



Universidade Federal de Ouro Preto
Escola de Minas
CECAU - Colegiado do Curso de
Engenharia de Controle e Automação



Erlane Silva Rezende

Análise da Eficácia de Redes Neurais Artificiais na Detecção de Lesões Cutâneas em Diferentes Fototipos

Monografia de Graduação

Ouro Preto, 2024

Erlane Silva Rezende

Análise da Eficácia de Redes Neurais Artificiais na Detecção de Lesões Cutâneas em Diferentes Fototipos

Trabalho apresentado ao Colegiado do Curso de Engenharia de Controle e Automação da Universidade Federal de Ouro Preto como parte dos requisitos para a obtenção do Grau de Engenheiro(a) de Controle e Automação.

Universidade Federal de Ouro Preto

Orientador: Prof. Jadson Castro Gertrudes, PhD.

Coorientador: Profa. Regiane de Sousa e Silva Ramalho, MSc.

Ouro Preto

2024



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA CONTROLE E
AUTOMACAO



FOLHA DE APROVAÇÃO

Erlane Silva Rezende

Análise da Eficácia de Redes Neurais Artificiais na Detecção de Lesões Cutâneas em Diferentes Fototipos

Monografia apresentada ao Curso de Engenharia de Controle e Automação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheiro de Controle e Automação

Aprovada em 09 de outubro de 2024

Membros da banca

Dr. Jadson Castro Gertrudes- Orientador - Universidade Federal de Ouro Preto
MSc. Regiane de Sousa e Silva Ramalho - Coorientadora - Universidade Federal de Ouro Preto
Júlia Eduarda Miranda de Sousa - Convidada - PPG em Ciência da Computação
Thiago Camargo da Costa - Convidado - PPG em Ciência da Computação

Jadson Castro Gertrudes, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 09/10/2024



Documento assinado eletronicamente por **Jadson Castro Gertrudes, PROFESSOR DE MAGISTERIO SUPERIOR**, em 18/10/2024, às 11:29, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0792628** e o código CRC **252DD0CC**.

Agradecimentos

Minha sincera gratidão e reconhecimento a todos que contribuíram para a realização deste Trabalho de Conclusão de Curso:

Agradeço profundamente ao meu orientador Jadson, cujo conhecimento, orientação e dedicação foram essenciais para a conclusão deste trabalho.

Aos meus pais, Júlio e Suzeléia, pelo apoio emocional e compreensão durante toda a minha jornada acadêmica.

À minha irmã, Júlia, cuja disposição em compartilhar seu conhecimento enriqueceu significativamente o estudo.

Aos meus amigos e colegas, que compartilharam experiências, forneceram ajuda mútua e incentivaram meu progresso. Suas contribuições amigáveis tornaram esta jornada mais significativa e prazerosa.

À UFOP pelo ensino público de qualidade, à equipe acadêmica e todos os profissionais envolvidos no meu percurso acadêmico.

Agradeço a todos que, de alguma forma, direta ou indiretamente, contribuíram para o sucesso deste projeto. Obrigada por fazerem parte desta jornada acadêmica e por tornarem possível a realização deste TCC.

Resumo

A eficácia das redes neurais artificiais (RNAs) na análise de lesões cutâneas é o objeto de estudo deste trabalho com foco na disparidade de desempenho relacionada à diversidade racial nos conjuntos de dados utilizados. Os experimentos foram organizados de acordo com categorias de tonalidades de pele segundo a escala de Fitzpatrick. Os resultados permitiram identificar limitações na capacidade de generalização do modelo. A precisão dos diagnósticos foi maior para lesões benignas e o modelo ViT apresentou desempenho superior ao EfficientNet V2 em várias métricas. Além disso, observou-se uma queda de desempenho em peles classificadas nos tipos 5 e 6 da escala de Fitzpatrick.

Palavras-chaves: Redes neurais artificiais. Lesões cutâneas. Diversidade racial. Viés racial. Diagnóstico assistido por computador. Inteligência artificial em saúde. ViT (Vision Transformer). EfficientNet V2. Escala de Fitzpatrick. Generalização de modelos. Tonalidades de pele.

Abstract

The effectiveness of artificial neural networks (ANNs) in analyzing skin lesions is the subject of this study, focusing on the disparity in performance related to racial diversity in the data sets used. The experiments were organized according to skin tone categories according to the Fitzpatrick scale. The results identified limitations in the model's ability to generalize. Diagnostic accuracy was higher for benign lesions and the ViT model outperformed EfficientNet V2 in several metrics. In addition, a drop in performance was observed in skins classified as types 5 and 6 on the Fitzpatrick scale.

Key-words: Artificial Neural Networks. Skin Lesions. Racial Diversity. Racial Bias. Computer-Assisted Diagnosis. Health Artificial Intelligence. Vision Transformer (ViT). EfficientNet V2. Fitzpatrick Skin Type Scale. Model Generalization. Skin Tone Categories.

Lista de ilustrações

| | |
|---|----|
| Figura 1 – Estágios de formação do câncer | 15 |
| Figura 2 – Diferenças entre tipos de tumores | 15 |
| Figura 3 – Metástase | 16 |
| Figura 4 – Incidência e mortalidade por continentes, referentes aos cânceres de pele não melanoma e melanoma em 2022. | 17 |
| Figura 5 – Incidência e mortalidade por países LAC, referentes aos cânceres de pele não melanoma e melanoma em 2022. | 17 |
| Figura 6 – Indicadores de saúde: atenção básica, recursos, mortalidade e morbidade e gastos. | 20 |
| Figura 7 – Inteligência artificial, aprendizado de máquina e aprendizado profundo. | 21 |
| Figura 8 – Imagem interpretada por uma máquina. | 22 |
| Figura 9 – Convolução entre imagem de entrada 6x6 e filtro 3x3. | 23 |
| Figura 10 – Mapa de ativação gerado com dimensões 4x4. | 23 |
| Figura 11 – Gráfico demonstrando a função não linear ReLU. | 24 |
| Figura 12 – Camada não linear aplicada em mapa de ativação. | 24 |
| Figura 13 – Camada de agrupamento máximo e médio. | 25 |
| Figura 14 – Aplicação de flattening. | 25 |
| Figura 15 – Camadas totalmente conectadas. | 26 |
| Figura 16 – Estrutura básica de uma rede neural convolucional. | 26 |
| Figura 17 – Visão Geral do Modelo. | 29 |
| Figura 18 – Amostras de imagens do conjunto de treinamento. | 41 |
| Figura 19 – Amostras de imagens do conjunto de teste. | 41 |
| Figura 20 – Predições do modelo EfficientNet_V2 no experimento 2.2. | 42 |
| Figura 21 – Predições do modelo ViT no experimento 2.2. | 42 |
| Figura 22 – Predições do modelo EfficientNet_V2 no experimento 3.2. | 43 |
| Figura 23 – Predições do modelo ViT no experimento 3.2. | 43 |

Lista de tabelas

| | |
|---|----|
| Tabela 1 – Fototipos na Escala Fitzpatrick. | 12 |
| Tabela 2 – Quantidade de imagens PAD-UFES-20 | 32 |
| Tabela 3 – Quantidade de imagens ISIC | 33 |
| Tabela 4 – Quantidade de imagens DDI | 34 |
| Tabela 5 – Resultado de métricas do experimento 1.1 | 36 |
| Tabela 6 – Resultado de métricas do experimento 1.2 | 36 |
| Tabela 7 – Resultado de métricas do experimento 2.1 | 37 |
| Tabela 8 – Resultado de métricas do experimento 2.2 | 37 |
| Tabela 9 – Resultado de métricas do experimento 2.3 | 38 |
| Tabela 10 – Resultado de métricas do experimento 3.1 | 38 |
| Tabela 11 – Resultado de métricas do experimento 3.2 | 39 |
| Tabela 12 – Resultado de métricas do experimento 3.3 | 39 |
| Tabela 13 – Resultado de métricas do experimento com imagens dos conjuntos PAD-UFES-20 e ISIC | 40 |
| Tabela 14 – Resultado de métricas do experimento com imagens dos conjuntos PAD-UFES-20, ISIC e DDI | 40 |
| Tabela 15 – Média de precisões por categoria de experimento | 41 |

Lista de abreviaturas e siglas

| | |
|------|--|
| RNAs | Redes Neurais Artificiais |
| ISIC | International Skin Imaging Collaboration |
| DDI | Diverse Dermatology Images |
| CBC | Carcinoma Basocelular |
| CEC | Carcinoma Espinoceleular |
| SBD | Sociedade Brasileira de Dermatologia |
| LAC | Latin America and the Caribbean |
| UV | Radiação Ultravioleta |
| IEPS | Instituto de Estudos para Políticas de Saúde |
| RNC | Rede Neural Convolutcional |
| ReLu | Rectified Linear Unit |
| ViT | Vision Transformer |
| PAD | Programa de Assistência Dermatológica |
| ISIC | International Skin Imaging Collaboration |
| IA | Inteligência Artificial |
| DDI | Diverse Dermatology Images |

Sumário

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 11 |
| 1.1 | Contextualização | 11 |
| 1.2 | Justificativas e Relevância | 12 |
| 1.3 | Materiais e Métodos | 12 |
| 1.4 | Objetivos | 13 |
| 1.5 | Organização e Estrutura | 13 |
| 2 | REVISÃO DE LITERATURA | 15 |
| 2.1 | O Câncer de Pele | 15 |
| 2.1.1 | Fatores de Risco | 17 |
| 2.1.2 | Campanha Dezembro Laranja | 18 |
| 2.2 | O Câncer na Pele Negra | 18 |
| 2.2.1 | Desigualdades Regionais no Sistema de Saúde Brasileiro | 19 |
| 2.3 | Aprendizado Profundo | 21 |
| 2.4 | Redes Neurais Convolucionais | 22 |
| 2.4.1 | EfficientNetV2 | 26 |
| 2.5 | Transformers Adaptados para Imagens | 27 |
| 2.5.1 | Vision Transformer (ViT) | 28 |
| 2.6 | Trabalhos Relacionados | 29 |
| 3 | METODOLOGIA | 31 |
| 3.1 | Bases de Dados | 31 |
| 3.2 | Ambiente de Implementação | 34 |
| 4 | AVALIAÇÃO DE DESEMPENHO | 35 |
| 4.1 | Experimento 1 | 35 |
| 4.2 | Experimento 2 | 36 |
| 4.3 | Experimento 3 | 38 |
| 4.4 | Síntese de Resultados | 39 |
| 5 | CONCLUSÃO | 44 |
| | Referências | 45 |

1 Introdução

1.1 Contextualização

Diversas pesquisas evidenciam a notável eficácia das redes neurais artificiais (RNAs) em sistemas de suporte a análise de lesões cutâneas, tanto para aplicações de classificação de padrões quanto para a automação da extração de atributos. No entanto, esses modelos de aprendizagem profunda podem encontrar limitações relacionadas à generalização.

Infelizmente, os conjuntos de dados de lesões cutâneas frequentemente carecem de diversidade, abrangendo principalmente amostras de pessoas com pele branca. Isso cria um problema significativo, pois pode resultar em modelos com preconceitos raciais, apresentando melhor desempenho no diagnóstico de lesões bem representadas em indivíduos brancos, enquanto podem ter dificuldades para diagnosticar com precisão lesões em pele negra (BARROS; AVILA, 2022).

Esse viés ocorre porque uma mesma lesão pode ter características distintas dependendo da cor da pele do indivíduo. Assim, se o diagnóstico é baseado nas características de lesões em pele branca, pode haver dificuldades em classificá-las corretamente em pele negra, impactando significativamente o diagnóstico e o tratamento dessas lesões (BARROS; CHAVES; AVILA, 2023).

Em pessoas com pele negra, os tumores frequentemente são diagnosticados em estágios mais avançados e apresentam prognóstico mais desfavorável. Diversos fatores contribuem para isso, incluindo a falta de experiência médica no diagnóstico e tratamento de doenças cutâneas em diferentes etnias, a crença errônea de que a pele negra é totalmente protegida contra o câncer de pele e o menor acesso a cuidados médicos e triagem preventiva devido a fatores socioeconômicos e culturais (EID; ALCHORNE, 2011).

No entanto, considerando a grande população negra no Brasil, é essencial que os médicos se familiarizem com as variações do câncer de pele em peles mais pigmentadas (EID; ALCHORNE, 2011).

Para avaliar modelos de aprendizado profundo no diagnóstico de câncer de pele em diferentes fotótipos, será utilizado a escala Fitzpatrick para classificar as imagens. Desenvolvida em 1976 pelo dermatologista americano Thomas B. Fitzpatrick, essa escala categoriza a pele em seis fotótipos com base na capacidade de bronzeamento, sensibilidade e propensão a vermelhidão quando exposta ao sol (SBD, s.d.).

