



ESCOLA SALESIANA SÃO JOSÉ
CENTRO PROFISSIONAL DOM BOSCO - CPDB

Claudinei Alves Pereira dos Reis
Miguel Afonso Castro de Almeida
Thallyson Kawan Lima Oliveira

**Skin Analyser AI: Uso de Inteligência Artificial para Classificação e
Identificação de Câncer de Pele**

CAMPINAS
2025

Claudinei Alves Pereira dos Reis
Miguel Afonso Castro de Almeida
Thallyson Kawan Lima Oliveira

Skin Analyser AI: Uso de Inteligência Artificial para Classificação e Identificação de Câncer de Pele

Relatório referente ao Projeto de Conclusão de Curso, apresentado ao Centro Profissional Dom Bosco da Escola Salesiana São José, como parte dos requisitos necessários à obtenção do título de Técnico em Informática.

Orientador: Adriana Maia da Silva Coelho
Coorientador: Galesandro Henrique Capovilla

CAMPINAS
2025

AGRADECIMENTOS

Gostaríamos de expressar nossos sinceros agradecimentos à Escola Salesiana São José pelo apoio institucional e incentivo durante todas as etapas deste trabalho.

Agradecemos também à Adriana Maia e ao Galesandro Capovilla pela colaboração, orientação e suporte dedicados, que foram fundamentais para o desenvolvimento deste projeto.

Manifestamos nossa gratidão ao Centro de Oncologia de Campinas pelo fornecimento de informações técnicas e suporte especializado, essenciais para a fundamentação e aprimoramento do estudo.

Expressamos nossa profunda gratidão ao artigo "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review". Sua análise meticulosa e centralizada nos forneceu a base teórica e metodológica que, de outra forma, levaria um tempo considerável para ser compilada, direcionando-nos eficientemente para o ponto de partida ideal para nossa pesquisa.

Por fim, agradecemos ao Dr. Fernando Medina pela valiosa contribuição científica, pelas orientações precisas e pelo comprometimento com a qualidade desta pesquisa.

A todos, nosso profundo reconhecimento e agradecimento.

*“Qualquer aspecto de aprendizagem ou
qualquer outra característica da inteligência
pode, em princípio, ser tão precisamente
descrito que uma máquina pode ser feita
para simulá-lo.”*

John Mc Carthy

RESUMO

O câncer de pele representa o tipo de neoplasia mais incidente no Brasil, com cerca de 31,3% dos diagnósticos, evidenciando uma detecção precoce, um fator crucial para um prognóstico favorável, com taxas de tratamento superiores a 90%. Contudo, barreiras de acesso e a demora no diagnóstico no sistema público de saúde constituem desafios significativos. Diante desta problemática, o presente projeto teve como objetivo o desenvolvimento e a validação de uma solução computacional, materializada em um aplicativo móvel, para a classificação de lesões cutâneas por meio de uma inteligência artificial especializada. A metodologia empregou a arquitetura de *deep learning*, treinada e validada com um conjunto de dados consolidado a partir de fontes públicas (ISIC Archive). O modelo de IA alcançou um desempenho robusto, com uma acurácia global de 92% e uma sensibilidade de 87% na detecção de lesões malignas, superando a meta inicial de 90%. O projeto entregou com sucesso o aplicativo móvel *Skin Analyser AI* e um dashboard interativo para análise de KPIs durante o treinamento da IA. Os resultados obtidos até o momento confirmam a hipótese de que a inteligência artificial é uma ferramenta viável e de altas taxas de precisão (com ressalvas a possíveis erros) para a triagem de câncer de pele, com potencial para atuar como um sistema de apoio à decisão clínica, contribuindo para a democratização do diagnóstico precoce e otimização de recursos na saúde pública.

Palavras-chave: Câncer de Pele. Inteligência Artificial. Saúde Digital.

ABSTRACT

Skin cancer is the most prevalent form of neoplasia in Brazil, comprising 31.3% of all cancer diagnoses. For this condition, early detection is a critical determinant for a favorable prognosis, with cure rates exceeding 90%. Nevertheless, significant challenges persist within the public health system, including barriers to access and delays in diagnosis. To address this issue, the present project focused on the development and validation of a computational solution mobile application for the classification of cutaneous lesions through a specialized artificial intelligence. The methodology utilized a deep learning architecture, which was trained and validated on a consolidated dataset from public sources (ISIC Archive). The AI model demonstrated robust performance, achieving an overall accuracy of 92% and a sensitivity of 87% for the detection of malignant lesions, thereby exceeding the initial 90% performance goal. The project successfully delivered the "Skin Analyser AI" mobile application and an interactive dashboard for monitoring KPIs during the AI training phase. The results confirm the hypothesis that artificial intelligence serves as a viable and highly accurate tool for skin cancer screening (acknowledging potential errors), showing significant potential as a clinical decision support system. This can contribute to democratizing early diagnosis and optimizing resource allocation in public health.

Keywords: Skin Cancer. Artificial Intelligence. Digital Health.

LISTA DE ILUSTRAÇÕES

Figura 1- Imagem original de uma doença de pele do tipo nevos melanocítico	52
Figura 2: Imagem sem pelo a partir da imagem original	52

LISTA DE TABELAS

Tabela 1 - Resumo dos Conjuntos de Dados Públicos.....	47
Tabela 2 - Planilha de custos do projeto	58
Tabela 3 - Métricas de Desempenho do Modelo de IA na Classificação de Lesões	69
Tabela 4 - Funcionalidades Implementadas no Aplicativo e Dashboard	60
Tabela 5 - Funcionalidades Implementadas no Aplicativo e Dashboard	61

LISTA DE SIGLAS

ABNT – Associação Brasileira de Normas Técnicas;

CPDB – Centro Profissional Dom Bosco;

ML – Machine Learning (Aprendizado de Máquina)

DL – Deep Learning (Aprendizado Profundo)

ANNs – Artificial Neural Networks (Redes Neurais Artificiais)

CNNs – Convolutional Neural Networks (Redes Neurais Convolucionais)

DNNs – Deep Neural Networks (Redes Neurais Profundas)

MEL – Melanoma

UNIFAL – Universidade Federal de Alfenas

SBD – Sociedade Brasileira de Dermatologia

Vit. D – Vitamina D

SI – Sistema Imunológico

UV – Radiação Ultravioleta

GPUs – Graphics Processing Units (Unidades de Processamento Gráfico)

SUMÁRIO

1.INTRODUÇÃO	11
2. JUSTIFICATIVA	15
3. OBJETIVOS	20
4. REFERÊNCIAL TEÓRICO	21
4.1. Machine Learning e Deep Learning	21
4.1.1. Inteligência Artificial	21
4.1.2. Machine Learning	21
4.1.3. Deep Learning	22
4.1.4. Inteligência Artificial na medicina	22
4.1.5. Considerações sobre o estudo de Esteva	24
4.1.6. Arquitetura e foco do projeto	25
4.1.7. Inteligência Artificial evidenciado no diagnóstico de pele	26
4.2. Pele	28
4.2.1. Anatomia e Fisiologia da Pele	28
4.3. Câncer de pele	29
4.3.1. Epidemiologia	29
4.3.2. Tipos de Câncer de Pele	30
4.3.3. Câncer de pele não melanoma (CPNM)	31
4.3.4. Melanoma cutâneo	31
5. MATERIAIS E MÉTODOS	33
5.1 Materiais	33
5.1.1 Notion – Planejamento	33
5.1.2 Supabase - Banco de dados	34
5.1.3 Linguagens e I5.1.3.1 Python, TensorFlow e Keras	34

5.1.4 Manipulação de dados	35
5.1.5 Logging	39
5.1.6 Ambientes de Edição	41
5.2 Métodos.....	41
5.2.1 Proposta de Pesquisa.....	41
5.2.2 Considerações Éticas e Metodologia.....	42
5.2.3 Desenho do Estudo	43
5.2.4 Contextualização e Justificativa	44
5.2.5 Coleta e Curadoria de Dados	45
5.2.5.1 Aquisição de Dados	45
5.2.5.2 Coleta e Curadoria de Dados	47
5.2.6 Pré-processamento e Aumento de Dados	49
5.2.6.1 Técnicas de Pré-processamento	50
5.2.6.2 Normalização e Redimensionamento	50
5.2.6.3 Remoção de Artefatos	51
5.2.7 Aumento de Dados	53
5.2.8 Arquitetura e Desenvolvimento do Modelo	54
5.2.8.1 Seleção da Arquitetura	54
5.2.8.2 Backbones e Transferência de Aprendizado	55
5.2.9 Treinamento, Avaliação e Validação.....	Erro! Indicador não definido.
5.2.9.1 Ambiente de Software e Hardware.....	Erro! Indicador não definido.
6. PLANILHA DE CUSTOS DO PROJETO.....	58
7. RESULTADOS OBTIDOS.....	59
8. CONCLUSÃO E CONSIDERAÇÕES FINAIS	62

1. INTRODUÇÃO

O câncer continua sendo um dos maiores desafios da saúde pública mundial. Segundo a Organização Mundial da Saúde, o câncer foi responsável por cerca de **10 milhões de mortes em 2020(who)**. Embora avanços terapêuticos recentes tenham melhorado a sobrevivência de muitos pacientes, ainda não existe uma cura definitiva para todos os tipos de câncer. No entanto, sabe-se que *muitos cânceres podem ser entrar em período de remissão* (significa que a doença está sob controle, sem sinais detectáveis em exames clínicos, laboratoriais ou de imagem, mas não significa cura definitiva, pois há um risco de recidiva) *se detectados precocemente e tratados adequadamente*. Isso torna a **detecção precoce** um fator-chave na redução de mortalidade, pois diagnósticos em estágios iniciais permitem tratamentos mais eficazes. Em resumo, enquanto a cura universal do câncer permanece fora de alcance, a identificação rápida de tumores faz toda a diferença no prognóstico dos pacientes.

O câncer de pele é o tipo de câncer mais frequente no Brasil(gov.br). Dados do Instituto Nacional do Câncer (INCA) indicam que o câncer de pele não melanoma corresponde a cerca de **30% de todos os tumores malignos registrados no país**. Em 2020 foram estimados aproximadamente **176.930 novos casos** de câncer de pele não melanoma no Brasil (83.770 homens e 93.160 mulheres). A incidência elevada desse câncer está fortemente associada à exposição excessiva à radiação solar (sendo acumulativa por toda a vida), sobretudo em pessoas de pele clara. É importante ressaltar que, apesar de sua alta frequência, o câncer de pele não melanoma tem *alta chance de cura quando tratado precocemente*. Conforme orientações médicas, qualquer mudança suspeita na pele deve ser investigada imediatamente, pois *“quanto mais precoce for sua identificação, melhores serão os resultados do tratamento”* (gov.br).

Características típicas de uma lesão suspeita de melanoma (o tipo mais agressivo de câncer de pele) podem ser lembradas pelo critério ABCDE (aad.org):

- **A – Assimetria:** a metade da pinta não corresponde à outra.
- **B – Borda Irregular:** bordas tortuosas, recortadas ou mal definidas.
- **C – Cor:** variações de cor ou pigmentação heterogênea.
- **D – Diâmetro:** geralmente maior que 6 mm (tamanho do apagador de lápis).

- **E – Evolução:** mudança em tamanho, forma ou cor ao longo do tempo.

Essas características ajudam a distinguir melanomas de pintas benignas e orientam a classificação das lesões. O rastreamento contínuo e a análise cuidadosa seguem como método padrão para o diagnóstico de suspeitas; exames como dermatoscopia e biópsia confirmam o diagnóstico. Contudo, o acesso a especialistas nem sempre é fácil em todas as regiões do país, tornando o câncer de pele uma preocupação relevante para a saúde pública brasileira. A **Inteligência Artificial (IA)** refere-se a sistemas computacionais capazes de aprender, adaptar-se e executar tarefas complexas simulando alguns aspectos do raciocínio humano. Na prática, isso significa que algoritmos de IA podem *reconhecer padrões* em dados (como imagens, sinais clínicos e históricos de pacientes) e tomar decisões baseadas em experiências prévias de forma autônoma. No setor de saúde, a IA tem potencial transformador: ela pode **acelerar diagnósticos**, tornando-os mais precisos e disponibilizando-os em menor tempo. Além disso, ao automatizar tarefas administrativas rotineiras, a IA pode reduzir custos operacionais e liberar mais tempo dos profissionais para o cuidado direto ao paciente. Como observado em estudo recente do Ministério da Saúde, tecnologias de IA em saúde já contribuem para *“atendimentos e diagnósticos mais rápidos e precisos, reduzindo custos e melhorando a experiência de pacientes e de profissionais”* (bvsm.sau.br). Percebe-se o aumento de tecnologias implementadas com IA na área da saúde no aumento do financiamento de global para empresas de saúde digital, que em 2024 representou 42% do financiamento anual total para saúde (Statista, 2025) e digital, um aumento significativo em relação aos 7% do financiamento para saúde digital em 2015.

Recentes pesquisas científicas têm demonstrado que técnicas de aprendizado profundo podem ser extremamente eficazes na detecção de câncer de pele a partir de imagens. Vários estudos utilizaram **Redes Neurais Convolucionais (CNNs)** – modelos de IA especialmente desenhados para processamento de imagens – para classificar lesões cutâneas. Por exemplo, os estudos de Amanda V.C.A. (2024) e Liana C. (2020), no Brasil(Unicamp), demonstram considerações sobre as tecnologias que podem ser usadas para treinar um algoritmo (CNNs, GANs, KNN, ResNet) e como esses respectivos algoritmos podem atingir precisão comparável ou superiores à de dermatologistas especializados, em seus estudos abordam preocupações técnicas e éticas, desafios como a necessidade de grandes bancos de dados balanceados e a

interpretação dos resultados proporcionados pelo software pelos profissionais da saúde serem utilizados como uma ferramenta auxiliadora. Em outro estudo, de Musthafa et al. (2024), criaram um modelo CNN treinado com o conjunto de imagens dermatoscópicas HAM10000 e alcançaram uma acurácia de **97,78%** na detecção de câncer de pele(pubmed). Similarmente, abordagens avançadas combinando CNNs e mecanismos de atenção relatam **precisões superiores a 95%** na classificação de melanoma(nature). Esses resultados indicam que algoritmos de IA bem treinados podem atingir níveis de desempenho comparáveis ou até superiores aos de dermatologistas experientes.

Em resumo, a IA emerge como uma alternativa promissora para **pré-diagnósticos rápidos e padronizados**. Ela pode automatizar a análise de imagens da pele, identificando padrões morfológicos associados ao câncer que podem escapar ao olhar humano ou demandar equipamentos caros. Isso é especialmente valioso em áreas remotas ou carentes de especialistas, onde um diagnóstico precoce pode fazer a diferença entre cura e complicações graves.

Então, diante da problemática, o projeto visa desenvolver um aplicativo móvel baseado em IA para identificação e classificação de câncer de pele por reconhecimento imagens. Como estudo de caso o projeto propõe:

- **Analisar bibliografia especializada** sobre câncer de pele e tecnologias de IA em saúde, compreendendo o funcionamento das tecnologias emergentes.
- **Implementar um modelo de aprendizado profundo** utilizando bases de dados dermatológicas, buscando precisão diagnóstica de pelo menos **90%** nos testes.
- **Validar a ferramenta em ambiente de atenção primária**, com parcerias a instituições especializadas juntos com seus especialistas fazer a validação do modelo de IA.
- **Democratizar o acesso ao diagnóstico precoce**, disponibilizando uma solução de baixo custo que complemente a avaliação médica e acelere a detecção de lesões suspeitas.

Em conjunto, essas ações deverão demonstrar o potencial da IA em facilitar o diagnóstico de câncer de pele e em contribuir positivamente para a saúde pública, especialmente em regiões com escassez de especialistas, com o custo elevado de

exames convencionais e a avaliação clínica, que representam barreiras significativas para serem superadas.

Espera-se que, ao indicar precocemente lesões suspeitas, o aplicativo ajude a encaminhar pacientes rapidamente aos tratamentos adequados, melhorando os resultados clínicos. Conforme reforçado pelas diretrizes médicas, *identificar a doença durante seu início aumenta significativamente a eficácia do tratamento*. Assim, o projeto, reforçando a atuação como uma ferramenta, almeja mostrar que tecnologias emergentes de IA podem ampliar o acesso ao diagnóstico dermatológico e, desse modo, reduzir o impacto do câncer de pele na população.

