

Supporte ao Diagnóstico de Câncer de Pele com Inteligência Artificial: Estudo-Piloto

<https://doi.org/10.32635/2176-9745.RBC.2026v72n1.5443>

Support for Skin Cancer Diagnosis Using Artificial Intelligence: Pilot Study

Apoyo al Diagnóstico de Cáncer de Piel con Inteligencia Artificial: Estudio Piloto

Eduarda Menezes da Silveira¹; Helena Cargnelutti Grimaldi²; Sandro da Silva Camargo³

RESUMO

Introdução: O câncer de pele constitui uma das neoplasias de maior incidência no Brasil, e o diagnóstico precoce representa um fator determinante para o êxito terapêutico e para a redução da morbimortalidade associada. Diante desse cenário epidemiológico, observa-se um contexto particularmente propício à incorporação de ferramentas computacionais complementares à avaliação clínica tradicional, com destaque para abordagens baseadas em inteligência artificial. **Objetivo:** Desenvolver e validar um modelo baseado em redes neurais convolucionais para a classificação automática de lesões cutâneas malignas e benignas. **Método:** Foram utilizadas 2.639 imagens da base pública *International Skin Imaging Collaboration* (ISIC), com anotações validadas por biópsia. O sistema computacional incluiu etapas de pré-processamento e treinamento supervisionado com arquitetura YOLOv11. O desempenho foi avaliado por validação interna e validação externa. **Resultados:** O modelo alcançou acurácia média de 80,53% e sensibilidade média de 80,44% na identificação de oito classes de lesões: melanoma, nevo, carcinoma de células basais, queratose actínica, queratose benigna, dermatofibroma, lesão vascular e carcinoma espinocelular. A implementação também resultou em uma base de imagens anotadas e em um fluxo de análise reproduzível. **Conclusão:** A aplicação de inteligência artificial no suporte ao diagnóstico de câncer de pele demonstrou desempenho promissor, com potencial aplicação em triagens clínicas. Estudos futuros devem considerar a expansão da base de dados e o desenvolvimento de interfaces para uso por profissionais da saúde.

Palavras-chave: Neoplasias Cutâneas/classificação; Avulsões Cutâneas/classificação; Aprendizagem Profunda; Redes Neurais Convolucionais; Processamento de Imagem Assistido por Computador.

ABSTRACT

Introduction: Skin cancer is one of the most prevalent neoplasms in Brazil, and early diagnosis is a key determinant of therapeutic success and reduction of associated morbidity and mortality. In this epidemiological context, there is a particularly favorable scenario for the incorporation of computational tools to complement traditional clinical evaluation, with emphasis on artificial intelligence-based approaches. **Objective:** To develop and validate a convolutional neural network-based model for the automatic classification of malignant and benign skin lesions. **Method:** A total of 2,639 images from the public International Skin Imaging Collaboration (ISIC) database, with biopsy-validated annotations, were used. The computational system included preprocessing steps and supervised training using the YOLOv11 architecture. Performance was assessed through internal and external validation. **Results:** The model achieved a mean accuracy of 80.53% and a mean sensitivity of 80.44% in the identification of eight classes of lesions: melanoma, nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma, vascular lesion, and squamous cell carcinoma. The implementation also resulted in an annotated image dataset and a reproducible analysis pipeline. **Conclusion:** The application of artificial intelligence to support skin cancer diagnosis demonstrated promising performance, with potential clinical screening applications. Future studies should consider expanding the dataset and developing user interfaces for healthcare professionals.

Key words: Skin Neoplasms/classification; Degloving Injuries/classification; Deep Learning; Convolutional Neural Networks; Image Processing, Computer-Assisted.

RESUMEN

Introducción: El cáncer de piel constituye una de las neoplasias de mayor incidencia en el Brasil, siendo el diagnóstico temprano un factor determinante para el éxito terapéutico y para la reducción de la morbimortalidad asociada. Ante este escenario epidemiológico, se observa un contexto particularmente propicio para la incorporación de herramientas computacionales complementarias a la evaluación clínica tradicional, con énfasis en enfoques basados en inteligencia artificial.

Objetivo: Desarrollar y validar un modelo basado en redes neuronales convolucionales para la clasificación automática de lesiones cutáneas malignas y benignas. **Método:** Se utilizaron 2639 imágenes de la base pública *International Skin Imaging Collaboration* (ISIC), con anotaciones validadas por biopsia. El sistema computacional incluyó etapas de preprocesamiento y entrenamiento supervisado con la arquitectura YOLOv11. El rendimiento fue evaluado mediante validación interna y validación externa. **Resultados:** El modelo alcanzó una exactitud media del 80,53% y una sensibilidad media del 80,44% en la identificación de ocho clases de lesiones: melanoma, nevo, carcinoma de células basales, queratosis actínica, queratosis benigna, dermofibroma, lesión vascular y carcinoma espinocelular. La implementación también resultó en una base de imágenes anotadas y en un flujo de análisis reproducible. **Conclusión:** La aplicación de inteligencia artificial en el apoyo al diagnóstico del cáncer de piel demostró un desempeño prometedor, con potencial aplicación en tamizajes clínicos. Estudios futuros deberán considerar la ampliación de la base de datos y el desarrollo de interfaces para su uso por profesionales de la salud.

Palabras clave: Neoplasias Cutáneas/clasificación; Lesiones por Desenguantamiento/clasificación; Aprendizaje Profundo; Redes Neuronales Convolucionales; Procesamiento de Imagen Asistido por Computador.

¹Universidade Federal do Pampa (Unipampa). Bagé (RS), Brasil. Serviço Nacional de Aprendizagem Industrial do Rio Grande do Sul (Senai-RS), *Center for Embedded Devices and Research in Digital Agriculture* (Cedra). São Leopoldo (RS) Brasil. E-mail: eduardasilveira.aluno@unipampa.edu.br. Orcid iD: <https://orcid.org/0009-0004-7829-9273>

²Pesquisadora-autônoma. Bagé (RS), Brasil. E-mail: hgrimaldi87@gmail.com. Orcid iD: <https://orcid.org/0009-0002-2762-4094>

³Unipampa, Programa de Pós-Graduação em Computação Aplicada (PPGCAP). Bagé (RS), Brasil. Senai-RS, Cedra. São Leopoldo (RS) Brasil. E-mail: sandrocamargo@unipampa.edu.br. Orcid iD: <https://orcid.org/0000-0001-8871-3950>

Endereço para correspondência: Sandro da Silva Camargo. Unipampa. Av. Maria Anunciação Gomes Godoy, 1650, Gabinete 3139 – Malafaia. Bagé (RS), Brasil. CEP 96413-172. E-mail: sandrocamargo@unipampa.edu.br



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições, desde que o trabalho original seja corretamente citado.