Tabela 1 – Fototipos na Escala Fitzpatrick.

| | | | | |
|-----|----------------------|------------------------|--------------------------|-----------------------------|
| I | Pele branca | Sempre queima | Nunca bronzeia | Muito sensível ao sol |
| II | Pele branca | Sempre queima | Bronzeia muito pouco | Sensível ao sol |
| III | Pele morena clara | Queima (moderadamente) | Bronzeia (moderadamente) | Sensibilidade normal ao sol |
| IV | Pele morena moderada | Queima (pouco) | Sempre bronzeia | Sensibilidade normal ao Sol |
| V | Pele morena escura | Queima (raramente) | Sempre bronzeia | Pouco sensível ao sol |
| VI | Pele negra | Queima (raramente) | Totalmente pigmentada | Minimamente sensível ao sol |

Fonte: Adaptado de (SBD, s.d.).

A melanina presente em peles mais escuras absorve e dispersa a energia ultravioleta, proporcionando uma maior tolerância à exposição solar. No entanto, isso não implica que outros danos relacionados à exposição ultravioleta, como aumento do risco de câncer de pele e envelhecimento prematuro, não possam ocorrer em pessoas com fotótipos mais elevados (SBD, s.d.).

1.2 Justificativas e Relevância

A eficácia das redes neurais artificiais na análise de lesões cutâneas é amplamente reconhecida, mas a falta de diversidade nos dados de treinamento pode comprometer sua aplicação prática. No Brasil, com sua rica diversidade étnica, a ausência de dados representativos pode resultar em diagnósticos imprecisos para pessoas com pele negra, levando a detecções tardias e tratamentos inadequados. Compreender e superar essas limitações é essencial para avançar na pesquisa e desenvolvimento de modelos de diagnóstico mais inclusivos, que possam proporcionar precisão e justiça na avaliação das lesões cutâneas para todos os grupos populacionais.

1.3 Materiais e Métodos

- Bases de Dados

Para a construção da base de dados deste trabalho, foram utilizadas imagens dos conjuntos:

1. PAD-UFES-20,
2. International Skin Imaging Collaboration (ISIC Archive) e
3. Diverse Dermatology Images (DDI).

- Método

Para uma melhor compreensão comparativa, este estudo conduziu três categorias de experimentos para avaliar o desempenho em diferentes fototipos. Na primeira

categoria, os modelos foram testados com imagens na escala Fitzpatrick 1-2. Na segunda categoria, as avaliações foram feitas com imagens da escala Fitzpatrick 3-4. Por fim, na terceira categoria, os testes foram realizados com imagens na escala Fitzpatrick 5-6.

- Modelos de Algoritmos

Para a avaliação do desempenho, foram utilizados dois modelos de algoritmos avançados de visão computacional:

1. EfficientNet_V2
2. Vision Transformer

1.4 Objetivos

O objetivo deste trabalho foi avaliar a eficácia de algoritmos de aprendizado de máquina na classificação de lesões cutâneas em diferentes fototipos de pele. A análise explora como esses algoritmos, treinados com imagens de pele branca, se comportam ao serem aplicados a lesões em fototipos variados, utilizando a escala Fitzpatrick como referência.

Especificamente, busca-se analisar a capacidade dos modelos em identificar e distinguir lesões cutâneas malignas e benignas através de diferentes categorias de pele. Assim sendo, o enfoque não foi minucioso na categorização dos tipos de câncer de pele, mas sim no reconhecimento de padrões normais e anormais em imagens de pele diversa.

1.5 Organização e Estrutura

Capítulo 2 - Revisão de Literatura: apresenta-se uma revisão da literatura sobre o câncer de pele e o comportamento das redes neurais artificiais.

Capítulo 2.6 - Trabalhos Relacionados: apresenta-se uma revisão de trabalhos relacionados à avaliação de modelos de redes neurais aplicados à análise de lesões de pele em diferentes fototipos.

Capítulo 3 - Metodologia: descreve as bases de dados empregada neste estudo e o ambiente de implementação.

Capítulo 4 - Avaliação de Desempenho: apresenta-se os resultados de desempenho dos modelos, incluindo métricas de desempenho como acurácia, precisão, recall, F1-score e coeficiente de correlação de Matthews, além de uma análise comparativa entre os diferentes modelos utilizados.

Capítulo 5 - Conclusão: apresenta-se as principais conclusões do estudo a partir dos resultados obtidos, direções para trabalhos futuros e reflexão sobre a contribuição do trabalho para a área de diagnóstico de câncer de pele.

2 Revisão de Literatura

2.1 O Câncer de Pele

O termo câncer refere-se a uma ampla variedade de doenças malignas, caracterizadas pelo crescimento anômalo e descontrolado de células, que se multiplicam rapidamente e exibem um comportamento agressivo, resultando na formação de neoplasias (INCA, 2022c).

As neoplasias, também conhecidas como tumores, surgem quando um grupo de células ultrapassa os mecanismos normais de controle do crescimento, passando a se replicar de maneira autônoma. Esse crescimento desregulado faz com que as células tumorais se comportem de maneira diferente das células normais, alterando as características estruturais e funcionais do tecido ou órgão afetado (ARAUJO et al., 2014). Veja na Figura 1 os estágios de formação do câncer.

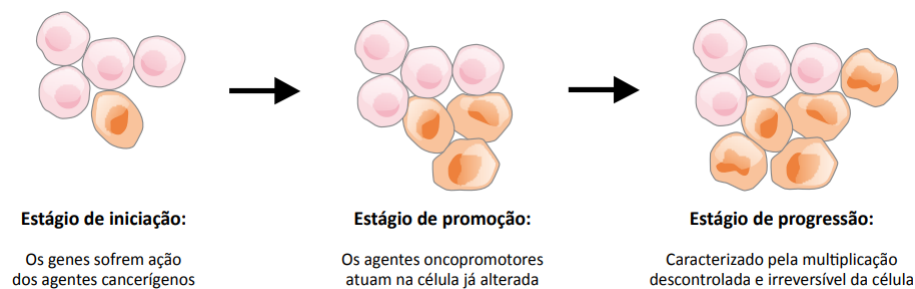


Figura 1 – Estágios de formação do câncer. Fonte: (INCA, 2011).

Os tumores podem ser benignos ou malignos, uma das características que os diferenciam é que os tumores malignos exibem um elevado grau de autonomia, como demonstra a Figura 2, podendo invadir tecidos próximos e desencadear metástases (ou seja, a disseminação do câncer para outros órgãos, veja a Figura 3). Eles têm a capacidade de resistir ao tratamento, podendo resultar na morte do hospedeiro. O câncer é uma neoplasia maligna (INCA, 2011).

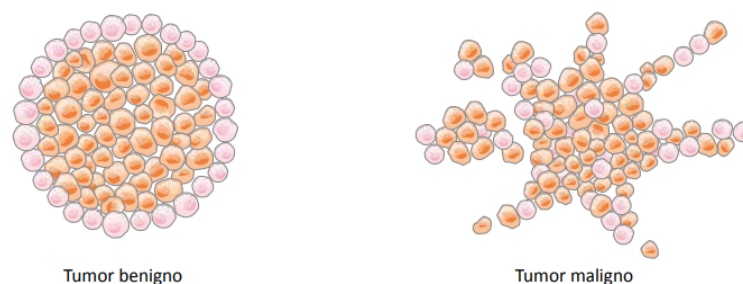


Figura 2 – Diferenças entre tipos de tumores. Fonte: (INCA, 2011).

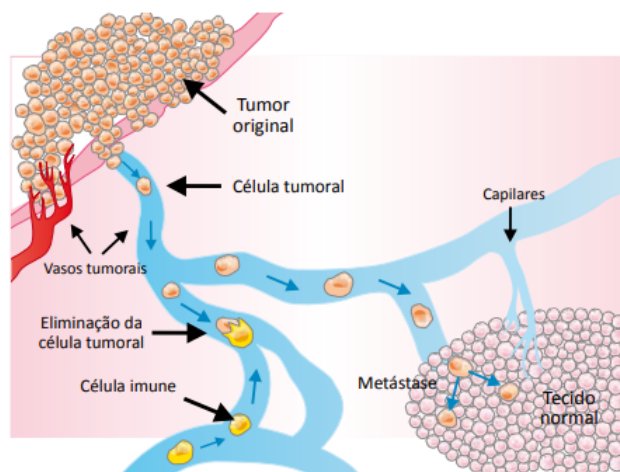


Figura 3 – Metástase. Fonte: (INCA, 2011).

No Brasil, o câncer de pele é o tipo mais comum, abrangendo cerca de 30% de todos os tumores malignos registrados no país (INCA, 2022a).

Ele é classificado em três tipos principais: carcinoma basocelular, carcinoma espinocelular e melanoma. O carcinoma basocelular (CBC) é o mais comum, ele se origina nas células basocelulares, localizadas nas camadas mais profundas da epiderme. Já o carcinoma espinocelular (CEC) se desenvolve nas células escamosas, que formam a superfície da pele. Ambos estão associados à exposição crônica ao sol, podendo se manifestar em áreas expostas ou não e são altamente curáveis quando diagnosticados e tratados precocemente (IBCC, 2023).

O melanoma, atinge as células produtoras de melanina, os melanócitos, que são responsáveis pela pigmentação da pele. Podendo manifestar-se em qualquer área do corpo: mucosas ou pele, incluindo as não expostas ao sol, como o couro cabeludo, as unhas, palmas das mãos e plantas dos pés, sob a forma de manchas, pintas ou sinais. Embora menos frequente, o melanoma é o tipo mais agressivo de câncer de pele e com pior prognóstico, devido à sua alta propensão para metastatizar e à maior taxa de mortalidade associada (INCA, 2022a) (IMB, s.d.).

No entanto, a taxa de sobrevivência tem melhorado significativamente nos últimos anos devido à detecção precoce do tumor, que permitem um tratamento mais eficaz e oportuno (INCA, 2022a). Segundo a Sociedade Brasileira de Dermatologia (SBD), o diagnóstico precoce pode elevar as chances de cura do melanoma para até 90% (IBCC, 2023) (SBD, 2023).

Nas figuras 4 e 5, apresenta-se a incidência absoluta e a taxa de mortalidade em 2022 para os cânceres de pele não melanoma e melanoma em cada continente, seguida pela relação específica dos cinco países com maior incidência da região LAC ("Latin America and the Caribbean" em inglês, ou seja, América Latina e Caribe).

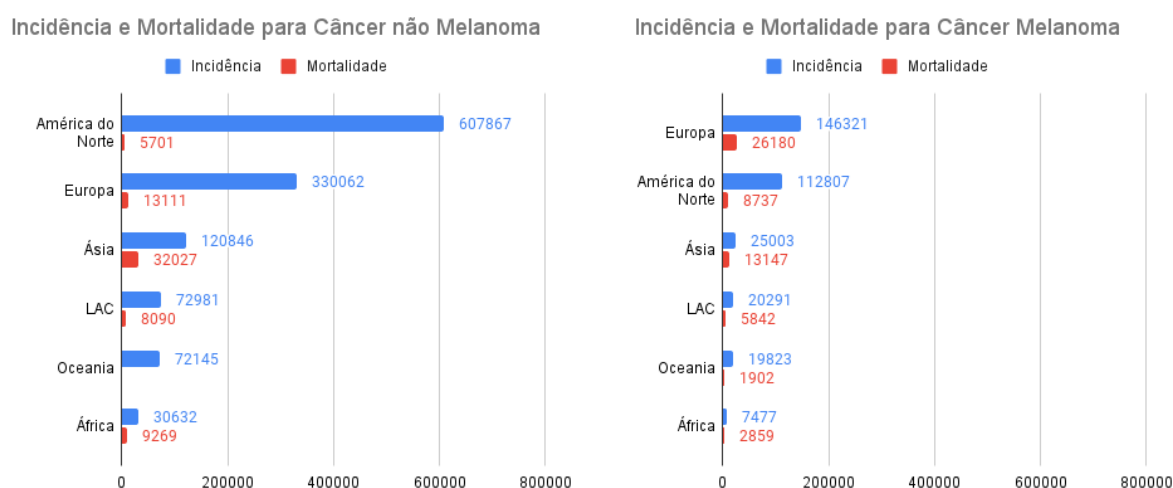


Figura 4 – Incidência e mortalidade por continentes, referentes aos cânceres de pele não melanoma e melanoma em 2022. Fonte: Adaptado de (IARC, 2022).

