Support for Skin Cancer Diagnosis Using Artificial Intelligence: Pilot Study

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

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

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

Eduarda Menezes da Silveira[1]; Helena Cargnelutti Grimaldi[2]; Sandro da Silva Camargo[3]

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.

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.

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.

[1]Universidade 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
[2]Pesquisadora-autônoma. Bagé (RS), Brasil. E-mail: hgrimaldi87@gmail.com. Orcid iD: https://orcid.org/0009-0002-2762-4094
[3]Unipampa, 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
Corresponding author: 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

INTRODUCTION

Skin cancer is the most incident malignant neoplasm in Brazil and worldwide, accounting for 30% of all tumors diagnosed in the country[1]. It is classified in two major groups: non-melanoma skin cancer, encompassing mainly basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), and melanoma, less frequent, but responsible for the majority of deaths due to the aggressiveness and metastatic potential[2,3]. In Brazil, the estimated incidence is approximately 60 cases per 100 thousand inhabitants, being BCC the most prevalent (70-75%), followed by SCC (15-20%)[2] presenting low mortality but locally invasive and elevated recurrence rate. Melanoma, on its turn, with 3% of the cases, concentrates the great portion of skin cancer related deaths[3]. Further to malignant neoplasms, other skin lesions as nevus, actinic keratosis, benign keratosis, dermatofibroma and vascular lesions are challenging to diagnose because they mimic malignant tumors[4-6].

In cases of melanoma, early detection is essential to reduce morbimortality. In this context, dermatoscopy has widened the diagnostic accuracy to 90% compared with 75-80% from single clinical investigation[7,8]. However, its effectiveness depends on the doctor's experience leading to variation of interpretation. In addition, poor access to experts, most of all in remote areas, reinforces the disparities of diagnoses and timely treatment[9]. AI (artificial intelligence)-based solutions emerge in this scenario as complementary tools of clinical practice with potential to speed up diagnoses, support medical decisions and expand healthcare access[10]. Convolutional neural networks (CNNs), in particular, have achieved remarkable performance in automated classification of skin lesions[11-18] if compared with skilled dermatologists. The algorithm You Only Look Once (YOLO) stands out because of real-time detection with elevated accuracy, a promising alternative for automated screening and diagnosis support[19].

The application of AI models as proposed herein is relevant for the National Health System (SUS) and cancer control national policies. An uneven geographic distribution of dermatologists is observed in Brazil concentrated in the South and Southeast regions which reduces the access to specialized diagnoses in large part of the country[9]. This disparity means late diagnoses and overload of high complexity services, in addition to limited availability of dermatoscopy devices at basic health units and racial-ethnicity diversity as additional challenges since the majority of public dermatological images databases contain samples of predominantly fair skin phototypes which can compromise the precision for darker skins.

In that sense, recent initiatives have explored AI as supporting tool for primary attention and oncology screening[10]. A few Brazilian researches already show the rising interest of the national scientific community in integrating intelligent systems into SUS line of care, expanding early detection and reducing regional disparities[10,17]. These efforts are consistent with the Non-Communicable Strategic Actions Coping Plan and goals of cancer control determined by the National Cancer Institute (INCA)/Ministry of Health (MS).

The objective of this study is to develop, test and evaluate a YOLO-based architecture computational model for automated detection of malignant and benign skin lesions in dermatoscopy images, attempting to provide support to medical diagnosis and contribute for early detection of skin cancer.

METHOD

Pilot-study configured as an applied research, quantitative approach and experimental design targeted to the development and evaluation of an AI-based computational support for the diagnosis of skin cancer. The investigation was grounded on the analysis of dermatoscopy images for automated detection of skin lesions through CNNs. The performance of the models was evaluated by standardized quantitative metrics, including accuracy and sensitivity to check its efficacy in identifying morphologic patterns associated with different types of skin lesions.

Dermatoscopy images have been extracted from the public bank International Skin Imaging Collaboration (ISIC) Archive 2019[19] containing 25,331 images with confirmed histopathological diagnosis as shown in Table 1. Inclusion criteria encompassed proper visual quality, absence of visible clinical interferences (biopsy scars or sutures) and unique imaging diagnosis. Anonymity is secured for all images in compliance with international ethical guidelines of use of clinical secondary data.

Images annotation is the process of lesion delimitation and indicate the diagnosis, performed manually with the tool LabelImg[20] labelling the lesions through bounding boxes and association of compatible classes labels with YOLO[21] required format. Histopathological reports available in conjunction with images base were utilized to define the classes of lesions[19].

Given that the process of annotation of lesions is manual and performed individually in each sample, this pilot-study was conducted over a fraction of the available dataset also presented in Table 1. The work is still ongoing with more images annotated and execution of new cycles of training and validation with the objective of improving progressively the indicators of the model performance.