INTRODUÇÃO

O câncer de pele é a neoplasia maligna mais incidente no Brasil e no mundo, representando cerca de 30% de todos os tumores diagnosticados no país¹. É classificado em dois grandes grupos: o câncer de pele não melanoma, que compreende principalmente o carcinoma basocelular (CBC) e o carcinoma espinocelular (CEC), e o melanoma, menos frequente, mas responsável pela maioria dos óbitos por causa da sua agressividade e do potencial metastático^{2,3}. No Brasil, a incidência estimada é de aproximadamente 60 casos por 100 mil habitantes, sendo o CBC o mais prevalente (70-75%), seguido pelo CEC (15-20%)². Embora apresentem baixa mortalidade, essas neoplasias possuem caráter invasivo local e elevada taxa de recorrência. Já o melanoma, apesar de corresponder a apenas 3% dos casos, concentra a maior parte das mortes relacionadas ao câncer de pele³. Além das neoplasias malignas, outras lesões cutâneas, como nevos, queratoses actínicas, queratoses benignas, dermatofibromas e lesões vasculares, também representam desafios diagnósticos, pois podem mimetizar tumores malignos⁴⁻⁶.

A detecção precoce é fundamental para reduzir morbimortalidade, especialmente nos casos de melanoma. Nesse contexto, a dermatoscopia tem ampliado a acurácia diagnóstica para cerca de 90%, em comparação a 75-80% obtidos pela inspeção clínica isolada^{7,8}. Contudo, sua efetividade depende da experiência do examinador, o que resulta em variabilidade na interpretação. Porém, a limitação do acesso a especialistas, sobretudo em áreas remotas, reforça as desigualdades no diagnóstico e tratamento oportunos⁹. Nesse cenário, soluções baseadas em inteligência artificial (IA) emergem como ferramentas complementares à prática clínica. A IA tem demonstrado potencial em acelerar diagnósticos, apoiar decisões médicas e ampliar o acesso ao cuidado em saúde¹⁰. Em particular, redes neurais convolucionais (RNC) têm alcançado desempenho comparável ao de dermatologistas experientes na classificação automática de lesões cutâneas¹¹⁻¹⁸. Entre os algoritmos disponíveis, o *You Only Look Once* (YOLO) destaca-se pela detecção em tempo real com elevada acurácia, configurando-se como alternativa promissora para triagem automatizada e apoio ao diagnóstico¹⁹.

A aplicação de modelos de IA como o proposto neste estudo tem especial relevância para o contexto do Sistema Único de Saúde (SUS) e das políticas nacionais de controle do câncer. O Brasil apresenta uma distribuição geográfica desigual de dermatologistas, concentrados nas Regiões Sul e Sudeste, o que limita o acesso a diagnósticos especializados em grande parte do território⁹. Essa disparidade se reflete em diagnósticos tardios e sobrecarga de serviços de alta complexidade. Ademais, a

disponibilidade limitada de dermatoscópios em unidades básicas de saúde e a diversidade étnico-racial da população brasileira representam desafios adicionais, uma vez que a maioria das bases públicas de imagens dermatológicas contém amostras predominantemente de fototipos claros, o que pode comprometer a precisão dos modelos em tons de pele mais escuros.

Nesse sentido, iniciativas recentes têm explorado o uso da IA como ferramenta de apoio à Atenção Primária e à triagem em oncologia¹⁰. Algumas pesquisas brasileiras já demonstram o interesse crescente da comunidade científica nacional em integrar sistemas inteligentes aos fluxos de cuidado do SUS, ampliando a detecção precoce e reduzindo desigualdades regionais^{10,17}. Tais esforços dialogam diretamente com as diretrizes do Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas não Transmissíveis e com as metas de controle do câncer estabelecidas pelo Instituto Nacional de Câncer (INCA)/ Ministério da Saúde (MS).

Assim, este estudo tem como objetivo desenvolver, testar e avaliar um modelo computacional baseado na arquitetura YOLO para a detecção automática de lesões cutâneas malignas e benignas em imagens dermatoscópicas, buscando fornecer suporte ao diagnóstico médico e, consequentemente, contribuir para a identificação precoce do câncer de pele.

MÉTODO

Estudo-piloto que se configura como uma pesquisa aplicada, de abordagem quantitativa e delineamento experimental, direcionada ao desenvolvimento e à avaliação de um sistema computacional de suporte ao diagnóstico de câncer de pele com uso de IA. A investigação fundamentou-se na análise de imagens dermatoscópicas para detecção automática de lesões cutâneas, por meio de RNC. O desempenho dos modelos foi avaliado por métricas quantitativas padronizadas, incluindo acurácia e sensibilidade, a fim de aferir sua eficácia na identificação de padrões morfológicos associados a diferentes tipos de lesões cutâneas.

Foram obtidas imagens dermatoscópicas extraídas do banco público *International Skin Imaging Collaboration (ISIC) Archive 2019*¹⁹, composto por 25.331 imagens com diagnóstico histopatológico confirmado, conforme distribuição apresentada na Tabela 1. Os critérios de inclusão envolveram qualidade visual adequada, ausência de interferências clínicas visíveis (como cicatrizes de biópsia ou suturas) e presença de um único diagnóstico por imagem. Todas as imagens estão anonimizadas, em conformidade com diretrizes éticas internacionais para uso secundário de dados clínicos.



A anotação das imagens é o processo que consiste em delimitar a lesão e indicar seu diagnóstico. Essas anotações foram realizadas manualmente utilizando a ferramenta LabelImg²⁰, com marcação das lesões por meio de caixas delimitadoras (*bounding boxes*) e associação de rótulos de classe compatíveis com o formato requerido pela arquitetura YOLO²¹. As classes das lesões foram definidas com base nos laudos histopatológicos disponibilizados juntamente com a base de imagens¹⁹.