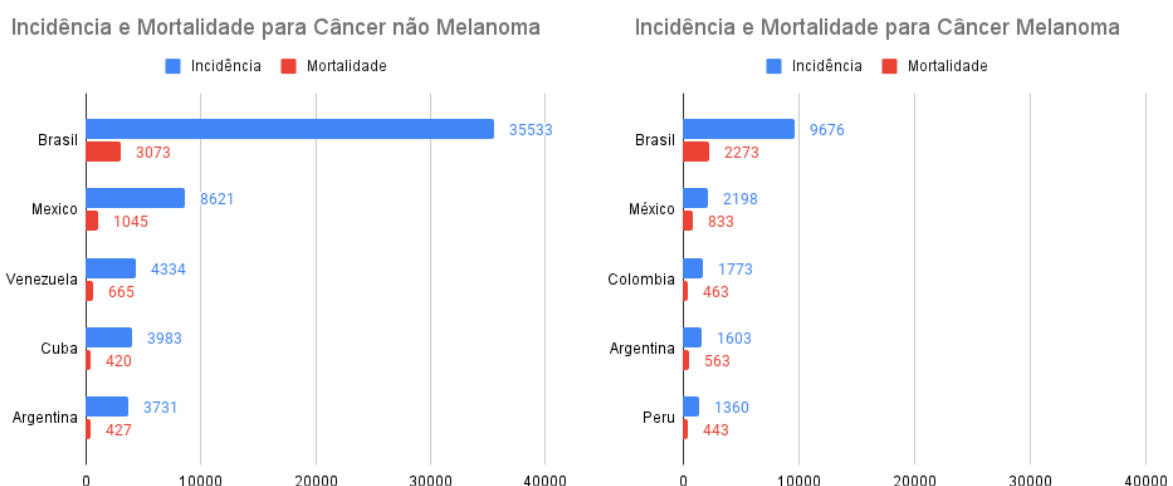


Figura 5 – Incidência e mortalidade por países LAC, referentes aos cânceres de pele não melanoma e melanoma em 2022. Fonte: Adaptado de (IARC, 2022).

No Brasil, tanto o câncer de pele não melanoma quanto o melanoma, são mais incidentes nas Regiões Sul, Sudeste e Centro-oeste, para ambos os sexos. As Regiões Nordeste e Norte, ocupam a segunda posição (INCA, 2023b) (INCA, 2022d) (INCA, 2022b).

2.1.1 Fatores de Risco

Os principais fatores associados ao surgimento do melanoma incluem a exposição solar intermitente, especialmente na forma de queimaduras durante a infância e adolescência, além de um histórico pessoal e familiar da doença. (BRAZIL, 2009). O uso de câmeras de bronzeamento artificial também é associado ao aumento da probabilidade (INCA, 2022a). Indivíduos com pele com fototipos mais claros (tipos de pele I e II de Fitzpatrick) (BRAZIL,

2009) e olhos claros, cabelos ruivos ou loiros, ou que são albinos, estão mais suscetíveis (INCA, 2022a).

Segundo o (IMB, s.d.): "pessoas com pintas congênicas (mais conhecidas como pintas de nascença), número elevado de pintas (50 ou mais pintas no corpo), sistema imunológico enfraquecido e idade mais avançada têm um risco aumentado de desenvolver o melanoma."

Profissionais expostos a agrotóxicos e solventes, como aqueles nos setores de agricultura, metalurgia, limpeza a seco, têxtil, borracha e outros que utilizam solventes, também estão mais propensos a desenvolver a doença (INCA, 2022a).

2.1.2 Campanha Dezembro Laranja

A educação em saúde é essencial para alertar tanto profissionais quanto a população sobre o câncer de pele e a importância do reconhecimento precoce de sinais de malignidade.

Marcado pelo início do verão nos países do hemisfério sul, o mês de dezembro foi escolhido pela SBD para instituir a campanha "Dezembro Laranja", buscando a prevenção e detecção precoce do câncer no maior órgão do corpo humano (INCA, 2023a).

Durante o Dezembro Laranja, orientações intensificam a conscientização sobre o câncer de pele, cuidados necessários e a necessidade de consultar um dermatologista para definir medidas de proteção personalizadas. As orientações devem considerar as diferenças individuais e contextos de vida, já que cada pessoa tem características e exposições distintas (SBD, 2023).

2.2 O Câncer na Pele Negra

O câncer de pele é menos prevalente em indivíduos com pele negra quando se comparado com indivíduos de pele clara, isso se deve à maior quantidade de melanina em suas peles (EID; ALCHORNE, 2011), que desempenha um papel fundamental na proteção contra a radiação ultravioleta (UV). Ela age absorvendo e dispersando a UV, o que diminui o risco de danos ao DNA e, conseqüentemente, reduz a probabilidade de desenvolvimento de câncer de pele (ANDRADE; BEM; TEIXEIRA, 2024). Ainda que a relação entre o câncer na pele negra e a exposição aos raios ultravioletas seja menos evidente; ainda assim, a proteção solar contínua é essencial e deve ser considerada como uma medida de prevenção para todos os tipos de pele (EID; ALCHORNE, 2011).

Os principais fatores de risco incluem cicatrizes crônicas, inflamações persistentes, albinismo, cicatrizes de queimaduras e lesões pigmentadas pré-existentes (EID; ALCHORNE, 2011) e ocorrem mais frequentes em áreas não expostas ao sol, como palmas, solas e mucosas (BARROS; CHAVES; AVILA, 2023) (EID; ALCHORNE, 2011).

Mesmo que menos comum em indivíduos negros, o câncer de pele quando ocorre, tende a ser diagnosticado tardiamente, em parte porque áreas acrais, regiões de alta incidência de neoplasias em negros, são frequentemente negligenciadas pelos dermatologistas durante avaliações físicas devido à menor exposição solar nessas áreas. A falta de representatividade de casos de pele negra no ensino médico e a crença equivocada de que a pigmentação escura oferece proteção completa reduzem a suspeita clínica e contribuem para diagnósticos tardios da doença. Isso resulta em estágios mais avançados e prognósticos piores, com a demora no diagnóstico sendo um fator crucial para a alta mortalidade observada em populações de pele negra (EID; ALCHORNE, 2011).

A dificuldade na detecção precoce ainda é agravada por barreiras socioeconômicas e culturais, que limitam o acesso a cuidados médicos e rastreios preventivos (EID; ALCHORNE, 2011).

2.2.1 Desigualdades Regionais no Sistema de Saúde Brasileiro

De acordo com a pesquisa do Instituto de Estudos para Políticas de Saúde (IEPS), entre 2010 e 2020, a região Norte e Nordeste do Brasil apresentou os piores índices de saúde. A Figura 6 ilustra essas disparidades no sistema de saúde brasileiro, organizadas em quatro blocos: atenção básica, recursos, mortalidade e morbidade, e gastos.

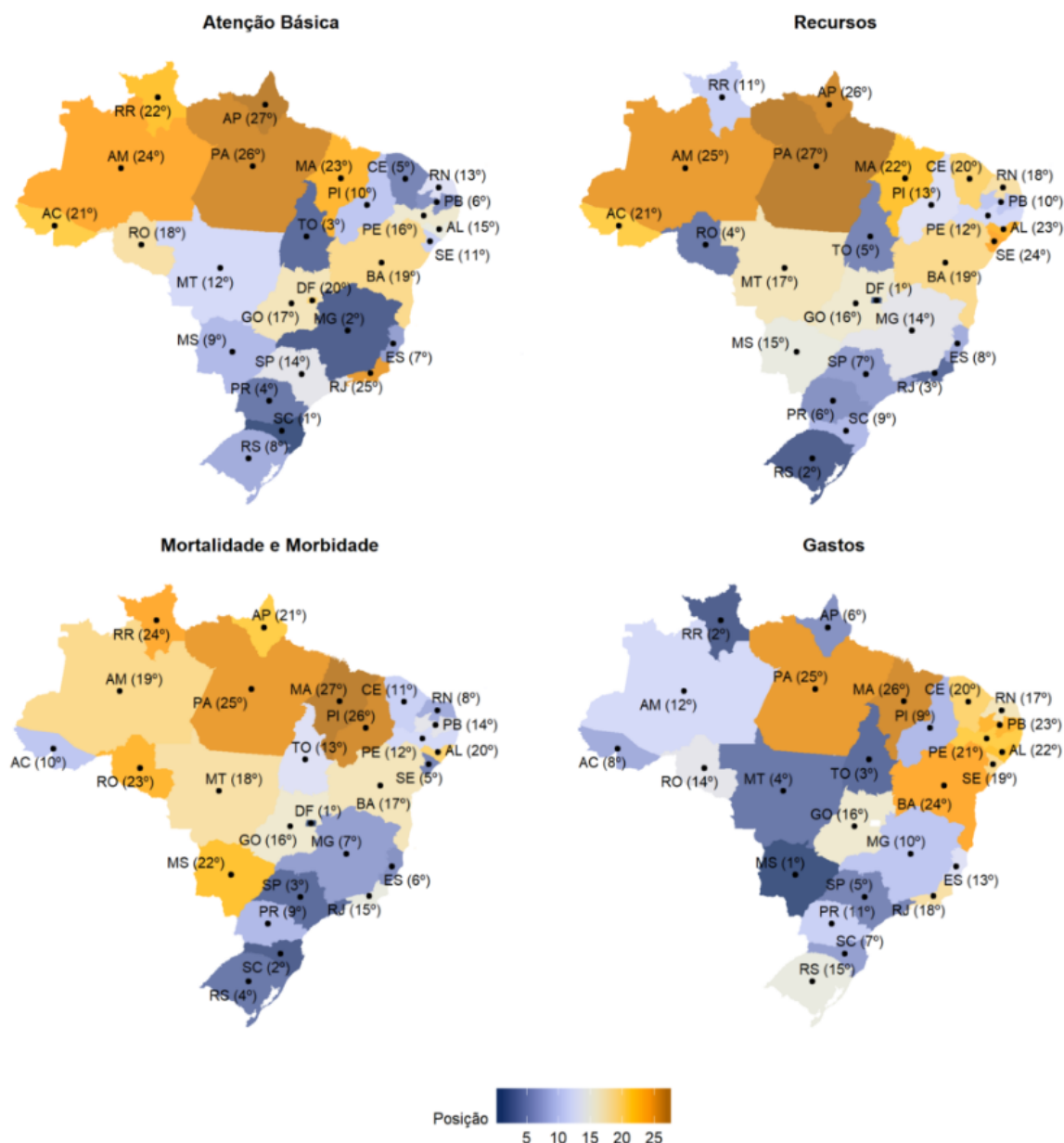


Figura 6 – Indicadores de saúde: atenção básica, recursos, mortalidade e morbidade e gastos. Fonte: (IEPS, 2022).

A região Nordeste do Brasil, que possui o maior percentual de população preta, (IBGEEDUCA, s.d.) enfrenta os piores indicadores de saúde, enquanto a região Sul, com o maior percentual de população branca (IBGEEDUCA, s.d.), destaca-se pelos melhores indicadores de saúde. Essa discrepância pode refletir não apenas diferenças no acesso e na qualidade dos serviços de saúde, mas também nas práticas de diagnóstico e prevenção de doenças, incluindo o câncer de pele. A interação desses fatores ressalta a necessidade de maior atenção do Poder Público nas abordagens de saúde pública, mais equitativas e adaptadas às características demográficas e regionais de cada área.

Embora as variáveis socioeconômicas sejam fundamentais para entender as desi-

gualdades de saúde, não explicam totalmente as disparidades entre grupos raciais. Fatores estruturais e institucionais, como discriminação racial, e fatores culturais e psicológicos, como o estresse crônico associado ao racismo, também desempenham papéis significativos. Além disso, diferenças na qualidade do atendimento, mesmo com acesso garantido, podem perpetuar essas desigualdades. Portanto, entender e abordar as disparidades de saúde requer uma análise mais ampla, que vá além das variáveis socioeconômicas e considere essas outras dimensões que influenciam a saúde de maneira desigual.

2.3 Aprendizado Profundo

O Aprendizado Profundo (do inglês, deep learning) é um subconjunto do Aprendizado de Máquina (do inglês, machine learning) e da Inteligência Artificial (do inglês, artificial intelligence) dedicado ao desenvolvimento de máquinas que podem aprender através da imitação de um padrão observado. Através de Redes Neurais inspiradas no cérebro humano, o aprendizado profundo é usado para lidar com tarefas em diversos campos, como visão computacional (imagem), processamento de linguagem natural (texto) e reconhecimento de fala automática (áudio) (TRASK, 2019). A Figura 7 ilustra o conceito de Aprendizado Profundo como um subconjunto do Aprendizado de Máquina e da Inteligência Artificial, conforme descrito acima.

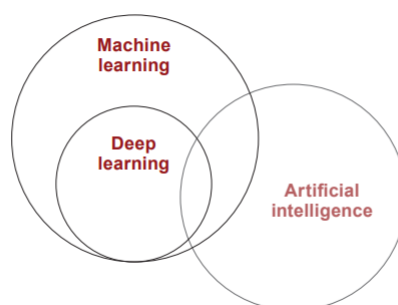


Figura 7 – Inteligência artificial, aprendizado de máquina e aprendizado profundo. Fonte: (TRASK, 2019).

Como a visão de uma máquina interpreta uma imagem?

Quando uma máquina recebe uma imagem como entrada, ela é confrontada com uma sequência de números que correspondem aos pixels. A máquina organizará esses pixels em uma matriz. Cada valor na matriz tem uma escala de 0 a 255, representando a intensidade de cor do pixel. Embora esses números possuam pouco significado para nós, seres humanos, constituem o único conjunto de dados acessível para as máquinas (CUNHA, 2020). A Figura 8 nos mostra a interpretação de uma imagem por uma máquina.