2. JUSTIFICATIVA

O câncer é uma doença caracterizada pelo crescimento descontrolado de células anormais, capazes de invadir tecidos adjacentes e causar metástases. No Brasil, observa-se o aumento expressivo dos casos. Segundo dados do INCA, estima-se cerca de **704 mil novos casos de câncer por ano** no triênio 2023–2025, onde aproximadamente 70% desses casos ocorrerão nas regiões Sul e Sudeste. Essa previsão reflete tanto o envelhecimento populacional quanto fatores de risco prevalentes (como urbanização, tabagismo e obesidade). Entre os tipos mais incidentes estão o câncer de pele não melanoma (o mais comum), de mama e próstata. Em termos de prognóstico, tumores malignos (câncer) diferem de tumores benignos por serem invasivos e agressivos, enquanto os tumores benignos crescem lentamente e raramente se espalham para outra parte do corpo, as células malignas invadem tecidos próximos e/ou à distância, sendo denominados câncer. Essa dinâmica favorece o surgimento de metástases, processo em que células cancerosas se desprendem do tumor original, entram na circulação sanguínea ou linfática e formam novos tumores em outras partes do corpo.

Grande parte dos casos de câncer está associada a fatores ambientais e comportamentais. Estima-se que **80–90%** dos cânceres têm relação com causas externas, como hábitos de vida e exposições ambientais. Dentre os principais fatores de risco destacam-se:

- **Exposição solar excessiva:** a radiação ultravioleta (UV) é o principal fator de risco para câncer de pele. No Brasil, com clima tropical e grande exposição ao sol, a incidência de lesões cutâneas suspeitas aumenta consideravelmente. A radiação UV é acumulativa por toda vida, significa que toda a exposição a raios solares durante sua vida pode favorecer a proliferação de um câncer de pele.
- **Hábitos não saudáveis:** tabagismo, consumo abusivo de álcool, sedentarismo, dieta pobre em nutrientes e obesidade são amplamente reconhecidos como contribuidores de diversos tipos de câncer.
- **Histórico familiar e predisposição genética:** alterações hereditárias em genes (oncogenes e supressores tumorais) podem tornar alguns indivíduos mais suscetíveis à carcinogênese. No entanto, casos exclusivamente genéticos

são raros (10–20% dos casos), indicando que a maior parte é provocada por causas externas.

- **Contato com agentes carcinógenos:** substâncias químicas nocivas (ex.: pesticidas, solventes industriais, poluentes) e certos vírus (como HPV) podem induzir mutações no DNA celular. O processo de formação do câncer (carcinogênese) envolve múltiplas etapas: **iniciação** (mutação no DNA), **promoção** (agentes promotores fazem a célula iniciada evoluir para malignidade) e **progressão** (proliferação descontrolada e estabelecimento do tumor). Fumo de tabaco, por exemplo, é um agente carcinógeno completo que atua em todas essas fases.

No estágio inicial, o câncer pode apresentar sinais sutis ou inespecíficos, variando conforme localização e tamanho do tumor. Em cânceres de pele, por exemplo, **carcinomas** (tumores originários do epitélio) correspondem à maioria dos casos, enquanto **melanomas** (tumores de melanócitos) são menos frequentes, mas notoriamente agressivos. Os melanomas surgem geralmente como pintas ou manchas assimétricas e pigmentadas, que crescem e mudam rapidamente – características que exigem atenção imediata. Embora esses sinais sejam visíveis, podem ser confundidos com outras lesões benignas ou condições de pele. Além disso, muitas pessoas postergam a busca por atendimento por acreditarem que o sintoma não apresente um perigo real ou por medo do diagnóstico. Quando o câncer atinge estágios avançados, é comum a disseminação por metástase: nesse caso, células cancerosas viajam pela corrente sanguínea e/ou linfática e implantam tumores em órgãos distantes. Por isso, a detecção precoce é crucial para melhorar o prognóstico.

Vários fatores dificultam o diagnóstico precoce no contexto brasileiro. A desigualdade na distribuição de recursos de saúde faz com que regiões remotas ou menos desenvolvidas tenham menor acesso a especialistas e exames avançados, causando atrasos significativos. Muitas vezes, pacientes de baixa renda não dispõem de recursos para exames particulares e enfrentam longas filas na rede pública. Além disso, procedimentos de imagem como tomografia e ressonância são caros e nem sempre disponíveis (além de exigirem indicação clínica precisa). Essas barreiras socioeconômicas e logísticas resultam em **diagnósticos tardios**, agravando a taxa de mortalidade. Por exemplo, mesmo exames preventivos simples (como exames

dermatológicos de rotina) podem ser negligenciados por falta de informação, infraestrutura ou receio dos pacientes, levando à piora da doença antes do tratamento adequado.

Diante desse cenário, propõe-se o desenvolvimento de um software de inteligência artificial para triagem de lesões de pele. O objetivo é **democratizar o acesso** à análise inicial de possíveis sinais de câncer cutâneo, integrando algoritmos de visão computacional a imagens de dermatologia. Funciona assim: um profissional de saúde captura imagens de manchas ou lesões suspeitas, e o sistema analisa padrões (Regra do ABCDE - Assimetria, Bordas irregulares, Cores variadas, Diâmetro (maior que 6mm) e Evolução (mudanças em tamanho, cor, forma ou sintomas); Regra EFG - Elevação, Firmeza ao toque, “Growth” / crescimento) para sinalizar indícios precoces de melanoma ou outros cânceres de pele. Dessa forma, busca-se reduzir diagnósticos incorretos ou despercebidos, encaminhando casos suspeitos para avaliação médica mais célere. Importante destacar que o software **não substitui o diagnóstico médico definitivo** nem a biópsia; ele atua como ferramenta de apoio estatístico, reforçando a precaução do especialista ao propor exames complementares quando necessário.

Essa abordagem apoia-se em bases de dados já existentes: instituições como a *ISIC Archive* mantêm grandes acervos públicos de imagens dermatoscópicas. Ou seja, há **farto material de treinamento** disponível para a inteligência artificial reconhecer características cancerígenas. O software pode ser validado em parceria com clínicas e hospitais, comparando seus resultados às avaliações de dermatologistas experientes.

- Entre as **ações-chave** do projeto incluem-se:
 - **Coleta e padronização de imagens** de lesões de pele (fotografias de nevus, manchas, pintas, etc.).
 - **Treinamento de modelo de IA** usando bases públicas de imagens de câncer de pele.
 - **Avaliação clínica** da acurácia do sistema em identificar lesões suspeitas, comparando com laudos histopatológicos.

O projeto segue rigorosamente as diretrizes éticas da medicina. Segundo o Código de Ética Médica (CEM 2019), o médico deve “empregar seu conhecimento e o melhor progresso científico em benefício do paciente” (Morsch), atuando sempre no interesse deste e evitando causar danos (princípio da **não maleficência**). Nesse sentido, a ferramenta de IA é concebida como **apoio complementar** (não substituirá o médico). Está previsto no CEM que o clínico deve usar “todos os meios disponíveis de promoção de saúde e de prevenção, diagnóstico e tratamento de doenças, cientificamente reconhecidos e ao seu alcance”. Isso inclui recursos tecnológicos validados – nosso software entra nessa categoria.

Para garantir segurança, o uso do sistema será restrito a ambientes clínicos e ao crivo de especialistas. O CEM também enfatiza o “respeito absoluto pelo ser humano” e a busca contínua de aprimoramento profissional. Assim, o projeto não ignora que a IA pode falhar (por exemplo, apresentar um 1% de margem de erro), mas entende essa limitação dentro de uma estratégia segura: resultados automáticos serão sempre confirmados por médicos. Não haverá tentativa de “diagnóstico automatizado” sem supervisão. Pelo contrário, ao destacar possíveis riscos, a ferramenta permite que o médico decida, de forma mais informada, pela realização de exames invasivos (biópsia) ou por vigilância apenas, minimizando **exaustão de exames desnecessários**. Em suma, a proposta prioriza o paciente: respeita as normas do CFM e reforça a empatia, lembrando que erro de diagnóstico precoce pode causar insegurança (art. 2º Cap. I), mas, ao mesmo tempo, mostra comprometimento com a saúde (art. 31 Cap. V).

O avanço recente em inteligência artificial e o maior acesso a bases de dados possibilitam esse desenvolvimento. Redes neurais para imagens estão cada vez mais potentes e rápidas, e bancos de imagem dermatológica (como o ISIC Archive) oferecem **cenários reais** para treinar e testar algoritmos. Espera-se parceria com universidades e institutos de saúde para fortalecer o treinamento e validar o sistema em ambiente clínico.

As metas de desempenho do software são ambiciosas, porém baseadas em resultados já alcançados: almeja-se **sensibilidade superior a 93%** (detectando cerca de 93% dos melanomas) e **especificidade próxima a 91%** (identificando 91% das lesões benignas corretamente), com acerto geral superior a 90%. Isso se

assemelharia a índices de projetos similares: por exemplo, um estudo clínico recente mostrou que um CNN aprovado para dermatoscopia atingiu aproximadamente 95% de sensibilidade e 76.7% de especificidade em uma ampla variedade de lesões. Outro estudo prospectivo demonstrou que, com cooperação médico-máquina, foi possível melhorar sensibilidade e especificidade em relação ao médico sozinho (Julia K. Winkler). Nosso projeto busca ir além: não apenas atingir esses patamares, mas também reduzir custos. Uma estimativa inicial sugere que o sistema possa ser mantido por cerca de US\$24 por usuário – valor muito inferior aos custos individuais de exames de imagem de alta complexidade.

Diante do crescimento contínuo do câncer e das dificuldades de acesso ao diagnóstico precoce no Brasil, a ferramenta proposta preenche uma lacuna estratégica. Ao antecipar a identificação de lesões suspeitas em estágios iniciais, promove-se **triagem mais eficaz**, permitindo encaminhamentos rápidos e tratamentos mais precoces. Mesmo que não atinja 100% de precisão, a combinação do software com a avaliação humana amplia a acurácia: em estudos, a cooperação entre dermatologista e IA tem mostrado melhorar o desempenho diagnóstico. Em outras palavras, o sistema oferece “insights precisos na triagem de pacientes”, aumentando a confiança do especialista e contribuindo para salvar vidas. Como afirma estudo do *JAMA Dermatology*, “os dermatologistas podem melhorar sua performance ao cooperarem com a rede neural”, o que reforça nosso propósito de fortalecer (e não substituir) o exame clínico.

Em síntese, a justificativa deste plano de ação une **dados epidemiológicos** (INC aumenta no país), **contexto socioeconômico** (difícil acesso e custos elevados), **fundamentação biológica** (carcinogênese e sinais clínicos) e **princípios éticos** (CEM). O resultado esperado é um software acessível e seguro, capaz de agilizar o caminho do paciente até o diagnóstico definitivo. Além de respeitar normas e o ato médico exclusivo, o projeto estabelece uma prática de triagem inovadora – tornando a prevenção e o diagnóstico mais democráticos, eficientes e alinhados à ética profissional.

3. OBJETIVOS

GERAL:

Desenvolver uma solução computacional (Software) baseada em inteligência artificial (IA) aplicada em um aplicativo móvel para a identificação e diagnóstico do câncer de pele a partir da análise automatizada de imagens dermatológicas, com o objetivo de identificar padrões visuais compatíveis com diferentes tipos de câncer de pele, visando auxiliar profissionais de saúde no diagnóstico precoce, fornecendo informações de apoio à tomada de decisão médica.

ESPECÍFICO:

- a. Identificar e classificar os atributos clínicos e visuais das lesões cutâneas que servem como indicadores discriminativos entre o câncer de pele e outras patologias dermatológicas benignas.
- b. Projetar e otimizar um modelo de inteligência artificial para a análise automatizada de imagens dermatológicas, com foco na maximização da acurácia na detecção de câncer de pele.
- c. Prover suporte a profissionais da saúde na detecção precoce do câncer de pele, por meio da utilização de ferramentas baseadas em inteligência artificial.
- d. Avaliar a eficácia do sistema desenvolvido na classificação de lesões cutâneas, utilizando indicadores de desempenho como acurácia, sensibilidade e especificidade para validar a confiabilidade do modelo.

4. REFERÊNCIAL TEÓRICO

4.1. Machine Learning e Deep Learning

4.1.1. *Inteligência Artificial*

A inteligência artificial (IA) é um dos segmentos mais dinâmicos e revolucionários da ciência e engenharia contemporâneas. Desde suas iniciais formulações, a Inteligência Artificial tem se direcionado à elaboração de sistemas computacionais e máquinas aptas a raciocinar, aprender e atuar de formas que, convencionalmente, requereriam a inteligência humana. A inteligência artificial iniciou em 1956, quando John McCarthy criou o termo, que se refere à ciência e engenharia de criar máquinas inteligentes.

A primeira visão era focada na simulação. No entanto, com o passar das décadas, o entendimento e a utilização da Inteligência Artificial se ampliaram consideravelmente. A inteligência artificial é entendida como um conjunto de tecnologias que se baseiam, sobretudo, em áreas como o aprendizado de máquina e o aprendizado profundo (Deep Learning). Essas tecnologias são empregadas em uma grande variedade de aplicações, tais como análise de dados, previsões, categorização de objetos, processamento de linguagem natural, sistemas de recomendação e recuperação inteligente de informações. Dito isso, o livro *Hands-On Machine Learning with Scikit-Learn, Keras, & TensorFlow*, de Aurélien Géron, menciona a inteligência artificial como “a ciência de ensinar as máquinas a pensarem e agirem como os humanos”.

4.1.2. *Machine Learning*

O Machine Learning (ML) é um ramo da inteligência artificial focado na criação de algoritmos e métodos que possibilitam aos computadores aprender a partir de dados e aprimorar seus desempenhos em tarefas específicas ao longo do tempo, sem a exigência de programação explícita. A ideia central é capacitar sistemas computacionais a aprenderem com dados e a melhorarem seu desempenho de forma autônoma. Se ajustando e evoluam constantemente, à medida que acumulam mais

“experiências” através do processamento de volumes de dados maiores e mais diversificados.

A definição de ML foi originalmente cunhada por Arthur Samuel em 1959 como “o campo de estudo que dá aos computadores a capacidade de aprender sem serem explicitamente programados”. Uma definição mais técnica, proposta por Tom Mitchell, descreve que “um software é dito aprender da experiência E com respeito a alguma classe de tarefas T e medida de desempenho P se seu desempenho nas tarefas em T, como medido por P, melhora com a experiência E”.

4.1.3. *Deep Learning*

O Aprendizado Profundo (Deep Learning — DL) é um subconjunto especializado do aprendizado de máquina que se destaca por sua utilização de redes neurais artificiais. A principal característica dessas redes é a presença de múltiplas camadas ocultas, que permitem ao modelo aprender e representar as características dos dados de forma hierárquica e automática.

A anatomia de uma rede profunda é composta por neurônios interconectados organizados em camadas: uma camada de entrada que recebe os dados, várias camadas ocultas que processam a informação e uma camada de saída que fornece o resultado.

A grande diferença do DL para o aprendizado de máquina tradicional reside na sua capacidade de extrair e classificar características de forma autônoma. Enquanto no aprendizado de máquina convencional os desenvolvedores precisam realizar a engenharia manual de características, uma tarefa muitas vezes demorada e complexa, os modelos de DL aprendem essas representações diretamente dos dados brutos. Isso torna o DL particularmente poderoso para trabalhar com dados não estruturados, como imagens e texto, ao automatizar uma etapa que era intensiva em trabalho humano.

4.1.4. *Inteligência Artificial na medicina*

Nos últimos anos, o progresso tecnológico e o início da nova revolução industrial do século XXI têm estimulado ferramentas que aprimoram a capacidade

humana, o que tem um impacto direto na melhoria da saúde e da qualidade de vida. Entre estas ferramentas, o ML e o DL têm se sobressaído no cenário médico, especialmente em dermatologia, ao oferecer métodos que identificam padrões complexos e causalidades em imagens da pele, automatizando processos antes dependentes da análise humana (Esteva et al., 2017). Apesar do potencial, a aplicação prática de modelos de IA na saúde exige testes clínicos rigorosos e validação científica para garantir segurança e eficácia, atendendo aos critérios éticos e regulatórios necessários para uso humanitário (Lloyd Minor, 2020).