Considerando que o processo de anotação das lesões nas imagens é manual e realizado individualmente em cada amostra, este estudo-piloto foi conduzido sobre uma fração do conjunto de imagens disponível, também apresentada na Tabela 1. O trabalho encontra-se em progresso, com a anotação de mais imagens e a execução de novos ciclos de treinamento e validação, com o objetivo de aprimorar progressivamente os indicadores de desempenho do modelo.

Tabela 1. Distribuição das imagens disponíveis por classe de lesão cutânea (ISIC 2019)¹⁹ e distribuição da quantidade de imagens por classe utilizadas nos experimentos

Classe de lesão cutânea	Número de imagens disponíveis	Número de imagens utilizadas
Melanoma	4.522	600
Nevo	12.875	600
Carcinoma basal	3.323	200
Queratose actínica	867	200
Queratose benigna	2.624	200
Dermatofibroma	239	239
Lesão vascular	253	200
Carcinoma espinocelular	628	400
Total	25.331	2.639

O *pipeline* de modelagem, representado na Figura 1, incluiu os seguintes passos:

1) Pré-processamento das imagens: todas as imagens foram padronizadas quanto ao tamanho e às anotações das lesões, garantindo consistência na análise e facilitando a comparação entre diferentes tipos de lesões. Essa etapa é equivalente à preparação cuidadosa de exames clínicos antes da interpretação diagnóstica.

2) Treinamento do modelo, com ajuste automático de hiperparâmetros, visando otimizar a capacidade de classificação das lesões. Esse processo permite que o modelo aprenda de forma otimizada a distinguir características clínicas relevantes, de maneira similar à experiência adquirida por um dermatologista ao revisar múltiplos casos.

3) Diagnóstico, ou validação interna do desempenho: durante o treinamento, o modelo foi continuamente avaliado em subconjuntos de dados reservados, medindo acurácia, sensibilidade e especificidade. Essa etapa garante que o desempenho do sistema seja confiável e consistente, similar à revisão por pares em um contexto clínico antes da implementação de um novo protocolo diagnóstico.

Foram treinadas arquiteturas da família YOLOv11 (versão 11)²⁰, com o objetivo de avaliar a acurácia e a eficiência da detecção automática de lesões cutâneas. Essa arquitetura foi escolhida em virtude de suas características de acurácia, velocidade e capacidade de generalização, especialmente em aplicações médicas^{22,23}. Tal escolha está alinhada ao objetivo de explorar redes neurais profundas como suporte ao diagnóstico precoce do câncer de pele. As arquiteturas da família YOLOv11 empregadas neste estudo foram:

- YOLOv11n (*nano*): versão compacta, otimizada para execução em dispositivos com capacidade computacional reduzida, tal como *smartphones*, *tablets* e sistemas embarcados em dispositivos médicos portáteis, onde há restrições de memória e processamento.
- YOLOv11s (*small*), YOLOv11m (*medium*), YOLOv11l (*large*): arquiteturas progressivamente mais complexas, com maior número de parâmetros e, portanto, maior capacidade de representação, à custa de aumento do tempo de processamento. Essas versões são mais indicadas para execução em estações de trabalho com Unidades de Processamento Gráfico (GPU) dedicadas, servidores clínicos e plataformas de pesquisa, contextos em que a precisão diagnóstica é priorizada em relação ao tempo de processamento.

Todos os modelos foram treinados mediante teste de diferentes hiperparâmetros, como número de épocas, com *batch size* adaptativo, parada antecipada e validação interna contínua durante o treinamento, o que permite otimizar o desempenho do modelo. Esse rigor metodológico, alinhado com as melhores práticas para treinamento e uso de modelos de classificação²⁴, assegura que o modelo aprenda padrões relevantes e minimize vieses na identificação das lesões.

A validação dos modelos de IA foi realizada em duas etapas complementares²⁵:

1) Validação interna, utilizando uma divisão estratificada aleatória de 90% das imagens para o treinamento do modelo e 10% para teste. Essa técnica consiste em avaliar o desempenho do modelo de forma controlada, utilizando os próprios dados do conjunto original, permitindo estimar a capacidade preditiva do modelo e garantindo que todas as classes de lesões cutâneas fossem representadas proporcionalmente nos conjuntos de treino e teste.



