



UNIVERSIDADE FEDERAL DO MARANHÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
Doutorado em Ciência da Computação Associação
UFMA/UFPI

Alan Rafael Ferreira dos Santos

**A Influência da Iluminação e Cor na Precisão de Detecção
de Câncer de Pele com Aprendizado Profundo**

Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires

Teresina - PI
junho, 2024

Alan Rafael Ferreira dos Santos

**A Influência da Iluminação e Cor na Precisão de Detecção
de Câncer de Pele com Aprendizado Profundo**

TESE DE DOUTORADO

Tese apresentada como requisito para
obtenção do título de Doutor em Ciência da
Computação, ao Doutorado em Ciência da
Computação, Associação UFMA/UFPI.

Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires

Teresina - PI
junho, 2024

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S237i Santos, Alan Rafael Ferreira dos

A influência da iluminação e cor na precisão de detecção de câncer de pele com aprendizado profundo. [recurso eletrônico] / Alan Rafael Ferreira dos Santos - 2024.

91 f.

1 Arquivo em PDF

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB

Aberto a pesquisadores, com restrições da Biblioteca

Trabalho de Conclusão de Curso (Tese) – Universidade Federal do Maranhão, Universidade Federal do Piauí, Doutorado da Ciência da Computação – Associação UFMA/UFPI, Teresina, 2024.

“Orientador : Prof. Dr. Kelson Rômulo Teixeira Aires”

1. Câncer de pele - classificação. 2. Lesões - cor. 3. Lesões - iluminação. 4. Aprendizado profundo. 5. Lesões malignas - detecção. II. Título.

CDD 616.994

Alan Rafael Ferreira dos Santos

A Influência da Iluminação e Cor na Precisão de Detecção de Câncer de Pele com Aprendizado Profundo

Tese apresentada como requisito para obtenção do título de Doutor em Ciência da Computação, ao Doutorado em Ciência da Computação, Associação UFMA/UFPI.

Aprovada pela banca examinadora:

Prof. Dr. Kelson Rômulo Teixeira Aires
Orientador
Universidade Federal do Piauí - UFPI

Prof. Dr. Rodrigo de Melo Souza Veras
Universidade Federal do Piauí - UFPI

Prof. Dr. Anselmo Cardoso de Paiva
Universidade Federal do Maranhão - UFMA

Prof. Dr. Leandro Augusto Frata Fernandes
Universidade Federal Fluminense - UFF

Prof. Dra. Andrea Gomes Campos Bianchi
Universidade Federal de Ouro Preto - UFOP

Certificamos que esta é a versão original e final da Tese de Doutorado que foi julgada adequada para obtenção do título de Doutor em Ciência da Computação.

Prof. Dr. Kelson Rômulo Teixeira Aires
Orientador

Prof. Dr. Rodrigo de Melo Souza Veras
Coordenador

Teresina - PI, 07 de junho de 2024

Dedico este trabalho a todos que amam tecnologia e que sempre buscam o inovador.

Agradecimentos

Agradeço a Deus por me proporcionar oportunidades de superar limitações e vencer os obstáculos da vida!

Aos meus pais, Antônio Leal dos Santos e Maria Irene Ferreira dos Santos, que sempre me ofereceram apoio incondicional em todas as etapas da minha vida, e ao meu irmão, Antônio Carlos Ferreira dos Santos, pelo apoio constante durante o período de mestrado e doutorado.

Não poderia deixar de agradecer à minha companheira, Maria das Graças Oliveira Araújo, que esteve sempre ao meu lado nos últimos doze anos, me dando todo o suporte emocional e apoiando minhas decisões.

Ao meu orientador, professor Kelson Rômulo Teixeira Aires, pelas orientações em todas as etapas de desenvolvimento deste trabalho, além da amizade e paciência durante os longos anos de pesquisa.

Ao CNPq, à UFMA e à UFPI, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos meus amigos e colegas que prestaram apoio em todos os momentos de dificuldade. Não citarei nomes, pois o espaço é curto e não quero arriscar esquecer alguém.

Aos professores membros da banca examinadora, que foram vitais para o ajuste e melhoria deste trabalho.

A todos os professores e funcionários do DCCMAPI pelos ensinamentos e pela colaboração.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

"A curiosidade é a luxúria da mente."(Thomas Hobbes)

Resumo

O diagnóstico precoce do câncer de pele maligno é essencial para evitar a metástase e definir um tratamento adequado. Nesse sentido, muitas pesquisas já propuseram abordagens automatizadas de aprendizado profundo para auxiliar na análise visual de imagens digitais de lesões de pele em busca de um diagnóstico preciso. Os aspectos de cor das lesões são essenciais para determinar a suspeita de malignidade. Em imagens digitais, a cor pode variar conforme as mudanças na intensidade e nos aspectos de iluminação. Nessa perspectiva, esta tese investiga a relevância das características de iluminação e cor na precisão da classificação do câncer de pele utilizando modelos de aprendizado profundo. O estudo analisa como variações nessas características podem influenciar a detecção de lesões malignas, destacando a importância da correção de iluminação e cor para melhorar a precisão em sistemas de detecção. A metodologia deste trabalho desenvolveu uma abordagem de validação que combina técnicas convencionais de correção, como *Contrast Limited Adaptive Histogram Equalization* (CLAHE), *Perfect Reflector* e *Shades of Gray* (SoG), com técnicas avançadas baseadas em aprendizado profundo, como *Local Color Distribution Prior Network* (LCDPNet), *Low-Light Neural Radiance Field* (LLNeRF) e *Deep Symmetric Network* (DSN). O processo de delimitação das regiões das lesões utiliza a técnica personalizada DeepLabV3+. Os modelos de aprendizado profundo analisados incluem Inception-V3, Xception, ResNet-50 e DenseNet-121, treinados utilizando transferência de aprendizado e aumento de dados com a técnica *Test-Time Augmentation* (TTA). Os experimentos mostraram que as técnicas de correção de iluminação e cor com aprendizado profundo superaram os métodos convencionais, melhorando a precisão na classificação de lesões malignas com os modelos avaliados. As arquiteturas de redes neurais profundas apresentaram sensibilidade variável às correções de iluminação e cor, sendo mais sensíveis aquelas com menor número de camadas e tamanhos menores de filtros. A correção das características de iluminação e cor produziu aumentos em todas as taxas de precisão nos modelos experimentados, embora alguns resultados não tenham sido expressivos. A correção dessas características é indispensável na classificação do câncer de pele com aprendizado profundo. A escolha adequada de técnicas analíticas pode aumentar a precisão na identificação de malignidades de pele. Além disso, o design e a estrutura da arquitetura do modelo influenciam a percepção da iluminação e da cor, tornando os modelos mais sensíveis às variações dessas características. Futuros estudos devem explorar novos métodos de correção com aprendizado auto-supervisionado e mecanismos de atenção para aprimorar a detecção de câncer de pele.

Palavras-chave: Câncer de Pele, Correção de Iluminação e Cor, Aprendizado Profundo, Redes Neurais Profundas.

Abstract

Early diagnosis of malignant skin cancer is essential to prevent metastasis and determine appropriate treatment. In this context, many studies have proposed automated deep learning approaches to assist in visualizing digital images of skin lesions in search of accurate diagnosis. The color aspects of lesions are crucial for determining the suspicion of malignancy. In digital images, color can vary according to changes in intensity and lighting aspects. From this perspective, this thesis investigates the relevance of lighting and color characteristics in the accuracy of skin cancer classification using deep learning models. The study analyzes how variations in these characteristics can influence the detection of malignant lesions, highlighting the importance of lighting and color correction to improve accuracy in detection systems. The methodology of this work developed a validation approach that combines conventional correction techniques, such as Contrast Limited Adaptive Histogram Equalization (CLAHE), Perfect Reflector, and Shades of Gray (SoG), with advanced deep learning-based methods, such as Local Color Distribution Prior Network (LCDPNet), Low-Light Neural Radiance Field (LLNeRF), and Deep Symmetric Network (DSN). The process of delimiting the lesion regions uses the customized DeepLabV3+ technique. The deep learning models analyzed included Inception-V3, Xception, ResNet-50, and DenseNet-121, and they were trained using transfer learning and data augmentation with Test-Time Augmentation (TTA) technique. Experiments showed that deep learning-based lighting and color correction techniques outperform conventional methods, improving accuracy in classifying malignant lesions with the evaluated models. The deep neural network architectures showed variable sensitivity to lighting and color corrections, with those having fewer layers and smaller filter sizes being more sensitive. Lighting and color characteristic correction increased all accuracy rates in the experimented models, although some results were insignificant. Correcting these characteristics is indispensable in classifying skin cancer with deep learning. The appropriate choice of analytical techniques can increase the accuracy of identifying skin malignancies. Additionally, the design and structure of the model architecture influence the perception of lighting and color, making the models more sensitive to variations in these characteristics. Future studies should explore new correction methods with self-supervised learning and attention mechanisms to enhance skin cancer detection.

Keywords: Skin Cancer, Lighting and Color Correction, Deep Learning, Deep Neural Networks.

Lista de ilustrações

Figura 1 – Variação de iluminação e cor em imagens de lesões de pele. Adaptado de (International Skin Imaging Collaboration 2020)	16
Figura 2 – Tipos de Câncer de Pele. Adaptado de (Tschandl et al. 2017).	35
Figura 3 – Arquitetura Convencional de Redes Neurais Convolucionais. Adaptado de (Zhou, Greenspan e Shen 2017)	37
Figura 4 – Arquitetura Redes Neurais Convolucionais Residuais (ResNet). Adaptado de (Hasan et al. 2021).	38
Figura 5 – Funcionamento das Redes Geradoras Adversárias. Adaptado de (Kelleher 2019).	39
Figura 6 – <i>Autoencoders</i> . Adaptado de (Kelleher 2019).	40
Figura 7 – Segmentação semântica. Adaptado de (Hong et al. 2023). (a) Imagem de Entrada. (b) Máscara do Especialista, (c) Máscara de Segmentação Semântica.	42
Figura 8 – Variação de tonalidade de cor conforme a iluminação.	43
Figura 9 – Correção de iluminação com histograma.	45
Figura 10 – Correção de branco com ajuste nos canais RGB.	46
Figura 11 – Correção de iluminação e cor com princípios de LCDs e aprendizado profundo. Adaptado de (Yang et al. 2021).	47
Figura 12 – Renderização com NeRF para recriação de imagens com aspecto de cor variável. Adaptado de (Martin-Brualla et al. 2021).	48
Figura 13 – Correção dos aspectos de iluminação e cor com princípios de DSN. Adaptado de (Zhao et al. 2021).	49
Figura 14 – Abordagem de validação para otimização de detecção do câncer de pele.	52
Figura 15 – Remoção de pelos com o método sugerido.	54
Figura 16 – Aplicação do método CLAHE para correção de iluminação e cor.	55
Figura 17 – Aplicação do método SoG para correção de iluminação e cor.	56
Figura 18 – Aplicação do método <i>Perfect Reflector</i> para correção de iluminação e cor.	57
Figura 19 – Aplicação do método LCDPNet para correção de iluminação e cor.	58
Figura 20 – Aplicação do método LLNeRF para correção de iluminação e cor.	59
Figura 21 – Aplicação do método DSN para correção de iluminação e cor.	60
Figura 22 – Definição da região de interesse.	61

Lista de abreviaturas e siglas

Acc	Acurácia
AUC	<i>Area Under Curve</i>
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i>
CNN	<i>Convolutional Neural Network</i>
DDI	<i>Diverse Dermatology Images</i>
FID	<i>Fréchet Inception Distance</i>
FN	Falsos Negativos
FP	Falsos Positivos
FrCN	<i>Full-resolution Convolutional Network</i>
GAN	<i>Generative Adversarial Network</i>
IA	<i>Artificial Intelligence</i>
IS	<i>Inception Score</i>
ISIC	<i>International Skin Imaging Collaboration</i>
LCDPNet	<i>Local Color Distribution Preserving Network</i>
LLNeRF	<i>Local Learning Neural Radiance Fields</i>
PAD-UFES	Programa de Assistência Dermatológica e Cirúrgica da Universidade Federal do Espírito Santo
P	Precisão
PSA-GAN	<i>Progressive Self-Attention Generative Adversarial Network</i>
R	<i>Recall</i>
ROC	<i>Receiver Operating Characteristic</i>
SoG	<i>Shades of Gray</i>
SVM	<i>Support Vector Machine</i>
SGD	<i>Stochastic Gradient Descent</i>

TVP	Taxa de Verdadeiro Positivo
TFP	Taxa de Falso Positivo
VP	Verdadeiros Positivos

Sumário

1	INTRODUÇÃO	14
1.1	Contexto e Motivação	15
1.2	Objetivos	16
1.3	Hipóteses	17
1.4	Questões de Pesquisa	18
1.5	Contribuições	18
1.6	Estrutura da Tese	19
2	REVISÃO DE LITERATURA	21
2.1	Metodologia de Revisão	21
2.2	Detecção do Câncer de Pele com Aprendizado Profundo	22
2.2.1	Aumento de Dados na Detecção do Câncer de Pele	28
2.3	Iluminação e Cor na Detecção do Câncer de Pele	30
2.4	Considerações Finais	33
3	FUNDAMENTAÇÃO TEÓRICA	34
3.1	Tipos de Câncer de Pele	34
3.2	Aprendizado Profundo	36
3.2.1	Redes Neurais Convolucionais	36
3.2.2	Redes Neurais Convolucionais Residuais	37
3.2.3	Redes Geradoras Adversárias	38
3.2.4	<i>Autoencoders</i>	40
3.2.5	<i>Aumento de Dados</i>	40
3.2.6	Transferência de Aprendizado e o Ajuste Fino	41
3.3	Segmentação Semântica	42
3.4	Iluminação e Cor em Imagens Digitais	43
3.4.1	Técnicas Baseada em Histogramas	44
3.4.2	Técnicas Baseada em Balanço de Branco	45
3.4.3	<i>Local Color Distributions</i>	46
3.4.4	<i>Neural Radiance Fields</i>	47
3.4.5	<i>Deep Symmetric Network</i>	49
3.5	Considerações Finais	50
4	ABORDAGEM DE CLASSIFICAÇÃO DO CÂNCER DE PELE	51
4.1	Metodologia e Avaliação do Estudo	51
4.2	Bases de Imagens	52

4.3	Remoção de Pelos	53
4.4	Técnicas de Correção de Iluminação e Cor	54
4.4.1	<i>Contrast Limited Adaptive Histogram Equalization</i>	54
4.4.2	<i>Shades of Gray</i>	56
4.4.3	<i>Perfect Reflector</i>	57
4.4.4	<i>Local Color Distribution Prior Network</i>	58
4.4.5	<i>Low-Light Neural Radiance Field</i>	59
4.4.6	<i>Deep Symmetric Network</i>	60
4.5	Segmentação e Delimitação da Área de Interesse	60
4.6	Test-Time Augmentation	61
4.7	Modelos de Aprendizado Profundo	62
4.8	Metodologia Experimental e Métricas de Avaliação	64
4.9	Considerações Finais	67
5	RESULTADOS E DISCUSSÕES	68
5.1	Especificações de Hiperparâmetros	68
5.2	Análise e Validação de Resultados	69
5.2.1	Resultados com Aumento de Dados no Conjunto de Teste	70
5.2.2	Resultados sem Aumento de Dados no Conjunto de Teste	73
5.2.3	Comparações com Trabalhos Semelhantes	75
5.3	Limitações do Estudo e da Sistemática de Classificação	76
5.4	Considerações Finais	78
6	CONCLUSÕES E TRABALHOS FUTUROS	79
	REFERÊNCIAS	82

1 Introdução

O câncer é um problema de saúde global, sendo uma das principais causas de morte e redução da expectativa de vida da sociedade (Cortez, Vasquez e Wei 2021). Nas últimas décadas, houve um aumento significativo na incidência de câncer, com uma previsão de 25 milhões de novos casos até 2030 em todo o mundo (Sung et al. 2021). Nos países com baixo Índice de Desenvolvimento Humano (IDH), verifica-se um aumento no número de casos de todos os tipos de câncer devido à ineficácia do diagnóstico e tratamento (Ferlay et al. 2021).

O câncer de pele tem maior taxa de prevalência em países de clima tropical. Embora essa doença seja causada por múltiplos fatores, a exposição excessiva ao sol e as predisposições genéticas são mais relevantes (Esteva et al. 2017). Em comparação com outros tipos de câncer, o câncer de pele apresenta um menor fator de incidência, porém a sua ocorrência desencadeia preocupações devido ao diagnóstico tardio e o tratamento inadequado (Marques et al. 2016).

Os tipos de câncer de pele que podem evoluir para situações malignas são os carcinomas basocelular, espinocelular e o melanoma. Entre todos esses tipos, o melanoma é o mais agressivo e apresenta o maior índice de mortalidade devido à sua alta capacidade de metástase (Bonalumi, Campos e Leal 2017). No Brasil, estima-se que até 2025 sejam registrados 704.080 novos casos dessa doença no país (Santos et al. 2023).

A detecção precoce do câncer de pele e o acompanhamento são fundamentais para encaminhar os pacientes ao tratamento eficaz e possibilitar a cura (Heibel, Hoey e Cockerell 2020). No entanto, o desafio reside no diagnóstico precoce, pois o paciente acometido tarda a buscar avaliação médica (Alchorne e Abreu 2008). Em muitos casos, as lesões malignas podem ser confundidas com lesões benignas. A metástase é o estágio mais avançado da doença, sendo caracterizada pela disseminação de células cancerígenas na corrente sanguínea, que podem se desenvolver em outros órgãos do corpo (Bonalumi, Campos e Leal 2017).

Um fator a ser observado no diagnóstico precoce é a característica da cor no interior da lesão. Isso quer dizer que lesões que possuem tonalidades de cor variável indicam anomalias na produção das células da pele, sendo indicativo de possível malignidade (Barata, Celebi e Marques 2014). Geralmente, a observação de tons de marrom e preto ou até mesmo cores incomuns é sinal de alerta para o dermatologista, que pode sugerir outras etapas de diagnóstico (Galdran et al. 2017).

As tecnologias exercem um papel importante na automatização do diagnóstico de doenças, permitindo uma identificação mais rápida e precisa (Chang e Guo-Ping 2019).

No caso do câncer de pele, a Inteligência Artificial (IA) impulsiona o desenvolvimento de sistemáticas automatizadas de acompanhamento e diagnóstico, com base em imagens digitais disponibilizadas gratuitamente (Esteva et al. 2021, Kawahara, BenTaieb e Hamarneh 2016). Essas abordagens empregam técnicas e algoritmos avançados que se baseiam em características relevantes extraídas de imagens de referência, visando melhorar a detecção do câncer maligno (Brunsgaard, Jensen e Grossman 2022).

1.1 Contexto e Motivação

O câncer de pele costuma ocorrer com maior frequência em pessoas de tons de pele claros, porém isso não é regra, visto que pessoas de tons escuros também podem ser acometidas pela doença (Ferlay et al. 2021). A primeira etapa do diagnóstico é realizada por meio de uma inspeção visual (Brunsgaard, Jensen e Grossman 2022). Em alguns casos, o médico dermatologista pode coletar imagens digitais ampliadas da lesão para análise posterior. O diagnóstico definitivo da doença somente é dado por meio de exames imuno-histoquímicos de materiais coletados em biópsia (Marques et al. 2016).

O avanço da IA, associado ao desenvolvimento de técnicas de aprendizado profundo (*Deep Learning*), trouxe mudanças significativas na criação de sistemas computadorizados de diagnóstico de doenças por imagens digitais médicas (Dorrell e Strowd 2019). Nos últimos cinco anos, muitas pesquisas demonstram a eficácia de sistemáticas que utilizam redes neurais profundas na identificação correta de vários tipos de câncer, inclusive o câncer de pele (Esteva et al. 2017). Além disso, pesquisas mais recentes reafirmam o potencial dessas redes na revolução do diagnóstico com maior eficácia e eficiência (Esteva et al. 2021).

A cor ainda é um atributo relevante para o diagnóstico preciso do câncer de pele maligno. No entanto, com o uso de aprendizado profundo, essa característica tem sido pouco explorada em novas abordagens com essa sistemática (Adegun e Viriri 2021). Além disso, o pré-processamento também é importante para ajustar as nuances de iluminação e cor em imagens digitais, melhorando a consistência e representação das imagens (Salvi et al. 2022). Um entendimento mais aprofundado da cor em imagens de lesões de pele poderia possibilitar um aumento significativo na precisão de identificação de lesões malignas (Barata e Marques 2019).

O treino de redes neurais profundas com bases de imagens pré-diagnosticadas permite a criação de modelos capazes de categorizar imagens não conhecidas (Barata, Celebi e Marques 2018). A qualidade dessas bases de imagens ainda é questionável, pois não há um processo de captura padrão (Salvi et al. 2022). A ausência de padronização nas condições de iluminação geram imagens com nuances significativas na representação das tonalidades de cor (Branciforti et al. 2023). A Figura 1 ilustra exemplos que destacam as diferenças de iluminação e cor para imagens lesões de pele.



Figura 1 – Variação de iluminação e cor em imagens de lesões de pele. Adaptado de ([International Skin Imaging Collaboration 2020](#))

Diante da complexidade e desafios descritos no diagnóstico do câncer de pele, especialmente em relação às características de iluminação e cor, se torna relevante a investigação da importância dessas características na precisão do diagnóstico do câncer de pele maligno, principalmente utilizando aprendizado profundo. Novos estudos devem ter foco em técnicas que permitam a correção da iluminação e da cor, seguindo sistemáticas diferentes, independente das condições de captura das imagens. Além disso, também é necessário investigar o quanto a correção das características de iluminação e cor podem influenciar no aprendizado e classificação de lesões de pele malignas.

1.2 Objetivos

O objetivo geral desta tese é apresentar um estudo e uma abordagem de validação para detecção do câncer de pele, com foco na identificação da relevância dos aspectos de iluminação e cor na precisão de classificação de malignidades em imagens digitais. O intuito é analisar o quanto as características de iluminação e cor podem influenciar na identificação correta do câncer de pele com o aprendizado profundo.

Para alcançar tal objetivo geral, foram propostos alguns objetivos específicos:

- Realizar uma revisão de literatura sobre a problemática do câncer de pele em imagens digitais, focando em aspectos negativos associados à iluminação e cor;
- Caracterizar as principais sistemáticas de identificação de lesões malignas que utilizam técnicas de aprendizado profundo;
- Selecionar e analisar métodos de correção de aspectos de iluminação e cor que possam ser utilizados em imagens de lesões de pele, explorando técnicas simples e avançadas no problema;
- Desenvolver estruturas para verificar, revisar e classificar lesões de pele em imagens digitais, corrigindo os principais problemas identificados na revisão bibliográfica;

- Avaliar a eficácia da estrutura classificação desenvolvida após o uso de técnicas de correção de iluminação e cor, utilizando as melhores métricas e comparando os resultados com o estado da arte;
- Elaborar uma análise objetiva de como diferentes modelos de aprendizado profundo respondem às correções de iluminação e cor, identificando quais modelos são mais adequados para lidar com essas variações;
- Validar a eficácia do estudo e da estrutura de classificação desenvolvida em diferentes contextos, utilizando bases de imagens públicas.

A essência desta tese está na investigação das limitações e desafios encontrados em abordagens atuais que lidam com as variações de iluminação e cor na detecção precisa do câncer de pele maligno. As questões relacionadas à constância da cor e aos problemas de sub-iluminação ou sobre-iluminação são abordadas como elementos principais, sendo analisadas na classificação correta de lesões com arquiteturas profundas.

1.3 Hipóteses

As hipóteses que direcionaram os objetivos descritos neste trabalho são:

- A cor é uma característica relevante no diagnóstico de lesões de pele malignas. O ajuste adequado deste aspecto em imagens digitais, mantendo a constância e a fidelidade com a aparência real, pode melhorar a precisão das sistemáticas de detecção do câncer de pele. A correção da iluminação resulta em uma melhor calibração de cores, tornando os métodos de classificação mais robustos e confiáveis.
- O processo de captura de imagens de lesões de pele em bases públicas insere uma discrepância na qualidade dos dados. A utilização de dispositivos diferentes e condições diversas de iluminação impactam negativamente na representação fidedigna das cores das lesões. A escolha da técnica de correção adequada pode mitigar as divergências, melhorando os resultados de classificação em sistemáticas de diagnóstico do câncer de pele maligno.
- A qualidade de uma imagem digital está intimamente ligada aos aspectos de iluminação e cor. As arquiteturas de aprendizado profundo podem obter um melhor desempenho de treinamento e classificação se treinadas com imagens de lesões de pele de alta qualidade. Isso proporciona um aprendizado de características invariantes, contribuindo para a classificação correta de lesões malignas.

- O uso de uma combinação adequada de técnicas que lidam com imagens do câncer de pele, aplicadas em uma estrutura de classificação automática, pode melhorar a identificação de malignidades, priorizando principalmente os aspectos de iluminação e cor. Uma abordagem que trata corretamente imagens capturadas de diversos dispositivos, sob condições de iluminações diferentes, pode gerar uma classificação sensível às variações intrínsecas dessas características.

As hipóteses descritas impulsionaram a tese deste trabalho, servindo como guias para a pesquisa e a combinação de técnicas aplicadas no desenvolvimento da abordagem de validação. Os artefatos implementados visaram superar desafios existentes e fornecer novos contextos e informações úteis no diagnóstico automatizado do câncer de pele maligno por imagens.

1.4 Questões de Pesquisa

A tese deste trabalho também tratou os seguintes questionamentos:

- Quais métodos de correção de imagens digitais podem ser aplicados para reduzir problemas causados por variações de iluminação e cor na captura de imagens de lesões de pele?
- As técnicas de correção de iluminação e cor que utilizam aprendizado profundo são melhores que técnicas convencionais?
- Qual é a relação entre as características de iluminação e cor e a detecção precisa do câncer de pele maligno usando modelos de aprendizado profundo para classificação?
- Em que medida um modelo de aprendizado pode melhorar a detecção de lesões malignas de pele, considerando as correções nas variações de iluminação e cor presentes em imagens capturadas por diferentes dispositivos e em diversas condições?

1.5 Contribuições

A tese deste trabalho investigou a detecção do câncer de pele em imagens digitais aplicando técnicas de aprendizado profundo na maior parte do processo. A partir do objetivo geral e dos objetivos específicos, foram obtidos as seguintes contribuições:

- O desenvolvimento de uma abordagem sinérgica que combina técnicas na identificação clínica de câncer de pele maligno, priorizando a iluminação e a cor.

Essa abordagem considera as variações na captura e no tratamento de imagens de lesões de pele em bases de dados públicas;

- A adaptação, aplicação e otimização de técnicas de aprendizado profundo nunca usadas para lidar com o viés de iluminação e cor, verificando e comparando o desempenho dessas técnicas outras convencionais, amplamente utilizadas no problema. A investigação não se limitou a novos métodos, mas aos potenciais de melhoria de resultados de classificação.
- A implementação de uma abordagem de validação que trata a relação entre a classificação por modelos de aprendizado profundo e as características de cada arquitetura na sensibilidade às variações de iluminação e cor. As correções por técnicas variadas serviram como base para avaliar as características individuais de cada modelo na precisão de classificação de lesões malignas.
- A produção de um comparativo confiável, verificando estratégias de tratamento de iluminação e cor e avaliando o impacto de cada uma delas na precisão de classificação de lesões malignas. Embora sejam usadas arquiteturas amplamente utilizadas no problema, as contribuições não estão na utilização dessas arquiteturas, mas no comparativo consistente. As informações produzidas servem para identificar novas combinações de técnicas, que influenciam na eficácia de novas abordagens voltadas à detecção automatizada do câncer de pele.

1.6 Estrutura da Tese

O conteúdo que envolve a tese deste trabalho está estruturado da seguinte forma:

- O Capítulo 2 descreve os principais trabalhos relacionados à problemática desta tese. Esses trabalhos serviram como base para a identificação e a caracterização das sistemáticas de detecção do câncer de pele por imagens, destacando os novos avanços e os aspectos positivos e negativos.
- O Capítulo 3 trata a fundamentação teórica deste trabalho, abordando os conceitos necessários para a compreensão da pesquisa, metodologia e abordagem desenvolvida. Além disso, também são retratados conteúdos relacionados às ferramentas utilizadas no desenvolvimento experimental da tese.
- O Capítulo 4 apresenta a abordagem de validação desenvolvida para classificação de imagens de lesões de pele. Essa abordagem é categorizada em etapas, sendo elas a seleção das bases de imagens, a remoção de pelos, a correção de iluminação e cor, a definição da região de interesse e as arquiteturas utilizadas no

processo de classificação. Além disso, também são descritas as metodologias experimentais e as métricas de avaliação utilizadas na análise de resultados.