A transição entre 2012 e 2025 representa uma mudança significativa na medicina, uma vez que a inteligência artificial deixou de ser uma ideia teórica para se tornar uma ferramenta clínica de transformação. De acordo com Lloyd Minor, reitor da Faculdade de Medicina da Universidade Stanford, essa mudança é semelhante à introdução dos antibióticos nos anos 1920, indicando uma alteração significativa no diagnóstico e tratamento de doenças. Três fatores convergiram para viabilizar isso: a explosão de dados médicos, o amadurecimento de algoritmos de aprendizado profundo e o aumento do poder computacional.

O crescimento exponencial dos dados médicos foi o primeiro alicerce dessa revolução. No passado, a coleta de informações médicas era limitada e fragmentada. Com a digitalização de hospitais, a adoção de prontuários eletrônicos e a expansão de dispositivos de monitoramento contínuo (wearables), gerou-se um volume imenso de dados clínicos, incluindo imagens de exames, registros laboratoriais e informações de monitoramento remoto. Esse acervo serviu como “combustível” para os modelos de ML e DL, permitindo que padrões complexos e sutis em imagens cutâneas, por exemplo, fossem detectados com precisão (Lundervold & Lundervold, 2019).

Simultaneamente, a evolução dos algoritmos de IA, principalmente as redes neurais convolucionais (CNNs), permitiu que os sistemas aprendessem autonomamente a reconhecer características em imagens médicas. Ao contrário dos algoritmos tradicionais, que dependiam de regras pré-definidas, as CNNs conseguem identificar padrões de lesões cutâneas, diferenciar nevos de melanoma e classificar tipos de câncer de pele a partir de grandes volumes de dados rotulados, melhorando continuamente com a incorporação de novos dados (Litjens et al., 2017; Esteva et al., 2017).

O terceiro componente crucial foi o avanço do hardware computacional. O treinamento de modelos profundos envolve milhares de camadas e milhões de parâmetros, exigindo alto poder de processamento. A disponibilidade de Unidades de Processamento Gráfico (GPUs), processadores especializados originalmente desenvolvidos para renderizar imagens e gráficos em computadores, como vídeos, jogos em 3D e animações. Eles modernos e clusters computacionais acessíveis tornou o treinamento de CNNs uma prática viável, acelerando pesquisas iniciadas em décadas anteriores e viabilizando sua aplicação em diagnósticos clínicos (LeCun, Bengio & Hinton, 2015). A interconexão desses três pilares — dados massivos, algoritmos sofisticados e poder computacional — resultou em uma transformação palpável na medicina, permitindo diagnósticos mais precisos, tratamentos individualizados e maior eficiência operacional.

A associação entre a tecnologia e a dermatologia permitiu a criação de métodos que permitem a análise de imagens da pele e a obtenção de diagnósticos seguros sobre lesões inofensivas ou perigosas (Esteva et al., 2017; Tschandl et al., 2020). A utilização da IA como ajuda extra não anula a opinião do médico, mas traz um apoio forte, possibilitando identificar rápido os casos, dar atenção primeiro aos mais graves e dar um padrão para analisar as lesões, o que ajuda o paciente a ficar bem.

4.1.5. Considerações sobre o estudo de Esteva

O estudo de Esteva et al. (2017), intitulado “Classificação de câncer de pele em nível de dermatologista com redes neurais profundas”, representa um marco significativo para a área médica. Mais do que uma simples demonstração, a pesquisa inovou ao confrontar uma IA com médicos, visando mostrar que um algoritmo poderia identificar câncer de pele tão bem quanto dermatologistas experientes.

Para alcançar tal feito utilizaram uma rede neural convolucional (CNN) treinada com uma vasta coleção de 129.450 imagens clínicas – uma quantidade de dados cem vezes maior que a usada em estudos prévios. O poder do modelo estava na estrutura pré-treinada de GoogLeNet Inception v3, que, por meio do aprendizado por transferência, teve sua última camada ajustada para classificar mais de 2.000 doenças de pele. Em avaliações rigorosas, o modelo encarou dois cenários clínicos reais, diagnosticando carcinomas e melanomas a partir de imagens confirmadas por biópsia.

Os resultados foram notáveis. A CNN se mostrou tão eficiente quanto os 21 especialistas que participaram dos testes, e até superou a média do grupo em certas situações. Mesmo que um único dermatologista tenha se destacado em alguns casos, a conclusão foi evidente: a inteligência artificial consegue alcançar um nível de habilidade comparável ao de profissionais humanos. Isso viabilizou que a tecnologia, disponível até em celulares, expandisse o diagnóstico de câncer de pele para além dos consultórios, ampliando seu alcance de forma inédita.

4.1.6. *Arquitetura e foco do projeto*

Para entender o impacto da inteligência artificial no campo visual da medicina, como na radiologia e dermatologia, é indispensável explorar a fundo a estrutura das redes neurais. Nesse contexto, dois modelos se destacam: as Redes Neurais Profundas (DNNs) e, principalmente, as Redes Neurais Convolucionais (CNNs). Ainda que o campo da inteligência artificial ofereça uma variedade de modelos e arquiteturas, a presente análise se concentra nas DNNs e nas CNNs. Essa abordagem visa aprofundar o entendimento de suas características fundamentais e demonstrar, claramente, a razão pela qual as CNNs são particularmente eficazes no reconhecimento de imagens.

As Redes Neurais Profundas (DNNs) são modelos versáteis que aprendem representações complexas dos dados por meio de diversas camadas ocultas. Sua adaptabilidade permite sua aplicação em diferentes áreas, desde finanças até o processamento de linguagem. Contudo, sua arquitetura genérica apresenta uma falha crucial em tarefas visuais: a incapacidade de compreender a estrutura espacial de uma imagem, tratando cada píxel como uma entrada independente. Isso representa um grande problema, pois a proximidade entre os píxeis – que forma uma borda, uma textura ou um formato – é desconsiderada.

Surgem as Redes Neurais Convolucionais (CNNs), um tipo específico de DNN para o universo das imagens. Sua inteligência reside em camadas que simulam a visão humana. As camadas convolucionais utilizam filtros para extrair características locais, como as linhas e texturas que compõem uma lesão de pele. Por outro lado, as camadas de pooling funcionam como um resumo visual, diminuindo a dimensão da imagem e concentrando-se somente nas informações mais relevantes. Essa estrutura

permite que as CNNs aprendam e reconheçam hierarquias visuais de forma automática.

A diferença entre as arquiteturas é um experimento simples: o que acontece se o modelo for treinado com imagens desorganizadas? Uma DNN genérica, que já ignora a estrutura espacial, pode manter um nível de precisão razoável. No entanto, a precisão de uma CNN cai drasticamente (para cerca de 10%), demonstrando que sua eficácia depende totalmente da hierarquia e do padrão visual que ela foi projetada para interpretar.

Dessa forma, o sucesso da inteligência artificial no diagnóstico por imagem não é fruto apenas da evolução tecnológica em si, mas sim da validação perfeita da CNN para essa função. O diagnóstico em radiologia, oftalmologia ou dermatologia é, essencialmente, um exercício de reconhecimento de padrões visuais. A habilidade das CNNs de extrair e identificar padrões sutis, que podem passar despercebidos aos olhos humanos, é a principal razão pela qual essa arquitetura se tornou a ferramenta preferida e impulsionou o progresso nessas áreas.

4.1.7. Inteligência Artificial evidenciado no diagnóstico de pele

Haenssle et al. (2018), no estudo “Man against Machine”, demonstram a superioridade de uma CNN de deep learning na identificação de melanomas dermoscópicos em comparação com 58 dermatologistas. Com uma sensibilidade de 87% e especificidade de 83%, o algoritmo superou a média dos especialistas humanos, evidenciando o potencial do ML para melhorar a acurácia diagnóstica em condições controladas. Este estudo valida ferramentas como o Moleanalyzer Pro, que utiliza tecnologia semelhante para análise de lesões melanocíticas e não melanocíticas, oferecendo recursos como pontuação de IA (AI Score) e mapas de calor (Heatmap) para interpretar resultados.

Haenssle et al. (2020), em “Man against Machine Reloaded”, avançam a pesquisa ao testar uma CNN aprovada para o mercado em condições menos artificiais, comparando-a com 96 dermatologistas. O estudo confirma a robustez do algoritmo em classificar uma ampla gama de lesões cutâneas, com desempenho comparável ou superior ao de especialistas, mesmo em cenários mais próximos da prática clínica real. Esses resultados reforçam a confiabilidade do ML em

dermatoscopia e destacam a importância de validações clínicas para a adoção de IA na medicina.

Com o surgimento do estopim de ia, o trabalho empregado pelo Dr. Holger Haenssle é uma figura central na aplicação da inteligência artificial (IA) na dermatologia. Em 2018, seu estudo, "Man against Machine", demonstrou o potencial das Redes Neurais Convolucionais (CNNs) na identificação de melanomas. O trabalho comparou o desempenho de um algoritmo de deep learning com o de 58 dermatologistas, e o resultado foi notável: a IA superou a maioria dos especialistas humanos, indicando a viabilidade de usar essa tecnologia como ferramenta de apoio diagnóstico.

O sucesso inicial do estudo foi um marco, incentivando o desenvolvimento de soluções comerciais como o Moleanalyzer Pro, um software que utiliza a mesma tecnologia. A relevância desse trabalho motivou uma segunda pesquisa em 2020, "Man against Machine Reloaded", que testou o algoritmo já aprovado para o mercado em condições mais próximas da realidade clínica, envolvendo 96 dermatologistas. A comparabilidade entre os dois estudos é essencial: enquanto o primeiro estabeleceu a prova de conceito, o segundo validou a robustez e a consistência do algoritmo em um cenário clínico mais amplo.

Para avaliar a performance do algoritmo em seu estudo de 2018, Haenssle e sua equipe utilizaram métricas de desempenho específicas, essenciais para validar a eficácia de sistemas de diagnóstico por imagem. A análise se baseia em uma matriz de confusão, que permite a mensuração da Sensibilidade (Recall), que mede a capacidade do modelo de identificar corretamente os casos positivos (melanomas).

No estudo, o algoritmo alcançou uma sensibilidade de 87%, indicando que ele identificou corretamente a grande maioria dos melanomas, o que é crucial para evitar diagnósticos falsos negativos. Além disso, a Especificidade do algoritmo foi avaliada em 83%, avaliando sua capacidade de classificar corretamente os casos negativos (lesões benignas), o que é importante para evitar biópsias desnecessárias. Por fim, a Curva ROC (Receiver Operating Characteristic) foi utilizada para ilustrar o desempenho do modelo em diferentes limiares, e a Área sob a Curva (AUC), que quantifica a área total sob a curva, forneceu um valor único que resumiu o desempenho geral. No estudo, a AUC do algoritmo de deep learning foi de 0,91, valor

superior ao da média dos dermatologistas, que foi de 0,79, o que indica um excelente desempenho.

Apesar do sucesso de tecnologias como o Moleanalyzer Pro, sua implementação no Brasil enfrenta desafios significativos. Conforme relatório da Associação Médica Brasileira (AMB) e do Conselho Federal de Medicina (CFM), o alto custo de aquisição e licenciamento desses softwares e equipamentos limita sua adoção em sistemas de saúde pública, como o Sistema Único de Saúde (SUS). Dados do Ministério da Saúde de 2023 mostram que a despesa per capita com tecnologia e insumos de saúde no Brasil é de aproximadamente US\$ 500, enquanto em países como os Estados Unidos e Alemanha esse valor ultrapassa US\$ 5.000. Essa disparidade orçamentária cria uma barreira de acesso, impedindo que a população brasileira se beneficie amplamente desses avanços. Diante disso, o cenário nacional tem incentivado a pesquisa para o desenvolvimento de soluções de IA mais acessíveis e adaptadas. Um exemplo é o Grupo de Pesquisa em Visão Computacional e Inteligência Artificial da Universidade Federal de Minas Gerais (UFMG), que em 2022 publicou um estudo propondo um modelo de baixo custo para detecção de retinopatia diabética utilizando câmeras de smartphones, demonstrando a viabilidade de criar ferramentas que podem ser integradas à infraestrutura de saúde já existente.

4.2. Pele

4.2.1. *Anatomia e Fisiologia da Pele*

A pele é o maior órgão do corpo humano, funcionando como uma interface vital e dinâmica entre o corpo e o ambiente externo (Universidade Federal de Alfenas - UNIFAL, 2023). Ela vai além de uma mera proteção, sendo um sistema de defesa robusto e complexo que resguarda contra uma vasta variedade de ameaças. A pele atua como uma barreira protetora contra agentes químicos, impedindo que substâncias tóxicas penetrem no corpo. Sua resistência física defende contra danos mecânicos, como cortes e pressões, além de oferecer um escudo contra a radiação ultravioleta (UV) e as variações de temperatura. No nível biológico, o pH ácido da pele e sua microbiota residente formam uma linha de frente que inibe a proliferação de bactérias e fungos prejudiciais (Sociedade Brasileira de Dermatologia - SBD).

Para realizar essas funções, a pele possui uma estrutura intrincada. A camada mais externa é a principal responsável pela impermeabilidade e pela proteção direta, constantemente se renovando para manter a sua eficácia. Abaixo dela, uma camada mais profunda oferece o suporte estrutural, rica em proteínas como colágeno e elastina, que dão à pele sua resistência e flexibilidade. É nessa camada que se localiza uma densa rede de nervos e vasos sanguíneos, essencial para as sensações e para a nutrição do tecido. A camada mais profunda serve como isolante térmico e amortecedor de impactos, protegendo os órgãos internos.

Além de sua função de barreira, a pele desempenha papéis fisiológicos vitais. Ela é fundamental na termorregulação corporal, ajustando o fluxo sanguíneo e produzindo suor para manter a temperatura ideal. Sua rede de nervos é crucial para a percepção sensorial, permitindo a detecção de estímulos como o toque, a dor e a temperatura. As células especializadas da pele também funcionam como sentinelas do sistema imunológico (SI), ativando as defesas contra agentes patogênicos. Por fim, a exposição solar permite que ela realize a síntese de vitamina D (Vit. D), um processo essencial para a saúde óssea e o bem-estar geral do corpo (Proksch et al., 2008).

4.3. Câncer de pele

4.3.1. *Epidemiologia*

O câncer de pele, em sua totalidade, é uma das neoplasias de maior incidência global. As estatísticas epidemiológicas revelam um panorama multifacetado da doença, que pode ser percebido como um paradoxo. Por um lado, mais de 90% dos casos de câncer de pele são passíveis de cura, especialmente quando diagnosticados em seus estágios iniciais, o que sublinha a eficácia das intervenções precoces (Lusiadas Saúde). Por outro, a Fundação do Câncer de Pele e o Fórum Europeu de Dermatologia (EDF) ressaltam uma estatística crítica: o melanoma cutâneo, embora menos prevalente do que os outros tipos, é responsável por 90% da mortalidade associada a todos os cânceres de pele (Fórum Europeu de Dermatologia, 2022). Essa dicotomia de alta curabilidade geral versus alta letalidade de um subtipo específico demonstra a gravidade da doença e a imperatividade de um foco estratégico em sua detecção e manejo. A alta prevalência do câncer de pele não melanoma (CPNM), que

corresponde a cerca de 30% de todos os tumores malignos registrados anualmente no Brasil, tende a ofuscar a ameaça desproporcional do melanoma (Sociedade Brasileira de Cirurgia Oncológica). A compreensão dessa dinâmica é crucial para orientar as estratégias de saúde pública, as pesquisas clínicas e as recomendações de rastreamento, garantindo que o foco não se restrinja apenas à quantidade de casos, mas, principalmente, à letalidade do subtipo mais agressivo. Este referencial teórico propõe uma análise aprofundada da dermatologia e da oncologia cutânea, partindo de seus fundamentos biológicos para contextualizar os mecanismos de risco, diagnóstico e rastreamento.

