

Dermalyze: uma aplicação para auxílio à triagem de lesões de pele baseado em aprendizado profundo

Eduarda Pylo Magesk
eduardapmagesk@gmail.com
Universidade Federal do Espírito Santo
Vitória, Brasil

Patricia H. L. Frasson
patricia.frasson@ebserh.gov.br
Universidade Federal do Espírito Santo
Vitória, Brasil

Luis Antônio de Souza Júnior
la.souza@inf.ufes.br
Universidade Federal do Espírito Santo
Vitória, Brasil

Andre G. Cardoso Pacheco
apacheco@inf.ufes.br
Universidade Federal do Espírito Santo
Vitória, Brasil

ABSTRACT

According to the Brazilian Institute of Cancer, skin cancer is the most common displasya in the country. To mitigate this situation and ensure quality of life for those affected by the disease, early diagnosis is fundamental. When treated early, 5-year survival rates can exceed 96%. In this context, we developed a skin lesion screening application to assist in the identification of skin cancer. This software, designed for use by healthcare professionals, leverages artificial intelligence to assess the cancerous potential of a lesion based on a clinical image captured by a mobile device. The developed model achieved a balanced accuracy of 80%, and the application in which it was embedded was tested in a clinical setting with a total of 82 users. Based on the predictions made during the case study, the model was able to correctly identify the severity of the lesions in approximately 70% of the cases. Additionally, about 92% of users reported the model to be useful during the screening process, describing their experience primarily with words such as *comfort* and *safety*.

KEYWORDS

skin cancer, artificial intelligence, healthcare, triage, mobile app

1 INTRODUÇÃO

O Instituto Nacional de Câncer (INCA) estima que cerca de 30% de todos os tumores malignos registrados no Brasil sejam câncer de pele. A realidade em alguns estados, como no Espírito Santo, é ainda mais grave, com uma estimativa de 41% [15]. Apesar da estimativa de sobrevida em 5 anos ser mais de 96% para aqueles diagnosticados precocemente [11], a doença, quando não tratada, pode levar a deformidades físicas e, em casos mais graves, à metástase, principalmente quando trata-se do câncer de pele do tipo melanoma [1]. Portanto, o diagnóstico precoce é a chave para assegurar ao paciente o melhor prognóstico possível. Contudo, a realidade brasileira é de que existam apenas 5 médicos especialistas em dermatologia para cada 100.000 habitantes [3]. Isso demonstra a necessidade de ferramentas tecnológicas como meio de impulsionar a detecção do câncer de pele.

Neste contexto, destacam-se o uso de sistemas de Diagnóstico Auxiliado por Computador (CAD - do inglês: *Computer-Aided diagnosis*), os quais são responsáveis por oferecer uma segunda opinião para médicos especializados no momento de diagnosticar uma doença. Na área de dermatologia, diversos modelos para a identificação do câncer de pele e/ou classificação de lesões de pele com o uso de *machine learning* já foram propostos [9, 19, 28]. Todavia, o problema central reside na discrepância entre a demanda por dermatologistas e a capacidade do sistema de saúde brasileiro em suprir essa necessidade, especialmente em municípios do interior do país. Essa desigualdade é evidenciada pelo fato de que apenas 32% dos profissionais da área estão localizados nessas regiões [3].

Uma solução para este problema é potencializar a capacidade de profissionais da saúde não especialistas em identificar possíveis casos de câncer de pele. O sistema de saúde, na maioria das vezes, funciona de modo que, antes que um paciente possa ser atendido por um profissional especialista, ele primeiro seja consultado por um médico clínico geral ou por algum outro profissional de atenção primária, dos quais destacam-se os agentes de saúde. Esses profissionais possuem como responsabilidade orientar a população sobre noções básicas de saúde por meio, principalmente, de visitas domiciliares regulares de monitoramento. Desse modo, cabe aos agentes de saúde identificarem possíveis riscos à saúde do paciente, incluindo a identificação de possíveis casos de câncer de pele, mesmo que não possuam treinamento específico na área de dermatologia.

A proposta da aplicação desenvolvida é de atuar justamente nessa triagem realizada por profissionais generalistas, sendo utilizada como uma ferramenta de auxílio capaz de intensificar a identificação das lesões de pele que possuem um maior potencial de serem tumores malignos. Apenas médicos especializados podem diagnosticar o câncer de pele [6]; porém, uma triagem eficiente permite que pacientes com maiores riscos de câncer consigam um atendimento prioritário, fato que aumenta as chances de um diagnóstico precoce e possibilita ao paciente iniciar o devido tratamento o quanto antes.

O aplicativo contém um modelo de inteligência artificial baseado na arquitetura *Convolutional Neural Network* (CNN) que classifica a imagem da lesão de pele enviada quanto ao nível de gravidade que ela representa. Os níveis variam de 1 a 5, sendo 1 a classificação referente às lesões com a maior chance de serem um câncer de pele grave e 5 referente a peles saudáveis e/ou com lesões comuns e não graves.

O modelo desenvolvido está embarcado em um aplicativo mobile que não precisa de acesso à internet para funcionar. Essa característica foi incluída tendo em consideração aqueles profissionais que trabalham em áreas remotas do país, onde o acesso à internet ainda não é garantido. Isso é necessário, sobretudo, quando agentes de saúde realizam visitas domiciliares a trabalhadores rurais, uma vez que, devido à alta exposição solar a que esses trabalhadores são condicionados pela sua profissão, eles são mais propensos ao desenvolvimento de câncer de pele. Como mencionado anteriormente, o público-alvo da aplicação são profissionais da saúde e a aplicação não está disponível para uso geral da população. Além disso, o objetivo da aplicação é ajudar na triagem de pacientes e não deve ser utilizada como forma de diagnóstico. Com o apoio do Programa de Assistência Dermatológica e Cirúrgica da Universidade Federal do Espírito Santo (PAD-UFES), foi possível realizar um estudo de caso como forma de avaliação da aplicação. O PAD-UFES realiza atendimentos mensais gratuitos com o objetivo de diagnosticar e tratar o câncer de pele no interior do estado do Espírito Santo. Durante o ano de 2024, a aplicação desenvolvida foi avaliada dentro do programa, que representa um ambiente clínico real, com o intuito de avaliar o desempenho do modelo de triagem e a percepção dos profissionais de saúde em relação ao uso da tecnologia. As principais contribuições deste artigo são resumidas a seguir:

- Foi realizada uma modelagem do problema da triagem de lesões de pele em cinco níveis de prioridade, os quais foram definidos de acordo com especialistas da área de dermatologia.
- Um modelo de aprendizado profundo foi treinado e embarcado em uma aplicação móvel, a qual funciona de forma totalmente local, sem necessidade de acesso à internet e/ou servidores externos.
- Foi realizado um estudo de caso ao longo de 1 ano com 82 participantes em 8 cidades diferentes para avaliar o desempenho da triagem fornecida pela aplicação.
- É realizada uma análise profunda em relação à usabilidade e utilidade da aplicação por parte dos usuários-alvos da mesma.