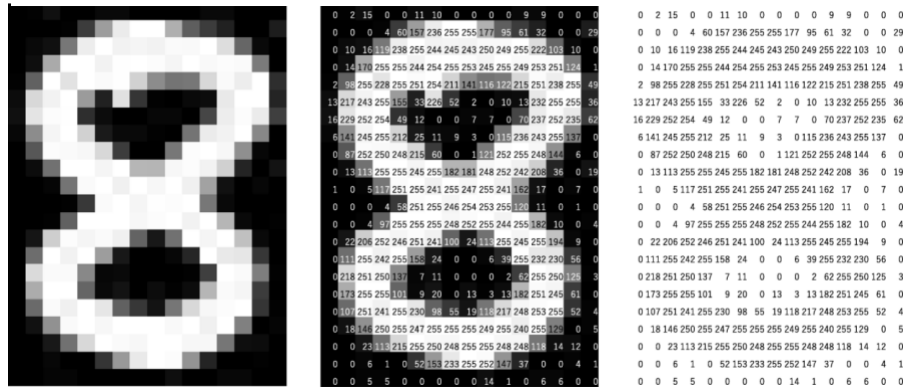


Figura 8 – Imagem interpretada por uma máquina. Fonte: (CUNHA, 2020).

O aprendizado profundo tem como objetivo gerar representações sofisticadas dos dados utilizando diversas camadas de processamento. Um modelo destacado nessa área é a rede neural convolucional (RNC), que foi utilizada por LeCun em 1998 para reconhecer dígitos manuscritos e mostrou-se eficaz em identificar objetos. Neste trabalho, implementamos uma abordagem de aprendizado profundo para desenvolver um sistema automatizado que classifica imagens clínicas de lesões na pele. (HAN et al., 2018).

2.4 Redes Neurais Convolucionais

Normalmente, a estrutura fundamental de uma RNC envolve a presença de camadas intercaladas de convolução e de agrupamento (do inglês, pooling), seguidas por uma ou mais camadas totalmente conectadas (CUNHA, 2020). Definiremos nas próximas seções o comportamento individual de cada uma dessas camadas.

A camada convolucional em uma RNC desempenha o papel crucial de extrair as denominadas características da entrada. O argumento inicial refere-se à imagem de entrada (do inglês, input), enquanto o segundo é denominado filtro (do inglês, kernel), que percorre os dados de entrada em largura, altura e profundidade, também conhecida como dimensão, realizando a operação de convolução nos dados (RODRIGUES, 2018). A saída, por sua vez, é conhecida como mapa de ativação (do inglês, feature map) (CUNHA, 2020).

Basicamente, convolução transforma todos os pixels dentro de sua área receptiva em um valor único. Por exemplo, ao realizar uma convolução em uma imagem, ocorre a redução do tamanho da imagem e a consolidação de todas as informações da área em um único pixel (DATABRICKS, s.d.). Veja na figura Figura 9.

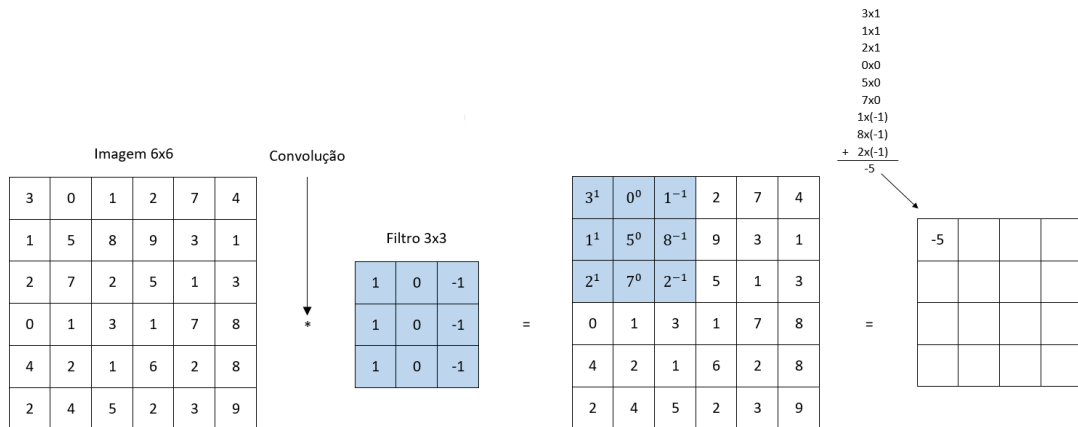


Figura 9 – Convolução entre imagem de entrada 6x6 e filtro 3x3. Fonte: Adaptado de (CUNHA, 2020).

Esse procedimento é repetido até que toda a imagem tenha sido explorada, resultando na formação de um mapa de ativação (Figura 10).

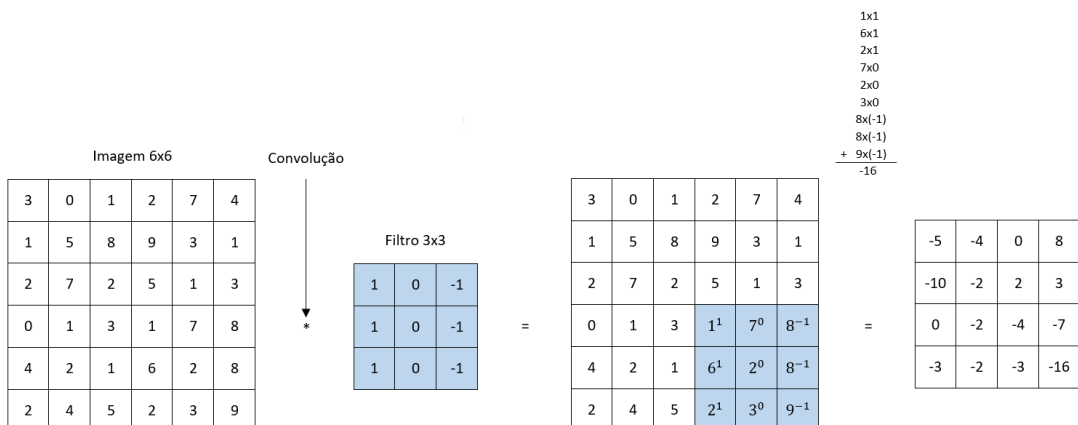


Figura 10 – Mapa de ativação gerado com dimensões 4x4. Fonte: Adaptado de (CUNHA, 2020).

Durante o treinamento da rede, a cada iteração de entrada, os filtros são ajustados de maneira a serem acionados quando a entrada contiver uma característica específica comum aos lotes de entrada, como, por exemplo, arestas, cores, entre outras características (RODRIGUES, 2018).

É importante destacar que os filtros são ajustados na etapa de treinamento da rede, não é uma decisão manual do projetista. A própria rede se encarregará de aprender os filtros que melhor extraem características da imagem através de algumas medidas e atributos fornecidos ao algoritmo (CUNHA, 2020).

Após as etapas de convolução, é comum aplicar uma camada não-linear (também chamada de camada de ativação) imediatamente em seguida. O objetivo dessa camada é introduzir não linearidade a um sistema que, em sua essência, realiza predominantemente

operações lineares na camada de convolução (como multiplicação e soma) (CUNHA, 2020).

Sendo uma das mais utilizadas, a função Rectified Linear Unit (ReLU), tem o intuito de intensificar a não-linearidade de uma entrada. O ativador ReLU emprega uma equação, que devolve o valor máximo entre zero e a própria entrada (SOUSA, 2023):

$$y = \max(0, x), \quad (2.1)$$

A Figura 11 ilustra a demonstração em gráfico da função ReLU.

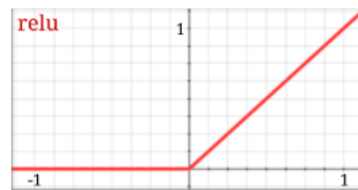


Figura 11 – Gráfico demonstrando a função não linear ReLU. Fonte: (CUNHA, 2020).

Em outras expressões, essa camada substitui todos os valores de ativação negativos resultantes da convolução por zero. Essa ação amplia as características não-lineares do modelo sem impactar as ativações da camada de convolução (CUNHA, 2020). Veja na Figura 12 uma demonstração da saída após a aplicação do ReLU.

| | | | | | | | | |
|----|----|----|----|----------|---|---|---|---|
| 9 | 3 | 5 | -8 | → ReLu → | 9 | 3 | 5 | 0 |
| -6 | 2 | -3 | 1 | | 0 | 2 | 0 | 1 |
| 1 | 3 | 4 | 1 | | 1 | 3 | 4 | 1 |
| 3 | -4 | 5 | 1 | | 3 | 0 | 5 | 1 |

Figura 12 – Camada não linear aplicada em mapa de ativação. Fonte: Adaptado de (CUNHA, 2020).

A Camada de Agrupamento, também referida como Camada de Pooling, tem como função diminuir as amostras dos mapas de características resultantes das operações de convolução, preservando sempre as mais proeminentes. Existem várias técnicas de Agrupamento, no entanto, duas delas são as mais prevalentes: Agrupamento Máximo e Agrupamento Médio (SOUSA, 2023).

O Agrupamento Máximo, sendo o mais utilizado, mapeia uma região específica e extrai seu valor máximo, como pode ser visto no exemplo da Figura 13, por outro lado, o Agrupamento por Média mapeia a mesma região e extrai seu valor médio.

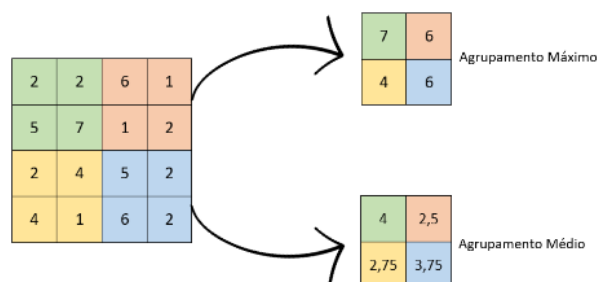


Figura 13 – Camada de agrupamento máximo e médio. Fonte: Autora.

A prática de agrupamento também contribui para o desempenho da rede neural, uma vez que diminui a quantidade de dados direcionados à camada subsequente, aprimorando a regularização da rede e, simultaneamente, reduzindo o uso de memória e processamento (RODRIGUES, 2018).

As camadas completamente conectadas geralmente desempenham o papel de camadas finais, divergindo das camadas convolucionais, pois todos os pesos estão integralmente vinculados à camada anterior (MASSUCATTO, 2018).

Antes de alcançar as camadas totalmente conectadas, é necessário aplicar uma operação de nivelamento (conhecido como flattening), para transformar as matrizes de características provenientes da camada de agrupamento em um vetor coluna, que será utilizado como entrada para a camada totalmente conectada (MASSUCATTO, 2018). Veja a operação flattening na Figura 14.

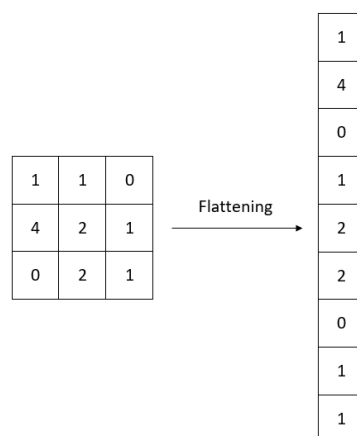


Figura 14 – Aplicação de flattening. Fonte: Adaptado de (MASSUCATTO, 2018).

A função da camada totalmente conectada em uma RNC é fornecer uma decisão sobre a classificação da imagem como um todo. Dessa forma, ela integra-se à estrutura das redes destinadas à classificação (CUNHA, 2020).

A camada totalmente conectada vai conectar todas as características e atributos extraídos da imagem por meio de uma estrutura neuronal convencional com camadas

ocultas, onde todos os neurônios da camada estão interligados (CUNHA, 2020). Veja a Figura 15.

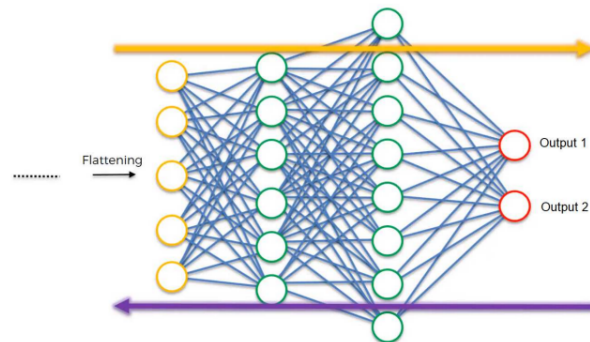


Figura 15 – Camadas totalmente conectadas. Fonte: (MASSUCATTO, 2018).

A Figura 16 ilustra a estrutura básica de uma RNC.