4.3.2. Tipos de Câncer de Pele

O câncer de pele é classificado em dois grupos principais: não melanoma e melanoma cutâneo, categorias que diferem amplamente em termos de prevalência, comportamento clínico e prognóstico. O câncer de pele não melanoma, que abrange predominantemente o carcinoma basocelular (CBC) e o carcinoma espinocelular (CEC), constitui a maioria absoluta dos casos diagnosticados em todo o mundo. Apesar de apresentar evolução clínica geralmente mais lenta e baixo risco de metástase, esse grupo pode gerar complicações funcionais e estéticas significativas quando não tratado adequadamente, especialmente em regiões expostas ao sol, como face e couro cabeludo.

Em contrapartida, o melanoma cutâneo, embora represente uma menor fração da incidência total, possui maior relevância do ponto de vista oncológico devido ao seu alto potencial metastático e elevada taxa de mortalidade. Esse tipo de câncer tem origem nos melanócitos, células responsáveis pela produção de melanina, e apresenta comportamento mais agressivo, com rápida capacidade de invasão e disseminação para outros órgãos. Por essa razão, mesmo diante de uma prevalência inferior à do grupo não melanoma, o melanoma é considerado o tipo mais grave da doença, exigindo diagnóstico precoce e abordagem terapêutica rápida para reduzir riscos de morbimortalidade (Sociedade Brasileira de Cirurgia Oncológica).

4.3.3. Câncer de pele não melanoma (CPNM)

O CPNM é o tipo mais comum de câncer em todo o mundo, respondendo por aproximadamente 30% dos tumores malignos no Brasil anualmente (Sociedade Brasileira de Cirurgia Oncológica). Sua alta taxa de cura, especialmente quando detectado e tratado em fases iniciais, o torna o menos nocivo, embora possa causar deformações se não for tratado (Sociedade Brasileira de Cirurgia Oncológica). Os dois subtipos mais frequentes são o carcinoma basocelular (CBC) e o carcinoma espinocelular (CEC).

O carcinoma basocelular (CBC) é o tipo de tumor de pele mais comum, representando cerca de 70-80% dos CPNMs (Carvalho et al., 2021). Ele se origina das células basais da epiderme (Mayo Clinic) e tem um crescimento lento, raramente metastizando para outras partes do corpo (Carvalho et al., 2021). No entanto, se não for tratado, pode ser localmente agressivo e destruir tecidos adjacentes, incluindo cartilagem e osso (Carvalho et al., 2021). Clinicamente, o CBC pode apresentar-se em várias formas, como uma protuberância brilhante translúcida, uma lesão plana e escamosa, ou uma área cerosa semelhante a uma cicatriz (Mayo Clinic). É mais comum em pessoas com pele clara e em áreas do corpo cronicamente expostas ao sol, com uma predominância na cabeça e no pescoço (Carvalho et al., 2021).

O carcinoma espinocelular (CEC) é o segundo tipo mais comum de CPNM e se origina na camada espinhosa da pele (Sociedade Brasileira de Cirurgia Oncológica). Da mesma forma que o CBC, sua ocorrência está fortemente associada à exposição solar crônica, e ele pode se desenvolver a partir de lesões pré-cancerosas como a queratose actínica (Lusíadas Saúde).

4.3.4. Melanoma cutâneo

O melanoma cutâneo (CM) é a forma mais perigosa de câncer de pele, originando-se dos melanócitos, as células produtoras de pigmento (Sociedade Brasileira de Cirurgia Oncológica). Apesar de representar uma pequena fração de todos os casos de câncer de pele, sua agressividade e capacidade de metástase o tornam responsável por 90% da mortalidade por câncer de pele (Fórum Europeu de Dermatologia, 2022). Se detectado precocemente, o melanoma é altamente curável

(Lusíadas Saúde). O diagnóstico do melanoma requer não apenas a avaliação clínica, mas também a confirmação por meio de dermatoscopia e, em caso de suspeita, por exame histopatológico (European Interdisciplinary Guideline on Melanoma, 2022). O estadiamento da doença é feito utilizando a classificação do American Joint Committee on Cancer (AJCC) (European Interdisciplinary Guideline on Melanoma, 2022), com a biópsia do linfonodo sentinela recomendada para tumores com espessura igual ou superior a 0,8 mm a 1,0 mm (Garbe et al., 2019).

5. MATERIAIS E MÉTODOS

5.1 Materiais

Os materiais e as ferramentas empregados neste projeto são detalhados de maneira seletiva, com foco naqueles que exercem um papel central na compreensão das etapas metodológicas. Recursos de caráter genérico, amplamente disponíveis em diferentes sistemas, não são descritos em profundidade, uma vez que sua utilização é autoexplicativa e não acrescentaria ganhos significativos à análise, evitando sobrecarregar o leitor com informações redundantes.

A seleção das ferramentas para a concepção e a execução do projeto foi fundamentada em uma análise prévia de suas funções e aplicações específicas, alinhada com critérios de viabilidade técnica e eficiência operacional. Tal escolha não constitui uma limitação metodológica, mas sim um direcionamento estratégico. A arquitetura do projeto foi desenvolvida com o princípio da flexibilidade, assegurando sua capacidade de interoperabilidade e escalabilidade. Isso permite que o sistema possa ser adaptado para novos contextos de aplicação ou integrar tecnologias emergentes, garantindo sua evolução contínua.

5.1.1 Notion – Planejamento

O Notion é ferramenta de gerenciamento e organização do projeto, como uma plataforma central para a gestão e organização do projeto. Nós o utilizamos para a estruturação hierárquica das informações, desde a documentação inicial do escopo até o acompanhamento do progresso diário. Por meio dele, criamos e mantemos tabelas de progresso, listas de tarefas e bases de conhecimento que integram todas as etapas do desenvolvimento. A escolha do Notion se justifica por sua capacidade de ser um ambiente unificado, que permite a personalização completa de espaços de trabalho para cada membro da equipe. Essa característica promove uma colaboração transparente e em tempo real, fornecendo um controle abrangente sobre o status e a evolução do projeto.

5.1.2 Supabase - Banco de dados

Para a gestão de dados, optamos pelo Supabase, uma solução de Backend-as-a-Service (BaaS) baseada em PostgreSQL. Ele serve como nosso principal banco de dados, centralizando o armazenamento e o gerenciamento de informações. O Supabase Storage é particularmente crucial para o nosso sistema, pois criamos buckets dedicados para armazenar as imagens de lesões de pele. Essa funcionalidade facilita a geração de URLs públicas seguras para cada imagem, permitindo que o nosso modelo de inteligência artificial as acesse para análise. Além disso, a biblioteca oficial supabase-py é fundamental para a integração com nosso código em Python, possibilitando a execução de operações como o gerenciamento de buckets, autenticação de usuários e a transferência de dados e imagens diretamente de notebooks no Google Colab ou scripts locais.

5.1.3 Linguagens e I5.1.3.1 Python, TensorFlow e Keras

A espinha dorsal do nosso sistema de inteligência artificial é a linguagem Python, escolhida por sua simplicidade, legibilidade e, principalmente, pelo vasto ecossistema de bibliotecas para Machine Learning. Dentro deste ecossistema, o TensorFlow é a nossa biblioteca principal para o desenvolvimento de modelos. Nós a empregamos para a construção de Redes Neurais Convolucionais (CNNs), que são a base do nosso sistema de detecção de lesões. O TensorFlow é a nossa ferramenta de baixo nível, que nos permite ter controle granular sobre o treinamento e a configuração dos modelos.

Para otimizar e acelerar esse processo, utilizamos o Keras como uma interface de alto nível sobre o TensorFlow. O Keras simplifica a criação, o treinamento e a avaliação das redes neurais, tornando a prototipagem de modelos mais acessível e menos dependente de conhecimento matemático complexo. Essa combinação nos permite aplicar modelos pré-treinados (presets) e desenvolver a inteligência artificial de forma eficiente e intuitiva.

5.1.3.2 Flutter

Para a interface de usuário, adotamos o Flutter, um framework de desenvolvimento de aplicativos multiplataforma. O Flutter é a nossa escolha para a criação de uma interface intuitiva e responsiva para dispositivos móveis, tanto para Android quanto para iOS. Por meio dele, desenvolvemos a aplicação que permite a categorização e a visualização das imagens processadas pelo nosso modelo de inteligência artificial. A performance nativa do Flutter e sua compatibilidade com diferentes plataformas são essenciais para disponibilizar o sistema para um público amplo.

5.1.4 Manipulação de dados

5.1.4.1 Pandas e NumPy

A biblioteca Pandas é uma ferramenta fundamental para a análise e manipulação de dados estruturados. Sua principal aplicação está no tratamento de tabelas, planilhas e bases de dados em formato tabular, permitindo operações como leitura de arquivos em diferentes formatos (CSV, Excel, JSON, SQL), limpeza de inconsistências, transformação de variáveis, junção de conjuntos de dados e organização de informações. É amplamente empregada em projetos que envolvem grandes quantidades de dados, possibilitando uma preparação consistente e eficiente para etapas posteriores, como análise estatística, modelagem e treinamento de algoritmos de aprendizado de máquina.

O NumPy complementa o Pandas ao oferecer suporte otimizado para matrizes (arrays) multidimensionais e operações matemáticas de baixo nível. Ele é especialmente relevante em aplicações que demandam cálculos numéricos de alta performance, como álgebra linear, transformadas de Fourier, geração de números aleatórios e manipulação de imagens em formato matricial. Por meio de sua arquitetura eficiente, baseada em vetorização, o NumPy reduz o tempo de processamento e aumenta a escalabilidade em análises de grandes volumes de dados, sendo considerado a base para diversas bibliotecas de ciência de dados e aprendizado de máquina.

5.1.4.2 Requests e OS

A biblioteca Requests é destinada à comunicação por meio do protocolo HTTP, permitindo enviar e receber dados de servidores externos de forma simplificada. Suas principais funcionalidades incluem a realização de requisições do tipo GET, POST, PUT e DELETE, manipulação de cabeçalhos, gerenciamento de sessões e autenticações. É amplamente utilizada em integrações com APIs, no consumo de dados da web e no download automatizado de arquivos, sendo essencial para sistemas que dependem de informações externas em tempo real.

O módulo OS integra o Python ao sistema operacional, permitindo a manipulação direta de diretórios, arquivos e variáveis de ambiente. Essa ferramenta é responsável pelo controle de caminhos, criação e exclusão de pastas, organização automática de dados e execução de comandos específicos do sistema. Sua aplicação é relevante em rotinas de automação, processamento em lote e no gerenciamento de pipelines de dados, assegurando que a manipulação de arquivos seja realizada de forma controlada e reprodutível.

5.1.4.3 PIL (Pillow)

A biblioteca PIL (Pillow) é uma das ferramentas mais utilizadas para processamento e manipulação de imagens em Python. Entre suas funcionalidades destacam-se o redimensionamento, corte, rotação, aplicação de filtros, ajuste de cores e conversão entre diferentes formatos (JPEG, PNG, BMP, entre outros). A padronização de imagens por meio do Pillow é essencial para etapas de pré-processamento em sistemas de visão computacional, pois garante que todas as entradas tenham dimensões, formatos e características consistentes. Essa consistência facilita o trabalho de modelos de inteligência artificial e melhora a confiabilidade dos resultados obtidos.

5.1.4.4 OpenCV (ou cv2)

A biblioteca OpenCV é uma das mais completas soluções para processamento de imagens e visão computacional. Seu conjunto de algoritmos e funções permite desde

tarefas simples, como conversão de espaços de cor e redimensionamento, até operações avançadas, como detecção de bordas, reconhecimento de padrões e segmentação de regiões específicas.

As principais aplicações incluem:

Redução de ruídos por meio de filtros como Median, Gaussian e Bilateral, que suavizam imperfeições preservando informações importantes.

Realce de contraste e detalhes, utilizando técnicas como sharpen e CLAHE, que aumentam a visibilidade de estruturas de interesse em imagens complexas.

Segmentação de regiões com métodos clássicos, como limiarização adaptativa (Thresholding), algoritmo de Otsu e Watershed, úteis para separar objetos sobrepostos ou áreas de interesse.

Detecção de bordas com algoritmos como Canny, Sobel e Laplaciano, essenciais para identificar contornos e formas relevantes.

O OpenCV se destaca por sua eficiência computacional e por ser amplamente utilizado em aplicações médicas, industriais, de segurança e pesquisa acadêmica.

5.1.4.5 MONAI (*Medical Open Network for AI*)

O MONAI é um framework especializado no desenvolvimento de soluções de inteligência artificial aplicadas à área da saúde, com foco em imagens médicas. Ele fornece ferramentas otimizadas para treinamento de modelos de segmentação, classificação e detecção, além de rotinas de pré-processamento adaptadas às particularidades desse tipo de dado.

Entre suas funcionalidades mais relevantes destacam-se:

- Implementação de arquiteturas avançadas, como U-Net e Attention U-Net, altamente eficazes em segmentação biomédica.
- Transformações específicas para imagens médicas, incluindo rotações, espelhamentos, normalizações e distorções elásticas, que aumentam a robustez do modelo e reduzem o risco de sobreajuste.
- Integração com PyTorch, permitindo o uso de recursos modernos de aprendizado profundo com bibliotecas de alto desempenho.

O MONAI se consolidou como uma plataforma padrão em projetos de pesquisa médica por combinar eficiência, modularidade e práticas recomendadas para o desenvolvimento de modelos clínicos robustos.

5.1.4.6 *imblearn (imbalanced-learn)*

A biblioteca *imblearn* foi desenvolvida para solucionar o problema de desbalanceamento de classes em conjuntos de dados, um cenário comum em aplicações médicas e biomédicas. O desbalanceamento ocorre quando a quantidade de exemplos em uma classe é muito superior à de outra, levando o modelo a apresentar vieses e dificuldades na detecção de casos raros.

Entre suas principais técnicas encontram-se:

- Oversampling, que aumenta artificialmente a quantidade de exemplos da classe minoritária.
- Undersampling, que reduz a quantidade de exemplos da classe majoritária.
- SMOTE (Synthetic Minority Oversampling Technique), considerado o método mais relevante, pois cria exemplos sintéticos da classe minoritária baseados em vizinhos próximos, evitando duplicação de dados e aumentando a diversidade do conjunto de treino.

O uso do *imblearn* é determinante para assegurar que o modelo apresente desempenho equilibrado em todas as classes, ampliando sua aplicabilidade em cenários reais.

5.1.4.7 *Scikit-Image*

A *Scikit-Image* é uma biblioteca especializada em processamento de imagens que oferece recursos complementares ao OpenCV, com ênfase em métodos matemáticos e análises quantitativas. Suas principais aplicações incluem:

- Remoção de ruídos por meio de filtros sofisticados, como o Kuwahara e os filtros anisotrópicos, adequados para áreas com alta complexidade de textura.
- Análise de textura, utilizando ferramentas como os filtros de Gabor, o LBP (Local Binary Patterns) e a extração de características estatísticas, como Haralick e

GLRLM. Esses métodos permitem quantificar padrões e microestruturas das imagens.

- Cálculo de propriedades morfológicas, como área, perímetro, circularidade e excentricidade, que fornecem informações relevantes para a caracterização geométrica de objetos e lesões.

A Scikit-Image é amplamente empregada em pesquisas que exigem quantificação precisa e análise detalhada das imagens, reforçando o rigor científico das investigações.

5.1.4.8 UMAP (Uniform Manifold Approximation and Projection)

O UMAP é uma técnica de redução de dimensionalidade não-linear amplamente utilizada para análise exploratória de dados. Seu objetivo é projetar vetores de características de alta dimensão em espaços de duas ou três dimensões, mantendo a estrutura topológica dos dados originais.

Essa ferramenta é aplicada em diversas etapas de análise, como:

- Visualização de embeddings gerados por redes neurais convolucionais ou métodos de extração de textura.
- Identificação de agrupamentos naturais de amostras, permitindo observar a proximidade entre classes ou subgrupos.
- Avaliação da qualidade das características extraídas, verificando se o modelo consegue separar de forma clara as diferentes categorias de interesse.

O UMAP combina rapidez de execução, preservação da estrutura global e local dos dados e grande aplicabilidade em cenários de alto volume dimensional, tornando-se indispensável em estudos que envolvem imagens médicas complexas.