- O Capítulo 5 descreve os principais resultados encontrados após a experimentação da abordagem de validação, juntamente com algumas discussões. Além disso, também é descrito um breve comparativo com outros trabalhos semelhantes e as limitações encontradas na abordagem implementada.
- Por fim, o Capítulo 6 apresenta as considerações finais sobre os resultados obtidos, os achados encontrados com a execução da abordagem, as repostas aos questionamentos de pesquisa, as observações sobre as limitações identificadas e as possibilidades de trabalhos futuros.

2 Revisão de Literatura

Este capítulo apresenta uma análise da literatura atual sobre a detecção automática de câncer de pele por meio de imagens digitais. A análise abrange a utilização de aprendizado profundo, não apenas para a classificação das lesões, mas também para ajustes relacionados à iluminação, cor e otimização. O capítulo destaca ainda o tipo, a metodologia e a relevância dos estudos, estabelecendo características associadas às técnicas convencionais e às novas metodologias utilizadas em sistemáticas de detecção do câncer de pele maligno.

2.1 Metodologia de Revisão

A metodologia de revisão utilizada para definir o escopo, buscar trabalhos, selecionar estudos, avaliar a qualidade, sintetizar evidências e discutir o tema do trabalho foi a *Rapid Review* (Tricco et al. 2015). Essa metodologia permitiu a identificação rápida de lacunas existentes, questões em aberto e debates sobre a problemática da classificação do câncer de pele utilizando aprendizado profundo.

As bases de dados utilizadas para a coleta e seleção de trabalhos relevantes foram: PubMed, IEEE Xplore, Scopus e Google Scholar. As palavras-chave foram definidas considerando a relevância, com a conjunção e disjunção dos termos. A melhor combinação foi determinada pela quantidade de trabalhos encontrados. Os principais termos utilizados foram: "*skin cancer detection*", "*deep learning*", "*image preprocessing*", "*color constancy*", "*color correction*" e "*illumination correction*".

A seleção e a avaliação da qualidade dos trabalhos escolhidos foram realizadas por meio da análise dos resumos e da relevância das propostas de pesquisa. No total, a pesquisa inicial encontrou 763 trabalhos que possuíam alguma ligação direta com o tema. O primeiro critério de seleção foi a identificação de metodologias que utilizavam técnicas de pré-processamento para correção dos aspectos de iluminação e cor, associadas ou não a técnicas de aprendizado profundo. O segundo critério de seleção foi a caracterização de técnicas de otimização utilizadas no processo de criação de modelos de aprendizado profundo para a classificação do câncer de pele.

Seguindo os critérios adotados, foram selecionados 38 trabalhos publicados entre os anos de 2019 e 2023, embora trabalhos de anos anteriores também tenham sido incluídos para reforçar a relevância do tema. Além disso, outras pesquisas que destacam particularidades na identificação de lesões malignas também foram incluídas. As próximas seções apresentam uma breve descrição de todos os trabalhos, categorizados por conteúdo. A descrição prioriza a aplicação do aprendizado profundo de forma geral, as

nuances e vieses das sistemáticas de detecção, e a evolução e os desafios da correção dos aspectos de iluminação e cor em imagens de lesões de pele.

2.2 Detecção do Câncer de Pele com Aprendizado Profundo

O uso de técnicas de aprendizado profundo é promissor na detecção de doenças por meio de imagens médicas (Esteva et al. 2021). Nos últimos cinco anos, a literatura destacou a capacidade dessas técnicas de abordar vieses associados ao diagnóstico do câncer de pele. As redes neurais profundas demonstraram desempenho superior na identificação de lesões malignas em comparação com dermatologistas experientes (Esteva et al. 2017). No entanto, essas técnicas ainda não substituem o médico dermatologista, mas auxiliam no acompanhamento e definição do diagnóstico (Guo et al. 2018).

As principais técnicas de aprendizado profundo presentes na maioria das pesquisas desenvolvidas são as Redes Neurais Convolucionais, Redes Neurais Residuais, Redes Adversárias Generativas e Redes Neurais Baseadas em Atenção, entre outras (Heibel, Hooey e Cockerell 2020). Essas técnicas podem ser empregadas nas etapas de pré-processamento, segmentação e classificação das lesões em imagens digitais (Dildar et al. 2021). Nesse contexto, é essencial descrever algumas pesquisas que se destacaram na literatura, por serem importantes para compreender o estado da arte e os possíveis direcionamentos para novas pesquisas.

A pesquisa de (Barata e Marques 2019) desenvolveu uma metodologia hierárquica para a detecção de câncer de pele, inspirada nos procedimentos de um dermatologista para determinar o diagnóstico. Inicialmente, decide se a lesão é melanoma ou não, e em seguida, identifica a classe correta, caso ela não seja melanoma. Para alcançar tal objetivo, foram utilizadas três arquiteturas convolucionais: uma multiclasse e duas hierárquicas. A DenseNet-161, pré-treinada com o ImageNet, foi selecionada para ambos os casos, implementando técnicas de transferência de aprendizado e ajuste fino. Os resultados mais expressivos obtidos para a base de imagens da *International Skin Imaging Collaboration* (ISIC) 2017 foram uma precisão de 59,8% na identificação de melanoma.

No estudo de (Mahbod et al. 2019), foi desenvolvida uma abordagem que emprega arquiteturas convolucionais pré-treinadas (AlexNet, ResNet-18 e VGG-16) para a extração de características, seguida pelo treinamento de um Support Vector Machine (SVM) multiclasse. Esta abordagem foi avaliada utilizando a base ISIC 2017, alcançando uma área sob a curva (AUC) de 97,55% para queratose seborreica e 83,83% para melanoma, respectivamente.

Em contrapartida, o estudo de (Majtner, Yildirim-Yayilgan e Hardeberg 2019) também propôs uma metodologia que utiliza arquiteturas de redes neurais convolucionais

pré-treinadas para a extração de características. Um diferencial entre os estudos de (Mahbod et al. 2019) e (Majtner, Yildirim-Yayilgan e Hardeberg 2019) está no processo de redução de dimensionalidade das imagens, utilizando as coordenadas de *bounding box*. Além disso, para essa redução, foi utilizado o *Linear Discriminant Analysis* (LDA). Após a redução da dimensionalidade, quatro classificadores distintos foram testados, dos quais o *K-Nearest Neighbors* (KNN) se destacou, alcançando uma precisão de 86% e uma especificidade de 99,9%.

A pesquisa conduzida por (Perez, Avila e Valle 2019) analisou os principais fatores que podem influenciar na escolha de arquiteturas de redes neurais profundas para análise de lesões de pele. Nesse estudo, foram avaliadas as arquiteturas DenseNet, Dual Path Nets, InceptionV4, InceptionResNetV2, MobileNetV2, PNASNet, ResNet, SEnet e Xception. Os experimentos utilizaram um conjunto de imagens contendo um total de 2.750 imagens (ISIC 2017). Os modelos avaliados obtiveram valores de área sob a curva ROC entre 84% e 91%, demonstrando um desempenho satisfatório na tarefa de classificação de lesões de pele.

O estudo realizado em (Harangi, Baran e Hajdu 2020) propôs uma abordagem para classificar o conjunto de imagens da base ISIC 2018 em sete classes, utilizando as arquiteturas GoogLeNet e InceptionV3. No método desenvolvido, as arquiteturas foram inicialmente treinadas em duas classes: malignas e benignas. Em seguida, foi utilizado o ajuste fino para aumentar o grau de confiança da rede nessas duas classes e nas outras cinco classes adicionais. O método proposto alcançou uma taxa de classificação de 67,7%. Esse resultado demonstrou ser possível melhorar a identificação de lesões de pele em múltiplas classes, considerando um treinamento inicial em apenas duas classes.

No trabalho desenvolvido por (Nahata e Singh 2020), foi proposta uma metodologia que utiliza diversas arquiteturas convolucionais para classificar imagens de lesões de pele. Essa abordagem se destaca pela aplicação de técnicas de aumento de dados e normalização dos aspectos das imagens, visando melhorar a qualidade e a diversidade do conjunto de treinamento. Um aspecto notável dessa metodologia é a adoção de um número maior de camadas *dropout* e a estratégia de transferência de aprendizado, ambas voltadas para evitar *overfitting*. As arquiteturas utilizadas nesse estudo foram InceptionV3, ResNet50, VGG16, MobileNet e InceptionResNetV2. O melhor resultado alcançado apresentou uma precisão média de 91%, com base em uma avaliação multiclasse realizada nos conjuntos de imagens ISIC 2018 e ISIC 2019.

No estudo de (Daghrir et al. 2020), foi implementada uma abordagem híbrida para a detecção de melanoma, combinando redes neurais convolucionais e algoritmos classificadores convencionais. Além disso, foram utilizadas características de bordas, cores e texturas, associadas a dados extraídos por arquiteturas convolucionais. A combinação dos resultados dos modelos gerados foi avaliada globalmente por meio do

voto majoritário. A metodologia envolveu pré-processamento, incluindo aprimoramento da qualidade da imagem com realce de cor no canal azul, remoção de pelos com filtro *Derivative of Gaussian* (DoG), *threshold* com OTSU, dilatação morfológica e *Inpainting*. Os experimentos demonstraram que o uso combinado de técnicas de aprendizado profundo e convencionais resultou em uma precisão de 88,4% na detecção correta do câncer de pele. Esse resultado foi alcançado utilizando uma personalização da base de imagens ISIC.

Os estudos de (Al-Masni, Kim e Kim 2020) resultaram em um *framework* de diagnóstico do câncer de pele que combina uma etapa de segmentação de bordas de lesões de pele e uma etapa de classificação de múltiplas lesões. O modelo proposto foi avaliado utilizando três conjuntos de dados independentes: ISIC 2016, ISIC 2017 e ISIC 2018, que contêm dois, três e sete tipos de lesões de pele, respectivamente. Nos experimentos, foram utilizados os modelos classificadores InceptionV3, ResNet50, Inception-ResNetV2 e DenseNet201. Os resultados demonstraram que, para duas classes, as acurácias médias obtidas foram de 81,27%. Para três classes, a média ponderada da acurácia foi de 73,44%. Por fim, para sete classes, a média ponderada da acurácia atingiu 89,28%. Nesse cenário, a ResNet50 apresentou o melhor desempenho entre os modelos avaliados.

A pesquisa de (Valle et al. 2020) trata da confiança de modelos de aprendizado profundo para a análise do câncer de pele. Nos estudos realizados, são investigados os impactos, *design* e arquiteturas do modelo, aprendizado por transferência, conjunto de dados, resolução das imagens, entre outros fatores. Os experimentos realizados mostraram que a quantidade de dados de treinamento é o fator mais influente, explicando parte da variação no desempenho. Outros fatores significativos incluíram aumento de dados no conjunto de teste e a resolução de entrada. O aprendizado por transferência mostrou-se importante para os melhores resultados. O estudo também revelou a eficácia de ensembles de modelos como uma opção confiável, oferecendo uma alternativa ao *design* e à otimização.

Em (Nawaz et al. 2021), foi implementada uma nova sistemática de detecção do câncer de pele em três etapas. A primeira etapa realiza o pré-processamento das imagens com normalização de iluminação e operações morfológicas para remoção de pelos. Um diferencial dessa metodologia na segunda etapa é a utilização de redes neurais convolucionais baseadas em regiões (Faster R-CNN) para realizar a localização e extração de características. Por fim, a terceira etapa utiliza um SVM para realizar a classificação das características obtidas no processo anterior. O *framework* proposto alcançou uma precisão média de 89,1%, com uma sensibilidade de 85,9% e especificidade de 87,0% na base ISIC 2016.

No estudo realizado por (Mazoure et al. 2022), foi desenvolvido um sistema web,

denominado DUNEScan, destinado à avaliação da precisão na classificação de lesões de pele. Esta ferramenta permite a análise e a variabilidade nas previsões de modelos de aprendizado profundo, empregando interpretabilidade de resultados com as técnicas UMAP e Grad-CAM. Os usuários da aplicação podem fazer o upload de imagens de lesões, recebendo em seguida um leque de informações sobre a similaridade entre lesões malignas e benignas. As arquiteturas utilizadas, que incluem InceptionV3, ResNet50, MobileNetV2, EfficientNet, BYOL e SwAV, foram previamente treinadas utilizando o conjunto de dados ISIC. Como conclusão, o estudo ressalta a importância das incertezas de classificação identificadas pelo aprendizado profundo como um meio de fornecer uma base mais sólida para o diagnóstico por parte de dermatologistas.

Em (Bissoto et al. 2022), foi realizado um estudo sobre os vieses encontrados na detecção do câncer de pele. O estudo aborda o desafio da dependência de modelos de aprendizado profundo em correlações irrelevantes, como artefatos presentes em imagens de lesões de pele, que podem prejudicar a precisão da análise. A metodologia incluiu a criação de ambientes baseados em artefatos para treinamento robusto e a aplicação de *debiasing* no momento do teste, melhorando a generalização do modelo em dados fora da distribuição. Os resultados indicam que técnicas para tornar modelos mais gerais podem ajudar a corrigir vieses, especialmente quando as diferenças entre os dados são pequenas.

No estudo de (Razmjoooy e Arshaghi 2023), foi apresentada uma metodologia chamada *Multi-Agent Fuzzy Buzzard*. Essa abordagem combina aprendizado profundo com meta-heurísticas, propondo uma integração de segmentação multinível baseada em histogramas e classificação que utiliza agentes de busca global e local. O objetivo dessa metodologia é aprimorar a eficiência de busca, tanto em âmbito global quanto local, para identificar as melhores regiões de interesse em imagens de lesões de pele. Além disso, visa otimizar os parâmetros de uma rede neural convolucional destinada à classificação das lesões como benignas ou malignas. Utilizando as bases de imagens do *Dermatology Information System* (DermIS) e DermQuest, o método alcançou um resultado relevante, atingindo 94% de acurácia na distinção entre lesões malignas e benignas.

Nas pesquisas de (Gomathi et al. 2023), são abordados aspectos de otimização de modelos de aprendizado profundo. Para isso, desenvolveu-se uma sistemática com dupla otimização para a detecção de câncer de pele maligno. Essa proposta integra a Otimização por Enxames de Partículas e a Otimização de Forrageamento Bacteriano. O pré-processamento é realizado com um filtro mediano adaptativo. A segmentação é aplicada por meio da arquitetura U-Net, com a dupla otimização. Por fim, uma arquitetura convolucional é utilizada para treinar e classificar as imagens. Os resultados apresentados na pesquisa mostram uma precisão de 98,76% na classificação da base de imagens HAM10000, demonstrando uma alta eficácia na detecção de câncer de pele.

No trabalho de (He et al. 2023), foi desenvolvida uma sistemática de detecção do câncer de pele em duas etapas: segmentação e classificação. A primeira etapa delimita as áreas das lesões usando uma Full-resolution Convolutional Network (FrCN). Após a segmentação, as regiões delimitadas são classificadas utilizando as arquiteturas profundas InceptionV3, ResNet50, Inception-ResNetV2 e DenseNet201. O estudo avaliou a sistemática desenvolvida utilizando as bases ISIC 2016, ISIC 2017 e ISIC 2018. Os experimentos demonstraram o melhor resultado com a arquitetura Inception-ResNetV2, atingindo uma precisão de 71,40% e acurácia de 81,79% para a base ISIC 2016. Uma conclusão apresentada no trabalho é que a integração da segmentação e classificação com o aprendizado profundo melhora a precisão e a confiabilidade no diagnóstico correto.

Uma metodologia semelhante à de (He et al. 2023) também foi desenvolvida por (Balaha e Hassan 2023). Nessa pesquisa, foram testadas as arquiteturas U-Net, U-Net++, Attention U-Net e SwinU-Net para segmentação semântica das lesões de pele. Na tarefa de classificação, foram avaliadas as arquiteturas VGG16, VGG19, MobileNet, MobileNetV2, MobileNetV3 Large, MobileNetV3 Small, NASNetMobile e NASNetLarge, otimizadas usando o *sparrow search algorithm* (SpaSA). Os conjuntos de imagens avaliados na proposta foram HAM10000, ISIC 2019, ISIC 2020 e PH². O melhor resultado alcançado, conforme detalhado no estudo, foi na classificação dos conjuntos ISIC 2019 e ISIC 2020 para melanoma e HAM10000, onde o modelo pré-treinado MobileNet atingiu uma precisão geral de 98,27% e 98,83%, respectivamente. A metodologia se destaca pela integração da otimização meta-heurística para melhorar significativamente o desempenho dos modelos.

Os estudos analisados mostram um panorama diversificado de abordagens para a detecção de câncer de pele utilizando aprendizado profundo. As metodologias variam desde abordagens hierárquicas e híbridas até técnicas avançadas de otimização e integração de segmentação e classificação. A precisão alcançada é variável, mas há uma tendência clara de que a combinação de várias técnicas como aumento de dados, normalização de aspectos (pré-processamento), técnicas de transferência de aprendizado e meta-heurísticas tende a melhorar significativamente o desempenho dos modelos. A Tabela 1 apresenta um resumo de todos os trabalhos analisado.

A eficácia de diferentes arquiteturas como DenseNet, ResNet, MobileNet e Inception foi demonstrada em vários contextos, com resultados variados, dependendo do conjunto de dados utilizado e das técnicas de pré-processamento e otimização aplicadas. Estudos que incorporam técnicas de *debiasing* e interpretabilidade, como (Bissoto et al. 2022) e (Mazoure et al. 2022), destacam a importância de lidar com vieses e melhorar a confiança nos diagnósticos assistidos por IA.

Referência	Bases de Imagens	Descrição do Estudo	Resultados/Perspectivas
(Barata e Marques 2019)	ISIC 2017	Deteção hierárquica de câncer de pele, semelhante à abordagem de um dermatologista. Desenvolve o ajuste estatístico de cor, segmentação por caixa delimitadora e classificação com a arquitetura convolucional DenseNet-161.	Recall para melanoma de 59,8%. Recall para queratose de 71,1%. Recall para nevus de 79,2%. Acurácia balanceada de 70%.
(Mahbod et al. 2019)	ISIC 2017	Uso das arquiteturas AlexNet, ResNet-18 e VGG16 pré-treinadas para extração de características, seguida pelo treinamento de um SVM multiclasse.	AUC de 97,55% para queratose seborreica. AUC de 83,83% para melanoma.
(Majtner, Yildirim-Yayilgan e Hardeberg 2019)	ISIC 2017	Utilização de arquiteturas convolucionais pré-treinadas, redução de imagens com <i>bounding box</i> , redução de dimensionalidade com LDA, treinamento de classificadores convencionais.	O KNN apresentou o melhor resultado, com precisão de 86% e especificidade de 99,9%.
(Perez, Avila e Valle 2019)	ISIC 2017	Avaliação do uso de arquiteturas profundas de duas formas na classificação: isoladamente e em conjunto (<i>ensemble</i>).	Arquiteturas isoladas mantêm um bom desempenho, mas os <i>ensembles</i> apresentam perspectivas de melhores resultados.
(Nahata e Singh 2020)	ISIC 2018; ISIC 2019	Implementação de aumento de dados, normalização e transferência de aprendizado com as arquiteturas InceptionV3, ResNet50, VGG16, MobileNet e InceptionResNet.	Precisão de 91% com a InceptionResNet.
(Harangi, Baran e Hajdu 2020)	ISIC 2018	Deteção hierárquica com as arquiteturas GoogLeNet e InceptionV3, aplicando ajuste fino em duas etapas: binária e multiclasse.	Precisão de 67,7%.
(Daghrir et al. 2020)	ISIC (não definido)	Desenvolvimento de uma abordagem híbrida, combinando dados de texturas, cores e bordas com características obtidas com arquiteturas convolucionais.	Precisão de 84%.
(Al-Masni, Kim e Kim 2020)	ISIC 2016; ISIC 2017; ISIC 2018	Framework de classificação utilizando as arquiteturas InceptionV3, ResNet50, Inception-ResNetV2 e DenseNet201.	Acurácia de 81,27% para maligno e benigno. Acurácia de 73,44% para três classes. Acurácia de 89,28% para sete classes.
(Valle et al. 2020)	ISIC 2017; Dermofit Image Library	Investigação sobre os impactos, design e arquiteturas do modelo, aprendido por transferência, conjunto de dados, resolução das imagens, entre outros fatores.	A quantidade de dados é o fator mais influente. A transferência de aprendizado proporciona os melhores resultados. Eficácia dos <i>ensembles</i> .
(Nawaz et al. 2021)	ISIC 2016	Uso da Faster-RCNN e SVM para deteção e classificação.	Precisão média de 89,1%.
(Mazoure et al. 2022)	HAM10000; ISIC (não definido)	Desenvolvimento de um sistema web para avaliação do diagnóstico de câncer de pele, utilizando arquiteturas como EfficientNet, InceptionV3, MobileNet, ResNet50, SwAV e BYOL. Implementação da auditoria de resultados com UMAP e GRAD-CAM.	Apresenta estimativas de incertezas na deteção de câncer de pele, comparando resultados entre técnicas.
(Bissoto et al. 2022)	ISIC 2019; PH ² ; Derm7pt; PAD-UFES	Análise de vieses e a dependência de modelos de aprendizado profundo em dados não correlacionados ao câncer de pele.	O uso de técnicas para mitigar vieses pode tornar os modelos de aprendizado profundo mais robustos.
(Razmjoooy e Arshaghi 2023)	DermIS; DermQuest	Desenvolvimento da técnica <i>Multi-Agent Fuzzy Buzzard</i> para segmentação e redes neurais convolucionais para classificação multiclasse.	Acurácia de 94%.
(Gomathi et al. 2023)	HAM10000	Implementação de pré-processamento e otimização por Enxames de Partículas e Otimização de Forrageamento Bacteriano.	Precisão de 98,76%.
(He et al. 2023)	ISIC 2016; ISIC 2017; ISIC 2018	Segmentação semântica com <i>Full Resolution Convolutional Network</i> (FrCN) e classificação usando InceptionV3, ResNet50, Inception-ResNetV2 e DenseNet201.	Precisão de 71,40% e acurácia de 81,79% para a ISIC 2016.
(Balaha e Hassan 2023)	ISIC 2019; ISIC 2020; PH ² ; HAM10000	Segmentação semântica com U-Net, U-Net++, Attention U-Net e SwinU-Net. Classificação com otimização de parâmetros usando a técnica SpaSA em arquiteturas pré-treinadas VGG16, VGG19, MobileNet, MobileNetV2, MobileNetV3 Large, MobileNetV3 Small, NASNetMobile e NASNetLarge.	Precisão geral de 98,27% e 98,83% para melanoma nas bases ISIC 2019 + ISIC 2020 e HAM10000.

Tabela 1 – Pesquisas sobre aprendizado profundo na deteção do câncer de pele.

Outra observação significativa, conforme relatado no estudo de (Perez, Avila e Valle 2019), é que a escolha da arquitetura não está diretamente correlacionada com as melhorias nos resultados de classificação, sugerindo que fatores não controlados podem ter um impacto maior no ganho de resultados do que na escolha da arquitetura. No geral, a integração de múltiplas etapas como pré-processamento, segmentação e classificação, com a aplicação de técnicas avançadas de aprendizado profundo e otimização, são necessárias uma melhor precisão na identificação do câncer de pele por imagens.

2.2.1 Aumento de Dados na Detecção do Câncer de Pele

Como mencionado por (Valle et al. 2020), o aumento de dados é uma estratégia necessária para a otimização de modelos de classificação com técnicas de aprendizado profundo. As bases de imagens públicas possuem limitações em relação à quantidade de imagens de classes malignas (Barata, Celebi e Marques 2018). Nesse sentido, é essencial descrever alguns trabalhos relevantes na literatura que utilizam técnicas de produção de dados sintéticos no problema de classificação do câncer de pele com aprendizado profundo.

O uso de Redes Geradoras Adversárias (*Generative Adversarial Networks - GAN*) tem sido alvo de discussões na literatura, visto que essa técnica passou a ter aplicação na criação de imagens sintéticas de imagens de câncer de pele (Qin et al. 2020). Inclusive, a utilização dessa técnica pode evidenciar uma melhora no processo de aprendizado de modelos classificadores de lesões de pele.

Em (Bissoto et al. 2018), as GANs foram apontadas como diferenciais no aumento de dados, corroborando com (Bisla et al. 2019). Essas pesquisas demonstraram a relevância dessa técnica na geração de imagens sintéticas para enfrentar o desafio do câncer de pele. Os experimentos realizados aplicaram arquiteturas de forma independente para balancear o conjunto de treinamento. Um modelo ResNet50, pré-treinado com ImageNet, foi ajustado com um conjunto de imagens expandido e melhorado, sendo utilizado para classificar lesões de pele em três categorias: melanoma, queratose seborreica e nevo. Os resultados alcançados mostram uma precisão de 86,1% com a base HAM10000, superando o desempenho do modelo ResNet50 de referência.

Outra metodologia para a geração de dados sintéticos de lesões de pele foi proposta por (Abdelhalim, Mohamed e Mahdy 2021), baseada na *Progressive Self-Attention Generative Adversarial Network (PSA-GAN)*. Os experimentos realizados indicaram que as imagens sintéticas permitiram alcançar uma precisão de 70,1%, em comparação com os 67,3% de precisão obtidos por um sistema não aprimorado, ambos usando as imagens originais de lesões de pele como base de comparação.

O estudo de (Qin et al. 2020) empregou uma GAN baseada em transferência de estilo para gerar imagens de alta qualidade na classificação de lesões de pele. Essas imagens sintéticas foram adicionadas ao conjunto de treinamento de um modelo ResNet50 pré-treinado, mostrando que a transferência de estilo supera outros métodos em termos de *Inception Score (IS)* e *Fréchet Inception Distance (FID)*, além de apresentar melhorias na acurácia, sensibilidade, especificidade e outros indicadores relevantes para a classificação de lesões cancerígenas.

Em (Bissoto, Valle e Avila 2021), explora-se o desenvolvimento de métodos de aumento de dados e anonimização utilizando GANs para a análise de lesões de pele. Os

experimentos realizados demonstram que a eficácia dessa técnica para lesões de pele é limitada, observando-se poucas melhorias em cenários fora da distribuição. Esse resultado apresenta críticas ao trabalho anterior de (Bissoto et al. 2018). A pesquisa conclui que o uso de GANs para aumento de dados pode intensificar vieses existentes nos conjuntos de imagens. Apesar do potencial dessas técnicas em contextos específicos, como o compartilhamento de conhecimento e a preservação da privacidade dos pacientes.

Os trabalhos analisados demonstram a eficácia do uso de GANs para o treinamento de modelos de aprendizado profundo, principalmente associados as imagens do câncer de pele. A Tabela 2 oferece um resumo dos estudos mencionados, destacando o uso de GANs em imagens de lesões de pele, as metodologias empregadas e as implicações dessas abordagens para o problema.

Referência	Descrição do Estudo	Resultados/Perspectivas
(Bissoto et al. 2018)	Realiza vários experimentos com GANs para aumento de dados. Avalia os resultados utilizando várias arquiteturas profundas.	Determina que as GANs são promissoras para geração de imagens sintéticas de lesões de pele.
(Bisla et al. 2019)	Destaca o uso de GANs para aumento de dados de imagens de lesões de pele. A proposta inclui um comparativo entre a utilização e a não utilização.	Apresenta melhoria de resultados na arquitetura ResNet50 com o uso de aumento de dados, alcançando uma precisão de 86,1%.
(Abdelhalim, Mohamed e Mahdy 2021)	Desenvolve a <i>Progressive Self-Attention Generative Adversarial Network</i> (PSA-GAN) para aumento de dados.	Apresenta resultados melhorados na arquitetura ResNet50 com o uso de aumento de dados, alcançando uma precisão de 70,1% comparado a 67,3% de um sistema não aprimorado.
(Qin et al. 2020)	Aplica a transferência de estilo com GANs na geração de imagens sintéticas de câncer de pele, comparando-as com as originais e utilizando-as no treinamento de modelos de aprendizado profundo.	As imagens sintéticas mantêm boa similaridade com as originais. O uso do aumento de dados com GANs proporciona ganhos de acurácia e precisão em todos os testes realizados.
(Bissoto, Valle e Avila 2021)	Analisa o aumento de dados e a anonimização com GANs.	Determina que os ganhos nos resultados são modestos e que o aumento de dados com essas técnicas pode reforçar vieses presentes nas bases de imagens.

Tabela 2 – Pesquisas relacionadas ao aumento de dados e geração de imagens sintéticas de lesões de pele com GANs.

Abordagens como o uso de PSA-GAN e transferência de estilo mostram que a geração de imagens sintéticas pode melhorar significativamente a precisão na classificação de lesões malignas. No entanto, os estudos também destacam algumas controvérsias e desafios, especialmente relacionados ao potencial reforço de vieses existentes nas bases públicas de imagens de câncer de pele. Isso é evidenciado pelas conclusões dos trabalhos de (Bissoto et al. 2018) e (Bissoto, Valle e Avila 2021), que apresentam relações opostas sobre os resultados descritos.

A necessidade de uma análise detalhada de novas técnicas de aumento de dados é evidente. Essas novas técnicas devem evitar a amplificação de vieses existentes nas bases de imagens públicas (Bissoto, Valle e Avila 2021). O uso de GANs pode ser

interessante em outros cenários relacionados ao pré-processamento de imagens, como na reconstrução de imagens ou na correção e ajuste de características desejáveis (Gomathi et al. 2023).

