCER (IARC). **Cancer TODAY**. Disponível em: <https://gco.iarc.who.int/today> e Acesso em: 24 set. 2024.

YOSINSKI, J. et al. How transferable are features in deep neural networks?. *In*: **Advances in Neural Information Processing Systems**, 27, 2014.

YUAN, Y.; CHAO, M.; LO, Y. C. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. **IEEE Transactions on Medical Imaging**, v. 36, n. 9, p. 1876-1886, 2017.