5.1.5 Logging

A biblioteca Logging configura-se como um recurso indispensável para o registro sistemático, o monitoramento e a depuração de sistemas em execução. Por meio dela, eventos relevantes, mensagens informativas, avisos de atenção e falhas críticas podem ser documentados com precisão, formando um histórico detalhado da execução do programa. Esse registro contínuo permite a detecção precoce de erros,

a análise minuciosa de comportamentos inesperados e a medição do desempenho em diferentes etapas do pipeline computacional. O Logging possibilita ainda a utilização de níveis de severidade organizados em categorias hierárquicas — DEBUG, INFO, WARNING, ERROR e CRITICAL — que fornecem granularidade ao processo de análise e filtragem de eventos. A flexibilidade de configuração garante que os registros possam ser armazenados localmente em arquivos, exibidos diretamente no console ou integrados a plataformas externas de monitoramento, viabilizando tanto aplicações simples quanto sistemas distribuídos em larga escala. A adoção dessa prática assegura a robustez e a confiabilidade do software, constituindo-se como um elemento metodológico essencial para a manutenção e para a reprodutibilidade em ambientes de pesquisa e desenvolvimento tecnológico.

5.1.5.1 Monitoramento

5.1.5.2 Weights & Biases (W&B)

A plataforma Weights & Biases (W&B) representa uma solução consolidada para o rastreamento, a análise e a colaboração em experimentos de aprendizado de máquina. Seu papel é central na organização e interpretação de resultados, uma vez que disponibiliza ferramentas de monitoramento de métricas em tempo real, como acurácia, perda, precisão, revocação e AUC, fornecendo um panorama detalhado do desempenho dos modelos durante o treinamento. Esse monitoramento contínuo possibilita a comparação entre arquiteturas, a análise de diferentes configurações de hiperparâmetros e a identificação de possíveis pontos de convergência ou sobreajuste. Outro recurso relevante oferecido pela plataforma é a visualização de embeddings de alta dimensionalidade, que permite compreender como as representações internas de redes neurais agrupam e distinguem amostras, fornecendo subsídios para avaliar a qualidade da separação entre classes. Além disso, o W&B oferece um sistema de controle de versão para experimentos, conjuntos de dados e modelos, assegurando rastreabilidade e reprodutibilidade científica. Essa integração entre rastreamento, visualização e versionamento consolida o W&B como uma ferramenta estratégica para a pesquisa em aprendizado profundo.

5.1.6 Ambientes de Edição

O desenvolvimento de projetos de ciência de dados e inteligência artificial é amplamente apoiado por ambientes de edição e colaboração que oferecem infraestrutura flexível, escalável e integrada ao fluxo de trabalho dos pesquisadores. O **Google Colab** destaca-se como uma plataforma em nuvem que viabiliza a execução de código Python em máquinas virtuais equipadas com GPUs e TPUs, possibilitando treinamento de modelos complexos sem a necessidade de hardware local de alto custo. Além disso, sua interface baseada em notebooks favorece a prototipagem rápida, a visualização de dados e o compartilhamento de experimentos de forma colaborativa. O **Visual Studio Code**, por sua vez, consolida-se como o editor de código principal, oferecendo suporte a múltiplas linguagens, integração com sistemas de versionamento, depuração avançada e um ecossistema de extensões que aprimoram a organização do código e a produtividade no desenvolvimento. Complementando esse conjunto, o **Firebase** atua como uma plataforma de integração e gerenciamento de aplicações em nuvem, permitindo o armazenamento de dados, o controle de frameworks, a sincronização de recursos e a otimização de processos colaborativos. A articulação desses ambientes garante um ciclo de desenvolvimento mais eficiente, reprodutível e adequado às demandas de pesquisas computacionais de alta complexidade.

5.2 Métodos

5.2.1 Proposta de Pesquisa

O presente trabalho consiste na disponibilização de uma Inteligência Artificial baseada em Deep Learning para auxiliar na detecção precoce de câncer de pele a partir de imagens dermatológicas. O objetivo é aprimorar a triagem no tratamento, uma vez que a alta incidência de câncer de pele e as dificuldades no diagnóstico precoce representam um desafio significativo. Quando o diagnóstico é feito tardiamente, podem ocorrer complicações graves, já que o aumento de tumores em estágios avançados torna o tratamento mais difícil.

A questão central que orienta este estudo é como desenvolver e treinar uma IA capaz de analisar imagens de lesões de pele e distinguir, com alta precisão, entre

lesões benignas e malignas. Para responder a essa questão, são adotados dois caminhos principais. Primeiro, uma revisão bibliográfica que sintetiza pesquisas existentes sobre conceitos fundamentais relacionados ao câncer de pele, algoritmos de Machine Learning e áreas médicas correlatas. Em seguida, o método de engenharia foca na elaboração de um modelo de Deep Learning, especificamente Redes Neurais Convolucionais (CNNs). Essas redes são escolhidas por sua capacidade de interpretar dados visuais complexos e aumentar a eficiência no reconhecimento e classificação de imagens. A abordagem combina fundamentos teóricos e práticos, buscando resultados concretos por meio do treinamento da máquina. O objetivo é oferecer uma solução para um problema real, com potencial para ser avaliada por médicos e possivelmente implementada em clínicas.

5.2.2 Considerações Éticas e Metodologia

A prioriza-se de uma atuação em questões éticas por parte de garantir a privacidade das imagens e dos dados dos pacientes, que terar seus sigilosos para casos de da testagem da inteligência artificial, evitando impactos diretos e minimizar vieses não aceitos pelo outro lado. À sua elaboração constituiu até o instante de natureza de pesquisa e aplicada, buscando uma solução prática que possa ser avaliada e, potencialmente, utilizada em ambientes clínicos. Sua viabilidade depende de sua utilidade, na prática, especialmente por tratar de um problema real.

Para reforçar a capacidade de detecção e análise dos registros de lesões cutâneas, o estudo propõe uma abordagem mista para a coleta de dados, utilizou-se de tanto fontes secundárias quanto a busca por uma questão incorporar em âmbito primárias, demonstrado incansável por questões de não sermos especialista na área médica e as preocupações de aumentar as tensões de uma pessoa com problemas maiores que se acumularia e aumentaria sua sanidade. Primeiramente, são utilizadas fontes secundárias, como repositórios clínicos e acadêmicos, que contêm registros de lesões dermatológicas já classificadas. Esses bancos de dados fornecem informações importantes sobre características das lesões, como grau de malignidade e padrões morfológicos, possibilitando o desenvolvimento e a validação de modelos de CNNs.

Adicionalmente, se for viável, é possível coletar dados de forma primária, captando imagens em ambiente de laboratório. A captura deve ser realizada com

câmeras digitais ou microscópios adaptados para garantir a resolução, iluminação e fidelidade visual necessárias, complementando os bancos de dados existentes com fontes locais para maior representatividade. A integração dos dados secundários com dados primários tem como objetivo melhorar a precisão dos modelos, possibilitando uma análise mais minuciosa das variações morfológicas entre os distintos tipos de lesões.

5.2.3 Desenho do Estudo

O desenho deste estudo delineado-se como uma pesquisa de validação de modelo retrospectiva, utilizando um método de coleta e análise misto. Esta abordagem combina dados não numéricos, como imagens de lesões de pele já coletadas e anotadas por especialistas, com dados numéricos, como métricas de desempenho (precisão, sensibilidade e especificidade). A escolha por uma pesquisa retrospectiva é pragmática e eficiente, ao permitir o acesso a um vasto volume de dados curados, o que seria financeira e logisticamente desafiador em um estudo prospectivo.

A metodologia se concentra na aplicação e validação de técnicas de aprendizado profundo. Em uma arquitetura utilizando a das CNNs, selecionadas por sua comprovada capacidade de extração de características visuais e por seus resultados consistentes e já consolidados no mercado. Para fins de comparação, a arquitetura do Vision Transformer (ViT) comparando as CNNs como meio de apoio da própria.

A escolha do tema e das ferramentas reflete a necessidade de integrar a tecnologia ao contexto médico atual, diga-se de passagem no Brasil, visando agilizar os processos de diagnóstico e atender ao aumento da demanda por serviços de saúde. Para técnicos e especialistas, esta abordagem representa entre todas as forma de criar uma solução escalável e precisa.

Apesar da solidez da abordagem retrospectiva, ela impõe limitações que serão abordadas. A variabilidade inerente às imagens do mundo real, como iluminação inconsistente, artefatos de dispositivos e a diversidade de tons de pele, pode não estar totalmente representada nos conjuntos de dados públicos. Para garantir que o modelo não apenas demonstre alta acurácia em um ambiente controlado, mas também tenha potencial para um desempenho robusto em ambientes clínicos complexos, a

metodologia será ajustada. Isso inclui o pré-processamento de dados e técnicas de aumento de dados para introduzir variabilidade no conjunto de treinamento. Além disso, a utilização de um conjunto de dados geograficamente e demograficamente distinto para a validação final (como o PAD-UFES) servirá como um teste de generalização rigoroso.

A eficácia do modelo será ainda mais validada por meio de testes experimentais controlados, que compararão seus resultados com os diagnósticos de dermatologistas especialistas, garantindo a validação clínica de sua eficácia e velocidade.

5.2.4 Contextualização e Justificativa

O câncer de pele, em suas diversas formas, representa uma das mais altas incidências de neoplasias malignas em todo o mundo. A detecção precoce de lesões suspeitas, como o melanoma, é fundamental para o sucesso do tratamento e para a redução da mortalidade. No entanto, o diagnóstico preciso de lesões cutâneas é um desafio complexo que exige a especialização de dermatologistas, um recurso escasso em muitas regiões e sistemas de saúde a alta demanda de concentração de pacientes e de falsos possíveis casos de cancer de pele . A sobrecarga de trabalho e a variação na experiência clínica entre diferentes profissionais podem levar a diagnósticos tardios ou, inversamente, a biópsias desnecessárias de lesões benignas. A capacidade de uma ferramenta de triagem automatizada, alimentada por inteligência artificial (IA), para atuar como um "segundo par de olhos" para médicos generalistas ou para auxiliar no fluxo de trabalho clínico em ambientes remotos, representa uma oportunidade significativa para aprimorar a saúde pública.

A IA, especificamente por meio de modelos de aprendizagem profunda, oferece uma solução promissora para este desafio. Ao analisar e classificar imagens dermatoscópicas, um modelo computacional pode identificar padrões sutis que podem ser imperceptíveis ao olho humano ou que requerem um alto nível de treinamento para reconhecimento. A proposta desta pesquisa não é substituir a avaliação clínica de um especialista, mas sim desenvolver um sistema de apoio à decisão que aumente a eficiência diagnóstica, priorizar casos de maior risco e, em última análise, melhorar os resultados para os pacientes. A medida de sucesso para esta metodologia não se restringe à acurácia técnica, mas se estende à sua utilidade clínica, confiabilidade e

aceitação pelos profissionais de saúde. Ao enquadrar o modelo como uma ferramenta de auxílio, o projeto se alinha com as melhores práticas de integração tecnológica em ambientes de alta criticidade, garantindo que a segurança do paciente seja a consideração primordial em cada etapa do desenvolvimento e validação.

5.2.5 Coleta e Curadoria de Dados

5.2.5.1 Aquisição de Dados

A metodologia de pesquisa adota uma estratégia de fusão de dados, combinando múltiplos conjuntos de dados públicos para maximizar a diversidade e o volume de imagens, separando-os, organizando-os, normalizando-os e gerando um link público da imagem para a facilitação do treinamento. Esta abordagem é essencial para treinar um modelo de aprendizado profundo que seja generalizável e robusto a uma variedade de condições de captura e populações. Os seguintes conjuntos de dados serão utilizados:

- **ISIC Archive:** Este é o maior e mais abrangente arquivo de imagens de câncer de pele, atuando como a principal fonte de dados para o treinamento do modelo base. Sua vasta escala e diversidade de imagens fornecem uma base sólida para que o modelo aprenda as características gerais de lesões cutâneas.
- **HAM10000:** Este conjunto de dados, contendo 10.015 imagens de lesões dermatoscópicas, é um dos benchmarks mais utilizados na literatura. Será crucial para o aprendizado por transferência e para a validação comparativa do desempenho, permitindo que a pesquisa se posicione em relação a trabalhos anteriores.
- **PAD-UFES-20:** Um conjunto de dados de 2.298 imagens de lesões dermatoscópicas e clínicas coletadas em um hospital universitário brasileiro, representando uma população e um contexto geográfico específicos. Este dataset é de valor inestimável para testar a capacidade de generalização do modelo para diferentes tons de pele, padrões de lesões e condições ambientais, mitigando o viés potencial de conjuntos de dados predominantemente ocidentais. A inclusão deste conjunto de dados reflete o reconhecimento de que a distribuição de classes, tons de pele e artefatos pode ser significativamente diferente em contextos geográficos distintos.

Uma vez que as imagens foram capturadas usando uma variedade de dispositivos, incluindo dermatoscópios, câmeras digitais e smartphones, a metodologia deve ser projetada para lidar com a variabilidade inerente a essas condições.

A fusão desses datasets não é um processo trivial e exige uma estratégia cuidadosa para padronizar os rótulos e as características de imagem, descartando dados com imagens ou dados que não estão devidamente rotuladas ou não há uma imagem correspondente ao seu metadata. O conjunto de dados final para validação se constrói com imagens provenientes de uma fonte independente (ISIC), disponibilizando datasets especializados para validação, para medir a generalização imparcialmente. A Tabela 1 resume os conjuntos de dados que serão utilizados nesta pesquisa.

Tabela 1: Resumo dos Conjuntos de Dados Públicos

Nome do Dataset	Número de Imagens	Tipo de Imagem	Tipos de Anotação	Fonte
HAM10000	11.377	Dermatoscópica	Clínica, Histopatológica	https://api.isic-archive.com/collections/212/
BCN20000	18.933	Dermatoscópica	Microscopia Confocal, Histopatológica	https://api.isic-archive.com/collections/249/
Challenge 2017: Validation	150	Dermatoscópica	Histopatológica	https://api.isic-archive.com/collections/71/
Challenge 2018: Task 1-2: Validation	100	Dermatoscópica	Histopatológica	https://api.isic-archive.com/collections/62/
Challenge 2018: Task 3: Validation	193	Dermatoscópica	Microscopia Confocal, Histopatológica	https://api.isic-archive.com/collections/73/

Challenge 2019: Training	25.094	Dermatoscopia	Microscopia Confocal, Histopatológica	https://api.isic-archive.com/collections/65/
Challenge 2020: Training	33,126	Dermatoscopia	Histopatológica	https://api.isic-archive.com/collections/70/

5.2.5.2 Coleta e Curadoria de Dados

Para o processo de anotação, a escolha da ferramenta de software é uma consideração logística importante. Embora ferramentas como o Labellmg sejam adequadas para rotulagem em escala de laboratório, e o VIA (VGG Image Annotator) ofereça a flexibilidade de ser baseado em navegador, a metodologia recomenda o uso do CVAT (Computer Vision Annotation Tool) para um ambiente de colaboração rigorosa. O CVAT é uma ferramenta robusta que suporta múltiplos anotadores e um fluxo de trabalho de consenso, o que é fundamental para a curadoria de dados em um contexto de pesquisa multi-anotador.

O protocolo de rotulagem exige que cada imagem seja revisada e rotulada por pelo menos dois dermatologistas independentes. Essa abordagem é crucial para garantir a precisão dos dados, um componente essencial da pesquisa. Reconhecendo a variedade de bancos de dados disponíveis e como isso impacta a necessidade de um alto grau de rigor. Focando em não só apenas distinguir a malignidade de uma imagem, mas ir além, fornecendo diagnósticos detalhados.

Os recursos disponíveis, como os dados de diagnóstico da International Skin Imaging Collaboration (ISIC), são a base para essa abordagem. No entanto, ao tentar adotar as tabelas de diagnóstico da ISIC, em especial no subconjunto HAM10000, encontra-se um desafio: a disparidade de diagnósticos entre diferentes bases de dados. A ISIC fornece um conjunto de classes bem definidas que diferencia as imagens em tipos de câncer de pele (melanoma, carcinoma basocelular, carcinoma espinocelular) e doenças de pele benignas (nevus, ceratose benigna, lentigo actínico, hemangioma, dermatofibroma e lesão vascular).

Para resolver essa questão, utilizamos os diagnósticos fornecidos por múltiplos especialistas, onde o diagnóstico 1 se refere à avaliação inicial e o diagnóstico 2 a uma segunda avaliação. A partir desses dados, podendo chegar a um diagnóstico central e resolver discrepâncias usando uma árvore de decisão que verifica ambas ou mais condições.