2.3 Iluminação e Cor na Detecção do Câncer de Pele

As variações nas condições de iluminação, seja por deficiência ou excesso de luz, podem comprometer a qualidade das imagens digitais (Rubegni et al. 2002). Esses fatores afetam diretamente a precisão das cores, criam problemas para qualquer sistema automatizado de detecção de objetos (Huang et al. 2023). Os métodos de correção são imprescindíveis para mitigar essas distorções e normalizar as condições de aquisição das imagens (Jones et al. 2022).

Muitos estudos já definiram a variação luminosa e seus efeitos nos atributos cromáticos como um desafio para o diagnóstico do câncer de pele por meio de imagens digitais (Barata, Celebi e Marques 2014). A investigação da cor e sua relevância na detecção de lesões malignas foi sugerida há muito tempo (Rubegni et al. 2002). Os tons de marrom e vermelho, por exemplo, contribuem para a identificação correta de lesões malignas (Barata, Celebi e Marques 2014). Conforme (Galdran et al. 2017), a variação das tonalidades de cor em imagens de câncer de pele requer uma correção analítica.

A iluminação e a cor são características essenciais para a precisão do diagnóstico automatizado de lesões malignas (Barata et al. 2013). Portanto, é relevante descrever um breve estado da arte sobre esse tema, especialmente em tempos em que o aprendizado profundo figura como tendência em novas abordagens de classificação. Essas características já tiveram grande importância em sistemáticas de detecção convencionais com *Machine Learning*, mas passaram a ser desconsideradas com o uso do aprendizado profundo (Salvi et al. 2022).

As alterações na iluminação modificam a representação das cores, fazendo com que as imagens de lesões malignas apresentem uma ampla variação de fotótipos (Adegun e Viriri 2021). Em (Barata, Celebi e Marques 2014), foi realizada uma avaliação comparativa entre abordagens que utilizam características locais e globais, evidenciando a superioridade das características de cor em relação às texturais na identificação de melanomas. A abordagem focada em características locais demonstrou maior eficácia, ressaltando a importância dos ajustes de cor para a precisão do diagnóstico automático.

As principais conclusões em (Barata, Celebi e Marques 2014) já haviam sido tratadas em (Barata et al. 2013), que também destaca a relevância dos ajustes de cor. É evidente que técnicas de correção de iluminação promovem efeitos positivos na representação da constância de cores. Um panorama abrangente sobre a extração de características é explorado em (Barata, Celebi e Marques 2018), discutindo quatro

classes de características: manuais, baseadas em dicionário, aprendizado profundo e clinicamente inspiradas. Este estudo trata a evolução das técnicas e as melhorias na precisão da detecção do câncer de pele maligno.

Já em (Barata e Marques 2019), observou-se uma transição das técnicas manuais para abordagens mais avançadas de aprendizado profundo no processo de diagnóstico automatizado do câncer de pele. Novas técnicas abrangem desde o pré-processamento, segmentação e classificação, como também otimização e auditoria. No entanto, desafios como a baixa qualidade das imagens em bases de dados descritas na literatura e a eficácia de generalização em diferentes tipos de pele são apontados como obstáculos a serem superados (Salvi et al. 2022).

Entre as abordagens comuns empregadas no pré-processamento ou normalização dos aspectos de iluminação e cor, destacam-se técnicas como *Gray World* (GW) (Buchsbaum 1980), *Shades of Gray* (SoG) (Finlayson e Trezzi 2004), maxRGB (Funt e Shi 2010), além de métodos baseados em histogramas e em algoritmos de agrupamento com *fuzzy*. Estas técnicas buscam ajustar a iluminação uniformemente, baseando-se em modelos estatísticos para garantir o padrão das cores (Salvi et al. 2022). A padronização das cores também é importante na segmentação de lesões de pele em imagens (Ng et al. 2019).

Os métodos estatísticos tradicionais para a correção de iluminação e cor vêm sendo substituídos por abordagens que utilizam o aprendizado profundo no pré-processamento de imagens de câncer de pele, conforme descrito em (Salvi et al. 2022). Essa mudança é observada nas atualizações da literatura, destacando a adoção de técnicas avançadas tanto em etapas de pré-processamento e segmentação quanto na aplicação de métodos de redes neurais profundas, incluindo aquelas que integram módulos de atenção (Branciforti et al. 2023). Além disso, a utilização do aprendizado profundo para ajustes na constância de cores tem se destacado na dermatologia, especialmente após extensas validações em bases de dados públicas (Salvi et al. 2024), reforçando sua relevância e eficácia na área.

A aplicação de arquiteturas de redes neurais profundas na histopatologia dermatológica destaca a necessidade de considerar a variabilidade de iluminação e cores no diagnóstico de tumores e cânceres (Tosta et al. 2023). Essa variabilidade é um desafio não apenas para a análise de imagens dermatológicas, mas também para outras imagens médicas, onde é crucial equilibrar a normalização das cores com a preservação dos detalhes das imagens.

O estudo de (Cong et al. 2022) sugere que, embora métodos estatísticos para correção cromática apresentem algumas melhorias, eles têm limitações significativas no tratamento de imagens de câncer de pele. Isso sublinha a necessidade de desenvolver abordagens mais sofisticadas e eficientes. As GANs são uma das técnicas de aprendizado

profundo aplicadas na correção de iluminação e cor, sendo usadas para gerar imagens com iluminação e cores ajustadas, podendo garantir a constância de cor por meio do ajuste uniforme da iluminação, seja com aprendizado supervisionado ou não supervisionado (Salvi et al. 2022).

Um exemplo de abordagem com GANs para correção de iluminação e cor em imagens clínicas de lesões de pele é apresentado em (Salvi et al. 2022), que desenvolveu a *Dermo Color Constancy-Generative Adversarial Network* (DermoCC-GAN). Essa arquitetura pré-processa as imagens das lesões de maneira heurística, aprimorando significativamente tanto a iluminação quanto a constância de cor. Os resultados obtidos com a correção implementada neste trabalho receberam avaliações positivas de dermatologistas, conforme descrito em (Branciforti et al. 2023).

A pesquisa realizada por (Venugopal et al. 2024) descreve o método *Illumination Equalization using Counter Exponential Transform* (IECET) para o tratamento de imagens de lesões cutâneas. Esse método aplica transformações contra exponenciais ao componente de luminosidade no espaço de cores HSV, e mantendo os componentes de matiz e saturação. Ao utilizar um regressor baseado na arquitetura profunda ResNet para a calibração de parâmetros, o IECET assegura a homogeneidade e a adequação da iluminação.

Outra pesquisa semelhante a (Venugopal et al. 2024) foi realizada por (Pandey et al. 2019), que implementa técnicas como *Contrast Limited Adaptive Histogram Equalization* (CLAHE) e *Multi Scale Retinex with Color Restoration* (MSRCR) proporcionaram uma pequena melhoria na precisão do conjunto de validação, indicando que a correção não causa alterações significativas nas imagens e oferece uma melhoria na qualidade em termos de iluminação e cor.

Ainda existem muitos desafios na utilização do aprendizado profundo para a correção de aspectos de iluminação e cor em imagens de câncer de pele (Tang et al. 2023). Os estudos apresentados destacam a importância de considerar a variabilidade de iluminação e cor no diagnóstico automatizado de lesões malignas, especialmente no contexto do câncer de pele. A transição de técnicas convencionais de *Machine Learning* para abordagens avançadas de aprendizado profundo trouxe melhorias significativas de precisão, mas também ressaltou novos desafios, como a necessidade de métodos mais sofisticados para correção de iluminação e cor.

As GANs e outras técnicas de aprendizado profundo mostraram-se promissoras na correção desses aspectos, proporcionando melhorias na qualidade das imagens e na precisão do diagnóstico. No entanto, a eficácia dessas técnicas depende da capacidade de equilibrar a normalização das cores com a preservação dos detalhes das imagens. Novas pesquisas devem focar na integração dessas técnicas em sistemas automatizados, tornando-os robustos e aplicáveis em diferentes condições de iluminação e tipos de pele.

2.4 Considerações Finais

Neste capítulo, foram explorados os principais trabalhos relacionados à proposta desta pesquisa, destacando a relevância do aprendizado profundo em abordagens de detecção de câncer de pele. Além disso, foram analisados estudos que ressaltam a importância de técnicas de correção de iluminação e cor na identificação de malignidades de pele, enfatizando a influência na precisão do diagnóstico. A revisão da literatura descrita revela um consenso do aprendizado profundo na melhoria dos sistemas de diagnóstico dermatológico, apontando para a necessidade crítica de abordar novas técnicas de correção para variações de iluminação e cor, além do aumento de dados para imagens clínicas de lesões de pele.

3 Fundamentação Teórica

Este capítulo abrange todos os conteúdos essenciais para a compreensão desta tese. Primeiramente, são apresentados conceitos básicos sobre câncer de pele, aprendizado profundo, aumento de dados, transferência de aprendizado, ajuste fino e segmentação semântica. Em seguida, são definidos os princípios relacionados à iluminação e à cor em imagens digitais. Por fim, são discutidas abordagens que impulsionam o desenvolvimento de técnicas convencionais e baseadas em aprendizado profundo para a correção das características de iluminação e cor.

3.1 Tipos de Câncer de Pele

O câncer de pele é causado pela combinação de diversos fatores, como exposição excessiva ao sol, pele e olhos claros, ou até mesmo fatores genéticos (Marques et al. 2016). Qualquer alteração nas características da pele, como o aparecimento de manchas, feridas que não cicatrizam, protuberâncias, alterações de cor ou tamanho em pintas já existentes, é motivo para buscar a avaliação de um dermatologista (Vieira et al. 2012). Os principais tipos de câncer de pele malignos são os carcinomas basocelular, espinocelular e o melanoma (Murad e Katz 1996).

O carcinoma basocelular é um tipo de câncer de pele que se origina nas células basais da epiderme. Essas células produzem a queratina, uma proteína que proporciona rigidez e proteção à pele. Esse tipo de câncer é comum em áreas expostas ao sol, com crescimento lento. As lesões desse tipo podem apresentar pequenos vasos sanguíneos visíveis (Bonalmi, Campos e Leal 2017).

O carcinoma espinocelular é caracterizado por lesões rugosas e com pouca elasticidade, que surgem nas camadas escamosas da epiderme. Esse tipo de câncer de pele se desenvolve em áreas expostas ao sol. No entanto, ele pode originar-se em outras partes do corpo, como cicatrizes, úlceras ou áreas de pele cronicamente danificada (Vieira et al. 2012).

O carcinoma de Merkel também é uma forma incomum de câncer de pele que se origina nas células de Merkel, encontradas nas camadas mais profundas da pele. Essas células estão localizadas próximas às terminações nervosas da pele. Geralmente, esse tipo de câncer se manifesta como nódulos vermelhos ou roxos na pele. Sua ocorrência pode ser agressiva e requer um diagnóstico precoce (Marques et al. 2016).

Já o melanoma é uma lesão que evolui, mudando de cor, formato e tamanho, podendo surgir em qualquer parte do corpo. Inicialmente, esse tipo de lesão se manifesta

como uma pigmentação anormal ou pequeno tumor. Essa doença também pode ocorrer ao redor ou nos olhos (Vieira et al. 2012). Dentre os tipos de câncer, ele é o mais agressivo, e sua ocorrência requer maior atenção médica (Marques et al. 2016). A Figura 2 apresenta exemplos visuais das principais lesões de pele malignas.



Figura 2 – Tipos de Câncer de Pele. Adaptado de (Tschandl et al. 2017).

As lesões com formato assimétrico, variação de cor, textura irregular e sangramento são indicativos de uma neoplasia maligna (Bonalmi, Campos e Leal 2017). No entanto, o diagnóstico preciso só é dado após a realização da biópsia e exame imuno-histoquímico. Esses procedimentos são recomendados pelo dermatologista para confirmar o tipo do câncer de pele (Brunsgaard, Jensen e Grossman 2022).

O tratamento do câncer maligno é determinado pelo seu estágio (Santos et al. 2023). Nos estágios iniciais, a excisão da lesão é comum, seguida de acompanhamento médico para garantir a cura. No entanto, em estágios metastáticos, uma cura definitiva torna-se desafiadora. O tratamento com quimioterapia e imunoterapia tem mostrado resultados promissores em pacientes com casos avançados (Bonalmi, Campos e Leal 2017). A quimioterapia é usada para combater as células cancerígenas em todo o corpo, enquanto a imunoterapia visa fortalecer o sistema imunológico do paciente para combater o câncer (Marques et al. 2016).

Embora a bibliografia médica possa fornecer distinções entre os diferentes tipos de câncer de pele, o diagnóstico das lesões é geralmente considerado maligno ou benigno (Bonalmi, Campos e Leal 2017). No geral, o prognóstico do câncer de pele não é uma tarefa fácil para o dermatologista, uma vez que existem vários tipos de câncer com

origens diversas (Murad e Katz 1996). Além disso, lesões malignas e benignas possuem características semelhantes que podem confundir e retardar a identificação. Portanto, diagnosticar o câncer de pele ainda é desafiador (Marques et al. 2016).

3.2 Aprendizado Profundo

O aprendizado profundo (*Deep Learning*) é uma subárea da Inteligência Artificial (IA) que desenvolve redes neurais artificiais com múltiplas camadas (profundas) (LeCun, Bengio e Hinton 2015). Essas redes conseguem identificar padrões complexos em dados como imagens, textos e sons, facilitando a extração de dados relevantes sem a necessidade de uma engenharia de características manual. Esse avanço permite que os modelos produzidos se ajustem de forma mais eficaz às características de entrada (Goodfellow, Bengio e Courville 2016).

As metodologias de aprendizado profundo tornaram-se tendências para tratar problemas complexos na medicina, principalmente voltados para imagens médicas (Adegun e Viriri 2021). Inclusive, essas metodologias já são exploradas no desenvolvimento de sistemas de acompanhamento e detecção do câncer de pele (Heibel, Hooley e Cockerell 2020). As técnicas principais do aprendizado profundo são Redes Neurais Convolucionais, Redes Neurais Residuais, Redes Geradoras Adversárias e *Autoencoders* (Yan, Yoshua e Geoffrey 2015). Nas próximas subseções, são descritos os conceitos fundamentais associados a cada uma dessas técnicas.

3.2.1 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs), também conhecidas como *Convolutional Neural Networks* (CNNs), constituem uma categoria de algoritmos de aprendizado profundo desenvolvida para reconhecimento de objetos em imagens e vídeos (Goodfellow, Bengio e Courville 2016). Inspiradas pela estrutura do córtex visual humano, essas técnicas conseguem extrair padrões relevantes, aprender e prever classes diretamente dos *pixels* (Zhou, Greenspan e Shen 2017).

A principal vantagem das redes neurais convolucionais é sua capacidade de aprender representações hierárquicas em níveis abstratos, começando por características simples, como bordas e texturas, e progredindo para características mais complexas e específicas (Yan, Yoshua e Geoffrey 2015). A arquitetura básica de uma rede desse tipo consiste em várias camadas, incluindo camadas de convolução, camadas de *pooling* e camadas totalmente conectadas (LeCun, Bengio e Hinton 2015).

A camada de convolução é responsável por aplicar filtros ou *kernels* às entradas de imagem, realizando operações de convolução para extrair características locais. Essas características são então passadas para a próxima camada por meio de funções de

ativação, que introduzem a não-linearidade no modelo (Goodfellow, Bengio e Courville 2016). A Figura 3 apresenta um exemplo de arquitetura convolucional.

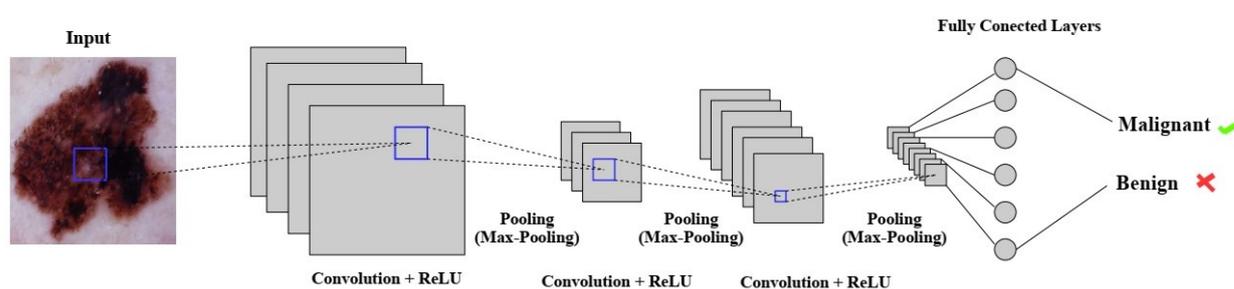


Figura 3 – Arquitetura Convencional de Redes Neurais Convolucionais. Adaptado de (Zhou, Greenspan e Shen 2017)

As camadas de *pooling* são utilizadas para reduzir a dimensionalidade da saída e tornar a representação mais compacta, preservando as características mais importantes (LeCun, Bengio e Hinton 2015). Geralmente, a operação de *pooling* mais utilizada é o *Max-Pooling*, que seleciona o valor máximo de uma região específica. Esse processo ajuda a melhorar a invariância a pequenas alterações na localização das características (Goodfellow, Bengio e Courville 2016). Já as camadas totalmente conectadas são responsáveis por combinar as características extraídas pelas camadas anteriores e realizar tarefas específicas, como classificação ou detecção de objetos (Kelleher 2019)

Durante o treinamento, as arquiteturas convolucionais se ajustam automaticamente por meio dos pesos das camadas para otimizar a função de perda, utilizando algoritmos de otimização, como o *Stochastic Gradient Descent* (SGD) (LeCun, Bengio e Hinton 2015). Neste caso, o processo de treinamento é supervisionado, sendo realizado por meio de exemplos rotulados, permitindo que o modelo aprenda a generalizar e fazer previsões precisas em dados não vistos anteriormente (Goodfellow, Bengio e Courville 2016). A capacidade de aprendizado de características relevantes e hierárquicas faz com que essas arquiteturas sejam poderosas na classificação de imagens médicas, principalmente em imagens de lesões de pele (Kawahara, BenTaieb e Hamarneh 2016).

3.2.2 Redes Neurais Convolucionais Residuais

As Redes Neurais Convolucionais Residuais, também conhecidas como *Residual Neural Networks* (ResNets), são uma variação das redes convolucionais convencionais. Esse tipo de arquitetura se destaca por sua capacidade de treinar redes mais profundas, superando os desafios de degradação de desempenho que podem ocorrer com inúmeras camadas convolucionais (He et al. 2015). O principal conceito por trás das ResNets é o uso de conexões residuais. Essas conexões permitem que as informações fluam diretamente através das camadas da rede (He et al. 2016).

As conexões residuais são consideradas atalhos na arquitetura das ResNets. Esses atalhos permitem que a rede aprenda representações residuais, ou seja, as diferenças entre o estado atual e o ideal (He et al. 2015). As representações obtidas são então utilizadas para atualizar as conexões residuais, melhorando a capacidade de aprendizado (Laina et al. 2016). A Figura 4 apresenta um exemplo de arquitetura desse tipo.

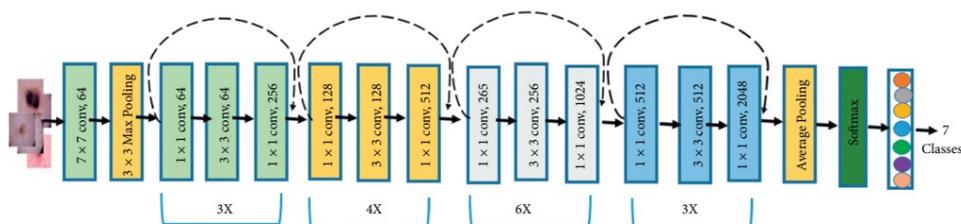


Figura 4 – Arquitetura Redes Neurais Convolucionais Residuais (ResNet). Adaptado de (Hasan et al. 2021).

A arquitetura de uma ResNet consiste em blocos residuais, que contêm várias camadas convolucionais e de ativação. Dentro de cada bloco, a entrada é alimentada nas camadas convolucionais, e a saída dessas camadas é adicionada à entrada original por meio das conexões residuais (He et al. 2016). A saída resultante é então tratada por uma função de ativação, antes de ser alimentada para o próximo bloco.

As conexões residuais permitem treinar arquiteturas com mais camadas, pois as informações relevantes são preservadas e propagadas de forma mais eficaz (He et al. 2016). Uma das vantagens das ResNets é a capacidade de alcançar um desempenho superior em tarefas de visão computacional, como reconhecimento de objetos e classificação de imagens. Além disso, aumentar o número de conexões residuais tem demonstrado melhorar a precisão dos modelos, permitindo a aprendizagem de características mais complexas e sutis (Laina et al. 2016).

Outra característica relevante das ResNets é a sua eficiência no uso de recursos computacionais, uma vez que a utilização de atalhos reduz a quantidade de parâmetros treináveis em comparação com outras técnicas de redes convolucionais de mesma profundidade (He et al. 2015). Por fim, essas redes também são uma abordagem que resolve o desafio da degradação do desempenho, pois as conexões residuais permitem o fluxo eficiente de informações, melhorando a capacidade de aprendizado e permitindo a construção de modelos mais precisos (Laina et al. 2016).

3.2.3 Redes Geradoras Adversárias

As Redes Geradoras Adversárias, também conhecidas como *Generative Adversarial Networks* (GANs), são métodos de aprendizado profundo utilizados na geração de imagens sintéticas. Esse método é composto por duas redes neurais

profundas: a primeira rede é conhecida como o gerador (*generator*) e a segunda como o discriminador (*discriminator*) (Goodfellow, Bengio e Courville 2016). A Figura 5 apresenta o funcionamento básico dessas redes.

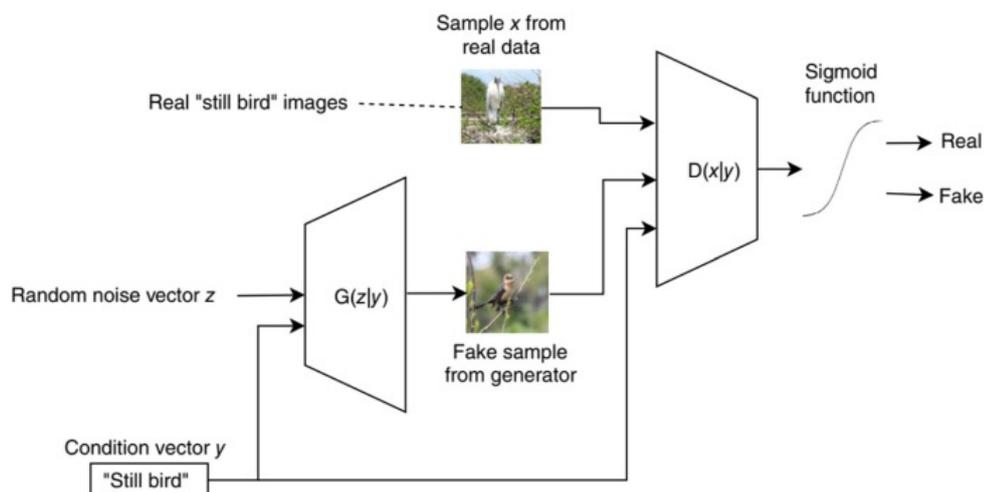


Figura 5 – Funcionamento das Redes Geradoras Adversárias. Adaptado de (Kelleher 2019).

O gerador é responsável por produzir imagens sintéticas que se assemelham às imagens reais. Esse gerador recebe uma entrada aleatória e a transforma em uma amostra de imagem sintética. O objetivo do gerador é aprender a mapear o espaço de entrada aleatória para o espaço de saída desejado, produzindo imagens que sejam indistinguíveis das reais (LeCun, Bengio e Hinton 2015). Já o discriminador atua como um classificador binário. Ele recebe como entrada as imagens reais e as imagens sintéticas produzidas pelo gerador e tenta diferenciá-las como reais ou *fakes*. O objetivo do discriminador é aprender a distinguir corretamente entre as imagens reais e as sintéticas (Goodfellow, Bengio e Courville 2016).

Durante o treinamento, o gerador e o discriminador são treinados simultaneamente em um processo conhecido como *minmax*. O gerador busca melhorar sua capacidade de enganar o discriminador, gerando imagens sintéticas que se tornam mais difíceis de serem distinguidas das reais (Kelleher 2019). Já o discriminador busca melhorar sua habilidade de diferenciar corretamente os dados reais e sintéticos. Essa competição entre o gerador e o discriminador resulta em um processo iterativo de aprimoramento, onde ambos se tornam cada vez melhores (Goodfellow, Bengio e Courville 2016).

As GANs se tornaram relevantes em aplicações que necessitam de geração de imagens sintéticas, tanto no processo de criação de imagens com iluminação e cores corrigidas, como na produção de imagens semelhantes para o balanceamento de bases de imagens (Salvi et al. 2022, Ferlay et al. 2021). Um problema ao utilizar essas redes está no treinamento, pois treinar arquiteturas desse tipo é uma atividade complexa, sendo necessário encontrar um equilíbrio entre o gerador e o discriminador. Além disso, o

treinamento consome uma quantidade significativa de recursos computacionais (Qin et al. 2020).

3.2.4 Autoencoders

Os *autoencoders* são uma técnica de aprendizado não supervisionado popular por sua capacidade de aprender representações compactadas de dados. A arquitetura de um *autoencoder* é dividida em codificador (*encoder*) e decodificador (*decoder*) (Kelleher 2019). O *encoder* é uma rede utilizada para reduzir os dados de entrada para uma representação de menor dimensionalidade, utilizando camadas que implementam transformações lineares e funções de ativação não lineares. Essa representação de menor dimensionalidade é processada pelo *decoder*, que tenta reconstruir a entrada original a partir do formato comprimido (Goodfellow, Bengio e Courville 2016). A Figura 6 apresenta a sistemática de funcionamento dessa metodologia.

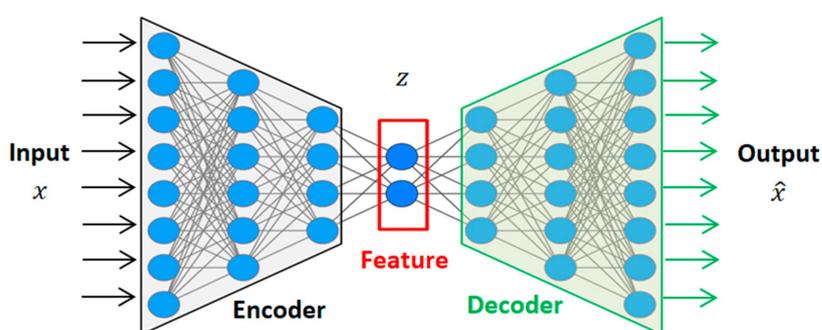


Figura 6 – Autoencoders. Adaptado de (Kelleher 2019).

Uma propriedade dos *autoencoders* é que eles conseguem aprender representações latentes significativas, mesmo em conjuntos de dados de alta dimensionalidade e complexidade (Goodfellow, Bengio e Courville 2016). Essa técnica pode ser utilizada para diversas tarefas, como redução de dimensionalidade, remoção de ruído, detecção de anomalias e até mesmo para geração de dados sintéticos (Kelleher 2019). Além disso, quando associada ao aprendizado profundo, oferece abordagens poderosas em representações eficientes e compactas, permitindo a extração de informações valiosas e a realização de tarefas complexas (LeCun, Bengio e Hinton 2015).

3.2.5 Aumento de Dados

As técnicas de aumento de dados (*Data Augmentation*) são amplamente utilizadas no campo do aprendizado profundo para diversificar os dados de treinamento. O objetivo dessas técnicas é melhorar a capacidade dos modelos de generalizar e obter um desempenho eficaz em diferentes cenários (Chlap et al. 2021). O funcionamento básico dessas técnicas é aplicar transformações em dados existentes e produzir novos dados sintéticos semelhantes aos originais (Rashid, Tanveer e Khan 2019). Essa aplicação pode

envolver operações como rotações, deslocamentos, espelhamento, alterações de brilho e contraste, além da utilização de filtros e transformações geométricas (Pham et al. 2018).

Ao expor uma arquitetura de aprendizado profundo a uma maior variedade de exemplos no treinamento, o aumento de dados ajuda a balancear o conjunto de treino e evitar o sobreajuste (*overfitting*) (Shorten e Khoshgoftaar 2019). Além disso, esse tipo de técnica pode auxiliar no aprendizado de características mais robustas e invariantes, úteis para a generalização em diferentes situações.

Os princípios do aumento de dados envolvem a preservação da semântica dos dados originais, a aplicação de transformações realistas e a criação de novas instâncias que representem variações encontradas nos dados reais (Shorten e Khoshgoftaar 2019). As técnicas desse tipo desempenham um papel importante no problema de desbalanceamento entre classes de imagens de lesões de pele utilizadas para treinamento, melhorando o desempenho e a generalização dos modelos de aprendizado profundo (Rashid, Tanveer e Khan 2019).