2 TRABALHOS RELACIONADOS

Conforme abordado previamente, existe uma diversidade de modelos de aprendizado profundo cujo objetivo é a classificação de doenças de pele ou do câncer de pele [9, 17, 19, 20, 25]. Contudo, poucos são os estudos que avaliam e acompanham o processo de como o modelo desenvolvido será utilizado na prática [26]. Quando trata-se da análise de potencial de risco de uma lesão, os estudos relacionados são ainda mais escassos.

Rosado *et al.* [24] apresentaram uma aplicação para detecção de câncer de pele. Nela, os pacientes submetiam uma imagem da lesão e algumas informações sobre ela para que, posteriormente, o algoritmo classificasse como risco alto ou baixo. Esse resultado era analisado por um dermatologista de uma comunidade online que, junto com essa informação, decidia se o paciente precisava de atendimento especializado. Contudo, a aplicação desse tipo de solução em larga escala torna-se inviável, uma vez que o número de dermatologistas disponíveis é limitado.

A ideia proposta neste artigo é baseada no trabalho de Castro *et al.* [4], que implementou um aplicativo de detecção de câncer de

pele para um contexto semelhante ao deste estudo, que é auxiliar profissionais generalistas da área da saúde na triagem de pacientes. Todavia, enquanto ele tinha como foco apenas a identificação do câncer de pele do tipo melanoma, classificando as amostras em "melanoma" ou "não melanoma", este trabalho propõe uma abordagem mais abrangente que envolve a triagem de diferentes lesões de pele, incluindo as de mais baixo risco.

Além disso, Chadwick *et al.* [5] avaliaram 5 dos principais aplicativos de detecção de câncer de pele disponíveis em lojas virtuais. Eles utilizaram 15 amostras com diagnóstico histopatológico conhecido para averiguar se o risco indicado pelos aplicativos condiz com a amostra analisada. A conclusão do estudo realizado é de que nenhum dos aplicativos disponíveis testados encontrava-se apto para ser utilizado em triagem de lesões de pele, uma vez que eram inconsistentes e possuíam baixo desempenho, além de apresentarem limitações quanto à avaliação de lesões não pigmentadas, como é o caso do câncer de pele carcinoma espinocelular.

Neste trabalho, como já mencionado, o objetivo é mais abrangente e envolve a validação da aplicação. O modelo proposto será avaliado não somente em relação aos dados de validação durante o treinamento, como também seu desempenho será avaliado após o uso em campo com imagens de amostras capturadas diretamente pelo aplicativo. Além disso, diferentemente da maioria dos trabalhos supracitados, a aplicação é desenvolvida para funcionar totalmente offline, o que é um requisito importante dado a realidade brasileira.

3 MÉTODOS

Nesta seção são apresentados os principais elementos para a construção da aplicação proposta. Primeiramente, o problema de triagem é modelado e os dados utilizados são descritos. Na sequência, o modelo de aprendizado profundo desenvolvido é apresentado e, posteriormente, é descrito como o desenvolvimento do aplicativo e como a implantação do modelo na aplicação foram realizados.

3.1 Modelagem do problema

As doenças de pele habitualmente retratadas em sistemas CAD de dermatologia são os cânceres de pele, melanoma, carcinoma espinocelular e carcinoma basocelular, sendo o melanoma o mais grave entre esses, e doenças como ceratose actínica, ceratose seborréica, nevo melanocítico (principalmente pois se assemelha bastante a um melanoma), demais nevos, dermatofibroma e lesão vascular [2, 9, 19, 20]. De acordo com as orientações dos especialistas em dermatologia do PAD-UFES, as lesões de pele foram agrupadas de acordo com suas semelhanças, riscos e a realidade da lesão no estado do Espírito Santo.

Desse modo, agruparam-se os tipos de lesões de pele selecionados em 5 níveis de gravidade, que vão desde uma pele saudável e/ou acneica até o câncer de pele mais grave (melanoma). A Tabela 1 mostra como ficou o agrupamento de doenças e seus respectivos CID (Classificação Internacional de Doenças) por prioridade e descrição macro de cada grupo. Assim, a cada imagem de lesão avaliada pela aplicação, o modelo de aprendizado profundo retorna como resposta um grau de gravidade, que pode ser **P1**, **P2**, **P3**, **P4** ou **P5**, sendo P1 a que demanda maior atenção.

Tabela 1: Descrição prioridades das lesões de pele

Grau de Prioridade	Possíveis diagnósticos	Descrição
P1	Melanoma (C43), Lentigo maligno (D03) ou Nevo melanocítico (D22).	A lesão pode ser perigosa e possui indícios de malignidade. Paciente precisa do maior nível de prioridade no atendimento.
P2	Carcinoma Basocelular (C80), Carcinoma Espinocelular (C44), Doença de Bowen (D04), Ceratoacantoma (D23) ou Corno Cutâneo (L75)	A lesão possui sinais de um tumor de pele ou pode se tornar um no futuro. O paciente precisa de um nível alto de atenção.
P3	Ceratose actínica (L57), Eczema (L30) ou Dermatite (L25, L30, L98)	A lesão possui sinais de uma doença de pele que pode ser tratada sem uma intervenção mais agressiva. O paciente precisa de nível médio de atenção.
P4	Ceratose Seborreica (L82) ou Nevo (L78)	A lesão aparenta ser um problema de pele mais simples de ser tratado. O paciente precisa de um nível baixo de atenção.
P5	Ausência de lesão ou caso muito simples (ex: acne)	A lesão é ausente ou muito simples, sem necessidade de atenção prioritária.

Tabela 2: Distribuição das Amostras

Prioridade	Quantidade
P1	624
P2	4.545
P3	5.979
P4	605
P5	1.816
Total	13.569

3.2 Dados

Para o treinamento do modelo de aprendizado profundo, utilizou-se os dados coletados dentro do próprio PAD-UFES através do SADE [18], software de gerenciamento e coleta implementado especificamente para o projeto em questão. Até o momento, essa base de dados com imagens de doenças de pele é privada e é uma extensão da base PAD-UFES-20 [22]. Para as imagens de pele saudável e pele acneica, as quais não possuíam representação no *dataset* privado, optou-se pelo uso de bases públicas [8, 13]. No total, 13.569 imagens foram utilizadas. A distribuição das imagens de acordo com a prioridade pode ser observada na Tabela 2.