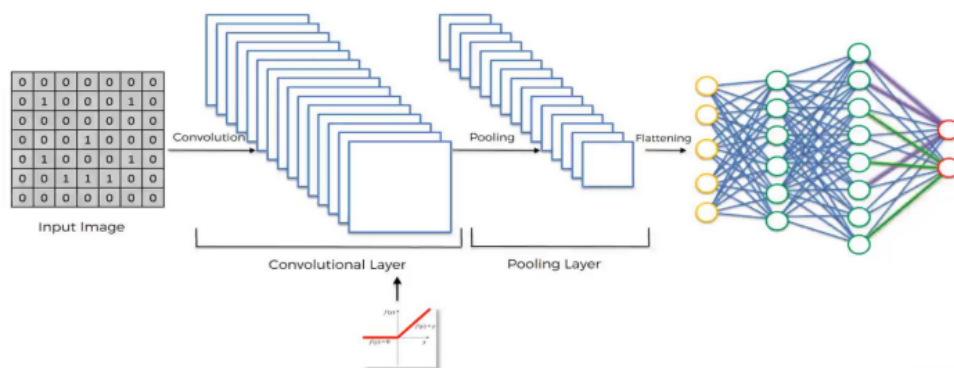


Figura 16 – Estrutura básica de uma rede neural convolucional. Fonte: (MASSUCATTO, 2018).

Por fim, após a RNA extrair e reduzir as características nas camadas de convolução e agrupamento, a camada totalmente conectada é empregada para a classificação de um conjunto de imagens. Essa camada recebe as características extraídas em formato vetorial como entrada, realiza um produto escalar com seus pesos e isso resulta na probabilidade de uma imagem pertencer a uma determinada classe (SOUSA, 2023).

2.4.1 EfficientNetV2

Para este trabalho, utilizou-se o modelo EfficientNetV2.

As versões anteriores do EfficientNet enfrentavam três principais desafios: a lentidão no treinamento quando trabalhava com imagens de grandes dimensões, a demora associada às convoluções profundas nas camadas iniciais e a eficiência limitada na ampliação uniforme das etapas de treinamento (TAN; LE, 2021).

O EfficientNetV2 foi desenvolvido para superar esses problemas de forma significativa. Ele oferece uma melhoria notável na velocidade de treinamento e na eficiência dos parâmetros, conseguindo acelerar o treinamento até 4 vezes mais e reduzir o tamanho dos parâmetros em até 6,8 vezes em comparação com os modelos anteriores (TAN; LE, 2021).

2.5 Transformers Adaptados para Imagens

A arquitetura Transformer foi inicialmente desenvolvida para processamento de linguagem natural, quando aplicada a tarefas de reconhecimento e análise de imagens, os Transformers são adaptados para lidar com dados visuais. Essas adaptações envolvem modificações na arquitetura original, ajustando-a para o reconhecimento e análise de imagens (ALEXEY, 2020).

Ao contrário das RNCs, que são especializadas em identificar características espaciais locais em dados estruturados, o Transformer se baseia exclusivamente em mecanismos de atenção. Essa abordagem permite que o modelo analise toda a sequência de dados simultaneamente, dispensando a necessidade de operações convolucionais para processamento (LIMA et al., 2023).

O Transformer processa a sequência de vetores utilizando uma série de camadas que alternam entre mecanismos de atenção multi-cabeça e mecanismos de auto-atenção (LIMA et al., 2023).

O mecanismo de **auto-atenção** avalia como cada vetor de entrada deve influenciar os outros na sequência, processando cada vetor e gerando três vetores distintos: o vetor de consulta (Q), o vetor de chave (K) e o vetor de valor (V) (LIMA et al., 2023).

A importância relativa de cada vetor em relação aos demais é calculada ajustando a representação de cada vetor com base nessas avaliações. Esse processo envolve determinar a contribuição de cada vetor para a construção da representação final, levando em conta a interação entre todos os vetores na sequência (LIMA et al., 2023). As pontuações entre os diferentes vetores de entrada são calculadas da seguinte forma:

$$S = Q \times K^T \quad (2.2)$$

Em seguida, as pontuações são normalizadas para estabilizar o gradiente e otimizar o processo de treinamento:

$$S_n = \frac{S}{\sqrt{d_k}} \quad (2.3)$$

Depois, são calculadas as probabilidades para as pontuações:

$$P = \text{softmax}(S_n) \quad (2.4)$$

Por fim, cada vetor de valor é ponderado pela soma das probabilidades, resultando na matriz de atenção.

$$Z = V \times P \quad (2.5)$$

Que pode ser resumida em:

$$\text{Atencao}(Q, K, V) = \text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) \times V \quad (2.6)$$

O mecanismo de **atenção multi-cabeça** surge como uma solução para a limitação dos mecanismos de atenção únicos, que podem ter dificuldades em se concentrar simultaneamente em várias posições relevantes (ALEXEY, 2020).

Com a atenção multi-cabeça, é possível explorar múltiplas representações de diferentes aspectos das entradas ao mesmo tempo. Essa abordagem utiliza múltiplos conjuntos de matrizes de consulta, chave e valor, permitindo que cada “cabeça” de atenção aprenda e processe informações em diferentes subespaços de representação. Isso permite que o modelo olhe para a entrada de várias maneiras diferentes ao mesmo tempo, capturando mais nuances e detalhes. Como resultado, o modelo pode entender melhor as relações entre diferentes partes da entrada, tornando-o mais poderoso e preciso na análise e na geração de respostas (LIMA et al., 2023).

2.5.1 Vision Transformer (ViT)

Para este trabalho, utilizou-se o modelo Vision Transformer (ViT).

O ViT adota uma abordagem inovadora ao processar imagens, dividindo-as em pequenos blocos ou “patches” de tamanho uniforme. Cada patch é transformado em um vetor de dimensão fixa através de uma projeção linear, resultando em uma sequência de vetores que representa a imagem, onde cada vetor corresponde a um patch específico (ALEXEY, 2020).

Para compensar a ausência de convoluções, o Transformer utiliza embeddings de posição que são adicionados aos vetores dos patches. Esses embeddings preservam a informação sobre a localização relativa dos patches na imagem original. Além disso, um vetor especial conhecido como “token de classificação” é incluído no início da sequência de vetores. Este token é crucial para gerar a representação final da imagem e para a realização da classificação (ALEXEY, 2020).

Em comparação com as redes neurais convolucionais tradicionais, os Vision Transformers conseguem alcançar resultados impressionantes com uma menor demanda de recursos computacionais durante o pré-treinamento (ALEXEY, 2020).

Na Figura 17 tem-se uma visão geral do modelo ViT.

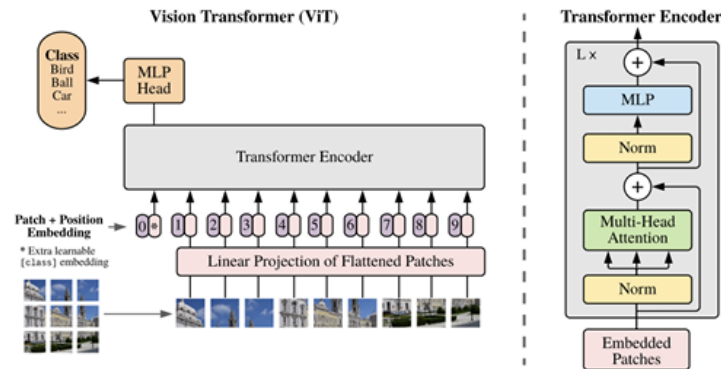


Figura 17 – Visão Geral do Modelo ViT. Fonte: Adaptado de (ALEXEY, 2020).

2.6 Trabalhos Relacionados

Os vieses raciais surgem em diversos sistemas de aprendizado de máquina, afetando negativamente a equidade e a justiça em suas aplicações. Além dos modelos de prognóstico médico, que prejudicam a precisão dos diagnósticos e tratamentos para populações marginalizadas, esses vieses também se manifestam em áreas como o reconhecimento facial, onde algoritmos frequentemente falham em identificar corretamente indivíduos de grupos racialmente sub-representados. A pesquisa (RUBACK; AVILA; CANTERO, 2021) destaca vieses de representação em sistemas de reconhecimento facial, relacionados à coleta de dados, resultando em conjuntos de treinamento não representativos. Os resultados indicaram que homens e indivíduos brancos costumam obter um desempenho melhor nos modelos, enquanto a classificação de mulheres negras demonstrou uma eficácia significativamente menor.

Além disso, a pesquisa identificou vieses de avaliação, que ocorrem quando os dados de teste não refletem adequadamente a diversidade da população. Também foram observados vieses de interpretação humana, referindo-se à maneira como as saídas dos modelos são analisadas, o que pode levar a decisões injustas. A situação é ainda mais grave quando se considera que 90,5% das prisões realizadas por reconhecimento facial no Brasil envolvem pessoas negras.

Para aprofundar a análise do viés racial, o estudo (FABI; XU; SA, 2022) utilizou um modelo de aprendizado de máquina focado na identificação de ativações de diferentes unidades de ação associadas a expressões faciais de dor. A pesquisa avaliou se o modelo

apresentava maior sensibilidade para rostos de indivíduos europeus em comparação com rostos africanos, especialmente quando treinado com conjuntos de dados que predominavam rostos brancos. Os achados revelaram variações na detecção de certas ativações musculares faciais, que podem ser atribuídas à cor da pele ou a características faciais específicas de cada raça. Apesar da predominância de dados favoráveis a rostos de aparência europeia, algumas ativações foram mais facilmente reconhecidas em rostos africanos. Dessa forma, nem a capacidade perceptual nem a familiaridade com rostos da própria raça explicam totalmente o preconceito racial observado na detecção da dor.

A pesquisa (MIRIKHARAJI et al., 2023), por sua vez, analisou 177 artigos sobre a segmentação de lesões cutâneas com aprendizado profundo. Os autores examinaram diferentes dimensões, como dados de entrada, arquitetura de modelos e aspectos de avaliação, e ressaltaram, dentre outros, a importância de investigar a adaptabilidade de modelos entre diferentes tonalidades de pele, dado que a maioria dos conjuntos de dados é dominada por imagens de pele clara.

Neste contexto, os estudos (BARROS; AVILA, 2022) e (BARROS; CHAVES; AVILA, 2023) focaram na análise do desempenho de modelos de RNAs na detecção de lesões cutâneas em diferentes fototipos, com ênfase em regiões acrais, que apresentam alta incidência de câncer em pessoas negras. Embora as autoras tenham excluído lesões malignas que não fossem melanomas, essa abordagem comprometeu a representatividade dos tumores malignos nas diferentes tonalidades de pele, resultando em uma base de dados desbalanceada.

Uma análise interessante apresentada por Barros e Avila (2022) destacou a dificuldade no diagnóstico das lesões, classificadas por médicos como de baixo, médio e alto grau de complexidade. Ao segmentar a base de dados conforme o nível de dificuldade, observou-se que os modelos reproduziram o comportamento dos médicos na gestão da complexidade do diagnóstico, com precisão maior para lesões de nível fácil e queda na precisão conforme a complexidade aumentava. Ambos os estudos concluíram que os modelos de aprendizado de máquina apresentaram desempenho insatisfatório na classificação de melanomas em pele negra, comparados a conjuntos de dados que continham apenas imagens de pele branca.

Vale ressaltar que antes mesmo de considerar abordagens de aprendizado profundo, técnicas simples já revelam a presença de vieses raciais, como a fotopletismografia (PPG), utilizada para medir variações no volume sanguíneo e estimar a saturação de oxigênio no sangue (SpO₂). A pesquisa (BERMOND et al., 2023) aponta que os oxímetros de pulso que empregam PPG tendem a ser menos precisos em indivíduos de pele mais escura, frequentemente superestimando os níveis de oxigênio. Essa ineficácia está ligada à presença de melanina na pele, o que dificulta a medição precisa da saturação de oxigênio.

3 Metodologia

3.1 Bases de Dados

Selecionaram-se conjuntos de dados que incluíssem informações sobre a tonalidade da pele e apresentassem lesões cutâneas em tons mais escuros, pois isso é crucial para avaliar a eficácia dos modelos neste contexto específico. Para isso, três conjuntos de dados foram utilizados: PAD-UFES-20, ISIC e DDI.

PAD-UFES-20: A criação desse banco de dados surgiu da necessidade de aprimorar diagnósticos utilizando inteligência artificial. O projeto focou na combinação de imagens clínicas coletadas com smartphones e informações clínicas dos pacientes, reconhecendo que esses dados poderiam enriquecer os modelos de aprendizado profundo e aumentar a precisão na identificação de lesões cutâneas (PACHECO; KROHLING, 2020).

Para desenvolvê-lo, uma equipe de pesquisadores colaborou com o Programa de Assistência Dermatológica (PAD) da Universidade Federal do Espírito Santo, coletando amostras de imagens clínicas, nas quais dermatologistas experientes forneceram diagnósticos. Informações relevantes sobre cada paciente foram organizadas cuidadosamente para facilitar análises posteriores. Durante a criação do banco de dados, técnicas de pré-processamento, como aprendizado por transferência e aumento de dados, foram consideradas para lidar com desafios típicos de conjuntos de dados médicos, incluindo desproporcionalidade nas classes e qualidade das imagens. Essas etapas foram fundamentais para garantir que o banco de dados representasse uma ampla variedade de lesões de pele e estivesse pronto para uso em modelos de aprendizado profundo (PACHECO; KROHLING, 2020).