3.2.6 Transferência de Aprendizado e o Ajuste Fino

A transferência de aprendizado (*transfer learning*) também é uma técnica necessária nas aplicações de aprendizado profundo. O principal conceito associado a essa técnica é o aproveitamento do conhecimento adquirido durante o treinamento anterior de uma arquitetura em uma tarefa específica, aplicando-o a um problema semelhante (Zhang et al. 2020). Essa sistemática permite que os modelos pré-treinados sirvam como base para novos treinamentos em problemas que possuem relações próximas (Torrey e Shavlik 2010).

Existem duas abordagens principais para a transferência de aprendizado: extração de características e ajuste fino (*fine tuning*). Na extração de características, as primeiras camadas da rede convolucional pré-treinada são usadas como um extrator de características (Swati et al. 2019). Essas camadas aprendem a detectar características de baixo nível, como bordas e texturas. Em seguida, essas características extraídas são utilizadas para alimentar um novo modelo, treinando-o especificamente para a tarefa desejada (Torrey e Shavlik 2010).

No ajuste fino, além de utilizar as características extraídas, as camadas finais da rede neural pré-treinada são adaptadas para uma nova tarefa (Zhang et al. 2020). Isso envolve continuar o treinamento do modelo usando um conjunto de dados diferente, rotulado especificamente para a tarefa em questão. O ajuste fino permite que o modelo refine suas representações de alto nível e se adapte às nuances da tarefa específica (Torrey e Shavlik 2010).

Na literatura, existem muitos modelos pré-treinados com diversas imagens. No

entanto, para o problema do câncer de pele, modelos pré-treinados com o ImageNet costumam se ajustar melhor na adaptação para a classificação correta (Morid, Borjali e Fiol 2021). Esse conjunto possui um tamanho considerável, sendo adaptável a qualquer problema, pois as imagens contidas nele são genéricas e relacionadas a substantivos do cotidiano (Kornblith, Shlens e Le 2019). O conhecimento dos modelos pré-treinados com essa base pode ser aproveitado como conhecimento inicial para o problema do câncer de pele, evitando um treinamento do zero (Kim et al. 2022).

A transferência de aprendizado e o ajuste fino no aprendizado profundo são essenciais para aproveitar o conhecimento prévio e acelerar o treinamento de arquiteturas profundas. Ao extrair características ou ajustar modelos pré-treinados, é possível melhorar a eficiência, a eficácia e a generalização na detecção de lesões de pele malignas (Perez, Avila e Valle 2019).

3.3 Segmentação Semântica

A segmentação semântica é uma técnica de processamento de imagens que categoriza objetos em imagens digitais (Taghanaki et al. 2021). Ao contrário de outras abordagens convencionais, que dependem de bordas e limites de regiões para destacar áreas de interesse, técnicas desse tipo identificam regiões específicas e atribuem um rótulo semântico a cada *pixel* que compõe a imagem. Isso quer dizer que há uma predição desses *pixels* em regiões que correspondem a objetos específicos, aumentando a compreensão da cena e uma melhor definição de regiões (Guo et al. 2018).

A aplicação da segmentação semântica fornece informações detalhadas sobre a composição de uma imagem, possibilitando destacar objetos com maior sutileza. Essa capacidade permite o uso de outras abordagens computacionais para a classificação dos *pixels* (Guo et al. 2018). Em particular, quando aplicada a imagens médicas, a segmentação semântica apresenta resultados promissores (Taghanaki et al. 2021). A Figura 7 apresenta um exemplo de aplicação para imagens de lesões de pele.

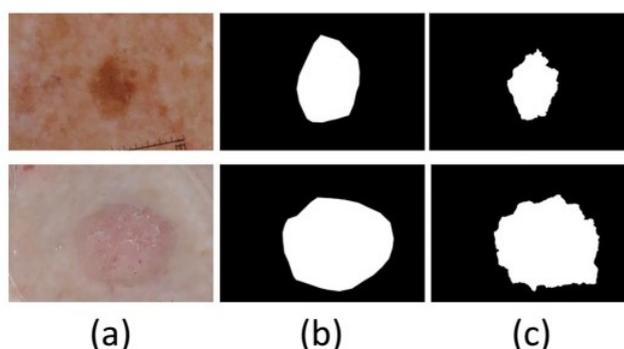


Figura 7 – Segmentação semântica. Adaptado de (Hong et al. 2023). (a) Imagem de Entrada. (b) Máscara do Especialista, (c) Máscara de Segmentação Semântica.

As técnicas de segmentação semântica são diversas, e muitas delas também se beneficiam de arquiteturas de aprendizado profundo, principalmente Redes Neurais Convolucionais e *autoencoders* (Guo et al. 2018). A utilização de modelos profundos permite uma melhor definição dos objetos nas imagens, que não são atrelados apenas a bordas e contornos. Isso faz com que o processo de segmentação seja analítico e preciso (Taghanaki et al. 2021, Silva et al. 2023).

3.4 Iluminação e Cor em Imagens Digitais

A iluminação e a cor são os principais componentes na percepção de imagens digitais, afetando significativamente como visualizamos e classificamos os objetos em cena (Penczek, Boynton e Splett 2014). A iluminação está relacionada à quantidade de luz disponível em um ambiente, influenciando diretamente a visualização de objetos e afetando cores, sombras, contraste e texturas (Gonzalez e Woods 2009).

Existem dois tipos principais de iluminação: natural e artificial. A iluminação natural, proveniente de fontes como o sol, varia ao longo do dia. Já a iluminação artificial é produzida por fontes criadas pelo ser humano, como lâmpadas e *flashes* de câmera (Gonzalez e Woods 2009). Cada fonte de luz possui características específicas, incluindo temperatura, intensidade e direção. A temperatura da fonte luminosa pode variar de tons quentes a frios, sendo importante para a interpretação visual. Essa característica pode realçar detalhes cromáticos, texturas e formas, garantindo uma representação da imagem fidedigna ao mundo real (Penczek, Boynton e Splett 2014).

Por outro lado, a cor é definida pela maneira como diferentes superfícies absorvem e refletem a luz. Nas imagens digitais, as cores são capturadas por sensores e representadas no espaço de cor RGB (*Red, Green, Blue*). Uma tonalidade é o resultado da combinação desses três componentes (Gonzalez e Woods 2000). Os aspectos de cor influenciam diretamente as tonalidades cromáticas, produzindo tons que variam de mais quentes a mais frios. Essa variação afeta a percepção e a representação das cores na imagem (Penczek, Boynton e Splett 2014). A Figura 8 ilustra como variações na iluminação podem alterar a tonalidade de cor percebida.



(a) Luz amarela.

(b) Luz branca.

Figura 8 – Variação de tonalidade de cor conforme a iluminação.

Outros termos ligados à iluminação e cor são contraste e brilho. O brilho é definido como uma medida que determina se uma imagem é clara ou escura. Já o contraste refere-se à diferença entre iluminação e cor, ou seja, se um objeto tem uma representação que pode ser distinguível dos demais. Em termos técnicos, o contraste é a diferença entre os elementos claros e escuros na imagem (Gonzalez e Woods 2009).

Além dos conceitos tratados, uma característica que também influencia a representação de iluminação e cores em imagens é o uso de dispositivos diferentes (Petschnigg et al. 2004). Ao capturar a mesma cena, mesmo sob condições de iluminação iguais, podem ocorrer discrepâncias nos aspectos de cores (Penczek, Boynton e Splett 2014). Essas diferenças podem ser devidas a fatores técnicos, de configuração ou de qualidade dos sensores. Também podem estar relacionadas à calibração de cor, à lente utilizada ou às configurações de software empregadas para processar a imagem (Poucke et al. 2010).

É possível utilizar técnicas para reduzir os impactos negativos causados pela variação da iluminação e os reflexos de cores. Essas técnicas são utilizadas para ajustar a representação da luminosidade, fazendo com que as cores sejam refletidas fielmente às condições dos objetos em cena (Gonzalez e Woods 2009). Dessa forma, é possível obter resultados mais precisos e confiáveis na análise e interpretação das imagens capturadas (Iyatomi et al. 2011).

Algumas técnicas convencionais para realizar a correção de iluminação e cor em imagens de lesões de pele são baseadas em histogramas e balanceamento de branco (Salvi et al. 2022, Barata, Celebi e Marques 2018). Essas técnicas aplicam modelos matemáticos para ajustar as propriedades de luminosidade, contraste e cor, partindo diretamente dos valores dos *pixels*, a fim de melhorar a qualidade visual (Gonzalez e Woods 2009).

Outras técnicas mais modernas também utilizam aprendizado profundo para realizar a correção de iluminação e cor, baseando-se nos princípios de *Local Color Distributions*, *Neural Radiance Fields* e *Deep Symmetric Network* (Wang, Xu e Lau 2022, Wang et al. 2023, Zhao et al. 2021). Essas técnicas podem adotar aprendizado supervisionado ou não supervisionado para recriar uma imagem corrigida a partir de uma imagem problemática. Nas próximas subseções, são tratados os conceitos e princípios de todas as sistemáticas de correção citadas.

3.4.1 Técnicas Baseada em Histogramas

A correção baseada em histogramas prioriza o ajuste da distribuição do brilho ou cor de uma imagem, visando melhorar o contraste e os aspectos de iluminação. O cálculo de ajuste é realizado a partir da análise de um gráfico com a distribuição dos valores dos

pixels da imagem, conhecido como histograma. A alteração dos valores da distribuição reflete na modificação do brilho e do contraste (Gonzalez e Woods 2009).

A equalização do histograma permite alcançar valores de distribuição uniforme para amenizar os efeitos de baixa iluminação ou excesso dela. Isso também reflete na mudança das tonalidades das cores nas imagens (Gonzalez e Woods 2009). A Figura 9 exemplifica a aplicação de equalização de histograma para uma imagem de lesão de pele.

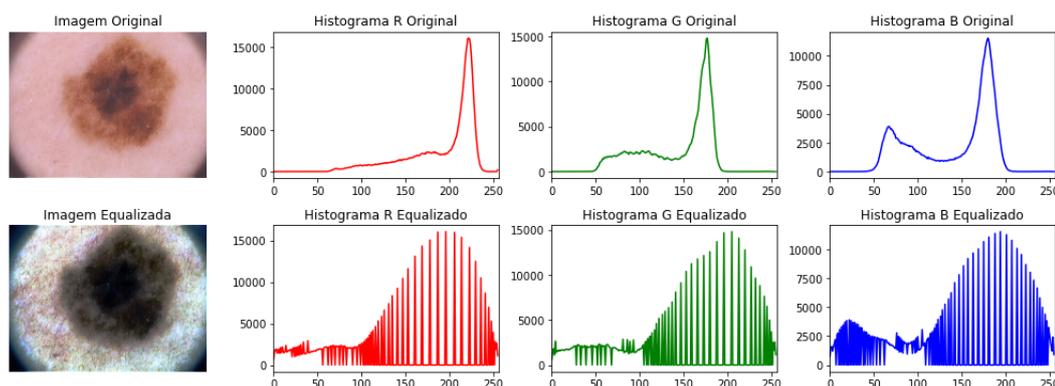


Figura 9 – Correção de iluminação com histograma.

Uma vantagem dos métodos baseados em histograma é a fácil implementação e a melhora expressiva de imagens que possuem baixo contraste. Em contrapartida, a correção com métodos desse tipo pode causar sobre-exposição ou sub-exposição em certas áreas da imagem, devido ao baixo poder analítico (Gonzalez e Woods 2009).

3.4.2 Técnicas Baseada em Balanço de Branco

As técnicas baseadas no balanço de branco priorizam a correção das cores em imagens digitais, para serem percebidas sob uma fonte de luz neutra. Portanto, esses métodos ajustam as cores para parecerem ter sido capturadas sob uma luz branca, neutralizando qualquer tonalidade indesejada (Afifi, Brubaker e Brown 2022). Quando uma imagem é capturada sob fontes de luz que não são brancas, as tonalidades de cor são alteradas, adicionando tons azuis, laranjas ou verdes (Gonzalez e Woods 2009).

As técnicas de correção do balanço de branco são utilizadas para identificar a temperatura da cor e compensar as divergências nos canais vermelho, azul e verde (RGB) da imagem. Isso neutraliza o excesso de cor e restabelece um equilíbrio mais natural (Afifi, Brubaker e Brown 2022). As principais técnicas incluem ajuste de ganho, *Perfect Reflector*, *Gray World*, *White Patch*, máxima entropia e balanço de branco por histograma (Fairchild 2013). A Figura 10 apresenta um exemplo de aplicação *Gray Words* na correção de cores para imagens de lesões de pele.

O ajuste de ganho modifica os canais RGB individualmente, buscando uma média de cor que equilibre a iluminação e, conseqüentemente, o aspecto de cor. O método *Gray*



(a) Imagem sem correção. (b) Imagem com correção.

Figura 10 – Correção de branco com ajuste nos canais RGB.

World baseia-se no princípio de que a média das cores em uma cena deve ser cinza, usando esse padrão para corrigir desvios cromáticos (Afifi, Brubaker e Brown 2022). Por outro lado, o *White Patch* se fundamenta no ponto mais brilhante da imagem, ajustando as demais cores em função desse referencial. A técnica de máxima entropia escolhe o balanço de branco que maximiza a entropia de cor, promovendo a maior variação possível para uma iluminação equilibrada (Fairchild 2013).

Existem diversas técnicas para correção dos aspectos de cor baseados no balanço de branco. No entanto, a maioria dessas técnicas lida com o ajuste individual de cada canal de cor, proporcionando um equilíbrio entre os *pixels* de uma imagem, melhorando tanto a sub-iluminação quanto a sobre-iluminação (Gonzalez e Woods 2009). No caso do câncer de pele, identificar uma técnica desse tipo que mantenha a qualidade da imagem é desafiador, dada a importância da cor para o diagnóstico do câncer de pele maligno (Barata, Celebi e Marques 2014).

3.4.3 Local Color Distributions

As Distribuições de Cores Locais (*Local Color Distributions* - LCDs) são um conceito utilizado na implementação de técnicas de correção de iluminação e cor com aprendizado profundo. Esse conceito indica que diferentes áreas de uma imagem podem ter distribuições de cores únicas sob variações de condições de iluminação (Nascimento, Amano e Foster 2016). Isso significa que cada região pode ter suas próprias características de cor, influenciadas por sombras, luz direta, reflexos, entre outros aspectos que distorcem a percepção visual (Wang et al. 2019).

As LCDs possibilitam o desenvolvimento de técnicas supervisionadas para identificar e corrigir problemas de iluminação em imagens que estão subexpostas (muito escuras) ou superexpostas (muito claras) (Wang, Xu e Lau 2022). O objetivo dessas técnicas é estabelecer uma correlação entre as condições extremas de iluminação para prever a situação ideal (Nascimento, Amano e Foster 2016). A Figura 11 apresenta um exemplo de aplicação de LCDs com uma arquitetura de aprendizado profundo.

A partir da identificação da luminosidade, as técnicas baseadas em LCDs



Figura 11 – Correção de iluminação e cor com princípios de LCDs e aprendizado profundo. Adaptado de (Yang et al. 2021).

podem aprimorar esse aspecto e melhorar a percepção de cores de maneira seletiva, corrigindo discrepâncias de exposição sem comprometer a qualidade da imagem. As distribuições são utilizadas como base para a correção de contraste e brilho, além de outros parâmetros para realçar detalhes, corrigir a exposição e melhorar a percepção visual (Nascimento, Amano e Foster 2016). Ao focar nas características locais de cor, as LCDs podem alcançar resultados mais naturais, mantendo a integridade da cena (Wang, Xu e Lau 2022).

A aplicação dos princípios de LCDs, por exemplo, com redes neurais convolucionais ou residuais, possibilita a criação de modelos que aprendem a mapear e distinguir distribuições de cores locais. O modelo analisa essas distribuições e aplica a transformação aprendida para corrigir a iluminação em cada bloco da imagem (Wang, Xu e Lau 2022). Para imagens de lesões de pele, a aplicação desse tipo de modelo resulta em uma imagem final onde a iluminação é mais uniforme e realista.

3.4.4 *Neural Radiance Fields*

Os Campos de Radiância Neural (*Neural Radiance Fields* - NeRFs) representam um avanço significativo na combinação de visão computacional, gráficos tridimensionais (3D) e aprendizado profundo (Xu et al. 2022). Essas técnicas combinam renderização para recriar cenas em duas dimensões (2D) para uma representação 3D. Uma característica dessas técnicas é a capacidade de extrair a complexidade da luz, permitindo recriar visualizações com alta definição, preservando a aparência real sob diversas condições de captura (Mildenhall et al. 2021).

Os NeRFs modelam o espaço como um campo volumétrico (Wang et al. 2023). Isso significa que cada ponto no espaço é definido por uma cor e uma densidade específicas, permitindo uma abordagem de problemas como dispersão de luz, neblina,

transparência e reflexão (Martin-Brualla et al. 2021). O aprendizado profundo é um componente essencial dessas técnicas, sendo a base para modelar o campo de radiância. Este, por sua vez, é uma função contínua que mapeia as coordenadas espaciais e as direções de visualização para valores de cor e densidade de luz.

Uma rede neural convolucional pode ser associada aos NeRFs para capturar características locais de um conjunto de imagens da mesma cena em ângulos diferentes (Wang et al. 2023). Uma função de perda lida com a comparação das cores previstas e reais, minimizando o erro entre elas. A técnica de retropropagação pode ser usada para calcular os gradientes das perdas em relação aos pesos da rede, permitindo a atualização dos mesmos. A cor de cada *pixel* na imagem de saída é determinada pela integração dos valores de radiância, considerando a atenuação causada pela densidade volumétrica (Martin-Brualla et al. 2021). A Figura 12 apresenta um exemplo de aplicação de um método NeRF para recriar uma imagem com múltiplas variações de aspecto de cor.



Figura 12 – Renderização com NeRF para recriação de imagens com aspecto de cor variável. Adaptado de (Martin-Brualla et al. 2021).

O uso de NeRFs em técnicas de correção de iluminação e cor permite ajustar arquiteturas profundas para prever essas características no processo de renderização (Wang et al. 2023). Inicialmente, o uso da radiância possibilita modelar tanto a cor quanto a densidade de um ponto, incluindo a iluminação (Mildenhall et al. 2021). No entanto, a formulação padrão do NeRF não separa os componentes de iluminação e cor dos objetos, pois eles já estão integrados (Xu et al. 2022).

A partir de um conjunto limitado de imagens sob vários ângulos e iluminações, o princípio do NeRF permite prever e recriar visualizações de uma mesma imagem sob novos pontos de vista, inclusive corrigindo detalhes finos (Wang et al. 2023). Apesar de todas as vantagens na correção de iluminação e cor, a renderização com NeRFs é volumétrica e consome um grande poder computacional, devido à quantidade

de pontos que devem ser processados. No entanto, os resultados alcançados são excepcionais (Wang et al. 2023).

3.4.5 *Deep Symmetric Network*

As Redes Simétricas Profundas (*Deep Symmetric Networks* - DSNs) incorporam conceitos de simetria e redes neurais reversíveis (*Invertible Neural Networks* - INNs) (Zhao et al. 2021). O princípio de reversibilidade refere-se à condição de que a rede possa ser revertida, ou seja, os dados de entrada podem ser recuperados a partir dos dados de saída. Já a simetria refere-se às propriedades dos dados ou das funções que permanecem invariantes sob certas transformações (Ardizzone et al. 2018).

O uso de INNs é indicado quando o problema exige a reversão exata da transformação dos dados. Para isso, são usadas camadas que dividem a entrada em dois conjuntos, sendo que cada um desses conjuntos passa por transformações diferentes. Um diferencial dessa abordagem é implementar o princípio da inversa, ou seja, cada camada da rede aprende transformações complexas, mas também aprende facilmente a inverter o resultado. Em suma, cada entrada é mapeada exatamente para uma saída e vice-versa (Ardizzone et al. 2018).

As propriedades de simetria e reversibilidade são úteis em uma vasta gama de problemas. Na correção de aspectos de iluminação e cor em imagens digitais, essas propriedades permitem modelar imagens, melhorando as condições dessas características (Ardizzone et al. 2018). Além disso, novas técnicas de DSNs com esse propósito também utilizam arquiteturas com mecanismos de atenção para identificar áreas em uma imagem que necessitam de mais foco durante o processo (Zhao et al. 2021). A Figura 13 apresenta um exemplo de aplicação de uma técnica desse tipo.



Figura 13 – Correção dos aspectos de iluminação e cor com princípios de DSN. Adaptado de (Zhao et al. 2021).

O uso de técnicas de DSNs para correção de iluminação e cor requer imagens em pares, ou seja, imagens subexpostas ou superexpostas e suas versões corrigidas. Geralmente, essas imagens são criadas por especialistas. Os modelos aprendem a

mapear os pares de imagens, ajustando seus parâmetros para minimizar as diferenças entre as imagens, resultando em versões de maior qualidade visual (Zhao et al. 2021). Portanto, essas técnicas oferecem uma solução robusta e eficiente para a melhoria de imagens e outras aplicações que exigem precisão e manutenção da integridade dos dados.

3.5 Considerações Finais

Este capítulo estabeleceu as bases teóricas da tese, proporcionando uma compreensão objetiva sobre os principais conceitos que envolvem a metodologia deste trabalho. Os principais conteúdos abordados foram os tipos de câncer de pele, técnicas de aprendizado profundo, segmentação semântica e correção de aspectos de iluminação e cor em imagens digitais. Além disso, também foram tratadas teorias sobre o aumento de dados, a transferência de aprendizado e o ajuste fino, primordiais no uso do aprendizado profundo. Em suma, este capítulo descreveu uma fundamentação teórica explorando todos os aspectos conceituais envolvidos na abordagem de validação, que será apresentada no próximo capítulo.

4 Abordagem de Classificação do Câncer de Pele

Este capítulo apresenta a abordagem de validação da tese, focando na detecção do câncer de pele em imagens digitais com correção de aspectos de iluminação e cor. A metodologia implementada segue etapas convencionais identificadas na revisão de literatura, que abrangem a seleção de bases de imagens, pré-processamento, segmentação, aumento de dados, desenvolvimento e análise de modelos de classificação.

4.1 Metodologia e Avaliação do Estudo

Como descrito na seção anterior, a pesquisa desenvolvida neste trabalho tem natureza aplicada com procedimentos experimentais. O objetivo é investigar e comprovar a importância dos aspectos de iluminação e cor na classificação do câncer de pele. A abordagem desenvolvida prioriza o uso de aprendizado profundo para validar as hipóteses, examinando o impacto dessas características na precisão do diagnóstico de lesões malignas.

O viés tratado na classificação está relacionado ao problema de aquisição de imagens de lesões de pele, uma vez que a detecção automatizada depende de bases de imagens capturadas com diversos dispositivos e sob diferentes condições de iluminação. A Figura 14 apresenta um fluxograma que expressa todas as etapas da implementação da abordagem.

A primeira etapa da abordagem trata da seleção das bases de imagens, com diversas classes de lesões previamente diagnosticadas por especialistas. A segunda etapa envolve o pré-processamento, que inclui a remoção do excesso de pelos das imagens e a aplicação de técnicas de correção de iluminação e cor. Já a terceira etapa é a segmentação das lesões, visando delimitar o melhor envoltório. É importante destacar que essa etapa é comum a todos os cenários testados, tratando as mesmas coordenadas para cortar as imagens, reduzindo a presença de pele saudável. Por fim, a última etapa trata da criação de modelos de aprendizado profundo, aplicando técnicas de aumento de dados, transferência de aprendizado e otimização de hiperparâmetros.

A validação da abordagem é feita pela análise e comparação de resultados entre os modelos gerados, definindo as melhores métricas conforme os cenários experimentais. Esses cenários foram tratados conforme a relevância dos fatores de iluminação e cor na identificação precisa de malignidades de pele. Para isso, a abordagem estabelece três cenários para destacar a importância desses fatores: sem correção, correção com

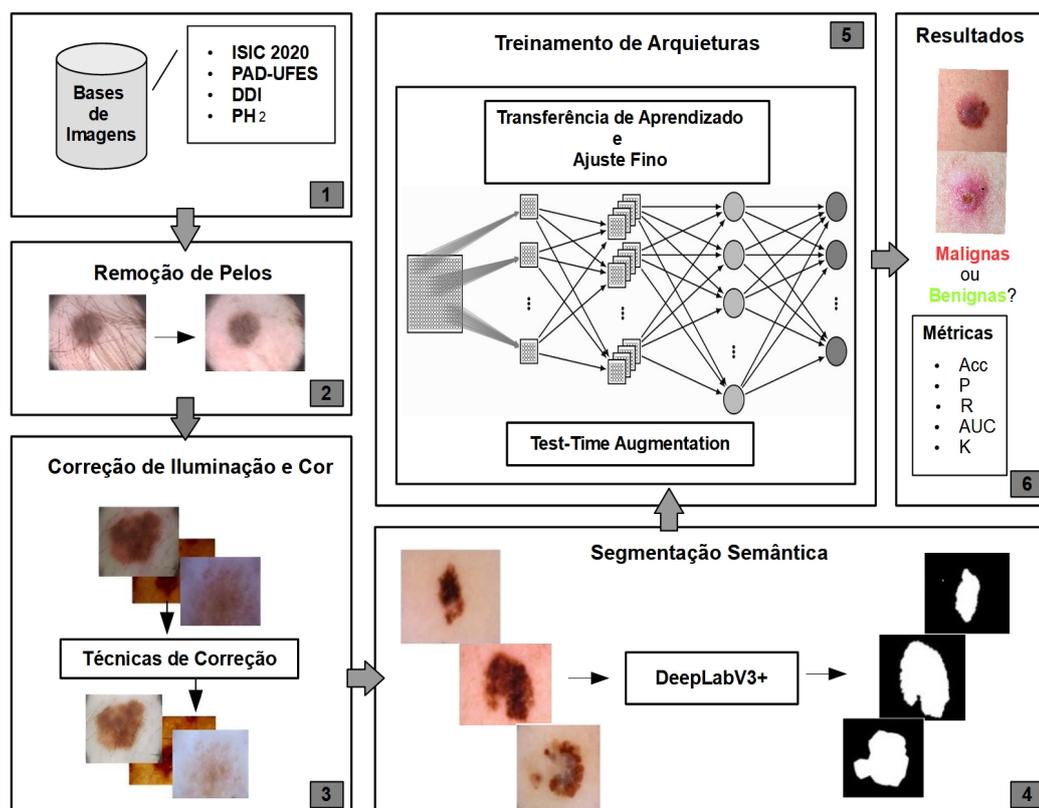


Figura 14 – Abordagem de validação para otimização de detecção do câncer de pele.

técnicas convencionais e correção com aprendizado profundo. As próximas seções detalham todas as metodologias aplicadas na execução da abordagem.

4.2 Bases de Imagens

As imagens sugeridas nesta abordagem fazem parte de bases públicas que já foram bem referenciadas na literatura, sendo elas: o *International Skin Imaging Collaboration* (ISIC 2020) ([International Skin Imaging Collaboration 2020](#)), o Programa de Assistência Dermatológica e Cirúrgica da Universidade Federal do Espírito Santo (PAD-UFES-20) ([Pacheco et al. 2020](#)), o *Diverse Dermatology Images* (DDI) ([Daneshjou et al. 2022](#)) e o PH² ([Mendonça et al. 2013](#)). Essas bases contêm imagens clínicas, dermatoscópicas e convencionais, capturadas por diferentes dispositivos. A Tabela 3 mostra um somatório entre quantidade de imagens malignas e benignas dessas bases.

Após a análise visual das imagens das bases, é perceptível uma variação nos aspectos de iluminação e cor. Essas variações constituem o viés do nosso estudo, no qual é explorado a importância desses atributos na classificação correta do câncer de pele utilizando modelos de aprendizado profundo. É relevante destacar que a abordagem desenvolvida trata a classificação binária, considerando o melanoma e os carcinomas

Bases de Imagens	Benigno	Maligno	Total
ISIC 2020	32.542	584	33.126
PAD-UFES	1.215	1.094	2.309
DDI	644	219	863
PH ²	160	40	200
Total	34.561	1.937	36.498

Tabela 3 – Quantitativo de imagens por classes de lesões benignas e malignas.

como malignos e as demais lesões como benignas.

Outro ponto a ser destacado sobre as bases de imagens é o desequilíbrio entre as classes malignas e benignas. Essa condição pode comprometer a eficácia de qualquer modelo classificador, favorecendo indevidamente a classe majoritária. Por isso, é necessário adotar técnicas que permitam um melhor balanceamento entre as classes e evitar resultados incorretos. Na seção 4.6, serão apresentadas e discutidas as sistemáticas adotadas na abordagem desenvolvida.