3.3 Arquitetura

Em trabalhos anterior

O modelo de inteligência artificial desenvolvido é uma rede neural convolutiva que baseia-se na arquitetura de *deep learning* conhecida como **MobileNetV2** [14]. Essa arquitetura foi selecionada por apresentar bons resultados em treinamentos realizados em estudos anteriores [21, 28], além de se destacar por sua arquitetura leve e por ser viável de implantar em dispositivos móveis. Utilizou-se da técnica de *fine-tuning* para aproveitar os pesos da rede pré-treinada na ImageNet [7] como ponto de partida para o treinamento. A

Tabela 3: Desempenho do modelo de acordo com a validação cruzada

Partição	Acurácia balanceada	Área sob a curva
Folder 1	0,810	0,954
Folder 2	0,809	0,957
Folder 3	0,789	0,949
Folder 4	0,773	0,941
Folder 5	0,783	0,947
Média	0,800	0,950
Desvio padrão	0,01	0,01

implementação foi feita utilizando-se o framework PyTorch ¹. O treinamento foi estipulado para ocorrer durante 120 épocas, mas com parada antecipada programada caso 15 épocas consecutivas ocorram sem melhora de desempenho. O otimizador SGD foi utilizado com uma taxa de aprendizado 0,0001. Todas as imagens foram redimensionadas para o tamanho 224 × 224 e técnicas padrão de *data augmentation* foram esporadicamente utilizadas, como, por exemplo, rotações horizontais e verticais, mudanças na saturação das imagens e a eliminação de pixels.

Como mencionado anteriormente, o classificador construído possui 5 possibilidades de resultados, os quais são referentes aos 5 níveis de prioridade que podem ser atribuídos a uma lesão. Como imagens do melanoma tendem a uma distribuição menor que as demais classes no dataset, fato que reflete sua ocorrência real na natureza [16, 23, 27], ajustes foram realizados de modo a minimizar falsos negativos para a prioridade 1, classe que contempla esse diagnóstico. Para isso, ponderaram-se os pesos das punições de forma inversamente proporcional ao número de imagens. Assim, em caso de erros na classificação da P1, o modelo será punido de forma mais rígida.

Para avaliação do modelo, utilizou-se a técnica de *cross-validation* para 5 partições. As métricas utilizadas foram acurácia balanceada e área sob a curva, que obtiveram valores respectivos de **0,80 +- 0,01** e **0,95 +- 0,01**. Os resultados obtidos encontram-se na Tabela 3.

3.4 Desenvolvimento da aplicação

Após a conclusão da construção de um modelo de classificação eficiente, deu-se início à próxima etapa - que é onde embarcar o modelo. Optou-se por desenvolver o aplicativo em Android usando Kotlin, uma vez que cerca de 90% dos brasileiros utilizam esse sistema operacional em seus dispositivos mobile [29], o que garante uma cobertura significativa. Nesta seção, serão descritas as etapas de planejamento e desenvolvimento do aplicativo.

A versão 8.0 do Android foi utilizada por ser aceita em cerca de 94% dos aparelhos deste sistema e ainda ter compatibilidade com as versões mais recentes. Seguindo as recomendações de implementação da Google, empresa responsável pelo suporte e atualização do Android, fez-se uso da arquitetura MVVM (*Model - ViewModel - View*) para garantir uma separação clara entre a interface do usuário, a lógica de apresentação e a lógica de dados [12].

Para o gerenciamento de sessões de usuário, utilizou-se a API do Google Firebase Authentication [10], a qual ficou encarregada de

¹Disponível em <https://pytorch.org/>.

cadastrar e, posteriormente, autenticar os usuários. Apesar do uso de internet ser necessário para a criação e/ou login do usuário, uma vez que a sessão foi iniciada, a menos que o usuário faça logout, o Firebase mantém a sessão aberta, o que permite o uso do aplicativo mesmo não tendo acesso a nenhuma rede de internet após o primeiro acesso ao aplicativo. Essa característica é fundamental, pois a aplicação deve permitir que agentes de saúde em áreas remotas consigam realizar a triagem mesmo não estando em regiões com cobertura de internet. Por essa mesma razão, optou-se por embarcar o modelo diretamente no aplicativo ao invés de acessá-lo remotamente por meio de API toda vez que uma predição precisasse ser realizada.

Para embarcar o modelo no Android, utilizou-se o padrão de representação ONNX², um formato aberto para troca e execução de modelos de aprendizado de máquina. O modelo treinado foi exportado para o formato ONNX por meio da função de conversão disponibilizada pela biblioteca do framework PyTorch. Em seguida, após configurar as dependências necessárias do ONNX Runtime no Android, foi empregada a API da biblioteca para realizar as inferências com o modelo carregado. No final, o modelo a ser embarcado ocupa apenas cerca de 8MB do total da aplicação (que possui 61 MB), o que a torna viável para funcionar em qualquer dispositivo mais recente.

3.5 Fluxo de telas

O primeiro contato dos usuários³ com o aplicativo é a tela de login, que pode ser observada na parte (a) da Figura 1. Em seguida, após efetuar o login, a aplicação apresenta a tela principal do aplicativo, na qual o usuário pode conferir uma contagem total de avaliações feitas por ele, assim como um gráfico mostrando a distribuição dessas avaliações por prioridade, como pode ser observado na parte (b) da Figura 1.

Ao selecionar a opção **Triagem** na barra de navegação inferior, uma tela com as instruções gerais da triagem será exibida. Além disso, esta tela conta com mais um alerta ao usuário sobre o uso adequado da aplicação, que deve servir apenas como uma ferramenta de suporte e não deve ser usada pelo público geral. Informações gerais de como realizar a triagem também estão presentes. A tela de instruções sobre a triagem está representada na Figura 2, parte (a).

Após clicar no botão **Iniciar triagem** na tela de informações, o paciente deverá responder ao usuário uma pergunta simples sobre a origem da lesão a ser avaliada e o usuário irá inserir essa resposta no aplicativo. Caso ela tenha surgido por uma causa conhecida, como uma pancada ou picada de inseto, a lesão não se enquadra no perfil de lesões que o modelo foi treinado para reconhecer. Por isso, em casos nos quais o paciente conhece a precedência da lesão, o aplicativo informa a este que a triagem não poderá ser realizada para essa circunstância, visto que a lesão não se encontra entre o quadro de diagnósticos contemplados pelo Software de Triagem. A Figura 2, parte (b), ilustra essa situação.

Por fim, caso o paciente não conheça a origem da lesão, o usuário terá duas opções: selecionar uma imagem da galeria do mobile ou

²Disponível em <https://onnx.ai/>.

³Destaca-se, novamente, que o usuário da aplicação é um profissional de saúde. Logo, sempre que *usuário* for mencionado é uma referência a este profissional e não ao público geral



Figura 1: (a) Tela de login e (b) Tela principal - Home.