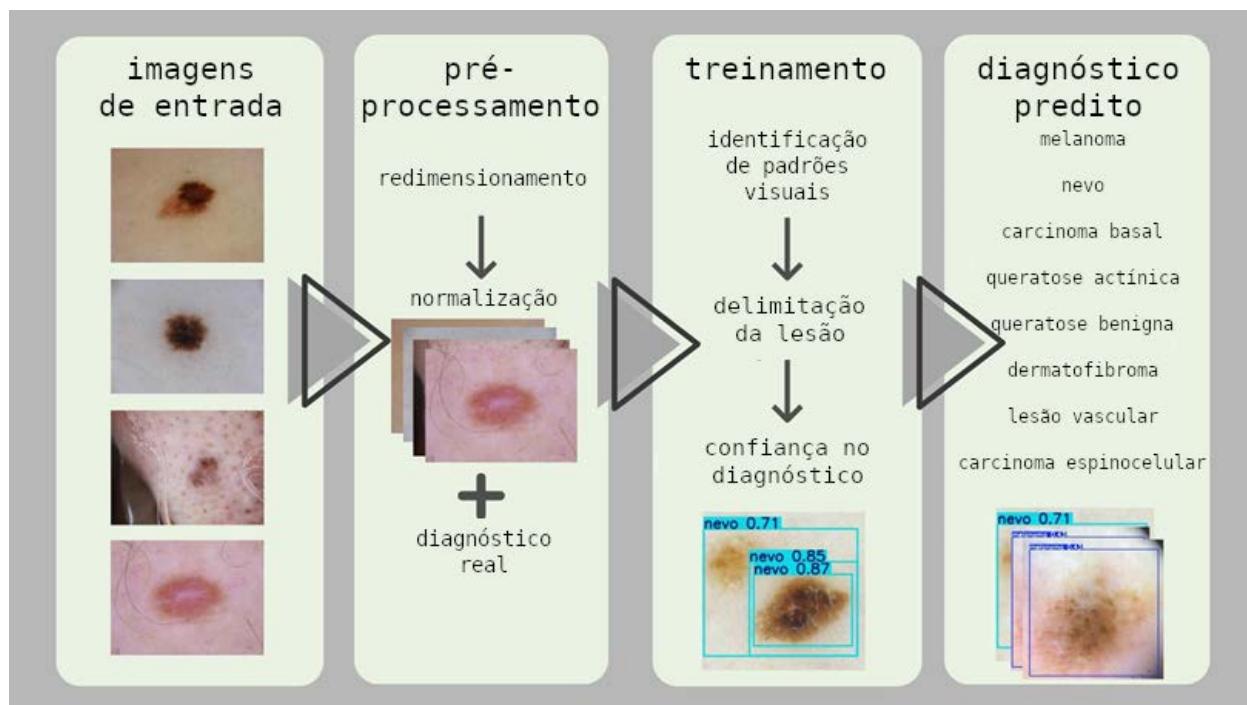


Figura 1. Pipeline de detecção de lesões cutâneas com YOLOv11

2) Validação externa, realizada com imagens inéditas, ou seja, não utilizadas anteriormente no treinamento, capturadas em um ambiente clínico real. Essas imagens foram processadas com o modelo treinado, permitindo verificar se o modelo mantém sua precisão em situações práticas, fora do ambiente de teste inicial.

Na validação interna, foram avaliadas as seguintes métricas de desempenho: acurácia (*mean Average Precision* – mAP), sensibilidade média, precisão média e tabela de desempenho diagnóstico.

A amostra da validação externa foi composta por 58 imagens dermatoscópicas obtidas em um ambulatório de dermatologia de um serviço privado de saúde. As imagens foram capturadas com Dermatoscópio DermLite DL4, com iluminação polarizada e não polarizada de 24 LEDs Brancos e *Pigment Boost Plus* (luz laranja), seguindo a rotina de triagem clínica de lesões suspeitas. Imagens com artefatos importantes (ex.: curativos, sangue, pelos espessos) foram excluídas.

Não houve balanceamento de classes na validação externa, de modo que a distribuição das lesões reflete a prevalência observada na prática clínica (maior presença de lesões benignas, como nevos e queratoses). Essa ausência de balanceamento pode favorecer o desempenho do modelo nas classes mais frequentes e prejudicar a avaliação em classes menos representadas, como melanoma e CEC, resultando em maior probabilidade de falsos-negativos nessas categorias. Essa limitação será tratada em estudos futuros mediante ampliação da amostra clínica e estratégias de balanceamento.

O treinamento foi realizado no ambiente de alto desempenho Saturno²⁶, equipado com oito GPU NVIDIA A100 e 2Tb de memória. O sistema foi implementado em *Python* 3.10, com uso das bibliotecas *PyTorch*, *Ultralytics*, *Pandas*, *NumPy*, *OpenCV* e *Matplotlib*, entre outras.

Considerando os aspectos éticos da pesquisa, destaca-se que não houve necessidade de submissão deste estudo ao Comitê de Ética em Pesquisa (CEP). Isso porque a investigação utilizou exclusivamente dados secundários de acesso público, disponíveis em base de dados aberta na Internet, sem qualquer possibilidade de identificação direta ou indireta de indivíduos. Tál enquadramento encontra respaldo na Resolução n.º 466/2012²⁷ do Conselho Nacional de Saúde (CNS), que estabelece diretrizes e normas para pesquisas envolvendo seres humanos e, de modo mais específico, na Resolução n.º 510/2016²⁸, a qual dispõe sobre as normas aplicáveis às pesquisas em ciências humanas e sociais e explicita que não necessitam de apreciação ética os estudos que utilizam informações de domínio público. Dessa forma, a realização desta pesquisa está em conformidade com os princípios éticos vigentes no Brasil.

RESULTADOS

Foram realizados testes preliminares com diferentes arquiteturas da família YOLOv11 (*nano*, *small*, *medium* e *large*) para avaliar a detecção de lesões cutâneas em imagens dermatoscópicas. Entre essas variantes, a arquitetura YOLOv11s (*small*) apresentou os melhores

resultados, os quais são aqui relatados. A Tabela de desempenho diagnóstico (Figura 2) ilustra os acertos e erros por classe, considerando no eixo *x* as classes reais e, no eixo *y*, as classes preditas. A análise dessa matriz permitiu identificar padrões de acerto e erro do modelo em diferentes tipos de lesões cutâneas.

O desempenho variou conforme a classe avaliada. Para lesões vasculares, a taxa de acerto foi de 93%, com 7% de falhas de reconhecimento, evidenciando elevada especificidade, mas sensibilidade limitada em casos marginais. Na queratose actínica, a taxa de acerto foi de 90%; contudo, 10% foram incorretamente classificadas como melanoma, o que denota risco clínico relevante em razão do caráter agressivo dessa neoplasia.

Nos nevos melanocíticos, 83% foram corretamente classificados, enquanto 16% foram reconhecidos equivocadamente como melanoma, sugerindo dificuldade do modelo em diferenciar lesões benignas e malignas. Para os dermatofibromas, a taxa de acerto foi de 78%, com 6% de erros como melanoma e 6% como nevos.

Quanto ao melanoma, 75% foram identificados corretamente, mas 12% foram confundidos com nevos e 7% com outras categorias, revelando a necessidade de maior precisão nessa classe. Nos carcinomas, observou-se desempenho heterogêneo: 67% de acurácia para o espinocelular e apenas 45% para o basocelular. A queratose benigna apresentou a menor taxa de acerto (33%), possivelmente em função da escassez de exemplos no conjunto de treinamento e da semelhança morfológica com outras lesões.

Os casos em que o modelo atribuiu a classe *background* corresponderam a falsos-negativos, ou seja, situações em que a lesão real não foi reconhecida. Esse tipo de falha tem relevância clínica crítica, pois pode contribuir para atrasar o diagnóstico e comprometer a conduta terapêutica.