4.3 Remoção de Pelos

O excesso de pelos é outro desafio que pode comprometer a precisão no diagnóstico automatizado (Barata, Celebi e Marques 2018). Na abordagem desenvolvida, o princípio da solução aplicada para amenizar o excesso desses artefatos nas imagens é tratá-los como estruturas curvilíneas. A técnica implementada para essa tarefa é baseada nos princípios de (Lee et al. 1997), porém, implementa algumas particularidades ao processo.

As operações de fechamento morfológico nas direções de 0°, 45° e 90° permitem identificar estruturas de pelos. Essas estruturas são tratadas em uma máscara que, ao final do processo, é aplicada à imagem de entrada. Essa aplicação faz com que a maioria dos pelos seja removida ou amenizada, melhorando o aspecto da imagem sem afetar a qualidade.

A técnica desenvolvida no processo funciona em três passos. O primeiro passo faz a decomposição da imagem em canais, seguidos da binarização de cada componente. Os filtros morfológicos de fechamento são aplicados no processo de rotação de cada componente, identificando curvas. O segundo passo é a combinação dos componentes para a formação de uma máscara dos pelos. Essa máscara é aplicada na imagem de entrada em um terceiro passo, possibilitando a exclusão dos *pixels* de cada componente que corresponde a pelos. Ao fim desse passo, são realizadas operações de interpolação bilinear para reconstruir as áreas afetadas. A Figura 15 ilustra um exemplo da aplicação do método.

Um aspecto diferencial do método de correção implementado, se comparado

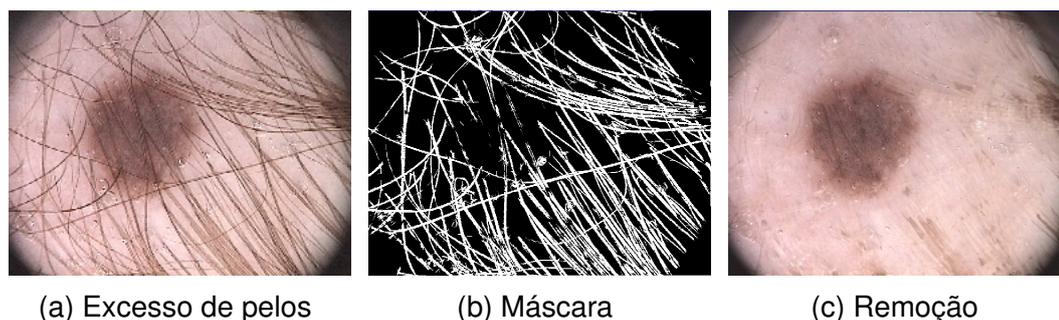


Figura 15 – Remoção de pelos com o método sugerido.

com (Lee et al. 1997), é a definição de um limite de tratamento para as imagens que possuem pelos. As imagens nas quais os pelos não predominam na maior parte ou aparecem em tons claros são excluídas do processo de correção. Por outro lado, as imagens que possuem uma quantidade significativa de pelos em tons escuros, próximos ao preto, são tratadas pelo método de correção.

Para selecionar as imagens, são utilizadas estatísticas de cores com limiares adaptativos. Isso visa restringir a correção apenas às imagens que realmente necessitam desse processo. Uma limitação do método é tratar pelos com uma fina espessura ou com base espessa. Em alguns casos, pode ocorrer a remoção parcial ou a não remoção desses tipos de pelos. A implementação da técnica sugerida neste trabalho foi realizada na linguagem Python, com o auxílio dos *frameworks* OpenCV, NumPy e SciPy.

4.4 Técnicas de Correção de Iluminação e Cor

Nessa parte da abordagem, são tratadas técnicas que permitem a correção das características de iluminação e cor nas imagens. Para isso, sugerimos duas sistemáticas de correção. Uma delas utiliza técnicas convencionais, amplamente aceitas na literatura para correção dessas características, enquanto a outra aplica novas técnicas de aprendizado profundo, analíticas no tratamento e correção da iluminação e cor.

As técnicas convencionais para correção da iluminação e cor aplicadas na abordagem foram o *Contrast Limited Adaptive Histogram Equalization*, *Shades of Gray* e *Perfect Reflector*. Já as técnicas de aprendizado profundo para lidar com essas características são baseadas em *Local Color Distributions (LCDs)*, *Neural Radiance Fields (NeRFs)* e *Deep Symmetric Network (DSN)*. As próximas subseções detalham cada técnica utilizada no processo de correção.

4.4.1 *Contrast Limited Adaptive Histogram Equalization*

O *Contrast Limited Adaptive Histogram Equalization (CLAHE)* é uma técnica de equalização de histograma projetada para aprimorar o contraste em imagens. Essa

metodologia é uma melhoria do *Adaptive Histogram Equalization* (AHE), que divide a imagem em regiões de tamanho predefinido, como 8x8 ou 16x16 *pixels*. Para cada uma dessas regiões, é realizada a equalização do histograma. Para suavizar as transições entre regiões e evitar bordas visíveis, essa técnica emprega interpolação bilinear (Reza 2004).

Uma consequência comum do processo de correção de iluminação e cor com o CLAHE é a amplificação de ruídos. Para mitigar esse problema, o método introduz um mecanismo de limitação do contraste. Quando um nível de cinza em uma região da imagem excede um limite pré-estabelecido, esse excesso é redistribuído de maneira uniforme ao longo do histograma. Esse procedimento evita o realce excessivo de áreas muito iluminadas da imagem, permitindo a correção de regiões subexpostas e superexpostas sem inserir excessos em ambas as condições, mantendo o contraste em limites que preservam a aparência real da cena (Musa, Rafi e Lamsani 2018).

Neste trabalho, a implementação do CLAHE foi realizada na linguagem Python com as ferramentas OpenCV e NumPy. Não utilizamos o espaço de cor RGB para tratar a correção de iluminação, e sim o espaço Lab*. Apenas o componente L é corrigido, mantendo intactos os demais componentes. Os parâmetros utilizados foram a divisão das regiões em tamanhos de 8x8 (*GridSize*), com limite de contraste (*TitleClipLimit*) igual a 2. A Figura 16 apresenta um exemplo de aplicação em imagens de lesões de pele.

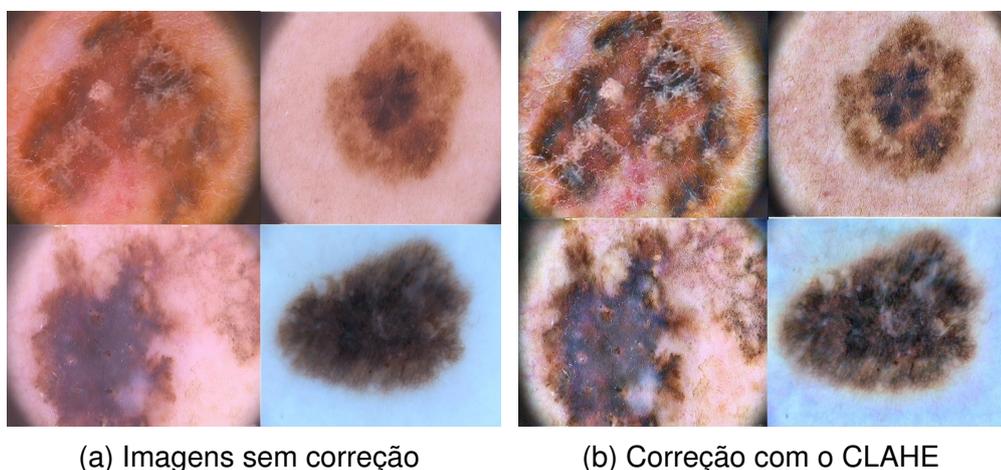


Figura 16 – Aplicação do método CLAHE para correção de iluminação e cor.

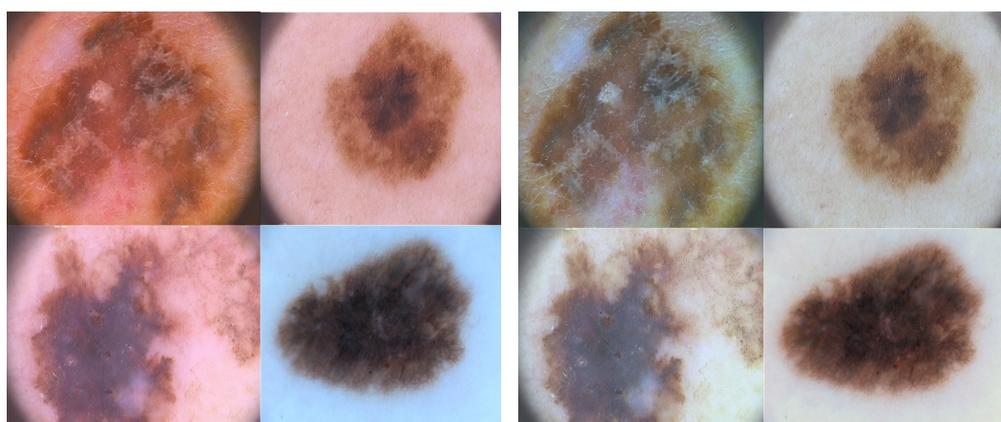
As vantagens de utilizar o CLAHE para correção de iluminação em imagens de lesões de pele incluem melhorias no contraste local, limitação da ocorrência de ruídos, melhorias na percepção visual de cores e a facilidade de personalização da técnica. A principal limitação deste método é o realce de artefatos indesejados, que podem estar ligados aos parâmetros utilizados.

4.4.2 Shades of Gray

O *Shades of Gray* (SoG) é um filtro estatístico para realizar o balanço de branco em imagens digitais. Esse filtro tem o objetivo de corrigir cores para que elas fiquem com aspecto natural sob diferentes condições de iluminação. Essa técnica define que uma imagem balanceada terá uma média de tons de cinza, considerada em todas as cores. Isso significa que a média de cinza deve ser igual em todos os componentes de cor da imagem (Finlayson e Trezzi 2004).

A aplicação do filtro SoG é dada pelo cálculo da média de todas as cores da imagem. Após isso, cada canal é ajustado para que a média se torne igual. Esse ajuste é dado por um fator de correção para cada canal, baseado no cinza-médio. Os fatores de correção são aplicados a cada *pixel* da imagem, ajustando a intensidade de cada componente para alcançar o equilíbrio de branco. Para evitar alterações que causem distorções nos aspectos de cor, é indicado limitar o ajuste dos canais, especialmente em imagens com alto contraste ou saturação.

A implementação do SoG neste trabalho foi realizada em Python, também utilizando as ferramentas OpenCV e NumPy. O método percorre os valores dos *pixels* da imagem e calcula as médias para cada componente de cor RGB. Uma característica do método desenvolvido é a utilização do expoente alfa (α) igual a 5, que permite calcular a opacidade da cor e manter um limiar de ajuste em cada canal. Esse valor também evita alterações extremas, mantendo uma média ponderada de valores por canal. A Figura 17 exemplifica a aplicação do método em imagens de lesões de pele.



(a) Imagens sem correção

(b) Correção com o SoG

Figura 17 – Aplicação do método SoG para correção de iluminação e cor.

O SoG é um método simples de correção de iluminação e cor para realizar o balanço de branco em imagens. Essa técnica mantém os aspectos regulares das imagens de lesões de pele, sendo uma estratégia relevante para manter a neutralidade de iluminação. O valor da média pode ser implementado de outras formas; no entanto, a média ponderada é uma alternativa viável associada ao parâmetro α . Um controle

inadequado da média pode fazer com que distorções de cores sejam inseridas na imagem, comprometendo a neutralidade da iluminação.

4.4.3 *Perfect Reflector*

O *Perfect Reflector* também é uma sistemática de correção baseada no princípio do balanço de branco. As técnicas desse tipo definem que partes de uma imagem superexposta devem ser corrigidas para que as condições de iluminação se tornem neutras (Eltayef, Li e Liu 2017). A correção por esse método é realizada neutralizando os desvios de cor, considerando a região de maior reflexão. Essa região é definida pela escolha do *pixel* de maior brilho e pela análise dos valores máximos por canal de cor (She et al. 2001).

O cálculo do fator de correção é uma característica de cada implementação. Esse cálculo pode considerar as diferenças entre intensidades ou um limiar de brilho. O fator de correção ajusta a intensidade das cores de forma que elas sejam percebidas como branco (Eltayef, Li e Liu 2017). O ajuste ocorre proporcionalmente para cada canal de cor, corrigindo os *pixels* identificados como reflexores, tornando-os mais próximos da cor branca (She et al. 2001). A Figura 18 exemplifica uma aplicação do método *Perfect Reflector* em imagens de lesões de pele.

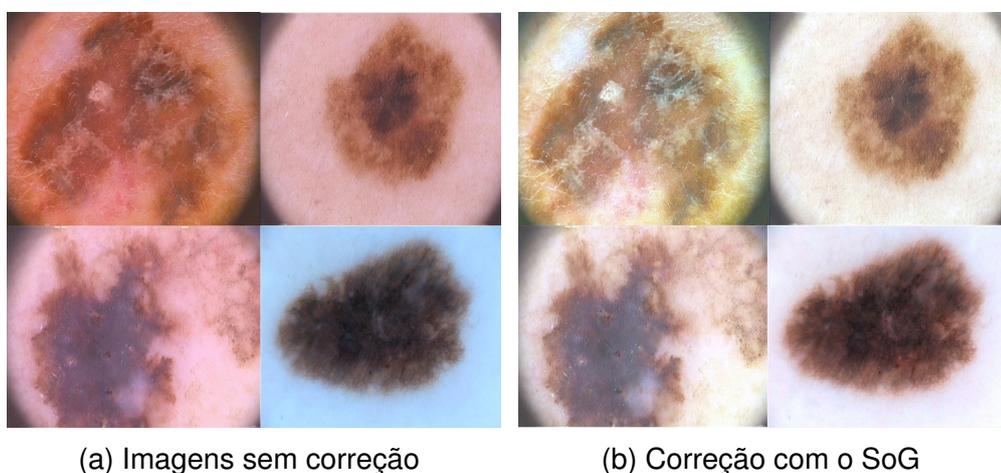


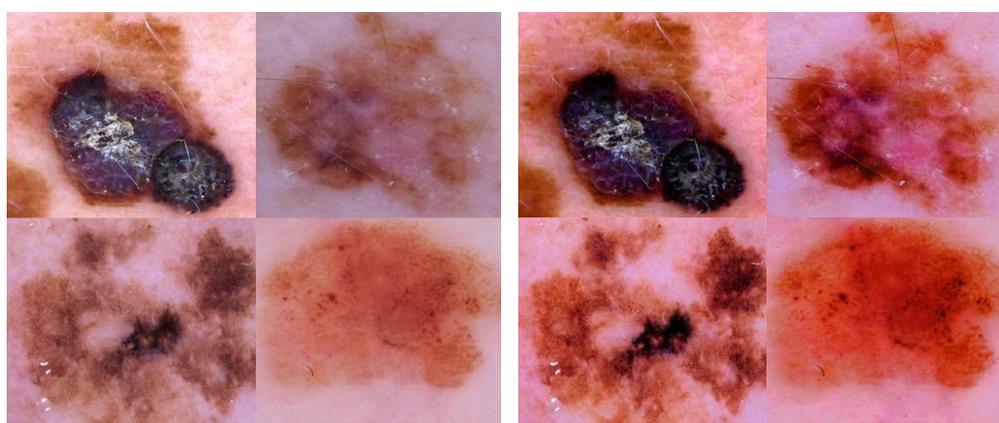
Figura 18 – Aplicação do método *Perfect Reflector* para correção de iluminação e cor.

Neste trabalho, a implementação do *Perfect Reflector* foi realizada em Python com o suporte do *framework* NumPy. A identificação dos *pixels* de reflexão é dada pelo valor máximo de cada canal, e o fator de correção é o percentil da divisão de 255 pelo valor máximo. Após o cálculo dos fatores, a correção é aplicada alterando os valores dos *pixels* por meio da multiplicação com o fator de correção. O resultado da aplicação desse método mostra uma uniformidade de iluminação; porém, em imagens com múltiplas fontes de luz, o método pode perder desempenho (Eltayef, Li e Liu 2017).

4.4.4 Local Color Distribution Prior Network

A *Local Color Distribution Prior Network* (LCDPNet) é uma metodologia de aprendizado profundo baseada em LCDs. Essa metodologia foi desenvolvida por (Wang, Xu e Lau 2022), visando tratar o problema de subexposição e sobre-exposição de iluminação em imagens digitais. A aplicação dessa técnica inclui um módulo denominado *Local Color Distribution Embedded* (LCDE), que extrai múltiplas LCDs por escalas. Esse módulo tem o intuito de modelar correlações entre diferentes regiões da imagem.

Outro mecanismo, denominado *Dual-Illumination Estimation*, desenvolve um aprendizado nas duas condições não uniformes de iluminação. Esse mecanismo melhora a estimação de sub-iluminação e sobre-iluminação. Para isso, faz a decomposição da imagem em um mapa de iluminação e em um mapa de reflexão com a imagem aprimorada. Porém, isso gera um problema de elevação dos valores dos *pixels*, tornando o método incapaz de melhorar a sobre-iluminação (Wang, Xu e Lau 2022). A solução desenvolvida na LCDPNet para esse problema é a implementação de um mecanismo de previsão de iluminação que considera uma imagem com sobre-iluminação como sendo uma sub-iluminação e uma imagem inversa. A Figura 19 apresenta um exemplo de aplicação às imagens de lesões de pele.



(a) Imagens sem correção

(b) Correção com o LCDPNet

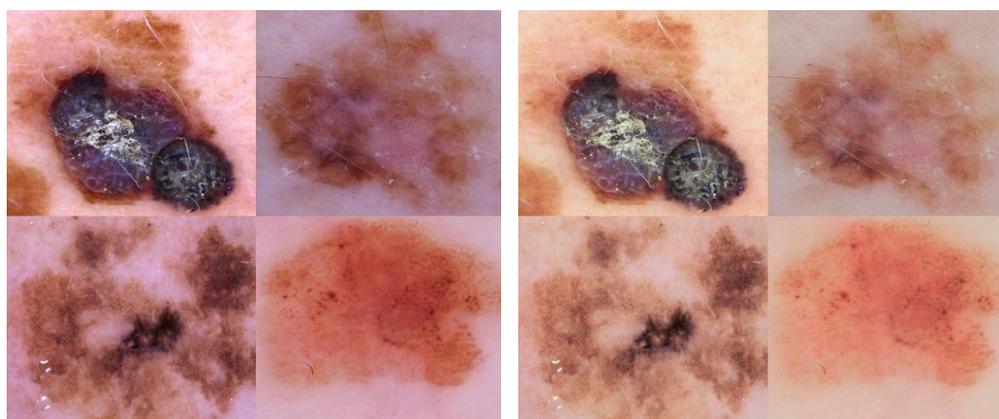
Figura 19 – Aplicação do método LCDPNet para correção de iluminação e cor.

Em outras palavras, a LCDPNet é uma metodologia que foca na identificação de distribuições locais de cores, que podem variar significativamente entre regiões, dependendo da iluminação. O seu objetivo é identificar e analisar características discriminativas em pequenas regiões das imagens. Outro fator que favorece essa arquitetura é o treinamento utilizando o conjunto MIT Adobe5k, personalizado por especialistas, que inclui imagens com sub-iluminação e sobre-iluminação (Wang, Xu e Lau 2022). A base LCDPNet para a implementação utilizada nesta etapa está disponível em (Wang, Xu e Lau 2022).

4.4.5 Low-Light Neural Radiance Field

O Low-Light Neural Radiance Field (LLNeRF) é uma técnica de aprimoramento de imagens digitais capturadas em condições de baixa luminosidade, desenvolvida por (Wang et al. 2023). Esta metodologia emprega uma abordagem de aprendizado profundo para modelar cenas complexas, permitindo a reconstrução de novas visões, geometria e características de reflexão da cena. Essa técnica é baseada em NeRF para criar imagens realistas, utilizando um processo de amostragem ao longo das dimensões e aplicando métodos tradicionais de renderização.

Diferente de outras técnicas de correção de iluminação, a LLNeRF utiliza os princípios de radiância para criar um mapeamento de iluminação, aprendendo ajustes complexos que não são obtidos do mapeamento entre imagens sub-iluminadas e sobre-iluminadas. É uma técnica que pode ser supervisionada ou não supervisionada para realce de luz. Uma função de perda personalizada foi desenvolvida para minimizar o erro de reconstrução, evitar o sobreajuste e garantir que as cores da imagem produzida sejam fidedignas às imagens originais (Wang et al. 2023). A Figura 20 apresenta um exemplo de aplicação em imagens de lesões de pele.



(a) Imagens sem correção

(b) Correção com o LLNeRF

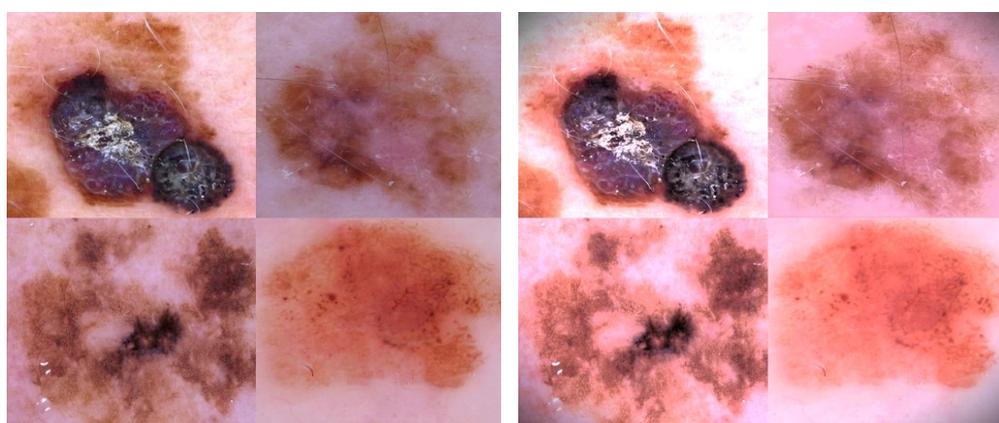
Figura 20 – Aplicação do método LLNeRF para correção de iluminação e cor.

A combinação dos aspectos dados na função de perda faz com que a arquitetura LLNeRF seja otimizada no processo de treinamento. Ao final do processo, a rede consegue prever mapeamentos de iluminação, produzindo imagens corrigidas e com aspectos melhorados de iluminação e cor (Wang et al. 2023). Essa arquitetura também foi treinada com um conjunto de imagens personalizado da base MIT Adobe5K, tornando-se um modelo capaz de recuperar detalhes claros, contraste distinto e cores naturais de imagens. A base para a implementação utilizada neste trabalho está disponível em (Wang et al. 2023).

4.4.6 Deep Symmetric Network

O *Deep Symmetric Network* (DSN), desenvolvido por (Zhao et al. 2021), é uma metodologia de aprendizado profundo diferenciada das anteriores. Essa metodologia sugere o uso de uma nova técnica denominada aprendizado recorrente com atenção, utilizando redes neurais simétricas para corrigir aspectos de sub-iluminação. Um ponto relevante dessa sistemática é fazer com que as características bidirecionais sejam aprendidas, evitando distorções de cores e recuperando o conteúdo da imagem com o módulo de atenção residual.

O DSN implementa um *Recurrent Residual-attention Module* (RRAM), que foi projetado para proporcionar ajustes incrementais de características de cores. Esse módulo permite um foco progressivo e aprimorado em áreas específicas da imagem. Em vez de aprender todas as correções de cores de uma única vez, divide a tarefa em várias etapas sequenciais. O *residual soft channel attention mechanism* é a parte crítica do RRAM, fazendo com que a arquitetura tenha foco nas informações estruturais independentemente do canal de cor. A Figura 21 apresenta um exemplo de aplicação do DSN em imagens de lesões de pele. A base utilizada para a implementação desta técnica neste trabalho está disponível em (Liu et al. 2023).



(a) Imagens sem correção

(b) Correção com o LLNeRF

Figura 21 – Aplicação do método DSN para correção de iluminação e cor.

4.5 Segmentação e Delimitação da Área de Interesse

A redução da imagem para representar somente a área da lesão na abordagem é realizada por meio de um processo de segmentação. Nessa etapa, o objetivo não é obter definições precisas de bordas, mas sim traçar boas coordenadas por meio de caixas delimitadoras (*bounding boxes*). As coordenadas definidas pelo envoltório permitem recortar as imagens de entrada em um tamanho menor, focando apenas no conteúdo da região da lesão e eliminando partes de pele saudável ou ruídos que afetem a detecção.

A técnica de segmentação sugerida nessa etapa é a arquitetura DeepLabV3+, desenvolvida por (Chen et al. 2018). Essa arquitetura foi especialmente projetada para tarefas de segmentação semântica, empregando a convolução dilatada (*atrous convolution*) para capturar informações contextuais em várias escalas, preservando a resolução original. O objetivo dessa técnica é ajustar o campo receptivo dos filtros de convolução, favorecendo o aprendizado de dependências espaciais de forma mais eficaz.

A aplicação do *Atrous Spatial Pyramid Pooling* (ASPP) no DeepLabV3+ permite que a rede capture contextos em diferentes escalas ao empregar a convolução dilatada com várias taxas de dilatação. Esse processo é realizado na estrutura *encoder-decoder*, onde o *encoder* extrai as características relevantes da imagem, enquanto o *decoder* aprimora essas características para gerar a segmentação final em alta resolução. A Figura 22 abaixo ilustra um exemplo desse processo de segmentação.

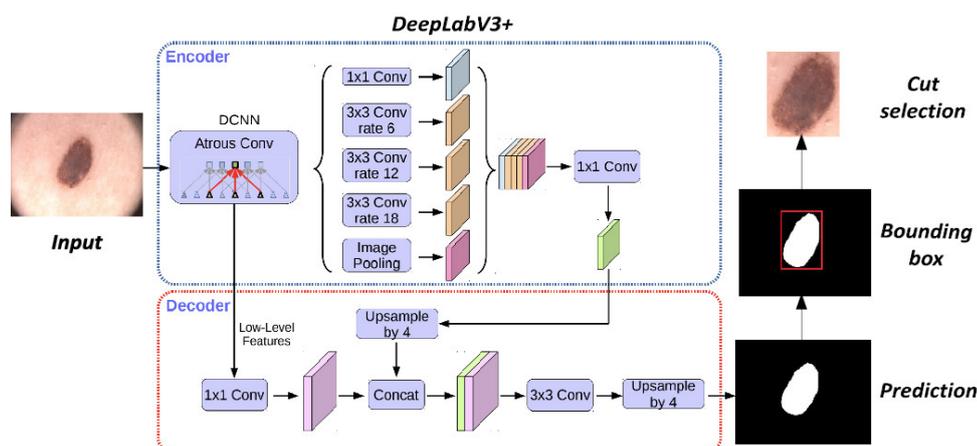


Figura 22 – Definição da região de interesse.

A técnica DeepLabV3+ aplicada na segmentação das imagens deste trabalho foi pré-treinada com as máscaras de segmentação das bases HAM10000, ISIC 2016, ISIC 2017 e ISIC 2018. O *backbone* utilizado é o ResNet-50. Os resultados alcançados com a aplicação dessa técnica na abordagem sugerida foram bem avaliados em (Silva et al. 2023), mostrando um excelente desempenho na segmentação de lesões de pele.

Após a definição das máscaras, é necessário aplicar uma operação de limiarização, pois as imagens resultantes não são binárias. Em seguida, é realizado o cálculo do *bounding box*. Os dados de coordenadas são utilizados para cortar a imagem. A base para codificação dessa técnica está disponível em (Paszke, Adam and Gross, Sam and Chintala, Soumith and Chanan, Gregory 2018).

4.6 Test-Time Augmentation

O aumento de dados é importante na abordagem desenvolvida, devido ao significativo desbalanceamento entre as classes de imagens descritas anteriormente. Para

mitigar os impactos negativos e os resultados incorretos decorrentes desse problema, é proposto na abordagem a adoção de uma política de criação de dados sintéticos, baseada na técnica Test-Time Augmentation (TTA). Esta política visa implementar uma sistemática para aprimorar a eficácia de classificação, intensificando o aprendizado dos modelos na etapa de treinamento.

A técnica TTA consiste na geração de múltiplas imagens sintéticas durante o treino, validação ou teste, assemelhando-se aos *ensembles* (Gonzalo-Martín, García-Pedrero e Lillo-Saavedra 2021). Em essência, essa técnica fornece ao modelo diferentes versões da mesma imagem de entrada, visando criar variantes de um modelo (Shanmugam et al. 2021). Após obter as previsões, é utilizado um método para combinar os resultados, que pode ser a média, o voto majoritário ou o grau de confiança.

Durante o treinamento de uma arquitetura profunda, o uso do TTA expande os dados, equilibrando as classes e aumentando a diversidade (Tursun et al. 2022). Na validação, aumenta a robustez do modelo em relação às variações, aprimorando a aprendizagem de características regulares. No teste, essa técnica possibilita ao modelo gerar previsões mais confiáveis (Shanmugam et al. 2021). No entanto, é necessário destacar que o aumento de dados nesse conjunto pode gerar resultados indesejados, pois previsões com aumento nesse conjunto podem transformar previsões corretas em incorretas.