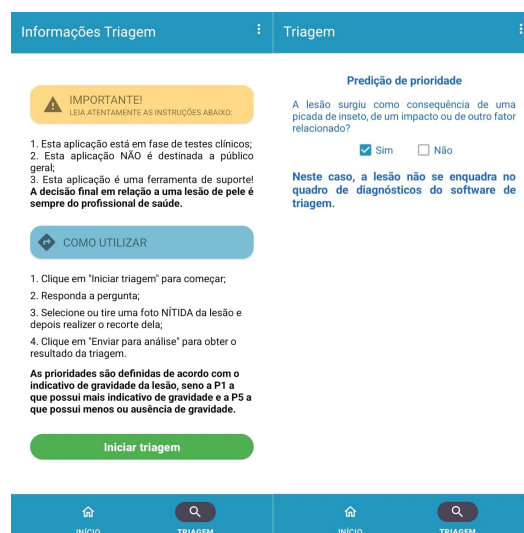


Figura 2: (a) Tela de instruções sobre a triagem e (b) Tela de triagem - origem de lesão de pele conhecida.

abrir a câmera e capturar uma foto. Após a seleção da imagem, aparecerá a opção de cortar a imagem enviada de modo a enquadrar na imagem, da melhor forma possível, a lesão que se deseja analisar. Após feito isso, o botão **Realizar triagem** estará disponível, e uma vez que o usuário clicar nele, a triagem será realizada. As informações sobre o grau de prioridade atribuído àquela lesão, assim como os possíveis diagnósticos e grau de confiança do modelo naquele resultado, serão exibidos. A Figura 3 representa as etapas de seleção de imagem e predição, respectivamente. A figura apresentada é apenas uma exemplificação de um resultado possível. As telas variam de acordo com a prioridade atribuída à lesão, sendo que cada prioridade possui uma cor atribuída a si. Como visto, o resultado P1, que é o mais grave, é representado pela cor vermelha.

P2 é representado pela cor laranja, P3 pelo amarelo, P4 pela cor azul e P5 pela cor verde. O padrão de cores escolhido possui como objetivo transmitir visualmente a predição dada pelo Software de Triagem.

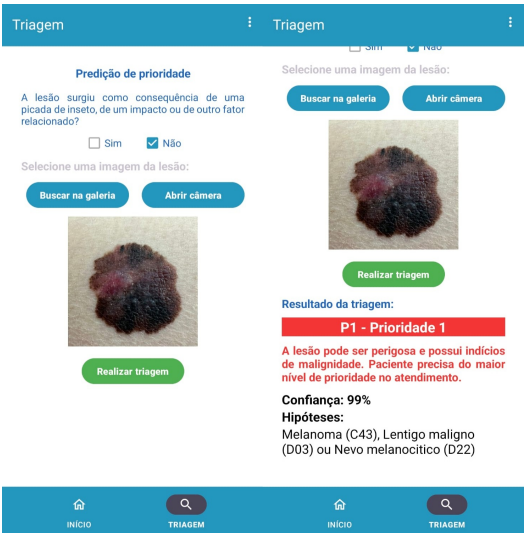


Figura 3: Tela de triagem – seleção imagem e predição grau de prioridade. A lesão apresentada é um melanoma⁴e esta amostra não foi vista pelo modelo durante o seu treinamento.

4 ESTUDO DE CASO

Esta seção descreve como o estudo de caso do Software de Triagem foi realizado, incluindo seu protocolo de estudo e os resultados obtidos a partir da avaliação do modelo de aprendizado profundo desenvolvido e da opinião de uso dos usuários.

4.1 Protocolo de estudo

Para avaliar o desempenho da aplicação descrita neste trabalho, tanto em termos de desempenho do modelo de triagem, quanto na avaliação de uso dos profissionais de saúde, foi realizado um estudo de caso ao longo do ano de 2024. A pesquisa ocorreu em diferentes cidades atendidas pelo PAD-UFES.

Durante os atendimentos, diversos profissionais da saúde estão presentes, cada um com diferentes níveis de experiência na área de dermatologia. Desse modo, para avaliar o potencial de uso do aplicativo, escolheu-se como principal público-alvo para a pesquisa agentes de saúde que atuam em postos de unidades básicas de saúde e estudantes de medicina que ainda não cursaram disciplinas relacionadas à dermatologia. O objetivo dessa escolha é avaliar como o uso do software impacta a decisão daqueles profissionais que possuem pouca/ninguma experiência em dermatologia. Porém, aqueles com experiência na área também foram incluídos na pesquisa, com o intuito de comparar as respostas de modo a avaliar se a experiência em dermatologia afetava a opinião de uso do usuário. Destaca-se que a pesquisa foi totalmente anônima para os voluntários e eles concordaram previamente com os termos de sua participação.

⁴A imagem encontra-se em domínio público e está disponível em <https://www.padrao.com.br/blog/melanoma-maligno>.

Tabela 4: Distribuição voluntários por cidade

Cidade	Qtd voluntários
Baixo Guandu	13
Domingos Martins	11
Laranja da Terra	17
Pancas	4
Santa Maria de Jetibá	11
São Gabriel da Palha	13
Vila Pavão	6
Vila Valério	7
Total	82

Tabela 5: Nível de experiência em dermatologia dos voluntários

Nível de experiência	Descrição	Nº de voluntários
Nível 1	Sem experiência	9
Nível 2	Estudou um pouco sobre, mas tem pouca experiência	33
Nível 3	Estudou sobre e consegue realizar alguns diagnósticos	28
Nível 4	Estudou e tem certa experiência na área	9
Nível 5	Possui mais de 3 anos de experiência na área	3

As cidades do interior do Espírito Santo que participaram do estudo de caso e a quantidade de voluntários recrutados em cada uma delas podem ser verificadas na Tabela 4. A experiência dos voluntários na área de dermatologia foi classificada em 5 níveis, sendo o primeiro nível referente àqueles sem experiência e o último dedicado àqueles com mais de 5 anos trabalhando diretamente com essa especialidade. A distribuição dos voluntários de acordo com os níveis de experiência na área de dermatologia e a descrição de cada nível estão presentes na Tabela 5. Os níveis 1 e 2 são majoritariamente formados por agentes de saúde e estudantes de Medicina que ainda não cursaram disciplinas de dermatologia. Os níveis 3 e 4 consistem basicamente em estudantes com alguma experiência prévia na área ou médicos generalistas, e o nível 5 são profissionais especializados em dermatologia.