Além da avaliação quantitativa, foi realizada análise qualitativa das previsões visuais em imagens dermatoscópicas. A Figura 3 mostra exemplos representativos, com destaque para a correta localização e classificação de diferentes tipos de lesões, evidenciando a aplicabilidade prática do modelo.

Para facilitar a visualização do desempenho discriminativo do modelo, as principais métricas por classe foram consolidadas na Tabela 2, permitindo observar diferenças de sensibilidade, precisão e perfis de erro entre os tipos de lesões.

DISCUSSÃO

Os resultados evidenciam o potencial do modelo proposto como ferramenta de apoio ao diagnóstico médico em oncologia cutânea, sobretudo em contextos de triagem

e suporte à decisão clínica. Apesar dos avanços observados, o desempenho heterogêneo entre classes aponta limitações relevantes. Classes com padrões morfológicos semelhantes, como nevos e melanomas, ainda apresentaram confusão significativa, indicando necessidade de aprimoramento na capacidade discriminativa do sistema.

A baixa performance em categorias específicas, como CBC e queratose benigna, sugere impacto tanto da desproporção amostral quanto da complexidade intrínseca dessas lesões. Estratégias de expansão e diversificação da base de dados emergem como caminhos prioritários para aumentar a robustez do modelo.

Do ponto de vista clínico, a ocorrência de falsos-negativos representa o maior desafio, uma vez que o não reconhecimento de uma lesão maligna pode ter repercussões severas para o paciente. Assim, futuros treinamentos devem priorizar a redução desse tipo de erro, buscando elevar a acurácia global para níveis superiores a 95%.

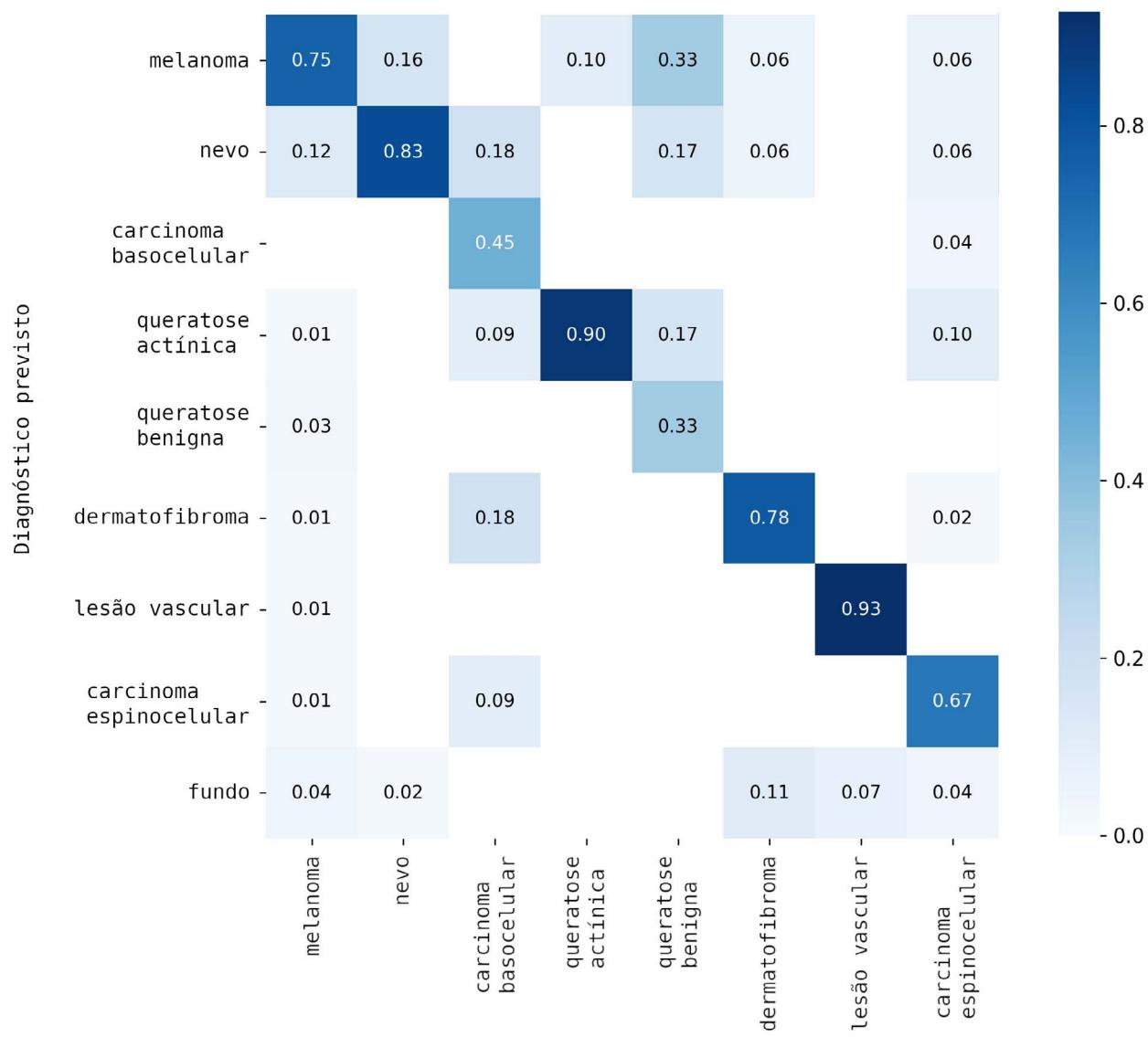
Quando comparado a estudos internacionais, o desempenho obtido neste trabalho é compatível com pesquisas recentes que utilizaram redes neurais profundas para detecção de melanoma. Estudos conduzidos com RNC ou modelos YOLO adaptados reportam sensibilidades entre 72% e 89% e acurácia média entre 78% e 91% na identificação de melanoma e CBC^{11,18}. No presente estudo, a sensibilidade para melanoma foi de 75% e a acurácia média geral foi de 80,53%, valores dentro da faixa observada na literatura, o que reforça a viabilidade técnica do modelo mesmo utilizando um conjunto reduzido de imagens anotadas manualmente.