Esse conjunto de dados inclui 2.298 imagens clínicas e, embora algumas imagens não contenham metadados sobre a escala Fitzpatrick, essas imagens foram mantidas no banco, uma vez que esse conjunto foi usado exclusivamente para treinamento dos modelos.

Na tabela 2, tem-se a relação de imagens pela escala Fitzpatrick para as classificações benignas e malignas.

Tabela 2 – Quantidade de imagens PAD-UFES-20

| Escala Fitzpatrick | Benignas | Malignas | Total |
|--------------------|----------|----------|-------|
| 1 | 33 | 120 | 153 |
| 2 | 213 | 663 | 876 |
| 3 | 120 | 272 | 392 |
| 4 | 33 | 29 | 62 |
| 5 | 5 | 5 | 10 |
| 6 | 1 | 0 | 1 |
| Não identificado | 804 | 0 | 804 |
| Total | 1209 | 1089 | 2298 |

Fonte: Elaborada pela autora.

International Skin Imaging Collaboration (ISIC): A ISIC tem como objetivo contribuir para a diminuição da mortalidade associada ao melanoma. Para isso, a iniciativa busca aprimorar a precisão e a eficácia na detecção precoce do melanoma, desenvolvendo padrões de imagem digital e promovendo uma colaboração estreita entre as áreas de dermatologia e ciência da computação. O uso de inteligência artificial (IA) é central nesse processo, visando melhorar o diagnóstico. Embora o foco inicial seja o melanoma, a ISIC também está comprometida com a ampliação dos conhecimentos na área de imagens de pele e IA, o que inclui câncer de pele não melanoma e outras condições dermatológicas (SOBRE... , s.d.).

Desde 2016, o projeto ISIC promove um desafio anual que envolve desenvolvedores de algoritmos de inteligência artificial (IA) no diagnóstico de melanoma. O desafio é dividido em duas etapas: a primeira utiliza um "conjunto de treinamento" de imagens ISIC para treinar os algoritmos, enquanto a segunda avalia a precisão diagnóstica com um "conjunto de teste" de imagens ISIC, permitindo uma comparação com os resultados de dermatologistas (ISIC... , 2023). Essa iniciativa do ISIC incentiva o desenvolvimento de algoritmos de IA, contribuindo para avanços significativos no diagnóstico precoce e preciso do melanoma.

Atualmente, o arquivo ISIC abriga mais de 482.781 imagens públicas, a maioria acompanhada de metadados clínicos revisados por especialistas. Para esse trabalho, as imagens utilizadas foram filtradas com base em atributos clínicos, selecionando tanto lesões malignas quanto benignas, além de incluir apenas imagens clínicas para garantir a padronização com outros conjuntos de dados utilizados no estudo. Imagens sem informações sobre a escala Fitzpatrick também foram mantidas à base de treinamento. Esse processo resultou em um conjunto final de 428 imagens, como pode ser visto na tabela 3.

Tabela 3 – Quantidade de imagens ISIC

| Escala Fitzpatrick | Benignas | Malignas | Total |
|--------------------|----------|----------|-------|
| 1 | 5 | 15 | 20 |
| 2 | 106 | 185 | 291 |
| 3 | 15 | 13 | 28 |
| Não identificado | 22 | 67 | 89 |
| Total | 148 | 280 | 428 |

Fonte: Elaborada pela autora.

Diverse Dermatology Images (DDI): A criação desse banco de dados foi motivada pela necessidade de melhorar o diagnóstico e a compreensão das condições dermatológicas em diferentes grupos étnicos, especialmente em pacientes autodescritos como negros, hispânicos ou latinos. Apesar do aumento no uso de tecnologias de imagem e inteligência artificial na dermatologia, observou-se que havia uma falta de representação adequada de peles de tons mais escuros nos bancos de dados existentes, o que poderia impactar a precisão dos diagnósticos. Para abordar essa questão, a coleta de imagens clínicas foi padronizada, com todas as lesões biopsiadas sendo fotografadas com câmeras de smartphones a uma distância específica. Essa abordagem permitiu a formação de um repositório consistente, que poderia ser utilizado para treinar algoritmos de aprendizado de máquina, visando melhorar o desempenho em diagnósticos para diferentes fototipos de pele. A inclusão de dados demográficos, como idade, sexo e tipo de Fitzpatrick, foi essencial para garantir que as análises considerassem as variações que podem influenciar os resultados clínicos (DANESHJOU et al., 2022).

Devido à diversidade de tons de pele presente no DDI, este banco de dados se destaca como o principal recurso utilizado neste trabalho, sendo o conjunto de teste composto exclusivamente por imagens desse banco. Ele é formado por 656 imagens clínicas que foram classificadas por dermatologistas com base na escala Fitzpatrick.

Para garantir que o conjunto de teste refletisse proporcionalmente a quantidade de imagens utilizadas no treinamento, foi necessário reduzir o conjunto a 357 imagens, conforme demonstrado na tabela 4.

Tabela 4 – Quantidade de imagens DDI

| Escala Fitzpatrick | Benignas | Malignas | Total |
|--------------------|----------|----------|-------|
| 1-2 | 53 | 49 | 102 |
| 3-4 | 81 | 74 | 155 |
| 5-6 | 52 | 48 | 100 |
| Total | 186 | 171 | 357 |

Fonte: Elaborada pela autora.

3.2 Ambiente de Implementação

Os experimentos foram realizados em uma máquina com 8 GB de RAM, processador intel core i7 e placa de vídeo NVIDIA Quadro FX 880M.

A implementação dos algoritmos foi realizada na plataforma Kaggle, utilizando a configuração de sessão com acelerador GPU P100. O ambiente de desenvolvimento foi definido com a linguagem Python, garantindo um suporte robusto para as bibliotecas e ferramentas necessárias para a análise.

O código implementado foi baseado em modelos já disponíveis na plataforma Kaggle ¹, o que possibilitou a reprodução dos experimentos e a validação dos resultados obtidos. Essa abordagem disponibilizada utiliza o *transfer learning*, que torna possível a adaptação dos modelos para as necessidades específicas deste estudo. Além disso, foram realizadas mudanças nos parâmetros, como o número de épocas (EPOCHS=20) e o tamanho do *batch* ajustado para o valor 16.

¹ <https://www.kaggle.com/code/matthewjansen/transfer-learning-skin-cancer-classification>

4 Avaliação de Desempenho

A análise do comportamento dos modelos de redes neurais foi feita a partir de diferentes experimentos, divididos pelos fototipos classificados na escala Fitzpatrick. Para avaliar a performance, considerou-se as métricas **accuracy**, que mede a proporção de previsões corretas em relação ao total de previsões feitas; **precision**, que mede a proporção de verdadeiros positivos em relação ao total de previsões positivas feitas; **recall**, que mede a proporção de verdadeiros positivos em relação ao total de casos positivos reais; **f1_score**, que mede a média harmônica entre precision e recall, balanceando as duas métricas e **coeficiente de correlação de Matthews**, que retorna a medida geral de desempenho que considera todos os quatro quadrantes da matriz de confusão (verdadeiros positivos, falsos positivos, verdadeiros negativos, falsos negativos).

Para definir o melhor modelo, o menor valor para o **trade-off** (T) foi considerado, para esse cálculo a **taxa de inferência** (i), a qual refere-se ao tempo necessário para um modelo processar e classificar uma nova amostra, é multiplicada pelo coeficiente de correlação de Matthews (c) daquele modelo.

$$T = i \times c \quad (4.1)$$

A base de treinamento dos modelos foi composta pelos bancos de dados ISIC e PAD-UFES-20, esse conjunto de treinamento será referido como base principal. Para as bases de teste foram usadas apenas as imagens do banco DDI, separadas conforme a escala Fitzpatrick, essas bases sofreram pequenas alterações de um experimento para outro e serão descritas em seguida.

4.1 Experimento 1

Comportamento dos modelos em imagens de lesões cutâneas nas escalas Fitzpatrick 1 e 2.

- **Experimento 1.1:** O modelo foi treinado utilizando a base principal e testado com imagens nas escalas Fitzpatrick 1 e 2 do conjunto DDI.

Tabela 5 – Resultado de métricas do experimento 1.1

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|----------|----------|-------------------|
| EfficientNet_v2 | 0.578431 | 0.626728 | 0.578431 | 0.547177 | 0.209533 |
| ViT_b16 | 0.588235 | 0.606748 | 0.588235 | 0.578070 | 0.198964 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.005809s/amostra +- 0.004721s/amostra para o modelo EfficientNet_V2 e 0.010162s/amostra +- 9.5e-05s/amostra para o modelo ViT_b16. Sendo definido o EfficientNet_V2 como o melhor modelo, com Trade-off igual a 0.7905, enquanto o ViT_b16 apresentou um Trade-off de 0.8011.

- **Experimento 1.2:** O modelo foi treinado com a base principal acrescida de 20 imagens previamente usadas para teste no Experimento 1.1. Dessas 20 imagens, 10 eram de lesões malignas e 10 de lesões benignas. Essas imagens foram retiradas do conjunto de teste e integradas somente ao de treinamento. Para teste foram as imagens das escalas Fitzpatrick 1 e 2 do conjunto DDI.

Tabela 6 – Resultado de métricas do experimento 1.2

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|----------|----------|-------------------|
| EfficientNet_v2 | 0.548780 | 0.561585 | 0.548780 | 0.540833 | 0.114996 |
| ViT_b16 | 0.573171 | 0.592528 | 0.573171 | 0.562505 | 0.170815 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.005877s/amostra +- 0.003136s/amostra para o modelo EfficientNet_V2 e 0.010174s/amostra +- 8e-05s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.8292, enquanto o EfficientNet_V2 apresentou um Trade-off de 0.8850.

4.2 Experimento 2

Comportamento dos modelos em imagens de lesões cutâneas nas escalas Fitzpatrick 3 e 4.

- **Experimento 2.1:** O modelo foi treinado com a base principal e testado com imagens das escalas Fitzpatrick 3 e 4 do conjunto DDI:

Tabela 7 – Resultado de métricas do experimento 2.1

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|----------|----------|-------------------|
| EfficientNet_v2 | 0.483871 | 0.496178 | 0.483871 | 0.438412 | -0.007735 |
| ViT_b16 | 0.522581 | 0.549433 | 0.522581 | 0.488275 | 0.080733 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.006475s/amostra +- 0.001618s/amostra para o modelo EfficientNet_V2 e 0.009768s/amostra +- 7.5e-05s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.9193, enquanto o EfficientNet_V2 apresentou um Trade-off de 1.0078.

- **Experimento 2.2:** O modelo foi treinado com a base principal, ampliada com 20 imagens de do conjunto de teste do Experimento 2.1, sendo 10 de lesões malignas e 10 de lesões benignas. Nesse experimento, essas 20 imagens foram excluídas do conjunto de teste e usadas somente para treinamento. O modelo foi treinado com imagens da escala 3 e 4 do conjunto DDI.

Tabela 8 – Resultado de métricas do experimento 2.2

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|---------|----------|-------------------|
| EfficientNet_v2 | 0.60000 | 0.634250 | 0.60000 | 0.584403 | 0.239916 |
| ViT_b16 | 0.62963 | 0.631223 | 0.62963 | 0.629874 | 0.260000 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.006792s/amostra +- 0.001963s/amostra para o modelo EfficientNet_V2 e 0.009857s/amostra +- 7.4e-05s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.7401, enquanto o EfficientNet_V2 apresentou um Trade-off de 0.7601.

- **Experimento 2.3:** O modelo foi treinado com a base principal, adicionada de 20 imagens do conjunto de teste do Experimento 2.1 (10 malignas e 10 benignas), além das imagens da escala Fitzpatrick 1 e 2 usadas em teste no Experimento 1.1. Assim, a base de treinamento incluiu também imagens das escalas Fitzpatrick 1 e 2 do conjunto DDI. O conjunto de teste foram as imagens na escala 3 e 4.

Tabela 9 – Resultado de métricas do experimento 2.3

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|----------|----------|-------------------|
| EfficientNet_v2 | 0.592593 | 0.591380 | 0.592593 | 0.589328 | 0.178914 |
| ViT_b16 | 0.637037 | 0.642016 | 0.637037 | 0.636678 | 0.279604 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.007774s/amostra +- 0.002363s/amostra para o modelo EfficientNet_V2 e 0.010213s/amostra +- 0.000217s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.7205, enquanto o EfficientNet_V2 apresentou um Trade-off de 0.8211.

4.3 Experimento 3

Comportamento dos modelos em imagens de lesões cutâneas nas escalas Fitzpatrick 5 e 6.