A pesquisa que fomentou o estudo de caso ocorreu com o suporte de um aplicativo *low-code* feito com o uso da ferramenta PowerApps da Microsoft. Um tablet com a aplicação era entregue ao voluntário para que ele pudesse responder às perguntas do questionário. A pesquisa ocorria de acordo com as seguintes etapas:

- (1) Escolha de voluntário que enquadrava-se no perfil anteriormente descrito.
- (2) O voluntário escolhia um paciente para ser avaliado.
- (3) O voluntário anonimamente respondia sobre seu nível de experiência em dermatologia e sua profissão no aplicativo com o questionário. Além disso, o número do SUS do paciente e informações básicas da lesão, como a região anatômica, eram mantidas para comparação com o diagnóstico clínico posterior.
- (4) O paciente apresentava a lesão pela qual foi encaminhado para o atendimento e o voluntário era responsável por avaliá-la e responder no questionário qual era o nível de gravidade

que ele atribuía aquela lesão e uma possível hipótese diagnóstica. Essa etapa ocorria **sem** o uso do aplicativo de triagem.

- (5) Agora, utilizando a aplicação de triagem, com o uso do Software de Triagem, uma foto da lesão era capturada e o modelo de inteligência artificial a classificava quanto ao grau de gravidade. O resultado obtido, assim como os possíveis diagnósticos, eram apresentados ao voluntário.
- (6) Com conhecimento sobre o resultado da triagem realizada pelo modelo desenvolvido, o voluntário respondia novamente no questionário qual a gravidade que ele atribuía aquela lesão e uma hipótese diagnóstica.
- (7) Ao finalizar a avaliação do paciente, o voluntário respondia as seguintes perguntas (com respostas fechadas) sobre sua experiência de uso e opiniões pessoais:
 - Qual sua opinião em relação ao uso deste Software para triagem de lesões de pele?
 - Você acredita que o Software foi útil para você?
 - Qual seu sentimento em relação a utilizar um Software para auxiliar na triagem de lesões?
 - Você é a favor de que um Software como este seja implantado no SUS para auxiliar profissionais de saúde na triagem de lesões em unidades em que não existem dermatologistas disponíveis?

Ao final do estudo de caso, 82 voluntários avaliaram um total de 101 lesões em 83 pacientes. Existia a possibilidade de um voluntário atender a mais de um paciente, o qual poderia ter mais de uma lesão.

4.2 Desempenho do modelo

Com as informações coletadas no questionário, foi possível mapear as lesões avaliadas durante a triagem para as respectivas lesões cadastradas durante a consulta com um especialista no sistema utilizado nos atendimentos [18]. Contudo, como alguns pacientes foram liberados do atendimento antes de serem cadastrados no sistema, 72,82% das lesões foram identificadas.

A Figura 4 apresenta a distribuição de resultados da aplicação ao contabilizar apenas as lesões identificadas. Nela, os resultados das predições estão divididos em três grupos: acerto, que representa quando o diagnóstico dado pelo especialista coincide com a prioridade atribuída pela aplicação; erro (priorizando), o diagnóstico clínico não coincide com nenhum dos diagnósticos possíveis associados à predição acusada e a prioridade atribuída indica que a lesão é mais grave do que ela realmente é, o que gera uma prioridade maior do que o paciente realmente precisa; erro (não priorizando), neste caso, além do diagnóstico clínico não se enquadrar nos diagnósticos possíveis da prioridade atribuída pela aplicação, a predição feita indica que a lesão é menos grave do que ela realmente é, o que gera uma prioridade de atendimento menor do que o paciente necessita.

Percebe-se que, ao comparar a prioridade atribuída a uma lesão pelo modelo com o diagnóstico posteriormente vinculado a ela pelo médico dermatologista, em torno de 71% das vezes o software acertava a prioridade de acordo com o diagnóstico e em aproximadamente 17% ele atribuía um nível de gravidade acima do real diagnóstico. Os resultados caracterizados como falsos negativos, que foram 12%, não desincentivam o atendimento do paciente, fato que suaviza esse percentual de equívoco. Todavia, nesses casos,

destaca-se que a intervenção do usuário é essencial para garantir que o paciente seja atendido de forma adequada. O papel da aplicação é fornecer subsídio; quem decide o desfecho é sempre o profissional de saúde. Em caso de dúvidas, o usuário deve, sempre que possível, encaminhar para um atendimento especializado. Ainda assim, os resultados corroboram para o potencial do software como ferramenta de subsídio a profissionais em saúde não especialistas.

Ademais, a diferença entre a acurácia balanceada média alcançada durante os treinos, que foi de cerca de 80%, e a taxa de acerto obtida em campo, cerca de 71%, está fortemente relacionada com a falta de lesões de pele saudável ou acneica, classificação P5, nas avaliações realizadas. Com os resultados obtidos durante o treinamento do modelo, se forem desconsideradas as amostras de prioridade 5, é possível estimar o nível médio de acerto com exceção desta classe. Neste caso, a acurácia balanceada média é de 73%. Este novo resultado condiz com os resultados obtidos pelo estudo de caso.

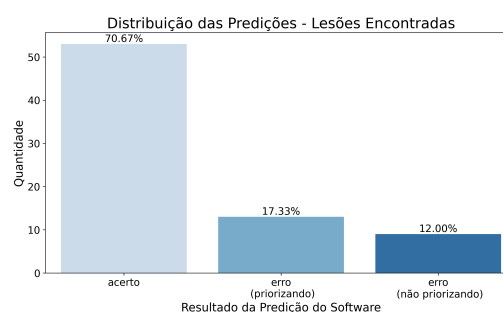


Figura 4: Desempenho do modelo de triagem.

4.3 Avaliação do usuário

O aplicativo desenvolvido tem como objetivo auxiliar profissionais da saúde sem experiência com a área de dermatologia a realizar uma triagem mais eficiente e, assim, potencializar as chances de diagnóstico precoce do câncer de pele. Porém, para que isso ocorra, é de suma importância que, primeiramente, os usuários se sintam confortáveis em utilizar a aplicação. Por isso, a primeira pergunta da pesquisa pós-utilização do aplicativo era **Qual sua opinião em relação ao uso deste Software para triagem de lesões de pele?** As respostas possíveis e os resultados obtidos podem ser observados na Figura 5. Percebe-se que, mesmo tendo opções que expressem uma experiência negativa de uso, todas as respostas obtidas indicam o contentamento do usuário com a aplicação.