Na abordagem desenvolvida, a aplicação do TTA baseou-se nas evidências descritas na literatura. A principal justificativa para o uso dessa técnica é seu potencial de melhoria na precisão de classificação, conforme descrito em (Shanmugam et al. 2021). Além disso, os trabalhos de (Bissoto et al. 2022), (Maron et al. 2021) e (Magalhães e Bernardino 2023) apresentam recomendações experimentais para melhorar a previsão de modelos com o uso dessa técnica, sendo seguidas e avaliadas ao final.

Nos experimentos realizados, a aplicação do TTA limitou-se à criação de imagens sintéticas da classe minoritária (lesões malignas) para balancear a quantidade de imagens nos conjuntos de treino, validação e teste. Uma particularidade da técnica é que a seleção das imagens é feita de forma aleatória, aplicando transformações de rotação, translação, zoom, espelhamento e ajuste de escala. Para evitar dúvidas sobre os resultados alcançados, também foi realizada a classificação do conjunto de teste sem a aplicação da técnica. Isso permitiu avaliar a ocorrência de alterações nas previsões ao final do processo.

4.7 Modelos de Aprendizado Profundo

Para avaliar os efeitos da correção dos aspectos de iluminação e cor na identificação do câncer de pele maligno, é recomendada a seleção de arquiteturas

profundas para a criação dos modelos de classificação. É importante destacar que a inovação da abordagem deste trabalho não está na criação dos modelos, mas na análise dos resultados de predição dos modelos desenvolvidos. O foco da classificação é apenas identificar a relevância na correção dos princípios de iluminação e cor com esses modelos, bem como o comportamento de cada um na classificação correta de lesões de pele.

Durante o processo de criação dos modelos, consideramos as arquiteturas profundas amplamente utilizadas no problema do câncer de pele, observando a quantidade de camadas, os conceitos empregados na construção e o desempenho em tarefas semelhantes. Para os experimentos, é indicado o uso das arquiteturas ResNet-50, Inception-V3, Xception e DenseNet-121. Essas arquiteturas possuem excelente aceitação em vários trabalhos que desenvolvem a detecção automática do câncer de pele (Esteva et al. 2021). Além disso, é importante considerar a estratégia de treinamento, que pode ser realizada a partir do zero ou com modelos pré-treinados.

O ResNet-50, desenvolvido por (He et al. 2016), baseia-se nos princípios das ResNets. Essa arquitetura é composta por 50 camadas organizadas em blocos residuais, cada um contendo camadas de convolução, normalização de *batch* e ativação não linear, visando resolver o problema do desaparecimento de gradientes em redes profundas. As conexões residuais permitem que a arquitetura mapeie a diferença entre a entrada fornecida e a saída desejada sem perder desempenho, mantendo o fluxo de informações relevante.

O Inception-V3, desenvolvido por (Szegedy et al. 2016), é uma arquitetura convolucional projetada para melhorar a precisão e a eficiência na classificação de imagens. Essa arquitetura possui 48 camadas e incorpora mecanismos de fatoração da convolução e regularização avançada de dados. Uma característica relevante desses mecanismos é a composição dos mesmos em módulos Inception, que combinam filtros de convolução de diferentes tamanhos em uma única camada, capturando informações em diversas escalas. A regularização nessa sistemática suaviza os rótulos verdadeiros, evitando o excesso de confiança dos modelos produzidos.

O Xception é uma extensão do Inception, porém com 71 camadas, desenvolvido por (Chollet 2017). Essa arquitetura se destaca por aprender representações complexas ao separar mapas de características por canal. Isso é feito substituindo os blocos convencionais por operações de convolução mais eficientes, executando-as em duas etapas separadas. Essa sistemática proporciona eficiência computacional e desempenho superior em determinadas situações, se comparado às redes antecessoras.

Por fim, a DenseNet-121, desenvolvida por (Huang et al. 2017), se diferencia das demais redes pelas conexões densas. Isso significa que cada camada é conectada às seguintes por blocos de conexões densas. Essa arquitetura possui 121 camadas, visando a eficiência em termos de parâmetros, permitindo o aprendizado de características

complexas em várias camadas sem a necessidade de repetir pesos. Ela se destaca pela facilitação do fluxo de gradientes durante o treinamento, possibilitando o uso de mais camadas sem perda de eficácia.

4.8 Metodologia Experimental e Métricas de Avaliação

A codificação da sistemática foi realizada na linguagem Python, utilizando o *framework* PyTorch para implementar as técnicas LCDPNet, LLNeRF, DSN e DeepLabV3+. As técnicas CLAHE, SoG e Perfect Reflector foram implementadas com os frameworks OpenCV e NumPy. O treinamento das arquiteturas selecionadas foi realizado com o *framework* TensorFlow. Todos os experimentos foram executados em um computador com processador Intel Core i9, 32 GB de memória RAM e GPU Gigabyte NVIDIA 3090 de 24 GB GDDR6.

Para a sistemática experimental, adotou-se a técnica de validação cruzada (*K-fold Cross-Validation*) para dividir as imagens processadas. Sugerimos uma divisão de 80% para o conjunto de treino e 20% para testes, sendo que 20% do conjunto de treino é destinado à validação. O processo seguiu uma repetição de K vezes, garantindo que cada divisão fosse usada exatamente uma vez. O valor de K foi definido como 5.

O aumento de dados também foi aplicado com a técnica TTA, ajustando a sistemática para o tempo de execução nos conjuntos de treino, validação e teste em cada *fold*, considerando o balanceamento da classe minoritária. Como descrito anteriormente, para validar os resultados e evitar problemas de mudanças de previsão com o aumento de dados no conjunto de teste, repetimos os experimentos sem a aplicação dessa técnica nesse conjunto.

As arquiteturas profundas Inception-V3, Xception, ResNet-50 e DenseNet-121 foram implementadas também utilizando o *framework* TensorFlow. As arquiteturas já haviam sido pré-treinadas com o conjunto de dados ImageNet (Krizhevsky, Sutskever e Hinton 2017), conforme indicado pelos bons resultados alcançados em (Morid, Borjali e Fiol 2021). Além disso, foi realizada a calibração de hiperparâmetros para utilizar parâmetros ótimos, proporcionando um comparativo igualitário. Nesta calibração, ajustamos a taxa de aprendizado, o tamanho do lote e a função de otimização, sem restrições sobre esses parâmetros.

Para a classificação binária, as camadas totalmente conectadas das arquiteturas sugeridas foram ajustadas. O treinamento dessas arquiteturas foi conduzido por meio da transferência de aprendizado. A reconstrução seguiu o design original de cada arquitetura, porém com o número de unidades e a função de ativação adaptadas para o número de classes. Além disso, o ajuste fino foi aplicado para garantir a integração das camadas pré-treinadas com as camadas adaptadas, otimizando o classificador.

As métricas sugeridas para a análise de resultados deste trabalho são baseadas na eficácia de classificação do câncer de pele maligno: acurácia (Acc), precisão (P), *Recall* (R) e *F1-score*. Essas métricas são calculadas a partir da matriz de confusão, que apresenta as amostras descritas em verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN) (James et al. 2013). Outras medidas importantes observadas incluem a área sob a curva e o coeficiente *Kappa* (K).

A acurácia (Acc) é a proporção de instâncias classificadas corretamente em relação ao total de instâncias avaliadas. É uma métrica amplamente usada, mas pode ser enganosa em casos de conjuntos de dados desbalanceados, onde uma classe tem uma prevalência muito maior que a outra. A Equação 4.1 apresenta o cálculo dessa métrica (James et al. 2013).

$$Acc = \frac{VP + VN}{VP + FP + VN + FN} \quad (4.1)$$

A precisão (P) mede a proporção de instâncias corretamente classificadas como verdadeiras positivas (VP) em relação ao total de instâncias classificadas como verdadeiras positivas e falsas positivas (FP). A precisão é útil quando o foco é minimizar os falsos positivos (James et al. 2013). A Equação 4.2 define o cálculo dessa métrica.

$$P = \frac{VP}{VP + FP} \quad (4.2)$$

O *recall* (R) é a proporção de instâncias corretamente classificadas como VP em relação ao total de instâncias VP mais falsos negativos (FN). A revocação é útil quando o objetivo é minimizar os falsos negativos (James et al. 2013). A Equação 4.3 demonstra a forma de calcular essa métrica.

$$R = \frac{VP}{VP + FN} \quad (4.3)$$

Já o *F1-score* é uma medida que visa equilibrar precisão e *recall*, fornecendo uma média harmônica entre essas duas métricas. Essa medida é útil quando é necessário avaliar uma minimização de falsos positivos e falsos negativos (James et al. 2013). A Equação 4.4 demonstra o cálculo dessa métrica.

$$F1 - Score = \frac{2 * Precisão * Recall}{Precisão + Recall} \quad (4.4)$$

A área sob a curva (*Area Under Curve* - AUC) é uma medida indicada para avaliar o desempenho de classificação binária. A curva descrita nessa métrica é as características operacionais do receptor (*Receiver Operating Characteristic* - ROC), sendo uma representação gráfica do desempenho entre a taxa de verdadeiro positivo (TVP) e

a taxa de falso positivo (TFP). As equações 4.5, 4.6 e 4.7 descrevem as equações referentes a essas métricas.

$$TVP = \frac{VP}{VP + FN}, \quad (4.5)$$

$$TFP = \frac{FP}{FP + VN}, \quad (4.6)$$

$$AUC = \sum_{i=1}^{n-1} \frac{(x_{i+1} - x_i) \cdot (y_{i+1} + y_i)}{2}, \quad (4.7)$$

onde $(x_{i+1} - x_i)$ representa as taxas de falsos positivos e $(y_{i+1} + y_i)$ as taxas de verdadeiros positivos.

O coeficiente *Kappa* é uma métrica utilizada para determinar a concordância dos resultados alcançados por um classificador. Isso quer dizer que esse coeficiente determina o grau de concordância entre duas avaliações independentes na classificação binária. Quando esse coeficiente se apresenta com valores menores que 0,4, os resultados de predição possuem pouca concordância ou nenhuma concordância (aleatório). Se esse valor estiver entre 0,4 e 0,6, existe uma concordância razoável. Caso esse coeficiente esteja entre 0,6 e 0,8, a concordância é boa ou ótima. Já os valores acima de 0,8 indicam uma excelente concordância. As equações 4.8 e 4.9 descrevem as fórmulas referentes ao coeficiente Kappa,

$$K = \frac{p_o - p_e}{1 - p_e}, \quad (4.8)$$

$$p_e = \sum \left(\frac{n_{i.} \times n_{.i}}{n^2} \right), \quad (4.9)$$

onde p_o é a concordância observada pelos avaliadores, p_e a proporção esperada de concordância, $n_{i.}$ o total de itens que o primeiro avaliador classificou na categoria i , e $n_{.i}$ o total de itens que o segundo avaliador classificou na categoria i . Esse cálculo de p_e considera a probabilidade de cada avaliador classificar aleatoriamente um item em cada categoria, somando-as para todas as categorias possíveis.

Após a definição das métricas, os melhores modelos são selecionados com base na precisão e no coeficiente *Kappa* do conjunto de validação. Além disso, como a sistemática de validação cruzada requer o cálculo da média das métricas, é necessário também calcular o desvio padrão para determinar a variação entre os resultados de cada *fold*. Esta medida estatística quantifica a diferença dos valores individuais de classificação em relação à média dos resultados. Embora não exista um valor ideal de desvio padrão, um valor próximo de zero indica uma uniformidade de resultados entre os *folds*.

4.9 Considerações Finais

Este capítulo apresentou uma abordagem para a detecção automática do câncer de pele utilizando técnicas avançadas de aprendizado profundo, com ênfase na correção de aspectos de iluminação e cor. Um conjunto de etapas foi definido, incluindo a seleção de bases de imagens, pré-processamento, segmentação, aumento de dados, treinamento de arquiteturas convolucionais com transferência de aprendizado e ajuste, e análise e avaliação de resultados. Cada etapa foi cuidadosamente integrada para assegurar a eficácia na identificação de lesões malignas. Essa abordagem é o objeto de validação do estudo da tese, abordando nuances da literatura em relação à variação dessas características e suas influências na detecção precisa de malignidades de pele. No próximo capítulo, serão apresentados os principais resultados alcançados com a execução da metodologia sugerida.

5 Resultados e Discussões

Neste capítulo, é delineada a execução da abordagem de detecção de câncer de pele descrita no capítulo anterior. Inicialmente, são apresentados os aspectos técnicos, como a definição dos hiperparâmetros e a descrição dos cenários de testes adotados nos experimentos. Em seguida, são apresentados os resultados alcançados, além de uma discussão crítica sobre a relevância da classificação de lesões malignas. Associados aos principais resultados, são retratados comparações com trabalhos relacionados, deficiências do estudo e limitações técnicas na validação da abordagem desenvolvida.

5.1 Especificações de Hiperparâmetros

Para o processo de otimização de hiperparâmetros, é utilizado o *GridSearch*, uma ferramenta que facilita o processo de otimização em Python. Embora existam diferenças entre as arquiteturas sugeridas, alguns parâmetros são comuns e aplicados em intervalos específicos. Para a otimização no treinamento, são tratados os parâmetros: taxa de aprendizado (*Learning Rate*) = $\{10^{-3}, 10^{-5}, 10^{-7}\}$, tamanho do lote (*Batch Size*) = $\{16, 32, 64, 128\}$, número de épocas (*Epochs*) = $\{1000\}$ e otimizador (*Optimizer*) = $\{\text{Adam, Adagrad, SGD e RMprop}\}$.

É importante destacar que é implementado o monitoramento de épocas com a ferramenta *EarlyStopping*, um pacote embutido no TensorFlow, com um comparativo da taxa de aprendizado a cada 50 épocas. Isso significa que, mesmo definindo a quantidade de épocas iguais para todas as arquiteturas, esse parâmetro pode ser variável conforme o melhor aprendizado. E como descrito anteriormente, a escolha do melhor modelo foi realizada utilizando as métricas de índice *Kappa* e a precisão do conjunto de validação. Entretanto, como o aumento de dados proporciona uma espécie de *ensembles*, o número de épocas podem manter uma variação específica não retratada.

Os testes realizados neste trabalho foram definidos em cenários distintos, sendo: Sem Correção de Iluminação e Cor, Correção Com a Técnica CLAHE, Correção com a Técnica *Perfect Reflector*, Correção com a Técnica SoG, Correção com a Técnica LCDPNet, Correção com a Técnica LLNeRF e Correção com a Técnica DSN. Considerando os modelos Inception-V3, Xception, ResNet-50 e DenseNet-121 produzidos em cada cenário, a Tabela 4 apresenta os hiperparâmetros ótimos encontrados para cada modelo, diante dos melhores resultados obtidos nos conjuntos de validação por *folds*.

Os parâmetros apresentados permitiram o melhor treinamento por transferência de aprendizado e ajuste fino, proporcionando uma adaptação eficaz dos modelos aos conjuntos de imagens de lesões de pele. Esses ajustes foram realizados utilizando as

Modelos	Taxa de Aprendizado	Tamanho de Lote	Épocas	Otimizador
Sem Correção de Iluminação e Cor				
Inception-V3	10^{-5}	64	265	Adam
Xception	10^{-3}	32	312	Adam
ResNet-50	10^{-5}	64	186	Adam
DenseNet-121	10^{-3}	16	213	Adam
Correção com a Técnica CLAHE				
Inception-V3	10^{-3}	32	184	Adam
Xception	10^{-3}	16	212	Adam
ResNet-50	10^{-3}	32	195	Adam
DenseNet-121	10^{-3}	32	135	Adam
Correção com a Técnica <i>Perfect Reflector</i>				
Inception-V3	10^{-5}	32	236	Adam
Xception	10^{-3}	32	302	Adam
ResNet-50	10^{-5}	64	198	Adam
DenseNet-121	10^{-3}	16	298	Adam
Correção com a Técnica SoG				
Inception-V3	10^{-3}	64	301	Adam
Xception	10^{-3}	32	206	Adam
ResNet-50	10^{-3}	32	231	Adam
DenseNet-121	10^{-3}	32	269	Adam
Correção com a Técnica LCDPNet				
Inception-V3	10^{-5}	32	198	Adam
Xception	10^{-3}	16	254	Adam
ResNet-50	10^{-5}	32	207	Adam
DenseNet-121	10^{-3}	32	254	Adam
Correção com a Técnica LLNeRF				
Inception-V3	10^{-3}	64	312	Adam
Xception	10^{-3}	16	179	Adam
ResNet-50	10^{-5}	32	367	Adam
DenseNet-121	10^{-3}	16	201	Adam
Correção com DSN				
Inception-V3	10^{-5}	32	288	Adam
Xception	10^{-3}	32	331	Adam
ResNet-50	10^{-3}	64	212	Adam
DenseNet-121	10^{-3}	32	406	Adam

Tabela 4 – Melhores parâmetros determinados nos cenários de testes.

técnicas de correção de iluminação e cor sugeridas, otimizando a precisão e a eficácia dos modelos.

5.2 Análise e Validação de Resultados

Para avaliar os resultados individualmente, é importante destacar a aplicação do aumento de dados, como apresentado na metodologia, a técnica TTA foi aplicada ao conjunto de teste. No entanto, os experimentos também foram realizados sem a aplicação dessa estratégia a esse conjunto. Desse modo, será apresentada a análise dos resultados de predição dos modelos com a aplicação do TTA no conjunto de teste e, em seguida, a análise dos resultados sem esse procedimento. O intuito disso é verificar a inversibilidade das predições dos modelos na precisão mediante imagens sintéticas.

As próximas subseções apresentam os resultados alcançados com a abordagem

desenvolvida. Essa abordagem segue a metodologia definida nos cenários descritos na seção anterior, tanto sem correção dos aspectos de iluminação e cor quanto com a aplicação das técnicas sugeridas para a correção dessas características.

5.2.1 Resultados com Aumento de Dados no Conjunto de Teste

Conforme a média das métricas de acurácia (Acc), precisão (P), *recall* (R), área sob a curva ROC (AUC), *Kappa* (K) e desvio padrão desses valores para cada *fold*, a Tabela 5 apresenta os resultados alcançados com as configurações definidas na metodologia e com os parâmetros descritos na seção anterior.

Models	Acc(%)	P(%)	R(%)	AUC(%)	K(%)
Sem Correção de Iluminação e Cor					
Inception-V3	89,3 ±0,016	91,4 ±0,031	87,1 ±0,041	91,8 ±0,078	82,7 ±0,033
Xception	87,8 ±0,076	83,6 ±0,071	89,8 ±0,068	85,6 ±0,053	79,6 ±0,046
ResNet-50	89,9 ±0,015	90,1 ±0,036	90,0 ±0,023	82,6±0,060	81,2 ±0,030
DenseNet-121	82,6 ±0,031	84,5 ±0,018	86,3 ±0,052	81,6±0,029	79,7 ±0,016
Correção com a Técnica CLAHE					
Inception-V3	77,1 ±0,057	72,0 ±0,061	73,2 ±0,078	71,8 ±0,048	70,6 ±0,046
Xception	70,4 ±0,078	68,8 ±0,099	68,3 ±0,088	65,6 ±0,093	66,4 ±0,097
ResNet-50	75,4 ±0,066	72,8 ±0,054	74,9 ±0,069	70,3 ±0,081	69,6 ±0,084
DenseNet-121	69,3 ±0,112	66,4 ±0,099	68,6 ±0,102	64,8 ±0,136	62,6 ±0,151
Correção com a Técnica <i>Perfect Reflector</i>					
Inception-V3	89,2 ±0,063	90,6 ±0,086	88,1 ±0,091	90,6 ±0,098	80,1 ±0,089
Xception	88,4 ±0,086	85,2 ±0,078	88,6 ±0,069	87,8 ±0,083	80,8 ±0,077
ResNet-50	90,5 ±0,072	91,2 ±0,066	91,8 ±0,061	86,6 ±0,070	82,9 ±0,082
DenseNet-121	83,6 ±0,098	86,2 ±0,086	88,6 ±0,072	82,3 ±0,099	79,2 ±0,093
Correção com a Técnica SoG					
Inception-V3	90,3 ±0,068	92,1 ±0,043	90,2 ±0,056	89,1 ±0,078	84,5 ±0,063
Xception	89,4 ±0,086	86,2 ±0,061	90,4 ±0,081	87,6 ±0,073	80,6 ±0,087
ResNet-50	91,5 ±0,065	89,1 ±0,052	88,0 ±0,083	86,6 ±0,090	80,1 ±0,060
DenseNet-121	83,1 ±0,044	85,8 ±0,038	83,6 ±0,042	80,1 ±0,029	81,4 ±0,056
Correção com a Técnica LCDPNet					
Inception-V3	90,1 ±0,023	93,2 ±0,034	88,3 ±0,048	90,6 ±0,058	81,3 ±0,023
Xception	86,5 ±0,043	84,2 ±0,062	86,4 ±0,068	85,3 ±0,043	76,2 ±0,057
ResNet-50	91,3 ±0,036	92,3 ±0,035	87,9 ±0,041	84,3 ±0,050	79,0 ±0,043
DenseNet-121	81,2 ±0,051	83,6 ±0,048	89,8 ±0,052	84,6 ±0,039	77,3 ±0,066
Correção com a Técnica LLNeRF					
Inception-V3	91,8 ±0,054	90,3 ±0,046	87,3 ±0,061	89,9 ±0,048	79,1 ±0,041
Xception	89,3 ±0,026	86,2 ±0,018	90,1 ±0,032	88,5 ±0,039	80,6 ±0,046
ResNet-50	92,9 ±0,025	90,6 ±0,027	89,1 ±0,041	90,3 ±0,040	78,8 ±0,052
DenseNet-121	84,6 ±0,037	86,2 ±0,048	89,6 ±0,052	89,3 ±0,059	80,2 ±0,043
Correção com a Técnica DSN					
Inception-V3	88,2 ±0,056	89,1 ±0,041	89,6 ±0,069	89,6 ±0,038	80,1 ±0,063
Xception	86,8 ±0,076	81,6 ±0,071	86,8 ±0,068	84,3 ±0,053	79,6 ±0,043
ResNet-50	93,1 ±0,015	91,1 ±0,036	90,0 ±0,023	83,1 ±0,070	80,2 ±0,030
DenseNet-121	86,5 ±0,061	87,9 ±0,073	86,3 ±0,076	83,1 ±0,059	77,3 ±0,081

Tabela 5 – Resultados de classificação com os cenários de correção dos aspectos de iluminação e cor (aplicação do TTA no conjunto de teste).

Ao analisar a Tabela 5, observa-se como diferentes técnicas de correção de iluminação e cor influenciam o desempenho dos modelos de classificação. Inicialmente, sem correção de iluminação e cor, os modelos apresentaram desempenhos relativamente

elevados. Nesse cenário, o melhor resultado de precisão foi alcançado pela Inception-V3, com uma taxa de 91,4% e AUC de 91,8%. Já a melhor acurácia foi obtida pelo ResNet-50, alcançando um valor de 89,9%. No entanto, a precisão nesse resultado foi de 90,1% e AUC de 82,6%. É notável que, nesse cenário, o Xception e o DenseNet-121 obtiveram resultados inferiores se comparados com o Inception-V3 e o ResNet-50.

A aplicação da correção com o CLAHE não melhorou o desempenho dos modelos. Ao invés disso, houve uma diminuição relevante em comparação com o cenário de não correção de iluminação e cor. O ResNet-50 obteve o melhor resultado com essa correção, alcançando uma precisão de 72,8%. Esse resultado é acentuadamente inferior a todos os resultados alcançados no cenário anterior, indicando que essa correção não é benéfica para a variedade de imagens utilizadas.

O uso da correção com o *Perfect Reflector* também não favoreceu significativamente a maioria dos modelos. Os resultados de precisão apresentam uma melhora marginal em comparação com o cenário sem correção. Para esse cenário, o ResNet-50 alcançou uma acurácia de 90,5% e precisão de 91,2%. Já o resultado de menor relevância foi obtido pela DenseNet-121, com acurácia de 83,6% e precisão de 86,2%.

A aplicação da técnica SoG na correção das imagens obteve um desempenho próximo aos resultados alcançados pelos modelos desenvolvidos com o *Perfect Reflector*. O resultado mais relevante nesse cenário foi do ResNet-50, atingindo uma acurácia de 91,5% e precisão de 89,1%. No entanto, é perceptível que o Inception-V3 teve um desempenho maior na precisão, com uma taxa de 92,1%.

A correção com os métodos CLAHE, *Perfect Reflector* e SoG não apresentou diferenças expressivas de desempenho nos modelos testados, principalmente em termos de precisão, quando comparados todos os resultados com o cenário sem correção de iluminação e cor. Embora sejam encontradas algumas melhorias pontuais para os modelos Inception-V3 e ResNet-50, a correção baseada em conceitos de histogramas e balanço de branco parece não favorecer diferenças expressivas para a identificação do câncer de pele maligno com modelos de aprendizado profundo.

A aplicação da técnica de correção LCDPNet mostrou uma tendência de melhoria nos resultados para a maioria dos modelos analisados em comparação com o cenário sem correção. Essa técnica aumentou a precisão dos modelos ResNet-50 e Inception-V3 para 93,2% e 92,3%, respectivamente. No entanto, observou-se uma diminuição no AUC do ResNet-50, o que pode indicar uma perda em sua capacidade de fazer previsões precisas. Por outro lado, o DenseNet-121 não apresentou melhorias significativas nas métricas, sugerindo que a correção não beneficiou o modelo.

A aplicação do método LLNeRF para correção melhorou as métricas em todos os modelos avaliados em comparação com o cenário sem correção. É notável um aumento

na precisão do Inception-V3 de 89,3% para 91,8%. No entanto, a precisão e o AUC tiveram uma leve diminuição. Em contraste, o Xception e o DenseNet-121 exibiram aumentos significativos em todas as métricas. Destacando todos os resultados após essa correção, o ResNet-50 foi o modelo que apresentou os maiores ganhos de desempenho.

A aplicação da técnica de correção DSN também mostrou resultados variados. O Inception-V3 diminuiu a precisão de 91,4% para 89,1%, e o AUC de 91,8% para 89,6%. O Xception também experimentou uma redução significativa na precisão, caindo de 85,6% para 81,6%. Por outro lado, o ResNet-50 e o DenseNet-121 melhoraram as métricas com aumento nas taxas de precisão e AUC. Notavelmente, a precisão do ResNet-50 aumentou de 89,9% para 93,1%.

Para tornar as variações entre os modelos analisados sem correção de iluminação e cor e as técnicas de correção avaliadas mais claras, a Tabela 6 apresenta os valores específicos de diferença na precisão com o desvio padrão ponderado. Essa sistemática permite uma comparação quantitativa do impacto de cada técnica de correção.

Modelos	P (%)					
	CLAHE	<i>Perfect Reflector</i>	SoG	LCDPNet	LLNeRF	DSN
Inception-V3	-19,4 ±0,099	-0,8 ±0,068	0,7 ±0,078	1,8 ±0,046	-1,1 ±0,046	-2,3±0,040
Xception	-14,8 ±0,096	1,6 ±0,077	2,6 ±0,049	0,6 ±0,094	2,6 ±0,018	-2,0 ±0,071
ResNet-50	-17,3 ±0,104	1,1 ±0,095	-1 ±0,081	2,2 ±0,050	0,5 ±0,027	1,0 ±0,036
DenseNet-121	-18,1 ±0,116	1,7 ±0,078	1,3 ±0,087	-0,9 ±0,051	1,7 ±0,048	3,4 ±0,077

Tabela 6 – Diferenças de precisão de classificação entre a não correção de iluminação com as técnicas para correção (Aplicação do TTA no conjunto de teste por *fold*).

Analisando a Tabela 6, a aplicação do CLAHE mostrou uma queda acentuada, tendo uma perda de precisão próxima de 20% para a Inception-V3. No mesmo cenário, esse comportamento de classificação foi unânime. Além disso, também é possível constatar que a aplicação do *Perfect Reflector* possui um impacto positivo na precisão para os modelos Xception, ResNet-50 e DenseNet-121. No entanto, a aplicação do SoG mostrou resultados variados, com um pequeno aumento de precisão para alguns modelos e uma leve queda para outros, como no ResNet-50.

O LCDPNet e o LLNeRF foram mais eficazes para os modelos Inception-V3 e ResNet-50, mostrando um aumento moderado na precisão. No entanto, o DenseNet-121 exibiu uma pequena redução de precisão, o que pode indicar uma baixa sensibilidade deste modelo às variações de iluminação e cor. A aplicação da técnica DSN também mostrou resultados mistos de precisão. O Inception-V3 e Xception tiveram uma queda na precisão nesse cenário, porém, uma particularidade de resultados foi encontrada nos resultados do DenseNet-121, mostrando uma melhoria significativa de precisão de 3,4%.