Além disso, trabalhos como SkinSage XAI de Munjal et al.¹² e o estudo prospectivo multicêntrico de Heinlein et al.¹⁸ demonstram que sistemas baseados em IA podem alcançar desempenho semelhante ou superior ao de dermatologistas em cenários de triagem, reforçando que a aplicação de modelos de visão computacional pode reduzir custos e acelerar o diagnóstico em contextos clínicos. Assim, o desempenho alcançado por este estudo aproxima-se da tendência global, reforçando sua relevância científica.

A perspectiva de implementação do modelo em dispositivos móveis ou plataformas *web* constitui um avanço estratégico, pois amplia significativamente a possibilidade de validação externa e aplicação prática. Essa portabilidade favorece a realização de testes colaborativos em diferentes Regiões do país, permitindo que especialistas de distintas instituições e contextos clínicos avaliem a ferramenta de forma sistemática. Tal abordagem potencializa a geração de evidências sobre a robustez e a generalização do modelo, além de contribuir para a consolidação de sua utilidade em cenários de telemedicina e oncologia de precisão.



Tabela de desempenho diagnóstico

**Figura 2.** Tabela de desempenho diagnóstico apresentando os diagnósticos reais e os diagnósticos preditos pelo modelo

Nota: No eixo horizontal, encontram-se as classes reais das lesões e, no eixo vertical, as predições do modelo. A diagonal principal representa os acertos. Quando o modelo classifica uma lesão maligna como benigna, ocorre um falso-negativo, situação de maior risco clínico por atrasar o diagnóstico. Quando uma lesão benigna é classificada como maligna, ocorre um falso-positivo, podendo levar a biópsias ou encaminhamentos desnecessários.

Assim, este estudo caracteriza-se como uma “prova de conceito” (*proof of concept*), demonstrando a viabilidade técnica e científica do uso do modelo proposto como ferramenta de apoio ao diagnóstico dermatológico. A criação de um *pipeline* completo (anotação → treinamento → inferência) e a validação com imagens reais do ambiente clínico reforçam o potencial de transferência tecnológica do modelo para o SUS, especialmente em localidades com baixa disponibilidade de dermatologistas e acesso limitado à dermatoscopia.

Apesar das limitações, sensibilidade ainda insuficiente em algumas classes e baixa representação de tons de pele mais escuros, os resultados mostram potencial clínico,

especialmente como ferramenta de triagem em atenção primária à saúde.

Em síntese, o modelo desenvolvido apresenta as seguintes potencialidades principais:

- viabilidade técnica comprovada para detecção automática de múltiplos tipos de lesões;
- desempenho compatível com estudos internacionais que utilizam *deep learning* em dermatoscopia;
- potencial de aplicação em triagem clínica, especialmente no contexto do SUS;
- estabelecimento de bases para estudos futuros com maior diversidade de dados, balanceamento e validação

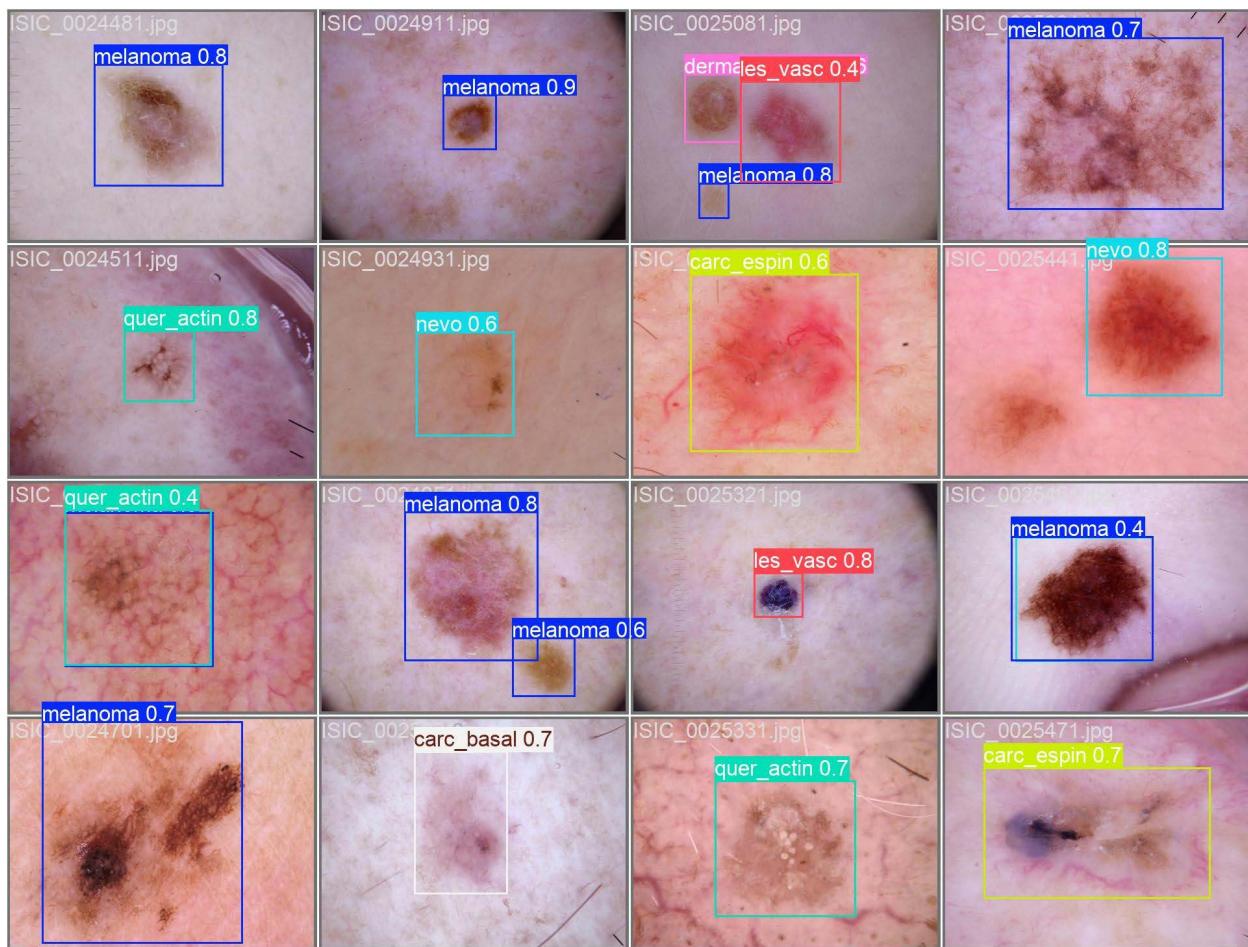


Figura 3. Exemplos de detecção automatizada de lesões cutâneas pelo modelo desenvolvido, acompanhados dos respectivos níveis de confiança atribuídos pelo sistema ao diagnóstico previsto