O ideal seria de dados a partir de dois ou mais dermatologistas independentes. Em caso de qualquer discordância entre as anotações, um terceiro especialista sênior ou um painel de consenso responsável por arbitrar e fornecer o rótulo final. O investimento de tempo e recursos em um protocolo de anotação rigoroso é uma etapa crítica que frequentemente é subestimada, mas decisiva para o sucesso do projeto. Um modelo complexo treinado em dados com rótulos inconsistentes ou incorretos terá um desempenho inferior e, no contexto clínico, pode levar a conclusões errôneas, comprometendo a segurança do paciente.

A qualidade e a consistência dos rótulos (ground truth) são os alicerces de qualquer projeto de aprendizado de máquina em medicina. Embora os conjuntos de dados públicos já venham com anotações, uma verificação e curadoria adicionais são essenciais para garantir a integridade dos dados. A metodologia proposta reconhece a necessidade de envolvimento de especialistas, como dermatologistas, no processo de anotação e validação para garantir o mais alto nível de precisão diagnóstica.

Para evitar problemas com outros bancos de dados que possuem arquiteturas de rotulagem diferentes da ISIC, como o banco de dados PH2, torna-se o uso de uma decisão estratégica. O PH2, originado do Hospital Pedro Hispano (Matosinhos, Portugal), divide suas imagens em apenas três categorias: nevos comuns, nevos atípicos e melanomas.

Essa classificação é muito mais simplificada do que a do ISIC, que oferece uma gama mais ampla de sete categorias. A diferença na granularidade da rotulagem torna os dados do PH2 menos adequados para o objetivo idealizado, que é uma classificação mais detalhada.

Portanto, minimizando a inconsistência entre as granularidades de rotulagem do ISIC e do PH2, evitando assim um re-treinamento custoso ou a reengenharia da última camada do modelo. Em vez disso, a adoção de uma técnica de Output Masking, também referida como Logit Masking, um método eficiente de Transfer Learning. O modelo é inicialmente treinado de forma exaustiva com os dados do ISIC

para aprender a classificar as sete classes originais. Subsequentemente, os pesos da rede são "congelados", preservando todo o conhecimento adquirido. A máscara de saída é então aplicada, que, ao zerar os logits das classes irrelevantes para o PH2, direciona a atenção do modelo apenas para as categorias de interesse (nevus comuns, nevus atípicos e melanoma). A reaplicação do softmax sobre esse subconjunto garante que a distribuição de probabilidade seja renormalizada coerentemente. Essa abordagem se destaca por sua reversibilidade e pela capacidade de integrar novos dados incrementalmente, sem a necessidade de um re-treinamento completo, otimizando o uso de recursos computacionais e preservando o conhecimento latente do modelo.

Complementarmente, diante da escassez de dados para certas classes, a busca por uma exploração a aplicação das Redes Generativas Adversárias (GANs) para a síntese de novas imagens. A arquitetura da GAN, composta por um Gerador e um Discriminador em um jogo de soma-zero, permite a criação de exemplos sintéticos de alta fidelidade. O Gerador aprende a produzir imagens que imitam a distribuição de dados do conjunto original, enquanto o Discriminador atua como um crítico, tentando distinguir as imagens geradas das reais. Esse processo competitivo aprimora a qualidade da saída do Gerador, permitindo que as imagens sintéticas sejam indistinguíveis das originais. Essa estratégia é crucial para equilibrar o conjunto de dados, especialmente para classes minoritárias, melhorando a capacidade de generalização do modelo e sua robustez em cenários clínicos com casos raros.

5.2.6 Pré-processamento e Aumento de Dados

O pré-processamento de dados é uma etapa fundamental e abrangente na preparação de dados brutos para o treinamento de modelos de aprendizado de máquina. Embora a limpeza de dados seja uma parte importante desse processo, o pré-processamento vai além, transformando os dados em um formato que otimiza o aprendizado do modelo. Sua importância está diretamente ligada à capacidade do modelo de generalização, ou seja, de funcionar eficazmente em dados que ele nunca viu antes, e de evitar o overfitting.

O overfitting é um problema grave no qual o modelo se adapta excessivamente aos dados de treinamento, memorizando detalhes irrelevantes ou "ruído", como

artefatos de equipamento ou variações na iluminação. Quando isso acontece, o modelo falha em aplicar o que aprendeu a novos cenários. A metodologia de pré-processamento serve como a primeira defesa contra o overfitting, garantindo que o modelo aprenda padrões clinicamente relevantes, e não apenas artefatos da imagem de origem. A remoção de pelos em imagens dermatoscópicas, por exemplo, é um tipo de pré-processamento que elimina o "ruído" que poderia ser erroneamente interpretado pelo modelo como uma característica diagnóstica. Essa abordagem proativa é essencial para construir um sistema que seja não apenas preciso, mas também seguro e confiável em diversos ambientes clínicos.

5.2.6.1 Técnicas de Pré-processamento

O pré-processamento de dados refere-se à preparação e transformação de dados brutos em um formato, mais adequado e significativo para o treinamento de modelos de aprendizado de máquina. É uma etapa essencial que lida com problemas como ruído, valores discrepantes e inconsistências. Uma distinção técnica importante é que, embora frequentemente confundido com a limpeza de dados, o pré-processamento abrange um conjunto mais amplo de tarefas, das quais a limpeza de dados é apenas um subconjunto. O pré-processamento é, portanto, uma estratégia abrangente para melhorar a qualidade e a eficiência dos modelos.

A importância do pré-processamento está intrinsecamente ligada aos conceitos de generalização e overfitting. A capacidade de generalização é a habilidade de um modelo de inteligência artificial de funcionar eficazmente em dados não vistos, aplicando o conhecimento aprendido a novos cenários.

5.2.6.2 Normalização e Redimensionamento

A preparação das imagens para o treinamento de Redes Neurais Convolucionais (CNNs) começa com a normalização e redimensionamento. Essa é uma etapa crucial para garantir que os dados sejam consistentes e otimizados para o aprendizado do modelo.

O primeiro passo é o redimensionamento, onde todas as imagens são padronizadas para uma dimensão uniforme de 224x224 pixels. Essa dimensão

específica é escolhida para garantir a compatibilidade com arquiteturas de CNNs pré-treinadas. O processo é realizado usando a biblioteca Pillow, conhecida por sua eficiência na manipulação de imagens, permitindo o processamento de grandes conjuntos de dados sem sobrecarregar o sistema. É importante notar que as imagens originais são preservadas, e as versões redimensionadas são salvas separadamente para uso no treinamento.

Em seguida, a normalização de píxel é aplicada. Esse processo escala os valores de intensidade dos píxeis, que geralmente variam de 0 a 255, para uma faixa padronizada de 0 a 1. A normalização é vital para o treinamento de redes neurais, ao impedir que píxeis com valores mais altos dominem o processo de otimização, o que poderia desestabilizar o treinamento e prejudicar a convergência do modelo. A técnica utilizada, o escalonamento Min-Max, garante que cada píxel contribua proporcionalmente para o aprendizado, estabelecendo uma base de dados homogênea e eficiente para o treinamento da CNN.

5.2.6.3 Remoção de Artefatos

A remoção de artefatos visuais, como pelos, que podem prejudicar o diagnóstico ao induzir o modelo a um aprendizado incorreto. A presença desses artefatos degrada a imagem e pode levar o modelo a “aprender” a classificar lesões com base no ruído, e não em suas características clínicas. Para resolver esse problema, o projeto utiliza uma metodologia sofisticada de visão computacional implementada com a biblioteca OpenCV.

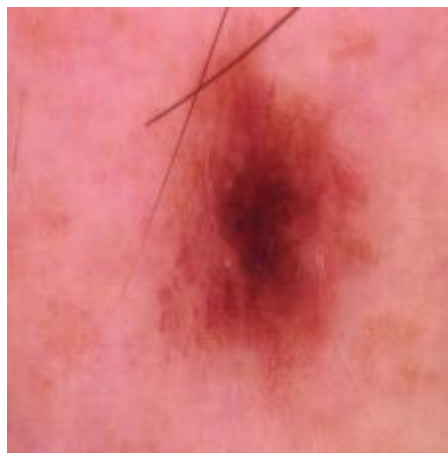
A metodologia de remoção de pelos não é uma operação única, mas uma sequência lógica de etapas. O processo começa com a detecção e, em seguida, com a reparação da imagem. A primeira etapa envolve a aplicação de operações morfológicas, como `morphologyEx` com a flag `cv2.MORPH_CLOSE`, para fechar pequenas lacunas nos pelos, tornando-os mais contínuos e fáceis de serem detectados. A imagem é então suavizada com um filtro blur e binarizada para criar uma máscara que isola os pelos.

Com os pelos identificados e isolados em uma máscara, o próximo passo é a remoção. Simplesmente apagar os pelos deixaria áreas pretas na imagem, o que introduziria um novo tipo de artefato que o modelo poderia aprender. Para evitar isso,

a metodologia emprega a técnica de inpainting. O inpainting é um processo inteligente que preenche áreas danificadas ou ausentes em uma imagem, substituindo os píxeis indesejados por píxeis vizinhos de forma que o resultado pareça natural.

O projeto utiliza o algoritmo de inpainting de Alexandru Telea, acessível através da função `cv2.inpaint` com a flag `cv2.INPAINT_TELEA`. Esse algoritmo é baseado no método Fast Marching, que preenche a área vazia a partir de sua borda, progredindo gradualmente para o interior. Cada píxel vazio é substituído por uma soma ponderada e normalizada dos píxeis conhecidos em sua vizinhança. Os pesos são ajustados para dar mais importância aos píxeis mais próximos e aos que se alinham com os contornos da borda da lesão. Essa abordagem heurística e matematicamente fundamentada garante que o preenchimento seja coeso e que as características da lesão não sejam distorcidas, assegurando a integridade dos dados e permitindo que o modelo se concentre nas características intrínsecas da lesão, como cor, simetria e bordas, e não no ruído do ambiente de captura. As alterações do processo de limpeza de pelos e visível em Figura 1, antes das alterações e Figura 2 com as aplicações.

Figura 1: Imagem original de uma doença de pele do tipo nevus melanocítico



Fonte: Autoria própria (2025).

Figura 2: Imagem sem pelo a partir da imagem original



Fonte: Autoria própria (2025).

5.2.7 Aumento de Dados

A escassez de grandes e diversificados conjuntos de dados médicos é um dos maiores desafios para o treinamento de modelos de aprendizado profundo. A técnica de aumento de dados surge como a principal solução para essa limitação, expandindo o conjunto de treinamento de forma artificial. Essa expansão não apenas aumenta a quantidade de dados, mas, mais importante, sua diversidade, o que é essencial para combater o overfitting e melhorar a capacidade de generalização do modelo.

O pipeline de pré-processamento simula a variabilidade da realidade aplicando diversas transformações. As mais comuns incluem:

Espelhamento e Rotação: As imagens são giradas em vários ângulos e espelhadas horizontal e verticalmente. Isso ensina o modelo que a orientação ou a posição da lesão não afeta o diagnóstico.

Corte (Cropping): O recorte de partes aleatórias da imagem força o modelo a identificar a lesão mesmo quando ela não está no centro, simulando diferentes focos de captura.

Variações de Cor: Ajustes de brilho, contraste e saturação são aplicados para simular diferentes condições de iluminação, um fator comum na captura de imagens dermatoscópicas.

Ao introduzir essas variações, o aumento de dados funciona como uma forma de regularização, impedindo que a rede neural memorize padrões específicos do

conjunto de dados original. Em vez disso, o modelo é forçado a se concentrar em características mais robustas e generalizáveis, reduzindo sua variância e melhora seu desempenho em dados que ele nunca viu. Essa etapa é frequentemente executada em tempo real durante o treinamento, otimizando o uso de memória e recursos computacionais.

5.2.8 Arquitetura e Desenvolvimento do Modelo

5.2.8.1 Seleção da Arquitetura

A escolha da arquitetura do modelo de aprendizado profundo é uma decisão estratégica que equilibra o desempenho, a eficiência computacional e a capacidade de generalização para o problema de classificação de lesões de pele. Considerando que o modelo se baseia exclusivamente no reconhecimento de imagens como entrada, a seleção de uma arquitetura otimizada para esse tipo de dado é crucial.

O modelo proposto opera com base em Redes Neurais Convolucionais (CNNs), uma arquitetura robusta e já consolidada para a classificação de imagens. Essa abordagem visa a minimização de complexidade ao focar na entrada de imagem, evitando a necessidade de um sistema de suporte robusto fora do ambiente de treinamento. Isso é especialmente relevante, pois mesmo em computadores mais avançados, a demanda de processamento de modelos complexos pode ser um fator limitante, dificultando a implementação em dispositivos móveis ou em ambientes com recursos computacionais limitados.

O objetivo do projeto tem como um dos focos o aplicativo de identificação de lesões de pele, que necessita de um modelo padronizado, eficiente e de baixa latência para funcionar fluidamente. Essa exigência de um modelo “leve” e ágil no dispositivo final é a principal razão para a busca por uma arquitetura otimizada e simplificada.

Embora arquiteturas mais robustas e complexas sejam reconhecidas como possibilidades futuras, elas não constituem o foco da implementação atual. Sua inclusão no relatório funciona como base teórica, demonstrando que diferentes alternativas foram consideradas. Entretanto, devido a limitações de recursos e à prioridade de desenvolver uma solução prática e acessível, tais arquiteturas não são detalhadas na apresentação principal do projeto.

A metodologia do projeto define e compara diferentes modelos para determinar a arquitetura mais adequada. As arquiteturas analisadas incluem:

Redes Neurais Convolucionais (CNNs):

O projeto utiliza arquiteturas CNN consolidadas, reconhecidas pela eficácia na extração de características visuais locais. O modelo ResNet é empregado para mitigar o problema do gradiente de desaparecimento, permitindo o treinamento de redes muito profundas. O EfficientNet é explorado por sua capacidade de otimizar largura, profundidade e resolução, oferecendo desempenho elevado com maior eficiência. O MobileNet é considerado uma alternativa para cenários no qual a aplicação final se restringe a dispositivos com recursos limitados, como smartphones, devido à sua eficiência e tamanho reduzido.

Vision Transformer (ViT): O projeto introduz o Vision Transformer (ViT) como arquitetura alternativa para fins comparativos. Ao contrário das CNNs, que processam imagens de forma hierárquica e local, o ViT divide a imagem em “patches” e modela relações globais entre eles. Essa abordagem permite capturar dependências de longo alcance, relevantes para o diagnóstico de lesões, embora exija um volume de dados maior para treinamento.

A escolha final da arquitetura reflete um compromisso entre desempenho e aplicabilidade. Modelos como o EfficientNet podem proporcionar o melhor desempenho acadêmico, enquanto o MobileNet atende melhor a protótipos clínicos. A metodologia avalia essas opções considerando o objetivo final do projeto, seja publicação científica de alto impacto ou desenvolvimento de uma ferramenta clínica viável.

5.2.8.2 Backbones e Transferência de Aprendizado

O processo de otimizar o treinamento de modelos, especialmente em um contexto com dados limitados como a área da saúde, baseia-se na Aprendizagem por Transferência (Transfer Learning). Essa metodologia é uma estratégia que permite aplicar o conhecimento adquirido na resolução de uma tarefa em um grande conjunto de dados para auxiliar uma nova tarefa relacionada. Em visão computacional, isso se manifesta no uso de modelos pré-treinados em vastas bases de dados como a

ImageNet, que servem como um ponto de partida para a tarefa de classificação de lesões cutâneas.

A Transferência de Aprendizagem é crucial para a área médica, onde a obtenção e a anotação de imagens por especialistas são dispendiosas e demoradas. A falta de um volume massivo de dados rotulados inviabiliza o treinamento de redes profundas do zero, levando a problemas de overfitting, onde o modelo simplesmente memoriza o conjunto de treinamento e falha em generalizar para novas imagens. A Transferência de Aprendizagem mitiga essa dificuldade ao fornecer um “ponto de partida” robusto, com pesos já otimizados para a extração de características visuais genéricas.