Todavia, opiniões positivas não necessariamente indicam que o software foi útil para a realização da triagem. Desse modo, a pergunta **Você acredita que o Software foi útil para você?** também foi adicionada ao questionário. As respostas estão presentes na Figura 6. O gráfico demonstra que os usuários viram utilidade no uso do aplicativo durante a triagem. Além disso, dos 82 voluntários entrevistados, apenas 5 acreditam que o software não possua alguma utilidade. Porém, ao analisar o perfil desses voluntários, percebeu-se que 4 deles se enquadram nos níveis de 3 a 5 em relação à experiência em dermatologia – Tabela 5. Esse é um comportamento esperado, uma vez que, para profissionais com conhecimento nesta área, o software de triagem é visto mais como uma ferramenta de

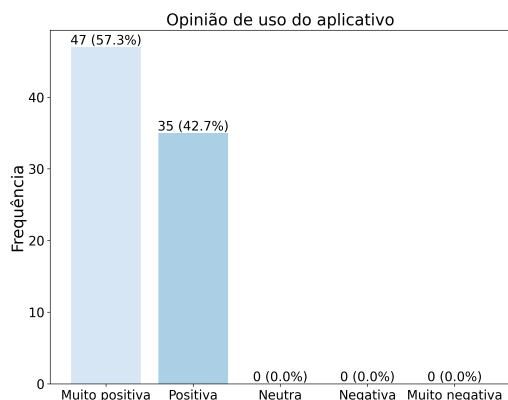


Figura 5: Opinião do usuário sobre o uso do aplicativo.

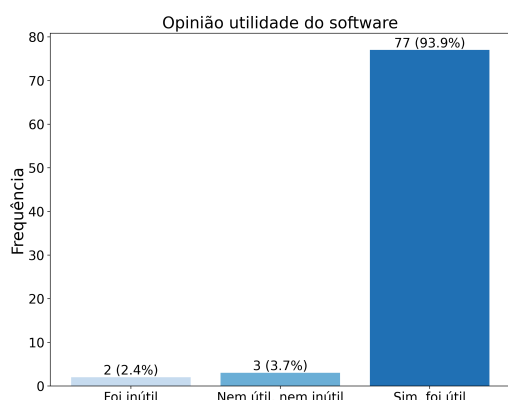


Figura 6: Opinião do usuário em relação a utilidade do software de triagem.

confirmação do que como uma ferramenta de suporte/apoio. Para quem consegue realizar diagnósticos clínicos de lesões de pele, triar os pacientes de acordo com a gravidade das lesões torna-se uma atividade relativamente trivial.

Ademais, buscou-se explorar qual percepção o uso do aplicativo durante a triagem trouxe para o usuário. O voluntário poderia escolher entre uma série de palavras qual melhor representava o seu "sentimento" em relação à utilização do Software de Triagem. A pergunta feita foi **Qual seu sentimento em relação a utilizar um Software para auxiliar na triagem de lesões?** e as palavras disponíveis como respostas eram: conforto, segurança, confirmação, precisão, neutralidade, dúvida, desconforto, insegurança e incerteza. A Figura 7 mostra quais foram as palavras mais escolhidas para representar a experiência com relação ao software. A ideia da aplicação ser utilizada como uma ferramenta de auxílio a profissionais da saúde é reforçada ao perceber que as palavras que mais se destacam são **Segurança, Confirmação e Conforto**.

Contudo, pode-se fazer um adendo interessante sobre como a percepção do aplicativo varia de acordo com o nível de experiência na área de dermatologia. Para verificar este comportamento, os usuários foram estratificados em dois grupos, sendo o grupo 1 aqueles com nenhuma ou pouca experiência em dermatologia (níveis 1

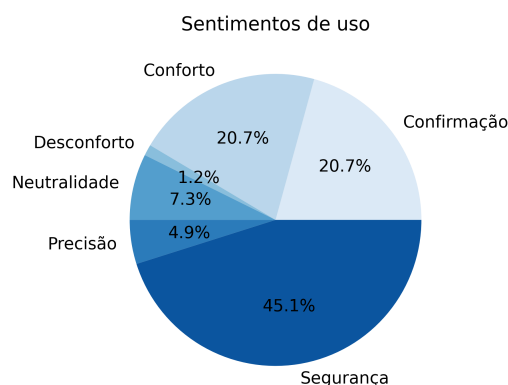


Figura 7: Sentimentos relacionados ao uso do aplicativo.

e 2 da Tabela 5) e o grupo 2 aqueles que, no mínimo, conseguem realizar alguns diagnósticos (níveis 3 - 5 da Tabela 5). A Figura 8 mostra a seleção de palavras de cada um dos grupos.

Observa-se que para aqueles com menos conhecimento em dermatologia, o sentimento de ter uma ferramenta de auxílio no momento de identificar quais pacientes precisam ser priorizados no atendimento é majoritariamente de **Segurança**, com uma taxa de 60% de escolha. Essa palavra aparece quase que o dobro de vezes para aqueles do Grupo 1 quando comparados com o Grupo 2. Uma possível causa dessa prevalência nítida do sentimento é o fato de que a falta de domínio na área de dermatologia gera insegurança e receio a esses profissionais no momento de avaliar uma lesão de pele. Apesar de realizarem seu trabalho com competência, agentes de saúde recebem treinamento apenas sobre noções básicas de saúde, o que, quase sempre, não inclui os conhecimentos necessários para a identificação do câncer de pele. Exemplificações disso ocorreram durante a seleção de voluntários para a participação na pesquisa, pois, apesar do estudo ser realizado de forma anônima, muitos não se sentiam confortáveis e preferiam não participar porque tinham receio de errar a triagem. Isso demonstra ainda mais o potencial que o Software de Triagem possui de não apenas ajudar na identificação do câncer de pele, mas como também de suavizar os desafios enfrentados por esses profissionais.

Além disso, para o grupo 2, a palavra **Confirmação** foi mencionada o dobro de vezes quando comparado ao grupo 1. Isso fortalece o que foi mencionado anteriormente sobre usuários que sejam especializados na área utilizarem o aplicativo principalmente como uma ferramenta de confirmação ao invés de suporte e apoio.