Em todos os experimentos realizados, os melhores resultados foram dados pelo ResNet-50 e Inception-V3, mostrando um melhor potencial desses modelos para lidar

com imagens de lesões de pele. Além disso, a correção com os métodos LCDPNet, LLNeRF e DSN proporcionaram um melhor desempenho geral de classificação, se comparado com as técnicas CLAHE, Perfect Reflector e SoG. Outro aspecto a ser observado é que, embora os modelos Xception e DenseNet-121 tenham obtido resultados inferiores, na maioria dos cenários houve uma certa estabilidade, com poucas perdas de taxas. Isso indica que essas arquiteturas podem ser menos sensíveis às variações de iluminação e cor.

5.2.2 Resultados sem Aumento de Dados no Conjunto de Teste

Para analisar a abordagem por outra perspectiva, como dito antes, a execução dos experimentos foram tratadas sem a execução do aumento de dados com a aplicação do TTA ao conjunto de teste. A Tabela 7 apresenta os resultados alcançados com a ausência dessa estratégia, servindo de comparativo com os resultados descritos na subseção anterior, em relação ao desempenho dos modelos com o conjunto de teste desbalanceado.

Ao comparar as Tabelas 5 e 7, observa-se que os resultados mantêm um padrão consistente, embora com perdas em quase todas as métricas. Notavelmente, as medidas de precisão em todos os experimentos apresentaram quedas acentuadas, com perdas que podem ultrapassar 5% em todos os cenários avaliados. No entanto, uma particularidade da ausência da estratégia de aumento de dados no conjunto de teste é a estabilização ou uma leve melhora nas métricas AUC e *Kappa*.

A correção com o CLAHE também resultou no pior desempenho entre os modelos. Nesta abordagem, a ResNet-50 obteve uma precisão de 71,6%, inferior à da sistemática anterior, que foi de 72,8%. Um ponto de diferença entre esses experimentos é a melhoria no *Kappa*, que aumentou de 69,6% na sistemática anterior para 70,7%. Esse resultado continua a indicar que essa correção não é vantajosa para a variedade de imagens utilizadas.

O uso da correção com o *Perfect Reflector* também não favoreceu a maioria dos modelos. Assim como nos experimentos anteriores, essa correção apresentou melhorias pouco significativas. Nesse cenário, o melhor resultado foi obtido pela ResNet-50, com precisão de 88,1%, inferior à sistemática anterior, que foi de 91,2%. No entanto, houve um aumento expressivo no *Kappa*, de 82,9% para 85,5%. O DenseNet-121 também manteve o resultado de menor relevância para essa correção.

A correção com o SoG apresentou resultados semelhantes ao *Perfect Reflector* em todos os modelos. Notavelmente, o Inception-V3 teve o melhor desempenho com essa correção, com uma taxa de precisão de 84,3%, bem inferior aos resultados anteriores. Isso também reforça a similaridade entre os métodos de correção, refletindo em taxas

Models	Acc(%)	P(%)	R(%)	AUC(%)	K(%)
Sem Correção de Iluminação e Cor					
Inception-V3	84,5 ±0,056	85,2 ±0,044	84,6 ±0,041	83,1 ±0,051	82,7 ±0,057
Xception	80,2 ±0,043	79,4 ±0,058	83,5 ±0,074	82,3 ±0,067	81,2 ±0,068
ResNet-50	85,7 ±0,065	84,3 ±0,076	87,1 ±0,055	83,5 ±0,060	82,4 ±0,041
DenseNet-121	79,2 ±0,061	80,6 ±0,057	81,5 ±0,044	82,4 ±0,079	80,6 ±0,066
Correção com a Técnica CLAHE					
Inception-V3	75,4 ±0,072	71,3 ±0,081	73,2 ±0,057	74,2 ±0,069	74,8 ±0,052
Xception	68,6 ±0,064	65,9 ±0,077	69,1 ±0,091	70,8 ±0,061	71,6 ±0,073
ResNet-50	72,7 ±0,058	71,6 ±0,049	70,8 ±0,066	72,4 ±0,061	70,7 ±0,071
DenseNet-121	66,1 ±0,098	65,8 ±0,100	64,4 ±0,084	65,2 ±0,096	69,9 ±0,086
Correção com a Técnica <i>Perfect Reflector</i>					
Inception-V3	85,4 ±0,053	87,2 ±0,066	84,6 ±0,071	87,0 ±0,054	84,3 ±0,043
Xception	81,6 ±0,066	83,3 ±0,069	85,1 ±0,051	84,4 ±0,061	83,6 ±0,068
ResNet-50	88,1 ±0,050	84,6 ±0,072	86,9 ±0,048	84,2 ±0,041	85,5 ±0,066
DenseNet-121	80,3 ±0,068	81,6 ±0,046	83,6 ±0,076	83,9 ±0,074	81,2 ±0,051
Correção com a Técnica SoG					
Inception-V3	84,2 ±0,065	86,7 ±0,042	82,6 ±0,053	84,6 ±0,039	86,0 ±0,047
Xception	82,6 ±0,071	83,6 ±0,086	84,5 ±0,084	85,6 ±0,047	84,6 ±0,070
ResNet-50	86,3 ±0,043	84,3 ±0,058	86,6 ±0,039	86,6 ±0,067	85,7 ±0,050
DenseNet-121	80,4 ±0,090	81,5 ±0,101	83,2 ±0,088	82,6 ±0,091	81,6 ±0,083
Correção com a Técnica LCDPNet					
Inception-V3	88,6 ±0,044	86,4 ±0,043	87,9 ±0,062	85,7 ±0,041	84,9 ±0,061
Xception	82,3 ±0,066	80,5 ±0,078	84,1 ±0,049	83,5 ±0,061	81,9 ±0,073
ResNet-50	88,1 ±0,056	87,6 ±0,077	88,6 ±0,039	85,9 ±0,043	84,9 ±0,076
DenseNet-121	80,1 ±0,073	82,0 ±0,068	83,7 ±0,079	80,6 ±0,061	81,6 ±0,046
Correção com a Técnica LLNeRF					
Inception-V3	86,4 ±0,036	86,3 ±0,049	84,7 ±0,059	86,1 ±0,068	82,9 ±0,055
Xception	81,3 ±0,072	83,4 ±0,058	85,6 ±0,071	84,6 ±0,051	81,3 ±0,066
ResNet-50	87,2 ±0,053	88,4 ±0,067	86,7 ±0,074	88,6 ±0,039	85,7 ±0,066
DenseNet-121	82,6 ±0,052	84,7 ±0,068	86,1 ±0,081	84,9 ±0,060	84,3 ±0,052
Correção com a Técnica DSN					
Inception-V3	87,6 ±0,069	85,3 ±0,077	86,4 ±0,041	85,8 ±0,060	83,9 ±0,040
Xception	83,8 ±0,050	84,6 ±0,039	82,7 ±0,088	82,6 ±0,066	81,1 ±0,073
ResNet-50	88,5 ±0,033	89,9 ±0,052	86,8 ±0,059	84,7 ±0,061	85,3 ±0,058
DenseNet-121	84,6 ±0,042	85,6 ±0,066	84,8 ±0,033	82,6 ±0,050	82,0 ±0,069

Tabela 7 – Resultados de classificação com os cenários de correção dos aspectos de iluminação e cor (Sem a aplicação do TTA no conjunto de teste por *fold*).

próximas em todas as métricas avaliadas.

O uso da correção LCDPNet trouxe resultados melhores para a maioria dos modelos, mantendo uma tendência de melhor desempenho quando comparado com a não correção da iluminação e cor. Nesta sistemática, a precisão dos modelos ResNet-50 e Inception-V3 se manteve acima de 86%. No entanto, um diferencial desse experimento é a harmonização das taxas de AUC e *Kappa*. Nos resultados anteriores, os valores dessas taxas apresentaram uma discordância maior em relação à acurácia.

A correção com o LLNeRF manteve um aumento na maioria dos modelos avaliados em comparação com o cenário sem correção. Neste experimento, também houve aumento na precisão da ResNet-50, de 84,3% para 88,4%. Uma particularidade deste experimento, em relação aos resultados anteriores, é a normalização das taxas de AUC e *Kappa*. Não foram percebidas discordâncias nas métricas, indicando uma melhor

consistência nos valores de classificação alcançados.

A técnica de correção DSN também apresentou resultados consistentes entre os modelos. A ResNet-50 obteve os melhores resultados nessa configuração, alcançando uma precisão de 89,9%. As taxas alcançadas com essa correção demonstram melhor concordância em comparação com os resultados anteriores. Embora esses experimentos não tenham produzido resultados de precisão excepcionais quando comparados às demais correções com aprendizado profundo, as taxas de classificação permaneceram estáveis.

Os resultados obtidos na maioria dos modelos com as correções de iluminação proporcionadas pelas técnicas LCDPNet, LLNeRF e DSN foram superiores aos alcançados com as técnicas CLAHE, Perfect Reflector e SoG. Isso reforça os resultados anteriores, evidenciando um padrão comportamental dos modelos em relação a essas características em ambos os cenários e configurações experimentados.

Além disso, os experimentos descritos nesta subseção comprovam uma diminuição em todas as métricas dos modelos devido ao desequilíbrio das classes. A aplicação do aumento de dados com o TTA ao conjunto de teste tende a melhorar os resultados, com poucas desvantagens para imagens de lesões de pele. A principal desvantagem é a diminuição das taxas de AUC e *Kappa*.

5.2.3 Comparações com Trabalhos Semelhantes

Para ressaltar a influência da correção de iluminação e cor com a abordagem sugerida, foi desenvolvido um comparativo com outras metodologias similares. Os critérios utilizados para essas metodologias foram a aplicação de alguma técnica de correção de iluminação e cor no pré-processamento, além da classificação das imagens com modelos de aprendizado profundo após a correção dessas características. A Tabela 8 apresenta um resumo das comparações conforme os melhores resultados encontrados.

Trabalhos	Acc (%)	P (%)	Comentários
(Galdran et al. 2017)	72.3	69	Resultado alcançado com o ResNet50. Aumento de dados com foco em Dual Estimation of Lighting.
(Nahata e Singh 2020)	85	86	Melhores resultados com o ResNet50. Correção por métodos estatísticos, com foco em ajustes globais de iluminação.
(Salvi et al. 2022)	79.2	-	Resultados alcançados com o DenseNet121. Utiliza GAN para correção de iluminação e cor, destacando o potencial desta técnica na segmentação e classificação. Dados de precisão não especificados.
Combinação desenvolvida	91.3	92.3	Resultados alcançados com o ResNet50, utilizando a técnica de correção LCDPNet.
	86.5	87.9	Melhores resultados com DenseNet121. Aplicação da técnica de correção DSN.

Tabela 8 – Comparação de desempenho da abordagem proposta com trabalhos semelhantes.

Ao analisar a Tabela 8, é possível validar o potencial do estudo e da abordagem de validação desenvolvida. No entanto, é importante destacar que os estudos comparados não seguiram as mesmas etapas deste trabalho. Isso quer dizer que eles compartilham apenas o uso de técnicas de correção de iluminação e cor, além da classificação com modelos de redes neurais profundas. Embora o objetivo principal dos trabalhos analisado seja melhorar a precisão da classificação através da correção de iluminação e cor, as comparações entre os estudos são limitadas. Essas limitações são decorrentes de:

- Bases de imagens utilizadas: Cada estudo empregou diferentes conjuntos de dados, o que pode influenciar diretamente nos resultados de classificação.
- Métodos de correção aplicados: As técnicas de correção variam consideravelmente, desde métodos estatísticos até o uso de GANs.
- Arquiteturas de classificação: Diferentes modelos de redes neurais foram utilizados, cada um com suas próprias características e *design*.
- Métricas de avaliação: Nem todos os estudos reportaram as mesmas métricas de desempenho, dificultando comparações diretas.

Os resultados observados podem ser atribuídos não somente às técnicas de correção, mas também às características intrínsecas de cada solução desenvolvida. Por exemplo, a combinação desenvolvida na abordagem deste trabalho demonstrou superioridade em termos de precisão e acurácia, destacando a eficácia das técnicas LCDPNet e DSN aplicadas.

Em suma, os resultados apresentados demonstram o potencial da sistemática metodológica desenvolvida. As métricas de precisão e acurácia são superiores em comparação com estudos similares. Apesar das limitações nas comparações, é evidente que correções estatísticas de iluminação e cor não garantem necessariamente bons resultados na classificação com redes neurais profundas. No entanto, a seleção e a aplicação adequadas das técnicas de correção de iluminação e cor podem aumentar significativamente a precisão dos modelos de aprendizado profundo, uma vez que esses modelos são sensíveis a variações nessas características.

5.3 Limitações do Estudo e da Sistemática de Classificação

Apesar dos resultados encontrados indicarem que a correção de iluminação e cor pode melhorar a precisão da classificação de lesões malignas com modelos de aprendizado profundo, é necessário destacar que ainda existem limitações, restrições e desafios a serem enfrentados. Por isso, é essencial descrever as limitações da metodologia adotada, principalmente para entender os resultados de maneira crítica.

A variação dos conjuntos de dados utilizados, especialmente sob condições específicas de iluminação e cor, ainda é uma questão importante. Cada base de imagens possui imagens coletadas com diferentes perspectivas de qualidade. Como cada base não foi testada individualmente, é possível que os resultados alcançados possam variar, especialmente associados à técnica de correção de iluminação e cor utilizada.

Existe uma dependência das técnicas de iluminação e cor na abordagem. Em alguns casos, é perceptível que a correção proporciona melhores resultados de precisão, porém, às custas da diminuição da AUC e do *Kappa*. Isso pode indicar falhas no processo de aprendizado dos modelos. Além disso, a aplicação da técnica de correção CLAHE apresentou um resultado discrepante em relação aos demais cenários de testes, indicando um possível sobreajuste dos modelos às condições específicas das imagens corrigidas.

As técnicas LCDPNet, LLNeRF e DSN não foram desenvolvidas especificamente para corrigir iluminação e cor em imagens de lesões de pele. Essas técnicas atendem a propósitos genéricos de correção em situações de sub-iluminação e sobre-iluminação. Embora as correções não tenham sido avaliadas por um especialista, é importante destacar que o uso dessas técnicas ainda é um desafio na literatura.

Apesar do sucesso de classificação com os modelos Inception-V3 e ResNet-50, cada arquitetura possui suas limitações, considerando a quantidade de camadas e o tamanho dos filtros de convolução. Filtros de convolução menores, como 3x3 ou 5x5, capturam características locais, enquanto filtros maiores, como 7x7, abrangem características mais globais. A quantidade de camadas permite uma melhor hierarquização de características, ou seja, aumentar o número de camadas permite que a rede processe informações visuais em múltiplos níveis de abstração.

Na identificação de aspectos de cores em lesões de pele, as diferenças sutis de tonalidade podem ser perdidas ao longo do treinamento, conforme o tamanho dos filtros e o número de camadas. A escolha do tamanho dos filtros e da quantidade de camadas deve ser considerada para otimizar a captura e o aprendizado das características visuais em imagens de câncer de pele, incluindo as nuances de cor essenciais para o diagnóstico preciso de malignidades. Por isso, uma explicação para os modelos Inception-V3 e ResNet-50 se saírem melhor nos experimentos é o menor número de camadas associado ao menor tamanho dos filtros de convolução, comparado com a Xception e DenseNet-121.

O uso do aumento de dados no conjunto de teste por *fold* com o TTA pode ter causado a inversão das predições dos modelos. Embora a aplicação dessa estratégia tenha melhorado consideravelmente os valores das métricas avaliadas, é necessário ter cautela ao utilizá-la, pois pode levar a taxas falsas de precisão, AUC e *Kappa*.

Outra limitação da abordagem é a otimização de parâmetros. Mesmo automatizando

esse processo com o *GridSearch*, pode ocorrer que essa ferramenta não alcance uma combinação ótima, devido à abundante quantidade de combinações possíveis de parâmetros. Além disso, dada a variação nos métodos de correção aplicados e nas arquiteturas utilizadas, a comparação direta com outros estudos é limitada. As diferenças nos conjuntos de dados, métodos de validação e métricas utilizadas podem resultar em discrepâncias nos resultados deste trabalho.

5.4 Considerações Finais

Este capítulo apresentou os principais resultados encontrados nos experimentos deste trabalho, além da análise e validação da abordagem de detecção do câncer de pele desenvolvida. Esses resultados demonstram como a aplicação de diferentes técnicas de correção de iluminação e cor pode influenciar significativamente o desempenho dos modelos de redes neurais profundas. Através do processo de ajuste de hiperparâmetros e da escolha das arquiteturas, como Inception-V3, Xception, ResNet-50 e DenseNet-121, foi possível explorar a sensibilidade desses modelos às variações de iluminação e cor. Os principais estudos comparados demonstram o potencial dessas características no diagnóstico do câncer maligno.

6 Conclusões e Trabalhos Futuros

A criação de sistemas automatizados para o diagnóstico precoce do câncer de pele continua sendo uma tarefa desafiadora. Muitas pesquisas destacam as particularidades e os vieses envolvidos na identificação de lesões malignas em imagens digitais. Embora a literatura mostre avanços nessas abordagens de diagnóstico, ainda não é possível definir um conjunto de técnicas eficazes para esse problema. No entanto, é possível direcionar novas pesquisas que podem levar a sistemáticas computadorizadas mais eficazes na identificação do câncer de pele maligno.

O aprendizado profundo é uma metodologia promissora para a identificação de doenças em imagens médicas. Esse fato já foi amplamente discutido em diversas pesquisas. Para o câncer de pele, muitos resultados impressionantes já foram apresentados, incluindo casos em que a precisão supera a do diagnóstico visual de dermatologistas. No entanto, muitas dessas técnicas precisam ser estudadas e ajustadas especificamente para lidar com imagens de câncer de pele. Além disso, a cor é uma característica que influencia significativamente a suspeita de lesões malignas.

Neste contexto, o estudo desenvolvido neste trabalho analisou a relação entre a iluminação, diretamente ligada à representação de cores em imagens digitais, e a classificação correta do câncer de pele utilizando modelos de aprendizado profundo. O objetivo foi investigar o quanto as características de iluminação e cor influenciam a precisão dos resultados em arquiteturas profundas para identificar câncer de pele maligno. A validação do estudo foi realizada por meio de uma abordagem sinérgica que emprega tanto técnicas de correção convencionais quanto novas, mais sofisticadas.

As bases de imagens disponíveis na literatura para a classificação do câncer de pele apresentam problemas significativos de iluminação e, conseqüentemente, de representações de cores. Para mitigar esses problemas, há uma variedade de técnicas de correção disponíveis. Porém, é importante identificar métodos que não distorçam as imagens das lesões. No estudo desenvolvido, constatamos que é possível aplicar várias técnicas de correção de iluminação e cor, sendo que as técnicas baseadas em histograma e balanço de branco tendem a não alterar significativamente essas características. Além disso, novas técnicas de aprendizado profundo podem ser aplicadas a esse problema, demonstrando uma capacidade superior no aprimoramento das imagens de lesões de pele.

Na prática, os experimentos realizados na abordagem de validação mostraram que a correção utilizando técnicas baseadas em histograma e balanço de branco, como CLAHE, *Perfect Reflector* e SoG, não produziram correções de iluminação e cor

superiores às técnicas baseadas em aprendizado profundo, como LCDPNet, LLNeRF e DSN. Isso permite afirmar que as técnicas de aprendizado profundo tendem a produzir melhores correções dessas características, melhorando o desempenho na classificação de lesões de pele malignas.

É importante destacar que a correção das características de iluminação e cor em imagens de lesões de pele é necessária devido aos problemas de captura e ao viés de confirmação. Na abordagem desenvolvida, a correção dessas características torna-se essencial devido à variabilidade das bases de imagens utilizadas. Nos experimentos realizados, observa-se que todos os cenários de correção produziram algum aumento na precisão da identificação correta de lesões malignas, em comparação à ausência de correção. Embora esses aumentos não tenham sido expressivos nos modelos avaliados, os ganhos em todas as taxas indicam que essas características também são percebidas em arquiteturas profundas.

Na maioria dos experimentos realizados, as arquiteturas Inception-V3 e ResNet-50, com menos camadas e filtros menores, mostraram-se mais sensíveis às variações de iluminação e cor em comparação com Xception e DenseNet-121. Isso significa que o *design* da arquitetura é essencial, pois uma maior quantidade de camadas, associada a filtros maiores, pode tornar um modelo treinado insensível a essas características. Essa sensibilidade destaca a importância da estrutura e organização das arquiteturas na percepção efetiva das variações desses aspectos.

Em suma, a correção dos aspectos de iluminação e cor não pode ser negligenciada na classificação do câncer de pele com modelos de aprendizado profundo. Embora as técnicas utilizadas não garantam melhores resultados com esses modelos, a escolha adequada de técnicas analíticas pode aumentar as chances de obter uma melhor precisão na identificação de malignidades. Além disso, a arquitetura do modelo também influencia a percepção dessas características. Um *design* e uma estrutura adequados podem tornar os modelos de classificação mais sensíveis às variações de iluminação e cor nas imagens digitais.

Outro aspecto que deve ser observado no estudo e na abordagem de validação é que o aumento de dados no conjunto de teste por *fold* com TTA produziu os melhores resultados de classificação. Esses resultados foram alcançados devido ao balanceamento do conjunto de teste. No entanto, o uso dessa estratégia também pode resultar na inversão de predições, diminuindo a concordância dos experimentos, especialmente nas métricas de AUC e *Kappa*. Na verificação de resultados, a inversão de predições não ocorreu com frequência significativa, indicando que essa técnica é uma alternativa viável para balancear bases de imagens de lesões de pele e melhorar o aprendizado dos modelos, mesmo quando aplicada ao conjunto de teste.

Por fim, após a apresentação das limitações do estudo e da abordagem descritas

neste trabalho, a comparação com outros estudos também revelou desafios na validação dos resultados alcançados, principalmente devido às diferenças entre as metodologias e bases de imagens utilizadas. Isso mostra que a abordagem de validação ainda necessita de aprofundamento e novos ajustes para obter um comparativo mais consistente. Além disso, são necessários estudos mais aprofundados sobre as interações entre iluminação e cor na classificação de lesões de pele, já que o uso de técnicas avançadas para a correção dessas características é limitado.

Para futuros estudos, é recomendado explorar outros métodos de correção de iluminação e cor, como métodos baseados em aprendizado auto-supervisionado, mecanismos de atenção e transferência de estilo. Um aprofundamento e personalização desses métodos podem proporcionar ajustes mais precisos, melhorando o aprendizado e a predição de modelos de aprendizado profundo. Outras arquiteturas profundas mais recentes também precisam ser avaliadas em relação às variações dos aspectos de iluminação e cor para o câncer de pele. Além disso, é incentivado o uso de arquiteturas com módulos de atenção para captar mais informações sobre as variações dessas características. Também é extremamente relevante desenvolver arquiteturas personalizadas que possam lidar melhor com a captura e o aprendizado de nuances de cor, frequentemente entendidas como texturas.

Referências

- ABDELHALIM, I. S. A.; MOHAMED, M. F.; MAHDY, Y. B. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. **Expert Systems with Applications**, Elsevier, v. 165, p. 113922, 2021. Citado 2 vezes nas páginas 28 e 29.
- ADEGUN, A.; VIRIRI, S. Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art. **Artificial Intelligence Review**, Springer, v. 54, p. 811–841, 2021. Citado 3 vezes nas páginas 15, 30 e 36.
- AFIFI, M.; BRUBAKER, M. A.; BROWN, M. S. Auto white-balance correction for mixed-illuminant scenes. In: **Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision**. [S.l.: s.n.], 2022. p. 1210–1219. Citado 2 vezes nas páginas 45 e 46.
- AL-MASNI, M. A.; KIM, D.-H.; KIM, T.-S. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. **Computer methods and programs in biomedicine**, Elsevier, v. 190, p. 105351, 2020. Citado 2 vezes nas páginas 24 e 27.
- ALCHORNE, M. M. d. A.; ABREU, M. A. M. M. d. Dermatologia na pele negra. **Anais Brasileiros de dermatologia**, SciELO Brasil, v. 83, p. 7–20, 2008. Citado na página 14.
- ARDIZZONE, L.; KRUSE, J.; WIRKERT, S.; RAHNER, D.; PELLEGRINI, E. W.; KLESSEN, R. S.; MAIER-HEIN, L.; ROTHER, C.; KÖTHE, U. Analyzing inverse problems with invertible neural networks. **arXiv preprint arXiv:1808.04730**, 2018. Citado na página 49.
- BALAHA, H. M.; HASSAN, A. E.-S. Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. **Neural Computing and Applications**, Springer, v. 35, n. 1, p. 815–853, 2023. Citado 2 vezes nas páginas 26 e 27.
- BARATA, C.; CELEBI, M. E.; MARQUES, J. S. Improving dermoscopy image classification using color constancy. **IEEE journal of biomedical and health informatics**, IEEE, v. 19, n. 3, p. 1146–1152, 2014. Citado 3 vezes nas páginas 14, 30 e 46.
- BARATA, C.; CELEBI, M. E.; MARQUES, J. S. A survey of feature extraction in dermoscopy image analysis of skin cancer. **IEEE journal of biomedical and health informatics**, IEEE, v. 23, n. 3, p. 1096–1109, 2018. Citado 5 vezes nas páginas 15, 28, 30, 44 e 53.
- BARATA, C.; MARQUES, J. S. Deep learning for skin cancer diagnosis with hierarchical architectures. In: IEEE. **2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)**. [S.l.], 2019. p. 841–845. Citado 4 vezes nas páginas 15, 22, 27 e 31.
- BARATA, C.; RUELA, M.; FRANCISCO, M.; MENDONÇA, T.; MARQUES, J. S. Two systems for the detection of melanomas in dermoscopy images using texture and color features. **IEEE systems Journal**, IEEE, v. 8, n. 3, p. 965–979, 2013. Citado na página 30.

- BISLA, D.; CHOROMANSKA, A.; BERMAN, R. S.; STEIN, J. A.; POLSKY, D. Towards automated melanoma detection with deep learning: Data purification and augmentation. In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops**. [S.l.: s.n.], 2019. p. 0–0. Citado 2 vezes nas páginas 28 e 29.
- BISSOTO, A.; BARATA, C.; VALLE, E.; AVILA, S. Artifact-based domain generalization of skin lesion models. In: SPRINGER. **European Conference on Computer Vision**. [S.l.], 2022. p. 133–149. Citado 4 vezes nas páginas 25, 26, 27 e 62.
- BISSOTO, A.; PEREZ, F.; RIBEIRO, V.; FORNACIALI, M.; AVILA, S.; VALLE, E. Deep-learning ensembles for skin-lesion segmentation, analysis, classification: Recod titans at isic challenge 2018. **arXiv preprint arXiv:1808.08480**, 2018. Citado 2 vezes nas páginas 28 e 29.
- BISSOTO, A.; PEREZ, F.; VALLE, E.; AVILA, S. Skin lesion synthesis with generative adversarial networks. In: SPRINGER. **OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis: First International Workshop, OR 2.0 2018, 5th International Workshop, CARE 2018, 7th International Workshop, CLIP 2018, Third International Workshop, ISIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 5**. [S.l.], 2018. p. 294–302. Citado na página 29.
- BISSOTO, A.; VALLE, E.; AVILA, S. Gan-based data augmentation and anonymization for skin-lesion analysis: A critical review. In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. [S.l.: s.n.], 2021. p. 1847–1856. Citado 2 vezes nas páginas 28 e 29.
- BONALUMI, A.; CAMPOS, E.; LEAL, F. **Oncologia Cutanea**. [S.l.]: Elsevier Brasil, 2017. Citado 3 vezes nas páginas 14, 34 e 35.
- BRANCIFORTI, F.; MEIBURGER, K. M.; ZAVATTARO, E.; VERONESE, F.; TARANTINO, V.; MAZZOLETTI, V.; CRISTO, N. D.; SAVOIA, P.; SALVI, M. Impact of artificial intelligence-based color constancy on dermoscopic assessment of skin lesions: A comparative study. **Skin Research and Technology**, Wiley Online Library, v. 29, n. 11, p. e13508, 2023. Citado 3 vezes nas páginas 15, 31 e 32.
- BRUNSGAARD, E.; JENSEN, J.; GROSSMAN, D. Melanoma in skin of color: Part ii. racial disparities, role of uv, and interventions for earlier detection. **Journal of the American Academy of Dermatology**, Elsevier, 2022. Citado 2 vezes nas páginas 15 e 35.
- BUCHSBAUM, G. A spatial processor model for object colour perception. **Journal of the Franklin institute**, Elsevier, v. 310, n. 1, p. 1–26, 1980. Citado na página 31.
- CHANG, Y.; GUO-PING, W. A review on image-based rendering. **Virtual Reality & Intelligent Hardware**, Elsevier, v. 1, n. 1, p. 39–54, 2019. Citado na página 14.
- CHEN, L.-C.; ZHU, Y.; PAPANDREOU, G.; SCHROFF, F.; ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: **Proceedings of the European conference on computer vision (ECCV)**. [S.l.: s.n.], 2018. p. 801–818. Citado na página 61.