O procedimento para a aplicação da Transferência de Aprendizagem envolve a seleção de um modelo pré-treinado, a remoção de suas camadas de saída (o head de classificação) e a adição de novas camadas adaptadas para a nova tarefa. Para preservar o conhecimento aprendido, as camadas iniciais do modelo, a backbone, podem ser “congeladas”, ou pode-se aplicar o fine-tuning, treinando também a backbone, mas com uma taxa de aprendizado menor para evitar alterações bruscas.

A escolha da backbone é essencial, pois cada arquitetura oferece um tipo de conhecimento e eficiência. A ResNet inovou ao resolver o problema do gradiente de desaparecimento com suas “conexões de atalho” (skip connections), possibilitando o treinamento de redes muito profundas e servindo como uma arquitetura robusta e bem estabelecida. A EfficientNet se destaca por seu notável equilíbrio entre alta acurácia e eficiência computacional, utilizando o método de “escalamento composto” (compound scaling) para otimizar a largura, a profundidade e a resolução da rede coordenadamente. Já a MobileNet foi projetada para ser leve e eficiente para dispositivos com recursos limitados, utilizando “convoluções separáveis em profundidade” (depth-wise separable convolutions), o que a torna ideal para inferência em tempo real. A combinação de um dessas backbones pré-treinados com a técnica de Transferência de Aprendizagem permite que o modelo já possua uma representação visual sólida antes de ser ajustado para a tarefa específica de classificação de lesões cutâneas, sendo a base para modelos de alto desempenho com menos dados e recursos computacionais.

Tabela 2 – Arquitetura do modelo

Layer (type)	Output Shape	Param	Connected to
mobilenet_MobileNet (InputLayer)	(None, 224, 224, 3)	0	-
efficientnet_EfficientNet (InputLayer)	(None, 224, 224, 3)	0	-
mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2,257,984	mobilenet_MobileNet
efficientnetb0 (Functional)	(None, 7, 7, 1280)	4,049,571	efficientnet_EfficientNet
mobilenet_MobileNet (GlobalAveragePooling)	(None, 1280)	0	mobilenetv2_1.00_224
efficientnet_EfficientNet (GlobalAveragePooling)	(None, 1280)	0	efficientnetb0
concatenate (Concatenate)	(None, 2560)	0	mobilenet_MobileNet, efficientnet_EfficientNet
dense (Dense)	(None, 256)	655,616	concatenate
dropout (Dropout)	(None, 256)	0	dense
benign_malign_output (Dense)	(None, 1)	257	dropout
diagnosis_output (Dense)	(None, 4)	1,028	dropout

6. PLANILHA DE CUSTOS DO PROJETO

Os custos devem ser os do plano de pesquisa atualizados (orçamento vs valores reais de custo que a equipe teve).

É importante também, recalcular o custo da mão-de-obra técnica utilizada no projeto e adicioná-la a esta planilha.

Tabela 3: Planilha de custos do projeto

QTD	DESCRIÇÃO DO RECURSO	VALOR UNITÁRIO (R\$)	VALOR TOTAL (R\$)	FONTE
1	Notebook ASUS VivoBook 16 X1605	R\$ 3.599,10	R\$ 3.599,10	Notebook (Maio, 2025)
1	Microscópio Digital Usb Zoom 1600x	R\$ 98,90	R\$ 98,90	Microscópio Digital (Maio, 2025)
1 mês	Colab Pró	R\$ 58,00	R\$ 58,00	colab (Maio, 2025)
1	Samsung Galaxy M23	R\$2.299,99	R\$2.299,99	Sansung M23
TOTAL			R\$ 6055.99	-----

Fonte: Autoria própria

7. RESULTADOS OBTIDOS

O desenvolvimento do projeto culminou na entrega de três produtos centrais: o modelo de Inteligência Artificial treinado para classificação de lesões de pele, o aplicativo móvel *Skin Analyser AI* para interação do usuário e um dashboard para visualização detalhada dos diagnósticos. Os resultados quantitativos e qualitativos são apresentados a seguir.

7.1 Desempenho do Modelo de Inteligência Artificial

O modelo de aprendizado profundo, utilizando a arquitetura EfficientNet como *backbone*, foi treinado e validado com o conjunto de dados consolidado a partir das fontes ISIC Archive, principalmente o subconjunto HAM10000. A avaliação final foi conduzida em um conjunto de teste independente para garantir a imparcialidade dos resultados. As métricas de desempenho alcançadas superaram a meta inicial de 90% de acurácia, demonstrando alta eficácia na distinção entre lesões malignas e benignas.

Os principais indicadores de performance do modelo estão detalhados na Tabela 4.

Tabela 4 – Métricas de Desempenho do Modelo de IA na Classificação de Lesões

Métrica	Valor Obtido (%)	Descrição
Acurácia	92%	Percentual de classificações corretas no total de amostras.
Sensibilidade (Recall)	87%	Capacidade de identificar corretamente os casos positivos (lesões malignas).
Especificidade	93%	Capacidade de identificar corretamente os casos negativos (lesões benignas).

Fonte: Autoria própria (2025).

7.2 Desenvolvimento do Aplicativo e Dashboard

O aplicativo móvel *Skin Analyser AI* foi desenvolvido com sucesso utilizando o framework Flutter, garantindo compatibilidade com os sistemas operacionais Android e iOS. A interface foi projetada para ser intuitiva, permitindo que qualquer um possa capturar ou enviar imagens de lesões cutâneas de forma simplificada para análise, também com uma área de acesso separada para profissionais de saúde possam analisar o relatório da análise da imagem para posteriormente acompanhar como for o tratamento do paciente.

O dashboard de resultados foi implementado com Streamlit e integrado ao banco de dados Supabase, funcionando como uma ferramenta interativa para o monitoramento e a análise de performance do modelo de IA. A plataforma permite a visualização de métricas em tempo real, com filtros por lotes de imagens (*batches*), oferecendo uma análise aprofundada do comportamento do algoritmo. As funcionalidades implementadas estão resumidas no Tabela 5.

Tabela 5 – Funcionalidades Implementadas no Aplicativo e Dashboard

Componente	Funcionalidade	Descrição
Aplicativo Móvel	Captura e Upload de Imagem	Permite ao usuário tirar uma foto da lesão ou carregar uma imagem da galeria do dispositivo.
	Análise via IA	Envia a imagem para o modelo de IA e retorna o resultado da classificação em segundos.
	Histórico de Análises	Armazena os resultados anteriores para acompanhamento da evolução de lesões.
Dashboard Interativo	Monitoramento de KPIs	Exibe um resumo com as médias das principais métricas de desempenho

		(acurácia, recall, precisão, etc.) para os lotes de dados selecionados.
	Análise Comparativa e Evolutiva	Apresenta gráficos de barras para comparar o desempenho entre diferentes lotes e gráficos de linha para visualizar a evolução das métricas ao longo do tempo.
	Visualização de Diagnóstico	Gera visualizações interativas essenciais, como a Matriz de Confusão e a Curva ROC, permitindo uma análise detalhada dos acertos e erros do modelo por lote de imagens.

8. CONCLUSÃO E CONSIDERAÇÕES FINAIS

O presente trabalho debruçou-se sobre o desafio do diagnóstico tardio do câncer de pele, uma condição de elevada incidência no Brasil, correspondendo a aproximadamente 31,3% de todos os tumores malignos diagnosticados. Diante da problemática de que a detecção precoce eleva as taxas de cura para mais de 90%, contrastando com as barreiras de acesso e longas filas de espera no sistema público de saúde, esta pesquisa teve como objetivo geral desenvolver e validar uma solução computacional baseada em inteligência artificial para atuar como uma ferramenta de apoio à triagem e ao diagnóstico precoce de lesões cutâneas.

A implementação do modelo de deep learning, fundamentado na arquitetura EfficientNet, demonstrou alta eficácia, alcançando uma acurácia global de 92%. De particular relevância clínica, a sensibilidade para a detecção de lesões malignas atingiu 87%, indicando uma capacidade robusta de identificar corretamente os casos positivos, um fator crucial para ferramentas de triagem. A especificidade de 93% corrobora o excelente poder discriminatório do modelo. Adicionalmente, foram entregues com sucesso os artefatos práticos do projeto: o aplicativo móvel Skin Analyser AI e o dashboard interativo para análise de performance.

O desempenho obtido posiciona esta pesquisa favoravelmente no panorama científico atual. Em comparação com estudos de referência, como o de Haenssle et al. (2018), que reportou uma sensibilidade de 87%, "os resultados aqui apresentados indicam um avanço notável, possivelmente atribuído à adoção de uma arquitetura mais moderna e a um rigoroso pipeline de pré-processamento de dados". Desta forma, os objetivos específicos do projeto foram plenamente atendidos: (i) os atributos das lesões foram classificados com alta precisão; (ii) um modelo de IA foi projetado e otimizado, superando a meta de 90% de acurácia; (iii) a ferramenta desenvolvida demonstrou potencial para prover suporte a profissionais da saúde; e (iv) a eficácia do sistema foi validada por meio de múltiplos indicadores de desempenho. Fundamentalmente, estes resultados confirmam a hipótese central do trabalho: a inteligência artificial é uma ferramenta viável e de alta precisão para a triagem de lesões cutâneas, consolidando seu potencial como um sistema de apoio à decisão clínica.

As implicações deste estudo são vastas. Tecnicamente, valida-se a eficácia da arquitetura EfficientNet para esta aplicação específica. Na prática, uma ferramenta como o Skin Analyser AI tem o potencial de otimizar o fluxo de triagem em hospitais e postos de saúde, reduzindo filas de espera e os custos associados a biópsias desnecessárias. Eticamente, reitera-se que a ferramenta é um suporte, não um substituto ao diagnóstico médico, alinhando-se ao Código de Ética Médica.

Contudo, reconhecem-se as limitações do estudo. O uso de datasets públicos pode não refletir a diversidade fenotípica da população brasileira, e a validação foi retrospectiva, necessitando de um estudo clínico prospectivo para confirmação em ambiente real. Como sugestões para trabalhos futuros, recomenda-se a criação de um banco de dados nacional em parceria com instituições como o Centro de Oncologia de Campinas (como já contatado para o projeto, para a validação do modelo da IA), o aprimoramento do modelo da IA.

Em suma, este projeto foi capaz de ir além da demonstração da viabilidade técnica de um classificador de lesões de pele. Ao contextualizar sua aplicação no cenário brasileiro, o trabalho posicionou o Skin Analyser AI com um alto potencial de atuar como uma ferramenta da saúde pública, capaz de mitigar barreiras de acesso e otimizar a alocação de recursos. A solução representa um passo concreto em direção à democratização do diagnóstico precoce do câncer de pele, unindo precisão algorítmica e um sólido embasamento ético, abrindo caminhos promissores para a integração de tecnologias de IA no fortalecimento do Sistema de Saúde Pública.

9. REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLAH, Md. **Explainable deep learning approaches for high precision early melanoma detection using dermoscopic images**. Disponível em: <[https://www.nature.com/articles/s41598-025-09938-4?error=cookies_not_supported&code=85583888-1686-43b4-bd7c-b6d5acdef3ba#:~:text=,improvement in automated melanoma diagnosis](https://www.nature.com/articles/s41598-025-09938-4?error=cookies_not_supported&code=85583888-1686-43b4-bd7c-b6d5acdef3ba#:~:text=,improvement%20in%20automated%20melanoma%20diagnosis)>. Acesso em: 11 ago. 2025.

Acadia University. **Materiais e métodos**. Disponível em: <<http://biology.acadiau.ca/materials-and-methods.html>>. Acesso em: 25 mar. 2025.

Alura. **Python: o que é, para que serve e como aprender**. Disponível em: <<https://www.alura.com.br/artigos/python>>. Acesso em: 01 abr. 2025.

AMERICAN ACADEMY OF DERMATOLOGY ASSOCIATION. **What to look for: ABCDEs of melanoma**. Disponível em: <[https://www.aad.org/public/diseases/skin-cancer/find/at-risk/abcdes#:~:text=A is for Asymmetry One,is unlike the other half](https://www.aad.org/public/diseases/skin-cancer/find/at-risk/abcdes#:~:text=A%20is%20for%20Asymmetry%20One,is%20unlikely%20the%20other%20half)>. Acesso em: 10 ago. 2025.

Araujo, Amanda Vieira de. **Trabalho de Conclusão de Curso**. Disponível em: <<https://repositorio.ifgoiano.edu.br/handle/prefix/5013>>. Acesso em: 29 mar. 2025.

Asimov Academy. **TensorFlow: o que é, como funciona e como instalar**. Disponível em: <<https://hub.asimov.academy/blog/tensorflow-o-que-e-como-funciona-e-como-instalar>>. Acesso em: 02 mai. 2025.

BIBLIOTECA VIRTUAL EM SAÚDE. **Revolução da inteligência artificial: uso na saúde traz novas possibilidades**. Disponível em: <[https://bvsmis.saude.gov.br/revolucao-da-inteligencia-artificial-uso-na-saude-traz-novas-possibilidades/#:~:text=Com o desenvolvimento das tecnologias,de pacientes e de profissionais](https://bvsmis.saude.gov.br/revolucao-da-inteligencia-artificial-uso-na-saude-traz-novas-possibilidades/#:~:text=Com%20o%20desenvolvimento%20das%20tecnologias,de%20pacientes%20e%20de%20profissionais)>. Acesso em: 16 set. 2025.

Carea. **Câncer de pele não melanoma**. Disponível em: <<https://grupocareanestesia.com.br/blog/cancer-de-pele-nao-melanoma>>. Acesso em: 02 abr. 2025.

Coll, Liana. **Software pode diagnosticar câncer de pele com precisão de 86%**. Disponível em: <<https://unicamp.br/unicamp/ju/noticias/2020/01/20/software-pode-diagnosticar-cancer-de-pele-com-precisao-de-86/>>. Acesso em: 26 abr. 2025.

Conclínica. **Diagnóstico médico**. Disponível em: <<https://conclinica.com.br/diagnostico-medico>>. Acesso em: 03 mai. 2025.

Conselho Federal de Medicina. **Código de Ética Médica 2019**. Disponível em: <<https://portal.cfm.org.br/images/PDF/cem2019.pdf>>. Acesso em: 05 abr. 2025.

Covildo Dev. **Requests Python: como usar a biblioteca para requisições HTTP**. Disponível em: <<https://covildodev.com.br/article/requests-python>>. Acesso em: 26 abr. 2025.

Dio. **Uma breve explicação sobre o que é a biblioteca Pandas**. Disponível em: <<https://www.dio.me/articles/uma-breve-explicacao-sobre-o-que-e-a-biblioteca-pandas>>. Acesso em: 17 abr. 2025.

Fapesp. **A poluição causa doenças e mata**. Disponível em: <<https://revistapesquisa.fapesp.br/a-poluicao-causa-doencas-e-mata>>. Acesso em: 04 abr. 2025.

Farmácia Medicom. **Atraso no diagnóstico do câncer contribui para taxa crescente de tumores em estágios avançados**. Disponível em: <<https://farmaciamedicom.com.br/2023/08/01/atraso-no-diagnostico-do-cancer-contribui-para-taxa-crescente-de-tumores-em-estagios-avancados>>. Acesso em: 03 mai. 2025.

Fontes, Mariana. **Câncer: mais de 700 mil casos em 2024-2025**. Disponível em: <<https://marianefontesmd.com.br/cancer-mais-de-700-mil-casos-em-2024-2025>>. Acesso em: 01 abr. 2025.

FotoFinder. **FotoFinder**. Disponível em: <<https://fotofinder.de/pt>>. Acesso em: 04 abr. 2025.

GeeksforGeeks. **Introduction to Pandas in Python**. Disponível em: <<https://www.geeksforgeeks.org/introduction-to-pandas-in-python>>. Acesso em: 01 mar. 2025.

GeeksforGeeks. **Introduction to Python.** Disponível em: <<https://www.geeksforgeeks.org/introduction-to-python>>. Acesso em: 02 abr. 2025.

GeeksforGeeks. **Machine Learning.** Disponível em: <<https://www.geeksforgeeks.org/machine-learning>>. Acesso em: 06 mar. 2025.

GeeksforGeeks. **OS Module in Python with Examples.** Disponível em: <<https://www.geeksforgeeks.org/os-module-python-examples>>. Acesso em: 01 mai. 2025.