A resistência contra o uso de inteligência artificial na área da saúde é um desafio real que as inovações nessa área precisam enfrentar [30]. Não apenas por parte dos médicos que relutam em compartilhar o seu espaço de trabalho com essa tecnologia emergente, como também por parte dos pacientes que, por vezes, descredita a capacidade de ajuda fornecida pela IA. Por isso, a última pergunta realizada na pesquisa foi **"Você é a favor de que um Software como este seja implantado no SUS para auxiliar profissionais de saúde na triagem de lesões em unidades em que não existem dermatologistas disponíveis?"**. Queria-se avaliar, em um futuro hipotético, qual seria a recepção dos profissionais

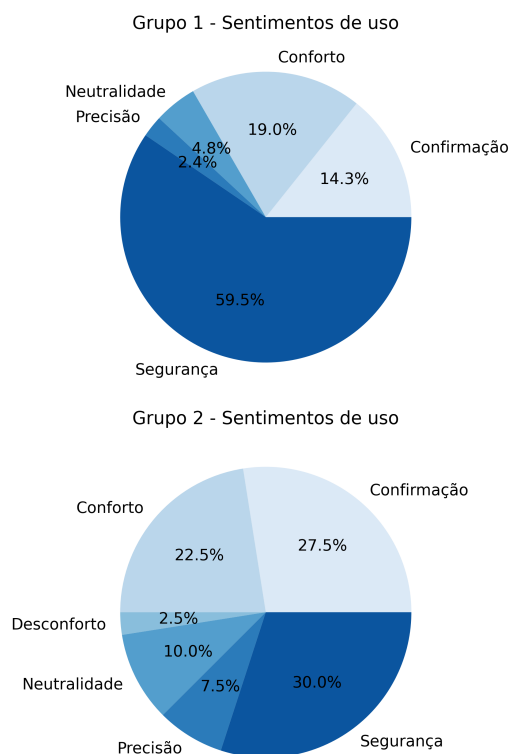


Figura 8: Sentimentos de uso do aplicativo de acordo com o nível de experiência em dermatologia. Grupo 1: Níveis 1 e 2, Grupo 2: Níveis 3 - 5.

da área ao receber no Sistema Único de Saúde um sistema como este, uma vez que isso implica uma convivência constante com tal ferramenta. A pergunta evidencia que a ideia por trás desse software não é substituir um profissional de saúde qualificado, apenas auxiliar quando um não estiver presente. Os resultados apresentados na Figura 9 foram bastante promissores quanto ao apoio de trabalhadores da área da saúde à inclusão do Software de Triagem em unidades de saúde carentes de dermatologistas.

5 CONCLUSÃO

Este trabalho teve como objetivo avaliar o desempenho do modelo de triagem de lesões de pele desenvolvido em campo. A recepção dos usuários, que são profissionais da saúde, também foi avaliada. Apesar do modelo de inteligência artificial possuir resultados teóricos positivos, com acurácia balanceada de aproximadamente 80%, um estudo de caso foi realizado para que fosse possível estudar como este modelo se comportaria na prática.

A pesquisa realizada demonstrou o alto potencial do Software de Triagem como ferramenta de auxílio a profissionais de saúde não especialistas, principalmente aqueles com conhecimento apenas sobre noções básicas de saúde, como é o caso de agentes de saúde. O modelo acertou a prioridade de cerca de 71% das lesões analisadas durante o estudo e que, posteriormente, receberam diagnósticos de especialistas. Esse fato afirma a capacidade do aplicativo em

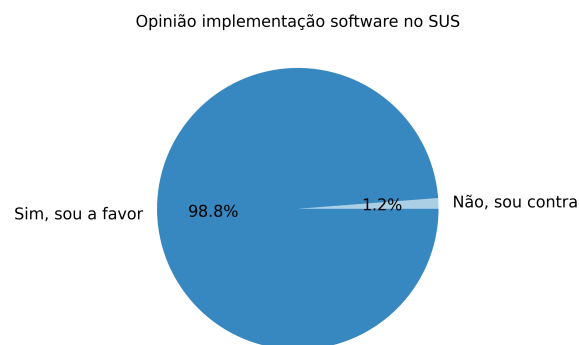


Figura 9: Opinião do usuário sobre a possibilidade de implantação do sistema no Sistema Único de Saúde (SUS).

potencializar a identificação do câncer de pele se utilizado em campo em conjunto com os profissionais da saúde.

Além disso, o estudo de caso revelou que profissionais da área da saúde com menos experiência em dermatologia se sentem mais seguros ao realizar a triagem de pacientes enquanto são auxiliados pelo aplicativo desenvolvido. Para aqueles com mais experiência em dermatologia, a percepção também é positiva. Porém, ao invés de usarem o aplicativo como suporte, a visão desses profissionais para com o Software de Triagem é a de uma ferramenta de confirmação, que possa servir como uma segunda opinião em caso de dúvida.

O estudo de caso realizado indica a capacidade do aplicativo de triagem de ser utilizado por profissionais de saúde não especialistas como forma de auxiliá-los na identificação precoce do câncer de pele e de outras doenças. Como, infelizmente, nem todos os pacientes podem ser atendidos por um profissional especialista em dermatologia, a solução é melhorar a triagem feita por generalistas. E a aplicação mostrou-se uma excelente ferramenta para tal tarefa.

Em trabalhos futuros, planeja-se investigar se o uso de imagens dermatoscópicas ao invés de imagens clínicas pode melhorar o desempenho do modelo de triagem. O uso do dermatoscópio possibilita uma imagem bem mais detalhada da lesão, o que permite que características não perceptíveis aos olhos sejam observadas.

DECLARAÇÃO DE ÉTICA

Este estudo foi aprovado pelo Comitê de Ética da Universidade Federal do Espírito Santo e pelo governo brasileiro por meio da Plataforma Brasil (CAEE nº 78728324.3.0000.5060). Todos os dados foram coletados com o consentimento dos pacientes.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES); o Instituto Capi-xaba de Ensino, Pesquisa e Inovação em Saúde (ICEPi); o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); o Ministério da Saúde (MS); e o Programa Nacional de Genômica e Saúde de Precisão (Genomas Brasil).