- **Experimento 3.1:** O modelo foi treinado com a base principal e testado com imagens das escalas Fitzpatrick 5 e 6 do conjunto DDI.

Tabela 10 – Resultado de métricas do experimento 3.1

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|--------|----------|-------------------|
| EfficientNet_v2 | 0.43 | 0.431257 | 0.43 | 0.423277 | -0.134603 |
| ViT_b16 | 0.45 | 0.452121 | 0.45 | 0.448844 | -0.096561 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.004338s/amostra +- 0.002708s/amostra para o modelo EfficientNet_V2 e 0.009884s/amostra +- 0.000303s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 1.0966, enquanto o EfficientNet_V2 apresentou um Trade-off de 1.1346.

- **Experimento 3.2:** O modelo foi treinado com a base principal, ampliada com 20 imagens usadas em teste no Experimento 3.1, sendo 10 de lesões malignas e 10 de lesões benignas. Essas imagens foram removidas do conjunto de teste e incluídas somente no conjunto de treinamento. As imagens usadas para treinamento foram as de escala 3 e 4 do conjunto DDI.

Tabela 11 – Resultado de métricas do experimento 3.2

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|--------|----------|-------------------|
| EfficientNet_v2 | 0.5000 | 0.494667 | 0.5000 | 0.492327 | -0.012926 |
| ViT_b16 | 0.5625 | 0.561989 | 0.5625 | 0.562157 | 0.121745 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.005142s/amostra +- 0.003916s/amostra para o modelo EfficientNet_V2 e 0.009826s/amostra +- 0.000116s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.8783, enquanto o EfficientNet_V2 apresentou um Trade-off de 1.0129.

- **Experimento 3.3:** O modelo foi treinado com a base principal, adicionada das 20 imagens do conjunto de teste do Experimento 3.1 (10 malignas e 10 benignas), e também das imagens das escalas Fitzpatrick 1 e 2 e 3 e 4 usadas nos conjuntos de teste dos Experimentos 1.1 e 2.1, respectivamente. Assim, a base de treinamento incorporou imagens das escalas Fitzpatrick 1 a 4 do conjunto DDI, enquanto as de escala 5 e 6 foram usadas somente para teste.

Tabela 12 – Resultado de métricas do experimento 3.3

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|----------|-----------|--------|----------|-------------------|
| EfficientNet_v2 | 0.5000 | 0.487461 | 0.5000 | 0.473333 | -0.025227 |
| ViT_b16 | 0.5125 | 0.506062 | 0.5125 | 0.499474 | 0.009264 |

Fonte: Elaborada pela autora.

Com taxa de inferência igual a 0.005809s/amostra +- 0.004721s/amostra para o modelo EfficientNet_V2 e 0.010162s/amostra +- 0.9.5e-05s/amostra para o modelo ViT_b16, sendo este definido como o melhor modelo, com Trade-off igual a 0.9908, enquanto o EfficientNet_V2 apresentou um Trade-off de 1.0252.

4.4 Síntese de Resultados

Em todos os experimentos, observou-se uma queda no desempenho ao transitar das escalas 3-4 para as escalas 5-6. No entanto, nem todos os modelos mostraram redução de desempenho ao passar das escalas 1-2 para 3-4, os quais até apresentaram resultados aproximados. Essa constatação pode ser atribuída à semelhança visual das imagens nas escalas 1 a 4, que predominam tonalidades semelhantes. Em contrapartida, as imagens das escalas 5-6 apresentam características distintas, o que pode ter influenciado os resultados.

Nas tabelas 13 e 14, são apresentadas as métricas de desempenho de dois experimentos: o primeiro foi treinado e testado exclusivamente com imagens dos conjuntos PAD-UFES-20 e ISIC, que são visualmente semelhantes. O segundo experimento, por sua vez, utilizou todas as imagens deste trabalho, isso inclui imagens dos conjuntos PAD-UFES-20, ISIC e DDI distribuídas proporcionalmente em treinamento e teste.

Tabela 13 – Resultado de métricas do experimento com imagens dos conjuntos PAD-UFES-20 e ISIC

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|-----------------|------------------|---------------|-----------|--------------------------|
| EfficientNet_v2 | 0.759531 | 0.774701 | 0.759531 | 0.754220 | 0.530319 |
| ViT_b16 | 0.807918 | 0.809452 | 0.807918 | 0.807206 | 0.615900 |

Fonte: Elaborada pela autora.

Tabela 14 – Resultado de métricas do experimento com imagens dos conjuntos PAD-UFES-20, ISIC e DDI

| modelo | accuracy | precision | recall | f1 | matthews_corrcoef |
|-----------------|-----------------|------------------|---------------|-----------|--------------------------|
| EfficientNet_v2 | 0.757683 | 0.761246 | 0.757683 | 0.752425 | 0.504747 |
| ViT_b16 | 0.775414 | 0.780758 | 0.775414 | 0.770265 | 0.542809 |

Fonte: Elaborada pela autora.

Estes experimentos evidenciaram que a abordagem dos experimentos anteriores, que treinaram com imagens dos conjuntos PAD-UFES-20 e ISIC e testaram com imagens do conjunto DDI, impactou negativamente os resultados das métricas. Isso ocorre porque as imagens de treinamento, embora clínicas, se assemelham com imagens dermatoscópicas, oferecendo uma visão aproximada das lesões, enquanto as imagens do conjunto de teste proporcionam uma visão mais ampla. Além disso, as imagens de teste incluem diversos atributos, como régua e marcações com canetas, que podem interferir na análise. As Figuras 18 e 19 ilustram amostras representativas dos conjuntos de treinamento e teste, respectivamente.

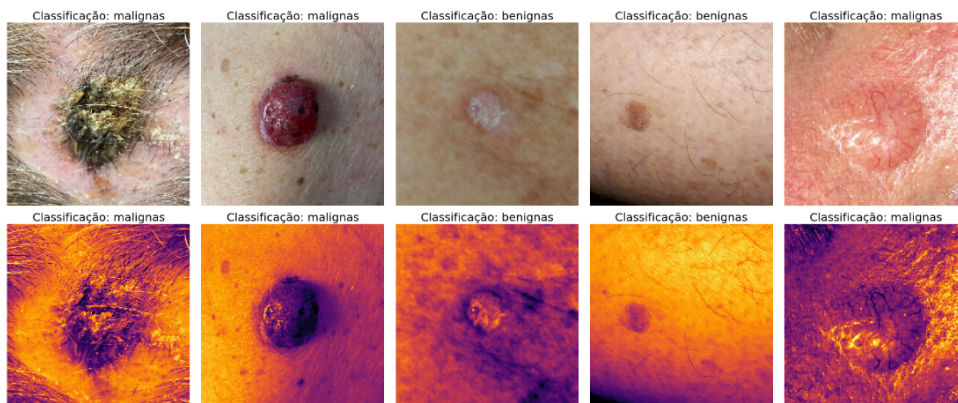


Figura 18 – Amostras de imagens do conjunto de treinamento. Fonte: Adaptado de (PAD-UFES-20... , 2020) e (ISIC... , s.d.).

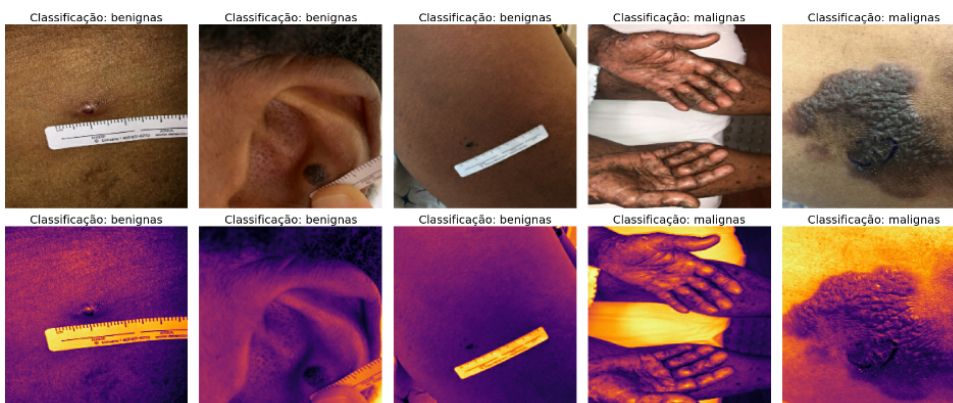


Figura 19 – Amostras de imagens do conjunto de teste. Fonte: Adaptado de (DDI... , 2022).

Observou-se também maior precisão na predição de imagens classificadas como benignas em todos os experimentos realizados nas seções 5.1, 5.2 e 5.3. Essa tendência foi consistente entre os diferentes modelos analisados. A Tabela 15 resume a média dessa métrica para cada categoria de experimento.

Tabela 15 – Média de precisões por categoria de experimento

| Experimento | Classificação | Precision | Recall | F1 |
|---------------|---------------|-----------|--------|------|
| Experimento 1 | benignas | 0.65 | 0.40 | 0.49 |
| | malignas | 0.54 | 0.76 | 0.63 |
| Experimento 2 | benignas | 0.63 | 0.46 | 0.52 |
| | malignas | 0.55 | 0.70 | 0.61 |
| Experimento 3 | benignas | 0.51 | 0.56 | 0.53 |
| | malignas | 0.47 | 0.42 | 0.44 |

Fonte: Elaborada pela autora.

Além disso, foi observado de forma geral que o modelo ViT apresentou um desem-

penho superior em comparação ao EfficientNet V2. Essa superioridade se manifestou em diversas métricas. O ViT demonstrou uma maior capacidade de generalização, inclusive em imagens complexas. Esse resultado reforça a eficácia da arquitetura Transformer na tarefa de classificação de imagens. Nas Figuras 20, 21, 22 e 23, apresentaremos algumas predições de cada modelo para alguns experimentos.

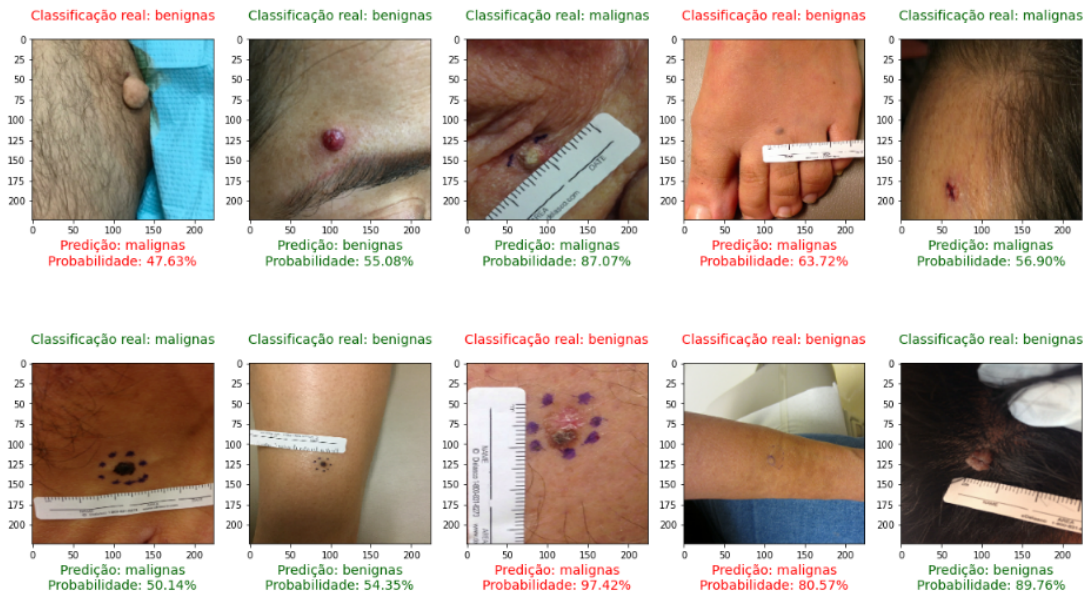


Figura 20 – Predições do modelo EfficientNet_V2 no experimento 2.2. Fonte: Adaptado de (DDI..., 2022).



Figura 21 – Predições do modelo ViT no experimento 2.2. Fonte: Adaptado de (DDI..., 2022).