Nota: Cada imagem apresenta a classe prevista e o nível de confiança (probabilidade) da predição. (a) Verdadeiro-positivo: o modelo identifica corretamente a lesão e delimita sua região. (b) Falso-positivo: o modelo classifica uma lesão benigna como maligna, o que pode levar à investigação desnecessária. (c) Falso-negativo: o modelo falha em identificar uma lesão maligna, o que representa risco clínico por possível atraso no diagnóstico. Esses exemplos ilustram como o modelo responde a diferentes padrões dermatoscópicos, como bordas irregulares, assimetria de pigmentação e presença de vasos sanguíneos.

Tabela 2. Desempenho do modelo por classe de lesão

Classe de lesão cutânea	Sensibilidade (%)	Precisão (%)	Principais erros observados
Melanoma	75	78	Falso-negativo: confundido com nevo
Nevo	83	81	Falso-positivo: confundido com melanoma
Carcinoma basal	45	72	Não detectado em lesões pequenas ou pouco pigmentadas
Queratose actínica	90	84	Confundida com melanoma em bordas irregulares
Queratose benigna	33	68	Confundida com outras queratoses
Dermatofibroma	78	70	Confundido com nevo ou melanoma
Lesão vascular	93	89	Baixa taxa de erro: padrão facilmente identificado
Carcinoma espinocelular	67	74	Confundido com queratose actínica



multicêntrica. Essa atividade apresentaria potencial significativo mediante expansão e fortalecimento por meio de parcerias institucionais estratégicas, especialmente com o INCA/MS. O apoio dessas entidades poderia viabilizar maior alcance, integração de bases de dados nacionais e aprimoramento metodológico, contribuindo para a consolidação de uma rede colaborativa voltada ao monitoramento e à pesquisa em oncologia no país.

CONCLUSÃO

Este estudo demonstrou avanços relevantes na aplicação da IA para o diagnóstico assistido do câncer de pele, com ênfase na utilização da arquitetura YOLOv11. A construção de um *pipeline* funcional, que abrange desde o pré-processamento das imagens até o treinamento do modelo com dados públicos anotados manualmente, evidencia a viabilidade técnica da proposta.

Entre as contribuições mais significativas, destacam-se: a implementação de um modelo de detecção automática de lesões cutâneas com múltiplas classes, a partir de imagens dermatoscópicas, ampliando seu potencial de aplicação clínica; o desenvolvimento de um conjunto de dados com anotações manuais, priorizando padrões visuais relevantes como vasos e bordas, com impacto direto na acurácia do modelo; a integração de tecnologias de visão computacional e IA em um fluxo de trabalho reproduzível e escalável.

Apesar dos avanços, o estudo apresenta limitações importantes: a acurácia do modelo, embora promissora com valor de 80,53%, ainda não atinge os patamares necessários para uso clínico direto. Essa limitação será abordada por meio de novos ciclos de treinamento, que estarão em execução nos próximos seis meses, com bases de dados expandidas e mais diversificadas; o conjunto de imagens utilizado tem baixa representatividade de tons de pele mais escuros, o que compromete a generalização do modelo à população brasileira. Tal lacuna evidencia a necessidade de iniciativas nacionais voltadas à coleta sistemática de imagens dermatológicas com diversidade étnico-racial; a validação externa, embora iniciada com imagens inéditas capturadas em ambiente clínico, requer detalhamento adicional quanto à amostragem, aos critérios de inclusão e aos parâmetros clínicos avaliados, para assegurar a extrapolação dos resultados aos serviços reais de saúde.

Por fim, essas considerações demonstram o potencial transformador do projeto, ao mesmo tempo em que evidenciam os desafios superados e traçam caminhos para um resultado mais robusto e aplicável na prática clínica.

AGRADECIMENTOS

À NVIDIA, pelo suporte computacional disponibilizado, fundamental para a execução dos experimentos e análises deste estudo.

CONTRIBUIÇÕES

Todos os autores contribuíram substancialmente na concepção e no planejamento do estudo; na obtenção, análise e interpretação dos dados; na redação e revisão crítica; e aprovaram a versão final a ser publicada.

DECLARAÇÃO DE CONFLITOS DE INTERESSE

Nada a declarar.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os conjuntos de dados gerados e analisados durante o estudo atual estão disponíveis no repositório <https://github.com/Sandrocamargo/publications/tree/main/rbc25a>

FONTES DE FINANCIAMENTO

Não há.

REFERÊNCIAS

1. Delgado LP, Regis MN, Caldas D, et al. Epidemiologia e manejo clínico do câncer de pele melanoma: uma revisão narrativa. *J Soc Issues Health Sci.* 2024;1(7):1-11. doi: <https://doi.org/10.5281/zenodo.17473590>
2. Rezende Filho AV, Yamamoto HG, Macedo JLS et al. Perfil epidemiológico de pacientes portadores de câncer de pele atendidos no Hospital Regional da Asa Norte/DF - Brasil. *Rev Bras Cir Plást.* 2020;35(3):316-21. doi: <https://doi.org/10.5935/2177-1235.2020RBCP0056>
3. Borges AL, Zalaudek I, Longo C, et al. Melanocytic nevi with special features: clinical-dermoscopic and reflectance confocal microscopic-findings. *JEADV.* 2014;28(7):833-45. doi: <https://doi.org/10.1111/jdv.12291>
4. Piccolo V, Russo T, Moscarella E, et al. Dermatoscopy of vascular lesions. *Dermatol Clin.* 2018;36(4):389-95. doi: <https://doi.org/10.1016/j.det.2018.05.006>
5. Álvarez-Salafranca M, Gómez-Martíns I, Bañuls J, et al. Dermoscopy of inflamed seborrheic keratosis: a great mimic of malignancy. *Australas J Dermatol.* 2022;63(1):53-61. doi: <https://doi.org/10.1111/ajd.13781>