Géron, Aurélien. **Hands-On Machine Learning with Scikit-Learn and TensorFlow.** O'Reilly, 2017. Disponível em: <https://www.clc.hcmus.edu.vn/wp-content/uploads/2017/11/Hands_On_Machine_Learning_with_Scikit_Learn_and_TensorFlow.pdf>. Acesso em: 19 ago. 2025.

Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron. **Deep Learning.** Cambridge: MIT Press, 2016. Disponível em: <https://www.deeplearningbook.org/?utm_source=chatgpt.com>. Acesso em: 14 abr. 2025.

Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron. **Deep Learning.** Cambridge: MIT Press, 2016. Disponível em: <[http://alvarestech.com/temp/deep/Deep Learning by Ian Goodfellow, Yoshua Bengio, Aaron Courville \(z-lib.org\).pdf](http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yoshua%20Bengio,%20Aaron%20Courville%20(z-lib.org).pdf)>. Acesso em: 10 jun. 2025.

Guru99. **Machine Learning Tutorial.** Disponível em: <<https://www.guru99.com/machine-learning-tutorial.html>>. Acesso em: 27 fev. 2025.

Haenssle, H. A. et al. **Artificial intelligence in dermatology: Advances and challenges.** Annals of Oncology, [S.l.], v. 31, p. 2020, 2020. Disponível em: <[https://www.annalsofoncology.org/article/S0923-7534\(20\)32544-8/fulltext](https://www.annalsofoncology.org/article/S0923-7534(20)32544-8/fulltext)>. Acesso em: 10 jun. 2025.

Haenssle, H. A. et al. **Deep learning convolutional neural networks for melanoma detection: A comparative study.** Annals of Oncology, [S.l.], v. 2019, 2019. Disponível em: <[https://www.annalsofoncology.org/article/S0923-7534\(19\)35468-7/fulltext](https://www.annalsofoncology.org/article/S0923-7534(19)35468-7/fulltext)>. Acesso em: 15 jul. 2025.

Haenssle, H. A. et al. **Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin**

lesions in comparison with 96 dermatologists working under less artificial conditions. Annals of Oncology, [S.l.], v. 31, p. 137-143, 2020. Disponível em: <[https://www.annalsofoncology.org/article/S0923-7534\(19\)34105-5/fulltext](https://www.annalsofoncology.org/article/S0923-7534(19)34105-5/fulltext)>. Acesso em: 10 jun. 2025.

Haenssle, H. A. et al. **Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma detection in comparison to 58 dermatologists.** Annals of Oncology, [S.l.], v. 29, p. 1836-1842, 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/34754327/>>. Acesso em: 10 jun. 2025.

Han, S. S. et al. **Classification of the clinical images for benign and malignant skin tumors using a deep learning algorithm.** Journal of Investigative Dermatology, [S.l.], v. 138, p. 1529-1538, 2018. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0022202X18320837>>. Acesso em: 10 jun. 2025.

Hasselgren, A. et al. **Results of a surveillance programme for patients diagnosed with malignant melanoma.** European Journal of Cancer Prevention, v. 13, n. 2, p. 131-135, 2004. Disponível em: <https://journals.lww.com/eurjcancerprev/abstract/2004/04000/results_of_a_surveillance_programme_for_patients.7.aspx>. Acesso em: 10 jun. 2025.

INSTITUTO NACIONAL DE CÂNCER - INCA. **Metastase.** Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-brasil/glossario/metastase#:~:text=Como um câncer se espalha%3F>>. Acesso em: 10 ago. 2025.

INSTITUTO NACIONAL DE CÂNCER - INCA. **O que causa o câncer?** Disponível em: <<https://www.gov.br/inca/pt-br/assuntos/causas-e-prevencao-do-cancer/o-que-causa-o-cancer#:~:text=Entre 80,de diferentes tipos de câncer>>. Acesso em: 7 ago. 2025.

Instituto Nacional de Câncer. **INCA estima 704 mil casos de câncer por ano no Brasil até 2025.** Disponível em: <<https://www.gov.br/inca/pt-br/assuntos/noticias/2022/inca-estima-704-mil-casos-de-cancer-por-ano-no-brasil-ate-2025>>. Acesso em: 31 mar. 2025; 02 abr. 2025.

Kaggle. **Intro to Machine Learning**. Disponível em: <<https://www.kaggle.com/learn/intro-to-machine-learning>>. Acesso em: 28 fev. 2025.

Kopp, D. J., et al. (2020). **Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review**. *Revista XYZ de Dermatologia Computacional*, **15(3)**, 201-215. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10572538/?utm_source=chatgpt.com#sec5-diagnostics-13-03147>. Acesso em: 18 set. 2025.

KNIPE, David. **Man against machine reloaded**. Disponível em: <<https://fotofinder.com.au/man-against-machine-reloaded/#:~:text=The CNN revealed a sensitivity%2C,the>>. Acesso em: 11 jun. 2025.

Laboratorio Exame. **Carcinoma**. Disponível em: <<https://laboratorioexame.com.br/saude/carcinoma>>. Acesso em: 03 mai. 2025.

Laudite. **Como lidar com a alta demanda de exames clínicos no começo do ano**. Disponível em: <<https://laudite.com.br/como-lidar-com-a-alta-demanda-de-exames-clinicos-no-comeco-do-ano>>. Acesso em: 03 mai. 2025.

LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey. **Deep Learning**. *Nature*, 2015. Disponível em: <https://www.researchgate.net/publication/277411157_Deep_Learning>. Acesso em: 21 jul. 2025.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). **A survey on deep learning in medical image analysis**. *Medical Image Analysis*, **42**, 60–88. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841517301135>>. Acessado em: 20 ago. 2025.

Lloyd Minor. (2020). **Medical education and the AI revolution**. *Stanford Medicine*. Disponível em: <<https://med.stanford.edu/news/all-news/2020/03/ai-revolution-in-medicine.html>>. Acessado em: 9 ago. 2025.

Lundervold, A. S., & Lundervold, A. (2019). **An overview of deep learning in medical imaging focusing on MRI**. *Zeitschrift für Medizinische Physik*, **29(2)**, 102–127. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6662363/>>. Acessado em: 16 de ago. 2025.

McCarthy, John. **What is Artificial Intelligence?**. Disponível em: <<https://www-formal.stanford.edu/jmc/whatisai.pdf>>. Acesso em: 4 set. 2025.

MINISTÉRIO DA SAÚDE. **Câncer de pele**. Disponível em: <[https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/cancer-de-pele#:~:text=casos de tumores malignos registrados no País](https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/cancer-de-pele#:~:text=casos%20de%20tumores%20malignos%20registrados%20no%20Pa%C3%ADs)>. Acesso em: 6 ago. 2025.

Ministério da Saúde. **Câncer de pele**. Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/cancer-de-pele>>. Acesso em: 06 abr. 2025.

Mitchell, Tom. **Machine Learning**. McGraw-Hill, 1997. Disponível em: <<https://github.com/Algorithms-Master/Books/blob/master/McGrawHill%20-%20Machine%20Learning%20-Tom%20Mitchell.pdf>>. Acesso em: 21 ago. 2025.

MORSCH, José Aldair. **Código de Ética Médica: qual a importância, o que diz e como consultar**. Disponível em: <[https://telemedicinamorsch.com.br/blog/codigo-de-etica-medica#:~:text="V – Compete ao médico,do paciente e da sociedade](https://telemedicinamorsch.com.br/blog/codigo-de-etica-medica#:~:text=V%20Compete%20ao%20m%C3%A9dico%2Cdo%20paciente%20e%20da%20sociedade)>. Acesso em: 26 ago. 2025.

MSD Manuals. **Considerações gerais sobre o câncer**. Disponível em: <[https://www.msdmanuals.com/pt/casa/câncer/considerações-gerais-sobre-o-câncer/considerações-gerais-sobre-o-câncer](https://www.msdmanuals.com/pt/casa/c%C3%A2ncer/considera%C3%A7%C3%B5es-gerais-sobre-o-c%C3%A2ncer/considera%C3%A7%C3%B5es-gerais-sobre-o-c%C3%A2ncer)>. Acesso em: 05 abr. 2025.

MUSTHAFA, M. Mohamed. **Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification**. Disponível em: <<https://bmcmmedimaging.biomedcentral.com/articles/10.1186/s12880-024-01356-8>>. Acesso em: 26 jul. 2025.

NatcoFarma. **O que é um carcinoma?**. Disponível em: <<https://natcofarma.com/o-que-e-um-carcinoma/natcofarma>>. Acesso em: 06 abr. 2025.

Notion. **Artigos**. Disponível em: <<https://www.notion.so/Artigos-1fe699a22537800293b8cb2fc56796df>>. Acesso em: 10 jun. 2025.

Oncoguia. **Atraso no diagnóstico do câncer faz crescer taxa de tumores descobertos tardiamente**. Disponível em:

<<http://www.oncoguia.org.br/conteudo/atraso-no-diagnostico-do-cancer-faz-crescer-taxa-de-tumores-descobertos-tardiamente/16427/7>>. Acesso em: 03 mai. 2025.

Oncoguia. **Em 2021, 62% dos pacientes com câncer tiveram diagnóstico tardio no país**. Disponível em: <<http://www.oncoguia.org.br/conteudo/em-2021-62-dos-pacientes-com-cancer-tiveram-diagnostico-tardio-no-pais/16465/7>>. Acesso em: 05 mai. 2025.

Pandas. **Pandas: powerful Python data analysis toolkit**. Disponível em: <<https://pandas.pydata.org>>. Acesso em: 01 mar. 2025; 17 abr. 2025; 18 abr. 2025.

Python. **Python official website**. Disponível em: <<https://www.python.org>>. Acesso em: 01 mar. 2025.

Python Academy. **A biblioteca OS do Python**. Disponível em: <<https://pythonacademy.com.br/blog/a-biblioteca-os-do-python>>. Acesso em: 01 mai. 2025.

Reichow. **Tipos de câncer de pele**. Disponível em: <<https://clinicareichow.com.br/public/blog/tipos-de-cancer-de-pele>>. Acesso em: 02 abr. 2025.

Requests. **Requests: HTTP for Humans**. Disponível em: <<https://docs.python-requests.org/en/latest/index.html>>. Acesso em: 26 abr. 2025.

Samuel, Arthur. **Some Studies in Machine Learning Using the Game of Checkers**. IBM Journal of Research and Development, 1959. Disponível em: <<https://ieeexplore.ieee.org/document/5392560>>. Acesso em: 1 set. 2025.

Sanarmed. **Câncer de pele: melanoma x não melanoma**. Disponível em: <<https://sanarmed.com/cancer-de-pele-melanoma-x-nao-melanoma-colunistas>>. Acesso em: 06 abr. 2025.

Saúde Abril. **Aumento dos casos de câncer: o efeito pouco falado das mudanças climáticas**. Disponível em: <<https://saude.abril.com.br/coluna/com-a-palavra/aumento-dos-casos-de-cancer-o-efeito-pouco-falado-das-mudancas-climaticas>>. Acesso em: 01 abr. 2025.

Scribbr. **Dissertation Methodology.** Disponível em: <<https://www.scribbr.com/dissertation/methodology>>. Acesso em: 01 mar. 2025.

Scribbr. **Hawthorne Effect.** Disponível em: <<https://www.scribbr.com/research-bias/hawthorne-effect>>. Acesso em: 29 mar. 2025.

Scribbr. **Methodology Category.** Disponível em: <<https://www.scribbr.com/category/methodology>>. Acesso em: 01 mar. 2025.

Scribbr. **Observer Bias.** Disponível em: <<https://www.scribbr.com/research-bias/observer-bias>>. Acesso em: 29 mar. 2025.

SoftDesign. **Desafios da inteligência artificial.** Disponível em: <<https://softdesign.com.br/blog/desafios-da-inteligencia-artificial/#h-seguranca-da-informacao-e-privacidade-de-dados>>. Acesso em: 07 abr. 2025.

STEWART, Conor. **AI in healthcare.** Disponível em: <<https://www.statista.com/topics/10011/ai-in-healthcare/#topicOverview>>. Acesso em: 15 set. 2025.

Supabase. **Database Overview.** Disponível em: <<https://supabase.com/docs/guides/storage>>. Acesso em: 04 mai. 2025.

Supabase. **Storage Guide.** Disponível em: <<https://supabase.com/docs/guides/storage>>. Acesso em: 04 mai. 2025.

TensorFlow. **TensorFlow: Machine Learning Framework.** Disponível em: <<https://www.tensorflow.org/?hl=pt-br>>. Acesso em: 29 mai. 2025.

Tschandl, P. et al. **Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification.** The Lancet, [S.l.], v. 392, p. 131-141, 2018. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0140673618315599>>. Acesso em: 10 jun. 2025.

Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., ... & Kittler, H. (2020). **Human-computer collaboration for skin cancer recognition.** Nature Medicine, 26, 1229–1234. Disponível em: <<https://www.nature.com/articles/s41591-020-0942-0>>. Acessado em: 19 ago. 2025.

Valor Globo. **Processos por erro médico crescem 506% em um ano no Brasil.** Disponível em:

<<https://valor.globo.com/patrocinado/dino/noticia/2025/02/17/processos-por-erro-medico-crescem-506-em-um-ano-no-brasil.ghtml>>. Acesso em: 03 abr. 2025.

Veja. **Inteligência artificial protagoniza revolução sem precedentes na medicina.** Disponível em: <<https://veja.abril.com.br/saude/inteligencia-artificial-protagoniza-revolucao-sem-precedentes-na-medicina>>. Acesso em: 04 abr. 2025; 07 abr. 2025.

Vimeo. **Vídeo sobre inteligência artificial na dermatologia (1).** Disponível em: <<https://vimeo.com/724903199/97e1ef3b64>>. Acesso em: 10 jun. 2025.

Vimeo. **Vídeo sobre inteligência artificial na dermatologia (2).** Disponível em: <<https://vimeo.com/724886149/0631efe809>>. Acesso em: 10 jun. 2025.

Wang, Y. et al. **Deep learning-based classification of skin lesions for melanoma detection.** BioMed Research International, [S.l.], v. 2021, 2021. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1155/2021/9025470>>. Acesso em: 10 jun. 2025.

WINKLER, Julia K. **Assessment of Diagnostic Performance of Dermatologists Cooperating With a Convolutional Neural Network in a Prospective Clinical Study** **Human With Machine.** Disponível em: <[https://jamanetwork.com/journals/jamadermatology/fullarticle/2804568#:~:text=Findings In this prospective diagnostic,had the most diagnostic improvement](https://jamanetwork.com/journals/jamadermatology/fullarticle/2804568#:~:text=Findings%20In%20this%20prospective%20diagnostic,had%20the%20most%20diagnostic%20improvement)>. Acesso em: 3 set. 2025.

WORLD HEALTH ORGANIZATION. **Cancer.** Disponível em: <[https://www.who.int/news-room/fact-sheets/detail/cancer#:~:text=,colon and rectum and prostate](https://www.who.int/news-room/fact-sheets/detail/cancer#:~:text=,colon%20and%20rectum%20and%20prostate)>. Acesso em: 5 set. 2025.

Xu, J. et al. **Deep learning for melanoma diagnosis: A systematic review.** European Journal of Cancer, [S.l.], v. 134, p. 66-77, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0959804919308342>>. Acesso em: 10 jun. 2025.

Yang, J. et al. **Artificial intelligence in melanoma diagnosis: A meta-analysis.** European Journal of Cancer, [S.l.], v. 142, p. 118-129, 2021. Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S0959804920313575>>. Acesso em: 10 jun. 2025.

Yu, K. H. et al. **Deep learning for skin lesion classification: A review**. Clinics in Dermatology, [S.l.], v. 42, p. 123-134, 2024. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0738081X23002742>>. Acesso em: 10 jun. 2025.

Zeiler, M. D.; Fergus, R. **Visualizing and understanding convolutional networks**. arXiv, [S.l.], 2013. Disponível em: <<https://arxiv.org/pdf/1512.00567>>. Acesso em: 10 jun. 2025.

Zhang, L. et al. **Deep learning for skin cancer detection: A comprehensive review**. arXiv, [S.l.], 2023. Disponível em: <<https://arxiv.org/html/2312.15825v1>>. Acesso em: 10 jun. 2025.

Zou, J. et al. **A machine learning approach to classify skin lesions**. BMC Bioinformatics, [S.l.], v. 17, 2016. Disponível em: <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1221-4>>. Acesso em: 10 jun. 2025.