REFERÊNCIAS

- [1] Ministério da Saúde 2025. *Câncer de pele*. Ministério da Saúde. Retrieved July 19, 2025 from <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/cancer-de-pele>
- [2] Olusoji Akinrinade and Chunglin Du. 2025. Skin cancer detection using deep machine learning techniques. *Intelligence-Based Medicine* 11 (2025), 100191. <https://doi.org/10.1016/j.ibmed.2024.100191>
- [3] AMB. 2023. Demografia Médica no Brasil 2023. Associação Médica Brasileira (AMB). Disponível em: https://amb.org.br/wp-content/uploads/2023/02/DemografiaMedica2023_8fev-1.pdf. Último acesso em: 05 de Julho 2025.
- [4] Pedro B. C. Castro, Breno Krohling, Andre G. C. Pacheco, and Renato A. Krohling. 2020. An app to detect melanoma using deep learning: An approach to handle imbalanced data based on evolutionary algorithms. In *2020 International Joint Conference on Neural Networks (IJCNN)*. 1–6. <https://doi.org/10.1109/IJCNN48605.2020.9207552>
- [5] Xavier Chadwick, Lois J. Loescher, Monika Janda, and H. Peter Soyer. 2014. Mobile Medical Applications for Melanoma Risk Assessment: False Assurance or Valuable Tool?. In *2014 47th Hawaii International Conference on System Sciences*. 2675–2684. <https://doi.org/10.1109/HICSS.2014.337>
- [6] Sociedade Brasileira de Dermatologia. 2025. Câncer da Pele. SBD. Disponível em: <https://www.sbd.org.br/doencas/cancer-da-pele/>. Último acesso em: 19 de Julho 2025.
- [7] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [8] DermNet. 2025. *The world's leading free dermatology website*. <https://dermnetnz.org/>. Acesso em: 19 jul. 2025.
- [9] Andre Esteva, Brett Kuppel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 7639 (2017), 115.
- [10] Google Firebase 2025. *Firebase Authentication*. Google Firebase. Retrieved July 09, 2025 from <https://firebase.google.com/docs/auth>
- [11] Skin Cancer Foundation. 2025. Skin Cancer Facts & Statistics. Skin Cancer Foundation. Disponível em: <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/>. Último acesso em: 05 de Julho 2025.
- [12] Google. 2025. Guide to app architecture. Google for Developers. Disponível em: <https://developer.android.com/topic/architecture>. Último acesso em: 10 de Julho 2025.
- [13] Matthew Groh, Caleb Harris, Luis Soenksen, Felix Lau, Rachel Han, Aerin Kim, Arash Koochek, and Omar Badri. 2021. Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset. arXiv:2104.09957 [cs.CV] <https://arxiv.org/abs/2104.09957>
- [14] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv:1704.04861 [cs.CV] <https://arxiv.org/abs/1704.04861>
- [15] INCA. 2023. Incidência do câncer no Brasil. Instituto Nacional do Câncer (INCA). Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf>. Último acesso em: 16 de Junho 2024.
- [16] National Cancer Institute. 2025. Cancer Stat Facts: Melanoma of the Skin. National Cancer Institute. Disponível em: <https://seer.cancer.gov/statfacts/html/melan.html>. Último acesso em: 09 de Julho 2025.
- [17] Muhammad Attique Khan, Khan Muhammad, Muhammad Sharif, Tallha Akram, and Victor Hugo C. de Albuquerque. 2021. Multi-Class Skin Lesion Detection and Classification via Teledermatology. *IEEE Journal of Biomedical and Health Informatics* 25, 12 (2021), 4267–4275. <https://doi.org/10.1109/JBHI.2021.3067789>
- [18] Andre Pacheco, Clayton Vicente, Eduarda Magesk, Gabriel Lucas, Guilherme Caldana, and Patricia Frasson. 2023. SADE: Software de Análise Dermatológica - Um sistema de coleta, gerenciamento e triagem de lesões de pele. In *Anais Estendidos do XXIX Simpósio Brasileiro de Sistemas Multimídia e Web (Ribeirão Preto/SP)*. SBC, Porto Alegre, RS, Brasil, 111–114. https://doi.org/10.5753/webmedia_estendido.2023.235905
- [19] Andre GC Pacheco, Abder-Rahman Ali, and Thomas Trappenberg. 2019. Skin cancer detection based on deep learning and entropy to detect outlier samples. In *Medical Image Computing and Computer Assisted Intervention (MICCAI) at Skin Lesion Analysis Towards Melanoma Detection (ISIC) challenge*. 1–6.
- [20] Andre GC Pacheco and Renato A Krohling. 2019. Recent advances in deep learning applied to skin cancer detection. In *Neural Information Processing Systems at Retrospectives workshop*. 1–8.
- [21] Andre GC Pacheco and Renato A Krohling. 2021. An attention-based mechanism to combine images and metadata in deep learning models applied to skin cancer classification. *IEEE journal of biomedical and health informatics* 25, 9 (2021), 3554–3563.
- [22] Andre GC Pacheco, Gustavo R Lima, Amanda S Salomão, Breno Krohling, Igor P Biral, Gabriel G de Angelo, Fábio CR Alves Jr, José GM Esgario, Alana C Simora, Pedro BC Castro, et al. 2020. PAD-UFES-20: a skin lesion dataset composed of patient data and clinical images collected from smartphones. *Data in Brief* 32 (2020), 1–10.
- [23] Amdad Hossain Roky, Mohammed Murshedul Islam, Abu Mohammed Fuad Ahasan, Md Saqline Mostaq, Md Zihad Mahmud, Mohammad Nurul Amin, and Md Ashiq Mahmud. 2025. Overview of skin cancer types and prevalence rates across continents. *Cancer Pathogenesis and Therapy* 3, 2 (2025), 89–100. <https://doi.org/10.1016/j.cpt.2024.08.002>
- [24] Luis Rosado, Maria João M. Vasconcelos, Fernando Correia, and Nuno Costa. 2015. A novel framework for supervised mobile assessment and risk triage of skin lesions. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. 266–267. <https://doi.org/10.4108/icst.pervasivehealth.2015.259254>
- [25] Marwen Sakl, Chaker Essid, Bassem Ben Salah, and Hedi Sakli. 2023. DL Methods for Skin Lesions Automated Diagnosis In Smartphone Images. In *2023 International Wireless Communications and Mobile Computing (IWCMC)*. 1142–1147. <https://doi.org/10.1109/IWCMC58020.2023.10183254>
- [26] Tobias Sangers, Suzan Reeder, Sophie van der Vet, Sharan Jhingoer, Antien Mooyaart, Daniel M. Siegel, Tamar Nijsten, and Marlies Wakkee. 2022. Validation of a Market-Approved Artificial Intelligence Mobile Health App for Skin Cancer Screening: A Prospective Multicenter Diagnostic Accuracy Study. *Dermatology* 238, 4 (Feb. 2022), 649–656. <https://doi.org/10.1159/000520474> _eprint: <https://karger.com/drm/article-pdf/238/4/649/3750042/000520474.pdf>
- [27] American Cancer Society. 2025. Key Statistics for Melanoma Skin Cancer. American Cancer Society. Disponível em: <https://www.cancer.org/cancer/types/melanoma-skin-cancer/about/key-statistics.html>. Último acesso em: 09 de Julho 2025.
- [28] Luis A. Souza, André G. C. Pacheco, Gabriel G. de Angelo, Thiago Oliveira-Santos, Christoph Palm, and João P. Papa. 2024. LiwTERM: A Lightweight Transformer-Based Model for Dermatological Multimodal Lesion Detection. In *2024 37th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. 1–6. <https://doi.org/10.1109/SIBGRAPI62404.2024.10716324>
- [29] Statcounter. 2025. Mobile Operating System Market Share in Brazil - June 2025. Statcounter GlobalStats. Disponível em: <https://gs.statcounter.com/os-market-share/mobile/brazil>. Último acesso em: 09 de Julho 2025.
- [30] Yikai Yang, Eric W.T. Ngai, and Lei Wang. 2024. Resistance to artificial intelligence in health care: Literature review, conceptual framework, and research agenda. *Information & Management* 61, 4 (2024), 103961. <https://doi.org/10.1016/j.im.2024.103961>