6. Thamm JR, Welzel J, Schuh S. Diagnosis and therapy of actinic keratosis. *J Dtsch Dermatol Ges.* 2024;22(5):675-90. doi: <https://doi.org/10.1111/ddg.15288>
7. Azulay RA, Azulay DR, Azulay-ABULAFIA L. Dermatologia. 7.ed. Rio de Janeiro: Guanabara Koogan; 2017.
8. Botton DV, Barbosa DGR, Cavalcante Junior CAC, et al. Relevância da dermatoscopia para o diagnóstico precoce de melanomas: uma revisão de literatura. *RISE.* 2020;1(2):159-74. doi: <https://doi.org/10.56344/2675-4827.v1n2a20209>
9. Schmitt JV, Miot HA. Distribution of Brazilian dermatologists according to geographic location, population and HDI of municipalities: an ecological study. *An Bras Dermatol.* 2014;89(6):1013-5. doi: <https://doi.org/10.1590/abd1806-4841.20143276>
10. Siqueira ASE, Santos Neto MF, Ferreira CBT, et al. Inteligência artificial nas ações de controle do câncer: solução ou problema? *Rev Bras Cancerol.* 2025;71(3):e-005291. doi: <https://doi.org/10.32635/2176-9745.RBC.2025v71n3.5291>
11. Phillips M, Marsden H, Jaffe W, et al. Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. *JAMA Netw Open.* 2019;2(10):e1913436. doi: <https://doi.org/10.1001/jamanetworkopen.2019.13436>
12. Munjal G, Bhardwaj P, Bhargava V, et al. SkinSage XAI: an explainable deep learning solution for skin lesion diagnosis. *Health Care Sci.* 2024;3(6):438-55. doi: <https://doi.org/10.1002/hcs2.121>
13. Cui X, Wei R, Gong L, et al. Assessing the effectiveness of artificial intelligence methods for melanoma: a retrospective review. *J Am Acad Dermatol.* 2019;81(5):1176-80. doi: <https://doi.org/10.1016/j.jaad.2019.06.042>
14. Jairath N, Pahalyants V, Shah R, et al. Artificial intelligence in dermatology: a systematic review of its applications in melanoma and keratinocyte carcinoma diagnosis. *Dermatol Surg.* 2024;50(9):791-8. doi: <https://doi.org/10.1097/DSS.0000000000004223>
15. Wang HH, Wang YH, Liang CW, et al. Assessment of deep learning using nonimaging information and sequential medical records to develop a prediction model for nonmelanoma skin cancer. *JAMA Dermatol.* 2019;155(11):1277-83. doi: <https://doi.org/10.1001/jamadermatol.2019.2335>
16. Chu YS, An HG, Oh BH, et al. Artificial intelligence in cutaneous oncology. *Front Med (Lausanne).* 2020;7:318. doi: <https://doi.org/10.3389/fmed.2020.00318>
17. Silva TAM, Souza EP, Bambil D, et al. Aprendizado de máquina aplicado ao diagnóstico de câncer de pele. In: 16 Anais do Congresso Brasileiro de Inteligência Computacional [Internet]; 2021 out 3-6; Joinville. Joinville: UDESC; 2021. doi: <http://dx.doi.org/10.21528/CBIC2021-55>
18. Heinlein L, Maron RC, Hekler A, et al. Clinical melanoma diagnosis with artificial intelligence: insights from a prospective multicenter study [Preprint]. *arXiv.* 2024. doi: <https://doi.org/10.48550/arXiv.2401.14193>
19. ISIC 2019: Skin lesion analysis towards melanoma detection [Internet]. Versão 2019. [sem local]: International Skin Imaging Collaboration; 2019 [acesso 2024 set 18]. Disponível em: <https://challenge2019.isic-archive.com/>
20. LabelImg [Internet]. Versão 1.8.6. [sem local]: Python Software Foundation; ©2025 [acesso 2024 set 18]. Disponível em: <https://pypi.org/project/labelImg/>
21. Ultralytics. Ultralytics YOLO [repositório GitHub] [acesso 2025 abr 10]. Disponível em: <https://github.com/ultralytics/ultralytics>
22. Terven J, Córdova-Esparza DM, Romero-González JA. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS. *Mach Learn Knowl Extr.* 2023;5(4):1680-716. doi: <https://doi.org/10.3390/make5040083>
23. Esteva A, Chou K, Yeung S, et al. Deep learning-enabled medical computer vision. *npj Digit Med.* 2021;4:5. doi: <https://doi.org/10.1038/s41746-020-00376-2>
24. Cechinel C, Camargo SS. Mineração de dados educacionais: avaliação e interpretação de modelos de classificação [Internet]. In: Jaques P, Pimentel M, Siqueira S, et al, editores. Metodologia de pesquisa científica em informática na educação: abordagem quantitativa. Porto Alegre: SBC; 2020. Disponível em: https://ceie.sbc.org.br/metodologia/wp-content/uploads/2019/11/livro2_cap12.pdf
25. Steyerberg EW, Harrell Jr FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* 2016;69:245-7. doi: <https://doi.org/10.1016/j.jclinepi.2015.04.005>
26. Saturn Cloud [Internet]. Nova Iorque: Saturn Cloud; ©2025 [acesso 2024 set 18]. Disponível em: <https://saturncloud.io/>
27. Conselho Nacional de Saúde (BR). Resolução nº 466, de 12 de dezembro de 2012. Aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos [Internet]. Diário Oficial da União, Brasília, DF. 2013 jun 13 [acesso 2025 ago 27]; Seção 1:59. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html



28. Conselho Nacional de Saúde (BR). Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução [Internet]. Diário Oficial da União, Brasília, DF. 2016 maio 24 [acesso 2024 abr 7]; Seção 1:44. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2016/res0510_07_04_2016.html

Recebido em 1/9/2025
Aprovado em 11/11/2025