CHLAP, P.; MIN, H.; VANDENBERG, N.; DOWLING, J.; HOLLOWAY, L.; HAWORTH, A. A review of medical image data augmentation techniques for deep learning applications. **Journal of Medical Imaging and Radiation Oncology**, Wiley Online Library, v. 65, n. 5, p. 545–563, 2021. Citado na página [40](#).

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 1251–1258. Citado na página [63](#).

CONG, C.; LIU, S.; IEVA, A. D.; PAGNUCCO, M.; BERKOVSKY, S.; SONG, Y. Colour adaptive generative networks for stain normalisation of histopathology images. **Medical Image Analysis**, Elsevier, v. 82, p. 102580, 2022. Citado na página [31](#).

CORTEZ, J. L.; VASQUEZ, J.; WEI, M. L. The impact of demographics, socioeconomic, and health care access on melanoma outcomes. **Journal of the American Academy of Dermatology**, Elsevier, v. 84, n. 6, p. 1677–1683, 2021. Citado na página [14](#).

DAGHRIR, J.; TLOG, L.; BOUCHOUICHA, M.; SAYADI, M. Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In: IEEE. **2020 5th international conference on advanced technologies for signal and image processing (ATSIP)**. [S.l.], 2020. p. 1–5. Citado 2 vezes nas páginas [23](#) e [27](#).

DANESHJOU, R.; VODRAHALLI, K.; NOVOA, R. A.; JENKINS, M.; LIANG, W.; ROTEMBERG, V.; KO, J.; SWETTER, S. M.; BAILEY, E. E.; GEVAERT, O. et al. Disparities in dermatology ai performance on a diverse, curated clinical image set. **Science advances**, American Association for the Advancement of Science, v. 8, n. 31, p. eabq6147, 2022. Citado na página [52](#).

DILDAR, M.; AKRAM, S.; IRFAN, M.; KHAN, H. U.; RAMZAN, M.; MAHMOOD, A. R.; ALSAIARI, S. A.; SAEED, A. H. M.; ALRADDADI, M. O.; MAHNASHI, M. H. Skin cancer detection: a review using deep learning techniques. **International journal of environmental research and public health**, MDPI, v. 18, n. 10, p. 5479, 2021. Citado na página [22](#).

DORRELL, D. N.; STROWD, L. C. Skin cancer detection technology. **Dermatologic clinics**, Elsevier, v. 37, n. 4, p. 527–536, 2019. Citado na página [15](#).

ELTAYEF, K.; LI, Y.; LIU, X. Detection of melanoma skin cancer in dermoscopy images. In: IOP PUBLISHING. **Journal of Physics: Conference Series**. [S.l.], 2017. v. 787, n. 1, p. 012034. Citado na página [57](#).

ESTEVA, A.; CHOU, K.; YEUNG, S.; NAIK, N.; MADANI, A.; MOTTAGHI, A.; LIU, Y.; TOPOL, E.; DEAN, J.; SOCHER, R. Deep learning-enabled medical computer vision. **NPJ digital medicine**, Nature Publishing Group UK London, v. 4, n. 1, p. 5, 2021. Citado 3 vezes nas páginas [15](#), [22](#) e [63](#).

ESTEVA, A.; KUPREL, B.; NOVOA, R. A.; KO, J.; SWETTER, S. M.; BLAU, H. M.; THRUN, S. Dermatologist-level classification of skin cancer with deep neural networks. **nature**, Nature Publishing Group, v. 542, n. 7639, p. 115–118, 2017. Citado 3 vezes nas páginas [14](#), [15](#) e [22](#).

FAIRCHILD, M. D. **Color appearance models**. [S.l.]: John Wiley & Sons, 2013. Citado 2 vezes nas páginas [45](#) e [46](#).

- FERLAY, J.; COLOMBET, M.; SOERJOMATARAM, I.; PARKIN, D. M.; PIÑEROS, M.; ZNAOR, A.; BRAY, F. Cancer statistics for the year 2020: An overview. **International journal of cancer**, Wiley Online Library, v. 149, n. 4, p. 778–789, 2021. Citado 3 vezes nas páginas 14, 15 e 39.
- FINLAYSON, G. D.; TREZZI, E. Shades of gray and colour constancy. In: SOCIETY FOR IMAGING SCIENCE AND TECHNOLOGY. **Color and Imaging Conference**. [S.l.], 2004. v. 2004, n. 1, p. 37–41. Citado 2 vezes nas páginas 31 e 56.
- FUNT, B.; SHI, L. The effect of exposure on maxrgb color constancy. In: SPIE. **Human Vision and Electronic Imaging XV**. [S.l.], 2010. v. 7527, p. 282–288. Citado na página 31.
- GALDRAN, A.; ALVAREZ-GILA, A.; MEYER, M. I.; SARATXAGA, C. L.; ARAÚJO, T.; GARROTE, E.; ARESTA, G.; COSTA, P.; MENDONÇA, A. M.; CAMPILHO, A. Data-driven color augmentation techniques for deep skin image analysis. **arXiv preprint arXiv:1703.03702**, 2017. Citado 3 vezes nas páginas 14, 30 e 75.
- GOMATHI, E.; JAYASHEELA, M.; THAMARAI, M.; GEETHA, M. Skin cancer detection using dual optimization based deep learning network. **Biomedical Signal Processing and Control**, Elsevier, v. 84, p. 104968, 2023. Citado 3 vezes nas páginas 25, 27 e 30.
- GONZALEZ, R. C.; WOODS, R. C. **Processamento digital de imagens**. [S.l.]: Pearson Educación, 2009. Citado 4 vezes nas páginas 43, 44, 45 e 46.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. [S.l.]: Editora Blucher, 2000. Citado na página 43.
- GONZALO-MARTÍN, C.; GARCÍA-PEDRERO, A.; LILLO-SAAVEDRA, M. Improving deep learning sorghum head detection through test time augmentation. **Computers and Electronics in Agriculture**, Elsevier, v. 186, p. 106179, 2021. Citado na página 62.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT press, 2016. Citado 4 vezes nas páginas 36, 37, 39 e 40.
- GUO, Y.; LIU, Y.; GEORGIU, T.; LEW, M. S. A review of semantic segmentation using deep neural networks. **International journal of multimedia information retrieval**, Springer, v. 7, p. 87–93, 2018. Citado 3 vezes nas páginas 22, 42 e 43.
- HARANGI, B.; BARAN, A.; HAJDU, A. Assisted deep learning framework for multi-class skin lesion classification considering a binary classification support. **Biomedical Signal Processing and Control**, Elsevier, v. 62, p. 102041, 2020. Citado 2 vezes nas páginas 23 e 27.
- HASAN, M. R.; FATEMI, M. I.; KHAN, M. M.; KAUR, M.; ZAGUIA, A. et al. Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks. **Journal of Healthcare Engineering**, Hindawi, v. 2021, 2021. Citado 2 vezes nas páginas 9 e 38.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 37, n. 9, p. 1904–1916, 2015. Citado 2 vezes nas páginas 37 e 38.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 770–778. Citado 3 vezes nas páginas 37, 38 e 63.

HE, X.; WANG, Y.; ZHAO, S.; CHEN, X. Co-attention fusion network for multimodal skin cancer diagnosis. **Pattern Recognition**, Elsevier, v. 133, p. 108990, 2023. Citado 2 vezes nas páginas 26 e 27.

HEIBEL, H. D.; HOOEY, L.; COCKERELL, C. J. A review of noninvasive techniques for skin cancer detection in dermatology. **American Journal of Clinical Dermatology**, Springer, v. 21, p. 513–524, 2020. Citado 3 vezes nas páginas 14, 22 e 36.

HONG, Y.; ZHANG, G.; WEI, B.; CONG, J.; XU, Y.; ZHANG, K. Weakly supervised semantic segmentation for skin cancer via cnn superpixel region response. **Multimedia Tools and Applications**, Springer, v. 82, n. 5, p. 6829–6847, 2023. Citado 2 vezes nas páginas 9 e 42.

HUANG, G.; LIU, Z.; MAATEN, L. V. D.; WEINBERGER, K. Q. Densely connected convolutional networks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 4700–4708. Citado na página 63.

HUANG, J.; ZHAO, F.; ZHOU, M.; XIAO, J.; ZHENG, N.; ZHENG, K.; XIONG, Z. Learning sample relationship for exposure correction. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2023. p. 9904–9913. Citado na página 30.

International Skin Imaging Collaboration. **SIIM-ISIC 2020 Challenge Dataset**. International Skin Imaging Collaboration, 2020. Disponível em: <<https://challenge2020.isic-archive.com/>>. Citado 3 vezes nas páginas 9, 16 e 52.

IYATOMI, H.; CELEBI, M. E.; SCHAEFER, G.; TANAKA, M. Automated color calibration method for dermoscopy images. **Computerized Medical Imaging and Graphics**, Elsevier, v. 35, n. 2, p. 89–98, 2011. Citado na página 44.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. et al. **An introduction to statistical learning**. [S.l.]: Springer, 2013. v. 112. Citado na página 65.

JONES, O.; MATIN, R.; SCHAAR, M. van der; BHAYANKARAM, K. P.; RANMUTHU, C.; ISLAM, M.; BEHIYAT, D.; BOSCOFF, R.; CALANZANI, N.; EMERY, J. et al. Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review. **The Lancet Digital Health**, Elsevier, v. 4, n. 6, p. e466–e476, 2022. Citado na página 30.

KAWAHARA, J.; BENTAIEB, A.; HAMARNEH, G. Deep features to classify skin lesions. In: IEEE. **2016 IEEE 13th international symposium on biomedical imaging (ISBI)**. [S.l.], 2016. p. 1397–1400. Citado 2 vezes nas páginas 15 e 37.

KELLEHER, J. D. **Deep learning**. [S.l.]: MIT press, 2019. Citado 4 vezes nas páginas 9, 37, 39 e 40.

KIM, H. E.; COSA-LINAN, A.; SANTHANAM, N.; JANNESARI, M.; MAROS, M. E.; GANSLANDT, T. Transfer learning for medical image classification: a literature review. **BMC medical imaging**, Springer, v. 22, n. 1, p. 69, 2022. Citado na página 42.

- KORNBLITH, S.; SHLENS, J.; LE, Q. V. Do better imagenet models transfer better? In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. [S.l.: s.n.], 2019. p. 2661–2671. Citado na página 42.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Communications of the ACM**, AcM New York, NY, USA, v. 60, n. 6, p. 84–90, 2017. Citado na página 64.
- LAINA, I.; RUPPRECHT, C.; BELAGIANNIS, V.; TOMBARI, F.; NAVAB, N. Deeper depth prediction with fully convolutional residual networks. In: IEEE. **2016 Fourth international conference on 3D vision (3DV)**. [S.l.], 2016. p. 239–248. Citado na página 38.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado 4 vezes nas páginas 36, 37, 39 e 40.
- LEE, T.; NG, V.; GALLAGHER, R.; COLDMAN, A.; MCLEAN, D. Dullrazor®: A software approach to hair removal from images. **Computers in biology and medicine**, Elsevier, v. 27, n. 6, p. 533–543, 1997. Citado 2 vezes nas páginas 53 e 54.
- LIU, Y.; JI, Z.; CEN, G.; SUN, H.; WANG, H.; ZHAO, C.; WANG, Z. L.; MAI, W. **Deep Symmetric Network Enhancement**. 2023. Acessado em 2 de dezembro de 2023. Disponível em: <<https://github.com/lin-zhao-resoLve/Deep-Symmetric-Network-Enhancement>>. Citado na página 60.
- MAGALHÃES, R.; BERNARDINO, A. Quantifying object detection uncertainty in autonomous driving with test-time augmentation. In: IEEE. **2023 IEEE Intelligent Vehicles Symposium (IV)**. [S.l.], 2023. p. 1–7. Citado na página 62.
- MAHBOD, A.; SCHAEFER, G.; WANG, C.; ECKER, R.; ELLINGE, I. Skin lesion classification using hybrid deep neural networks. In: IEEE. **ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.], 2019. p. 1229–1233. Citado 3 vezes nas páginas 22, 23 e 27.
- MAJTNER, T.; YILDIRIM-YAYILGAN, S.; HARDEBERG, J. Y. Optimised deep learning features for improved melanoma detection. **Multimedia Tools and Applications**, Springer, v. 78, p. 11883–11903, 2019. Citado 3 vezes nas páginas 22, 23 e 27.
- MARON, R. C.; HAGGENMÜLLER, S.; KALLE, C. von; UTIKAL, J. S.; MEIER, F.; GELLRICH, F. F.; HAUSCHILD, A.; FRENCH, L. E.; SCHLAAK, M.; GHORESCHI, K. et al. Robustness of convolutional neural networks in recognition of pigmented skin lesions. **European journal of cancer**, Elsevier, v. 145, p. 81–91, 2021. Citado na página 62.
- MARQUES, C.; BARRETO, C.; MORAIS, V.; JˆNIOR, N.; MORAIS, A.; CARVALHO, A. **Oncologia: uma abordagem multidisciplinar**. [S.l.]: Carpe Diem, 2016. Citado 5 vezes nas páginas 14, 15, 34, 35 e 36.
- MARTIN-BRUALLA, R.; RADWAN, N.; SAJJADI, M. S.; BARRON, J. T.; DOSOVITSKIY, A.; DUCKWORTH, D. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2021. p. 7210–7219. Citado 2 vezes nas páginas 9 e 48.

MAZOURE, B.; MAZOURE, A.; BÉDARD, J.; MAKARENKOV, V. Dunescan: a web server for uncertainty estimation in skin cancer detection with deep neural networks. **Scientific Reports**, Nature Publishing Group UK London, v. 12, n. 1, p. 179, 2022. Citado 3 vezes nas páginas 24, 26 e 27.

MENDONÇA, T.; FERREIRA, P. M.; MARQUES, J. S.; MARCAL, A. R.; ROZEIRA, J. Ph 2-a dermoscopic image database for research and benchmarking. In: IEEE. **2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)**. [S.l.], 2013. p. 5437–5440. Citado na página 52.

MILDENHALL, B.; SRINIVASAN, P. P.; TANCIK, M.; BARRON, J. T.; RAMAMOORTHY, R.; NG, R. Nerf: Representing scenes as neural radiance fields for view synthesis. **Communications of the ACM**, ACM New York, NY, USA, v. 65, n. 1, p. 99–106, 2021. Citado 2 vezes nas páginas 47 e 48.

MORID, M. A.; BORJALI, A.; FIOL, G. D. A scoping review of transfer learning research on medical image analysis using imagenet. **Computers in biology and medicine**, Elsevier, v. 128, p. 104115, 2021. Citado 2 vezes nas páginas 42 e 64.

MURAD, A. M.; KATZ, A. Oncologia: bases clínicas do tratamento. In: **Oncologia: bases clínicas do tratamento**. [S.l.: s.n.], 1996. p. 435–435. Citado 2 vezes nas páginas 34 e 36.

MUSA, P.; RAFI, F. A.; LAMSANI, M. A review: Contrast-limited adaptive histogram equalization (clahe) methods to help the application of face recognition. In: IEEE. **2018 third international conference on informatics and computing (ICIC)**. [S.l.], 2018. p. 1–6. Citado na página 55.

NAHATA, H.; SINGH, S. P. Deep learning solutions for skin cancer detection and diagnosis. **Machine learning with health care perspective: machine learning and healthcare**, Springer, p. 159–182, 2020. Citado 3 vezes nas páginas 23, 27 e 75.

NASCIMENTO, S. M.; AMANO, K.; FOSTER, D. H. Spatial distributions of local illumination color in natural scenes. **Vision research**, Elsevier, v. 120, p. 39–44, 2016. Citado 2 vezes nas páginas 46 e 47.

NAWAZ, M.; MASOOD, M.; JAVED, A.; IQBAL, J.; NAZIR, T.; MEHMOOD, A.; ASHRAF, R. Melanoma localization and classification through faster region-based convolutional neural network and svm. **Multimedia Tools and Applications**, Springer, v. 80, n. 19, p. 28953–28974, 2021. Citado 2 vezes nas páginas 24 e 27.

NG, J. hua; GOYAL, M.; HEWITT, B.; YAP, M. H. The effect of color constancy algorithms on semantic segmentation of skin lesions. In: SPIE. **Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging**. [S.l.], 2019. v. 10953, p. 138–145. Citado na página 31.

PACHECO, A. G.; LIMA, G. R.; SALOMAO, A. S.; KROHLING, B.; BIRAL, I. P.; ANGELO, G. G. de; JR, F. C. A.; ESGARIO, J. G.; SIMORA, A. C.; CASTRO, P. B. et al. Pad-ufes-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. **Data in brief**, Elsevier, v. 32, p. 106221, 2020. Citado na página 52.

PANDEY, P.; SAURABH, P.; VERMA, B.; TIWARI, B. A multi-scale retinex with color restoration (msr-cr) technique for skin cancer detection. In: SPRINGER. **Soft Computing for Problem Solving: SocProS 2017, Volume 2**. [S.l.], 2019. p. 465–473. Citado na página 32.

Paszke, Adam and Gross, Sam and Chintala, Soumith and Chanan, Gregory. **DeepLabV3 Model Architecture**. 2018. Acessado em 5 de junho de 2023. Disponível em: <<https://pytorch.org/vision/main/models/deeplabv3.html>>. Citado na página 61.

PENCZEK, J.; BOYNTON, P. A.; SPLETT, J. D. Color error in the digital camera image capture process. **Journal of digital imaging**, Springer, v. 27, p. 182–191, 2014. Citado 2 vezes nas páginas 43 e 44.

PEREZ, F.; AVILA, S.; VALLE, E. Solo or ensemble? choosing a cnn architecture for melanoma classification. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops**. [S.l.: s.n.], 2019. p. 0–0. Citado 3 vezes nas páginas 23, 27 e 42.

PETSCHNIGG, G.; SZELISKI, R.; AGRAWALA, M.; COHEN, M.; HOPPE, H.; TOYAMA, K. Digital photography with flash and no-flash image pairs. **ACM transactions on graphics (TOG)**, ACM New York, NY, USA, v. 23, n. 3, p. 664–672, 2004. Citado na página 44.

PHAM, T.-C.; LUONG, C.-M.; VISANI, M.; HOANG, V.-D. Deep cnn and data augmentation for skin lesion classification. In: SPRINGER. **Intelligent Information and Database Systems: 10th Asian Conference, ACIIDS 2018, Dong Hoi City, Vietnam, March 19-21, 2018, Proceedings, Part II 10**. [S.l.], 2018. p. 573–582. Citado na página 41.

POUCKE, S. V.; HAEGHEN, Y. V.; VISSERS, K.; MEERT, T.; JORENS, P. Automatic colorimetric calibration of human wounds. **BMC medical imaging**, Springer, v. 10, p. 1–11, 2010. Citado na página 44.

QIN, Z.; LIU, Z.; ZHU, P.; XUE, Y. A gan-based image synthesis method for skin lesion classification. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 195, p. 105568, 2020. Citado 3 vezes nas páginas 28, 29 e 40.

RASHID, H.; TANVEER, M. A.; KHAN, H. A. Skin lesion classification using gan based data augmentation. In: IEEE. **2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)**. [S.l.], 2019. p. 916–919. Citado 2 vezes nas páginas 40 e 41.

RAZMJOOY, N.; ARSHAGHI, A. Application of multilevel thresholding and cnn for the diagnosis of skin cancer utilizing a multi-agent fuzzy buzzard algorithm. **Biomedical Signal Processing and Control**, Elsevier, v. 84, p. 104984, 2023. Citado 2 vezes nas páginas 25 e 27.

REZA, A. M. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. **Journal of VLSI signal processing systems for signal, image and video technology**, Springer, v. 38, p. 35–44, 2004. Citado na página 55.

RUBEGNI, P.; BURRONI, M.; PEROTTI, R.; FIMIANI, M.; ANDREASSI, L.; CEVENINI, G.; DELL'EVA, G.; BARBINI, P. Digital dermoscopy analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: a retrospective study.

Journal of investigative dermatology, Elsevier, v. 119, n. 2, p. 471–474, 2002. Citado na página [30](#).

SALVI, M.; BRANCIFORTI, F.; MOLINARI, F.; MEIBURGER, K. M. Generative models for color normalization in digital pathology and dermatology: Advancing the learning paradigm. **Expert Systems with Applications**, Elsevier, v. 245, p. 123105, 2024. Citado na página [31](#).

SALVI, M.; BRANCIFORTI, F.; VERONESE, F.; ZAVATTARO, E.; TARANTINO, V.; SAVOIA, P.; MEIBURGER, K. M. Dermocc-gan: A new approach for standardizing dermatological images using generative adversarial networks. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 225, p. 107040, 2022. Citado 7 vezes nas páginas [15](#), [30](#), [31](#), [32](#), [39](#), [44](#) e [75](#).

SANTOS, M. de O.; LIMA, F. C. de S. de; MARTINS, L. F. L.; OLIVEIRA, J. F. P.; ALMEIDA, L. M. de; CANCELA, M. de C. Estimativa de incidência de câncer no Brasil, 2023-2025. **Revista Brasileira de Cancerologia**, v. 69, n. 1, 2023. Citado 2 vezes nas páginas [14](#) e [35](#).

SHANMUGAM, D.; BLALOCK, D.; BALAKRISHNAN, G.; GUTTAG, J. Better aggregation in test-time augmentation. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)**. [S.l.: s.n.], 2021. p. 1214–1223. Citado na página [62](#).

SHE, Z.; DICKSON, M.; FISH, P.; DULLER, A.; BAMBER, J. Specular reflection reduction in context of skin pattern analysis. In: **Medical Image Understanding and Analysis**. [S.l.: s.n.], 2001. p. 16–19. Citado na página [57](#).

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. **Journal of big data**, SpringerOpen, v. 6, n. 1, p. 1–48, 2019. Citado na página [41](#).

SILVA, J. V. M. da; AIRES, K. R.; SANTOS, A. R. dos; VERAS, R. de M.; NETO, L. de S.; SOUSA, L. P. de; FILHO, F. das C. Segmentação semântica do câncer de pele utilizando aprendizado profundo. In: SBC. **Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde**. [S.l.], 2023. p. 304–315. Citado 2 vezes nas páginas [43](#) e [61](#).

SUNG, H.; FERLAY, J.; SIEGEL, R. L.; LAVERSANNE, M.; SOERJOMATARAM, I.; JEMAL, A.; BRAY, F. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, Wiley Online Library, v. 71, n. 3, p. 209–249, 2021. Citado na página [14](#).

SWATI, Z. N. K.; ZHAO, Q.; KABIR, M.; ALI, F.; ALI, Z.; AHMED, S.; LU, J. Brain tumor classification for mr images using transfer learning and fine-tuning. **Computerized Medical Imaging and Graphics**, Elsevier, v. 75, p. 34–46, 2019. Citado na página [41](#).

SZEGEDY, C.; VANHOUCHE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 2818–2826. Citado na página [63](#).

TAGHANAKI, S. A.; ABHISHEK, K.; COHEN, J. P.; COHEN-ADAD, J.; HAMARNEH, G. Deep semantic segmentation of natural and medical images: a review. **Artificial**

Intelligence Review, Springer, v. 54, p. 137–178, 2021. Citado 2 vezes nas páginas 42 e 43.

TANG, H.; ZHU, H.; FEI, L.; WANG, T.; CAO, Y.; XIE, C. Low-illumination image enhancement based on deep learning techniques: A brief review. In: MDPI. **Photonics**. [S.l.], 2023. v. 10, n. 2, p. 198. Citado na página 32.

TORREY, L.; SHAVLIK, J. Transfer learning. In: **Handbook of research on machine learning applications and trends: algorithms, methods, and techniques**. [S.l.]: IGI global, 2010. p. 242–264. Citado na página 41.

TOSTA, T. A. A.; FREITAS, A. D.; FARIA, P. R. de; NEVES, L. A.; MARTINS, A. S.; NASCIMENTO, M. Z. do. A stain color normalization with robust dictionary learning for breast cancer histological images processing. **Biomedical Signal Processing and Control**, Elsevier, v. 85, p. 104978, 2023. Citado na página 31.

TRICCO, A. C.; ANTONY, J.; ZARIN, W.; STRIFLER, L.; GHASSEMI, M.; IVORY, J.; PERRIER, L.; HUTTON, B.; MOHER, D.; STRAUS, S. E. A scoping review of rapid review methods. **BMC medicine**, Springer, v. 13, p. 1–15, 2015. Citado na página 21.

TSCHANDL, P.; HOFMANN, L.; FINK, C.; KITTLER, H.; HAENSSLE, H. Melanomas vs. nevi in high-risk patients under long-term monitoring with digital dermatoscopy: do melanomas and nevi already differ at baseline? **Journal of the European Academy of Dermatology and Venereology**, Wiley Online Library, v. 31, n. 6, p. 972–977, 2017. Citado 2 vezes nas páginas 9 e 35.

TURSUN, O.; DENMAN, S.; SRIDHARAN, S.; FOOKES, C. Learning test-time augmentation for content-based image retrieval. **Computer Vision and Image Understanding**, Elsevier, v. 222, p. 103494, 2022. Citado na página 62.

VALLE, E.; FORNACIALI, M.; MENEGOLA, A.; TAVARES, J.; BITTENCOURT, F. V.; LI, L. T.; AVILA, S. Data, depth, and design: Learning reliable models for skin lesion analysis. **Neurocomputing**, Elsevier, v. 383, p. 303–313, 2020. Citado 3 vezes nas páginas 24, 27 e 28.

VENUGOPAL, V.; NATH, M. K.; JOSEPH, J.; DAS, M. V. A deep learning-based illumination transform for devignetting photographs of dermatological lesions. **Image and Vision Computing**, Elsevier, p. 104909, 2024. Citado na página 32.

VIEIRA, S. C.; LUSTOSA, A. M. L.; BARBOSA, C. N. B.; TEIXEIRA, J. M. R.; BRITO, L.; SOARES, L. F. M.; FERREIRA, M. A. T. Oncologia básica. **Teresina, MA Fundação Quixote**, 2012. Citado 2 vezes nas páginas 34 e 35.

WANG, H.; XU, K.; LAU, R. W. **LCDPNet: Local Color Distribution Preservation Network**. 2022. Acessado em 16 de setembro de 2023. Disponível em: <<https://github.com/onpix/LCDPNet>>. Citado na página 58.

WANG, H.; XU, K.; LAU, R. W. Local color distributions prior for image enhancement. In: SPRINGER. **European Conference on Computer Vision**. [S.l.], 2022. p. 343–359. Citado 4 vezes nas páginas 44, 46, 47 e 58.

WANG, H.; XU, X.; XU, K.; LAU, R. W. Lighting up nerf via unsupervised decomposition and enhancement. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision**. [S.l.: s.n.], 2023. p. 12632–12641. Citado 5 vezes nas páginas 44, 47, 48, 49 e 59.

WANG, H.; XU, X.; XU, K.; LAU, R. W. **LLNeRF: Lightweight Local Neural Radiance Fields**. 2023. Acessado em 20 de novembro de 2023. Disponível em: <<https://github.com/onpik/LLNeRF>>. Citado na página 59.

WANG, R.; ZHANG, Q.; FU, C.-W.; SHEN, X.; ZHENG, W.-S.; JIA, J. Underexposed photo enhancement using deep illumination estimation. In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. [S.l.: s.n.], 2019. p. 6849–6857. Citado na página 46.

XU, Q.; XU, Z.; PHILIP, J.; BI, S.; SHU, Z.; SUNKAVALLI, K.; NEUMANN, U. Point-nerf: Point-based neural radiance fields. In: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. [S.l.: s.n.], 2022. p. 5438–5448. Citado 2 vezes nas páginas 47 e 48.

YAN, L. C.; YOSHUA, B.; GEOFFREY, H. Deep learning. **nature**, v. 521, n. 7553, p. 436–444, 2015. Citado na página 36.

YANG, M.; XIE, K.; LI, T.; YANG, Z. Color constancy based on deep residual learning. In: IEEE. **2021 IEEE 4th International Conference on Electronics and Communication Engineering (ICECE)**. [S.l.], 2021. p. 109–112. Citado 2 vezes nas páginas 9 e 47.

ZHANG, L.; GAO, H. J.; ZHANG, J.; BADAMI, B. Optimization of the convolutional neural networks for automatic detection of skin cancer. **Open Medicine**, De Gruyter, v. 15, n. 1, p. 27–37, 2020. Citado na página 41.

ZHAO, L.; LU, S.-P.; CHEN, T.; YANG, Z.; SHAMIR, A. Deep symmetric network for underexposed image enhancement with recurrent attentional learning. In: **Proceedings of the IEEE/CVF international conference on computer vision**. [S.l.: s.n.], 2021. p. 12075–12084. Citado 5 vezes nas páginas 9, 44, 49, 50 e 60.

ZHOU, K.; GREENSPAN, H.; SHEN, D. **Deep learning for medical image analysis**. [S.l.]: Academic Press, 2017. Citado 3 vezes nas páginas 9, 36 e 37.