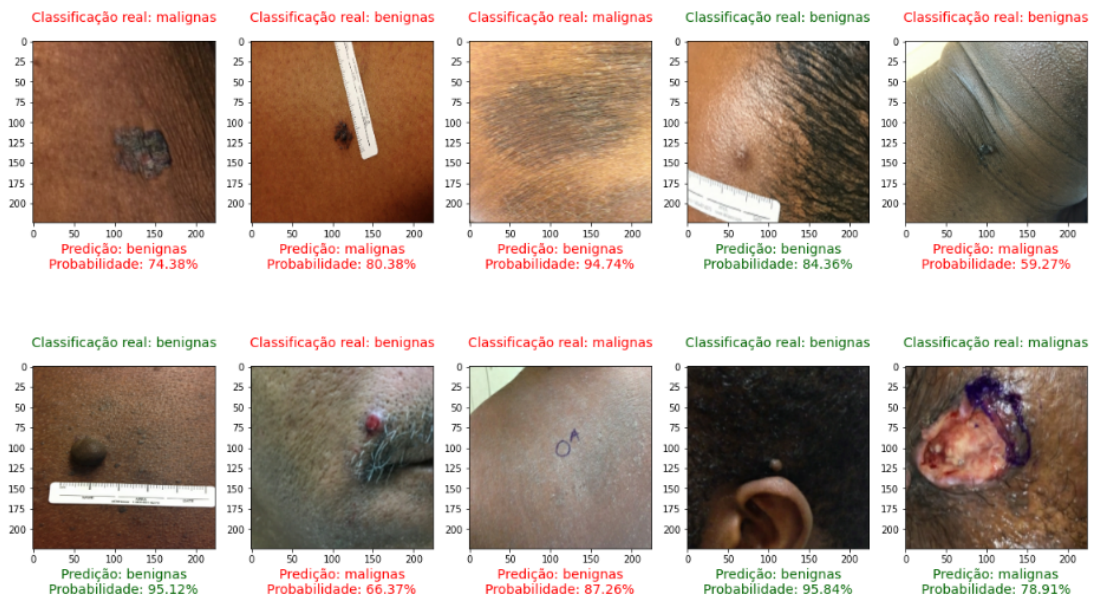


Figura 22 – Predições do modelo EfficientNet_V2 no experimento 3.2. Fonte: Adaptado de (DDI..., 2022).



Figura 23 – Predições do modelo ViT no experimento 3.2. Fonte: Adaptado de (DDI..., 2022).

5 Conclusão

Este estudo evidencia a necessidade de desenvolver soluções de inteligência artificial que sejam justas e equitativas. As redes neurais artificiais demonstraram um desempenho significativo na classificação e análise de lesões cutâneas, especialmente em amostras de pele branca. No entanto, a análise dos resultados revelou que a generalização desses modelos é comprometida em relação a imagens de indivíduos com pele negra, ressaltando a importância de conjuntos de dados diversificados.

Futuras pesquisas devem priorizar a criação de conjuntos de dados abrangentes, colaborando com instituições de saúde e comunidades para garantir uma representação justa. Campanhas de sensibilização são essenciais para incentivar a participação de indivíduos de diversas origens, promovendo um envolvimento mais amplo na coleta de dados. Além disso, a aplicação de técnicas avançadas de aumento de dados e algoritmos de geração adversária pode ser fundamental para a criação de imagens sintéticas que reflitam a diversidade dos tons de pele.

A utilização de transfer learning pode também ser uma estratégia promissora, permitindo que modelos treinados em populações mais amplas sejam adaptados para grupos sub-representados, melhorando assim a precisão diagnóstica. A elaboração de modelos compreensíveis é essencial, pois não apenas fornece previsões, mas também ajuda os profissionais de saúde a entender as razões por trás das decisões do modelo. Visualizações que mostram como as características de diferentes tipos de pele influenciam as classificações podem ser ferramentas valiosas nesse processo.

Em suma, a formação de parcerias multidisciplinares e a realização de estudos longitudinais são cruciais para avaliar o impacto das tecnologias assistidas por IA em populações diversas. Esses esforços garantirão que a inovação em saúde não apenas avance tecnicamente, mas também promova a equidade e inclusão, beneficiando a sociedade como um todo.

Referências

- ALEXEY, Dosovitskiy. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv: 2010.11929*, 2020. Citado 5 vezes nas páginas 27–29.
- ANDRADE, Livia Silveira; BEM, Yuri Almeida de; TEIXEIRA, Gabrielle Isis Alcântara. Análise dos diferentes tipos de pele e sua suscetibilidade ao câncer de pele sob a exposição aos raios ultravioleta no Brasil. In: CONGRESSO Médico Acadêmico UniFOA. 2024. v. 10. Citado 1 vez na página 18.
- ARAUJO, Izabella Costa et al. Melanoma cutâneo: aspectos clínicos, epidemiológicos e anatomopatológicos de um centro de formação em Belo Horizonte. *Revista Brasileira de Cirurgia Plástica*, SciELO Brasil, v. 29, n. 4, p. 497–503, 2014. Citado 1 vez na página 15.
- BARROS, Luana; CHAVES, Levy; AVILA, Sandra. Assessing the Generalizability of Deep Neural Networks-Based Models for Black Skin Lesions. In: SPRINGER. IBEROAMERICAN Congress on Pattern Recognition. 2023. P. 1–14. Citado 3 vezes nas páginas 11, 18, 30.
- BARROS, Luana Felipe de; AVILA, Sandra Eliza Fontes de. *Análise de Modelos Baseados em Redes Neurais Profundas para Lesões de Pele Negra*. 2022. Citado 3 vezes nas páginas 11, 30.
- BERMOND, Matteo et al. Reducing racial bias in SpO₂ estimation: The effects of skin pigmentation. In: IEEE. 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). 2023. P. 1–5. Citado 1 vez na página 30.
- BRAZIL, SKIN MELANOMA IN. Melanoma cutâneo no Brasil. *Arquivos Catarinenses de Medicina*, v. 38, Suplemento 01, p. 14, 2009. Citado 2 vezes na página 17.
- CUNHA, Leonardo Cardoso da. *Redes neurais convolucionais e segmentação de imagens: uma revisão bibliográfica.*, 2020. Citado 8 vezes nas páginas 21–26.
- DANESHJOU, Roxana 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, eabq6147, 2022. Citado 1 vez na página 33.
- DATABRICKS. *Camada convolucional*. Acessado em: jan. 2024. databricks. Disponível em: <https://www.databricks.com/br/glossary/convolutional-layer>. Citado 1 vez na página 22.
- DDI. Acessado em: jan. 2024. 2022. Disponível em: <https://stanfordaimi.azurewebsites.net/datasets/35866158-8196-48d8-87bf-50dca81df965>. Citado 0 vez nas páginas 41–43.

- EID, Renata Trefiglio; ALCHORNE, Maurício Mota de Avelar. Câncer na pele negra. *Rev Bras Clin Med. São Paulo*, v. 9, n. 6, p. 418–22, 2011. Citado 8 vezes nas páginas 11, 18, 19.
- FABI, Sarah; XU, Xiaojing; SA, Virginia de. Exploring the racial bias in pain detection with a computer vision model. In: 44. PROCEEDINGS of the Annual Meeting of the Cognitive Science Society. 2022. v. 44. Citado 1 vez na página 29.
- HAN, Seung Seog et al. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, Elsevier, v. 138, n. 7, p. 1529–1538, 2018. Citado 1 vez na página 22.
- IARC. *Data visualization tools for exploring the global cancer burden in 2022*. Acessado em: jun. 2024. International Agency for Research on Cancer. 2022. Disponível em: <https://gco.iarc.who.int/today/en>. Citado 0 vez na página 17.
- IBCC. *Câncer de pele representará 31,3% dos casos da doença em 2023, aponta estudo*. Acessado em: ago. 2024. Instituto Brasileiro de Controle do Câncer. 2023. Disponível em: <https://ibcc.org.br/cancer-de-pele-representara-313-dos-casos-da-doenca-em-2023-aponta-estudo/>. Citado 2 vezes na página 16.
- IBGEEDUCA. *Cor ou raça*. Acessado em: ago. 2024. IBGEeduca. Disponível em: <https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18319-cor-ou-raca.html>. Citado 2 vezes na página 20.
- IEPS. *Norte e Nordeste apresentam piores índices do sistema de saúde, segundo pesquisa do IEPS*. Acessado em: abr. 2024. Instituto de Estudos para Políticas de Saúde. 2022. Disponível em: <https://ieps.org.br/norte-e-nordeste-apresentam-piores-indices-do-sistema-de-saude-segundo-pesquisa-do-ieps/>. Citado 0 vez na página 20.
- IMB. *O que é melanoma*. Acessado em: jan. 2024. Instituto Melanoma Brasil. Disponível em: <https://www.melanomabrasil.org/new-oquee/>. Citado 2 vezes nas páginas 16, 18.
- INCA. *ABC do Câncer - Abordagens Básicas para o Controle do Câncer*. Acessado em: jan. 2024. Instituto Nacional de Câncer. 2011. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/abc_do_cancer.pdf. Citado 1 vez nas páginas 15, 16.
- INCA. *Câncer de pele melanoma*. Acessado em: jan. 2024. Instituto Nacional de Câncer. 2022. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/pele-melanoma>. Citado 6 vezes nas páginas 16–18.
- INCA. *Dezembro Laranja: prevenção e detecção precoce do câncer de pele*. Acessado em: ago. 2024. Instituto Nacional de Câncer. 2023. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-brasil/prevencao-ao-cancer/dezembro-laranja-prevencao-e-deteccao-precoce-do-cancer-de-pele>. Citado 1 vez na página 18.

- INCA. *Melanoma maligno da pele (taxas brutas)*. Acessado em: jun. 2024. Instituto Nacional de Câncer. 2022. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/numeros/estimativa/por-neoplasia-taxas-brutas/pele-melanoma>. Citado 1 vez na página 17.
- INCA. *O que é câncer?* Acessado em: jan. 2024. Instituto Nacional de Câncer. 2022. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/o-que-e-cancer>. Citado 1 vez na página 15.
- INCA. *Pele não melanoma (taxas brutas)*. Acessado em: jun. 2024. Instituto Nacional de Câncer. 2022. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/numeros/estimativa/por-neoplasia-taxas-brutas/pele-nao-melanoma>. Citado 1 vez na página 17.
- INCA. *Síntese de Resultados e Comentários*. Acessado em: jun. 2024. Instituto Nacional de Câncer. 2023. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/numeros/estimativa/sintese-de-resultados-e-comentarios>. Citado 1 vez na página 17.
- ISIC. Acessado em: out. 2024. 2023. Disponível em: <https://dermosclopedia.org/w/index.php?title=09-ISIC&oldid=->. Citado 1 vez na página 32.
- ISIC Archive. Acessado em: jan. 2024. Disponível em: <https://www.isic-archive.com/>. Citado 0 vez na página 41.
- LIMA, Daniel Lopes Soares et al. Classificação de imagens de exames de endoscopia por cápsula utilizando transformers. Universidade Federal do Maranhão, 2023. Citado 5 vezes nas páginas 27, 28.
- MASSUCATTO, J. Aplicação de conceitos de redes neurais convolucionais na classificação de imagens de folhas. *trabalho de conclusão de curso, Universidade Tecnológica Federal do Paraná*, 2018. Citado 2 vezes nas páginas 25, 26.
- MIRIKHARAJI, Zahra et al. A survey on deep learning for skin lesion segmentation. *Medical Image Analysis*, Elsevier, v. 88, p. 102863, 2023. Citado 1 vez na página 30.
- PACHECO, Andre GC; KROHLING, Renato A. The impact of patient clinical information on automated skin cancer detection. *Computers in biology and medicine*, Elsevier, v. 116, p. 103545, 2020. Citado 2 vezes na página 31.
- PAD-UFES-20. Acessado em: jan. 2024. 2020. Disponível em: <https://data.mendeley.com/datasets/zr7vgbcyr2/1>. Citado 0 vez na página 41.
- RODRIGUES, Diego Alves. Deep Learning e redes neurais convolucionais: reconhecimento automático de caracteres em placas de licenciamento automotivo. Universidade Federal da Paraíba, 2018. Citado 3 vezes nas páginas 22, 23, 25.

RUBACK, Lívia; AVILA, Sandra; CANTERO, Lucia. Vieses no aprendizado de máquina e suas implicações sociais: Um estudo de caso no reconhecimento facial. In: SBC. WORKSHOP sobre as Implicações da Computação na Sociedade (WICS). 2021. P. 90–101. Citado 1 vez na página 29.

SBD. *Cada um com a sua prevenção*. Acessado em: ago. 2024. Sociedade Brasileira de Dermatologia. 2023. Disponível em: <https://sbd.org.br/campanha/dezembrolaranja/>. Citado 2 vezes nas páginas 16, 18.

SBD. *Classificação dos fototipos de pele*. Acessado em: ago. 2024. Sociedade Brasileira de Dermatologia. Disponível em: <https://www.sbd.org.br/cuidados/classificacao-dos-fototipos-de-pele/>. Citado 3 vezes nas páginas 11, 12.

SOBRE o ISIC. Acessado em: out. 2024. Disponível em: <https://www.isic-archive.com/mission>. Citado 1 vez na página 32.

SOUSA, Alexandre Santana. Análise comparativa de redes neurais convolucionais para a detecção de câncer de pulmão em tomografias computadorizadas, 2023. Citado 3 vezes nas páginas 24, 26.

TAN, Mingxing; LE, Quoc. Efficientnetv2: Smaller models and faster training. In: PMLR. INTERNATIONAL conference on machine learning. 2021. P. 10096–10106. Citado 2 vezes nas páginas 26, 27.

TRASK, Andrew W. *Grokking deep learning*. Simon e Schuster, 2019. Citado 1 vez na página 21.