Table 1. Distribution of images available per class of skin lesion (ISIC 2019)19 and distribution of the quantity of images per class utilized in the experiments

Class of skin lesion	Number of lesions available	Number of images utilized
Melanoma	4,522	600
Nevus	12,875	600
Basal cell carcinoma	3,323	200
Actinic keratosis	867	200
Benign keratosis	2,624	200
Dermatofibroma	239	239
Vascular lesion	253	200
Squamous cell carcinoma	628	400
Total	**25,331**	**2,639**

Figure 1 depicts the steps of the modelling pipeline:

1) Images pre-processing: all the images were normalized for size and annotations of the lesions, ensuring consistency of the analysis and allowing the comparison among different types of lesions. It is equivalent to careful preparation of clinical exams prior to diagnostic interpretation.

2) Training of the model with automated adjustment of hyperparameters to optimize the capacity of lesions classification. This process allows the model to optimize learning to distinguish the relevant clinical characteristics similar to the experience acquired by a dermatologist while reviewing multiple cases.

3) Diagnosis or internal validation of the performance: during training, the model was continuously evaluated in reserved data subsets measuring accuracy, sensitivity and specificity. This step ensures that the performance of the system is reliable and consistent, similar to peer-review in a clinical context before the implementation of a new diagnostic protocol.

Architectures of the family YOLOv11 (version 11)[20] were trained to evaluate the accuracy and efficiency of the automated detection of skin lesions. This architecture was chosen due to its characteristics of accuracy, speed and capacity of generalization, particularly in medical applications[22,23]. The choice is aligned with the objective of exploring deep neural networks as support for early diagnosis of skin cancer. The architectures of the family YOLOv11 adopted herein were:

- YOLOv11n (nano): complete version, optimized for performance in reduced computational devices as smartphones, tablets and systems docked in portable medical devices where there are memory and processing restrictions.
- YOLOv11s (small), YOLOv11m (medium), YOLOv11l (large): progressively more complexes architectures with high number of parameters and, therefore, high capacity of representation at the expense of increased processing time. These versions are indicated more for performance in working stations as dedicated Graphic Processing Units (GPU), clinical servers and research platforms, contexts where diagnostic precision is prioritized in relation to processing time.

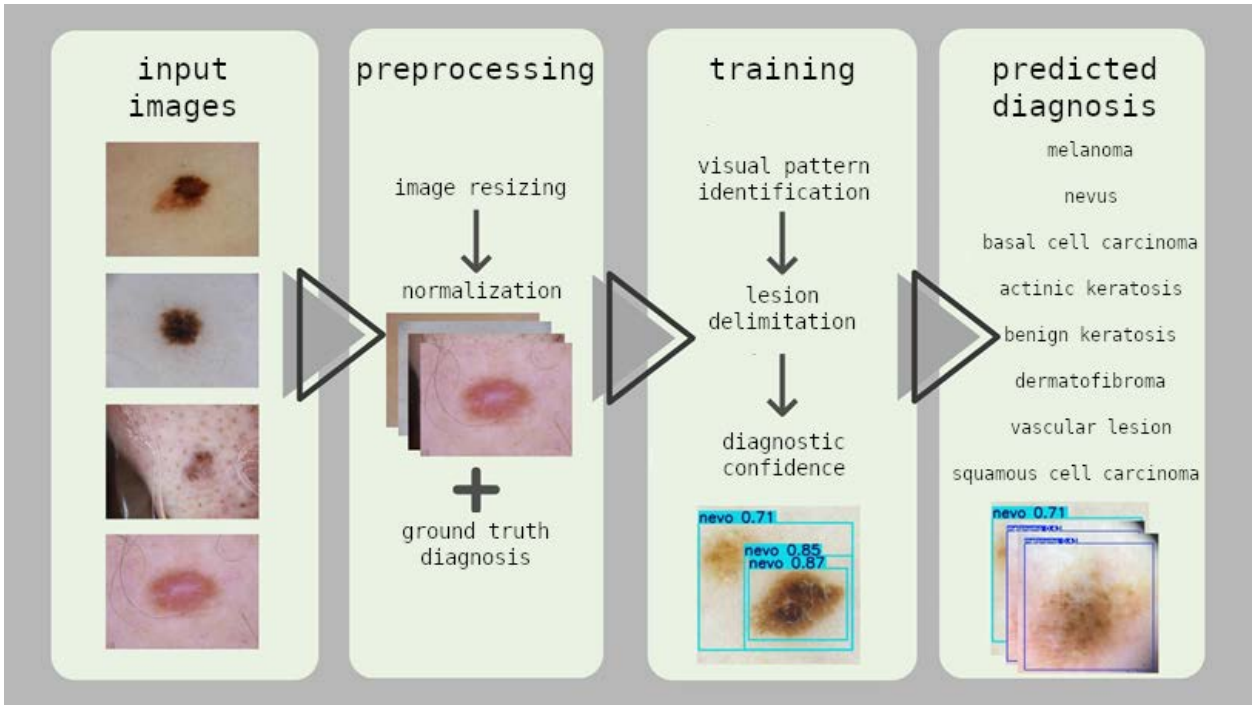

Figure 1. Pipeline of detection of skin lesions with YOLOv11

All the models were trained through tests of different hyper-parameters as number of epochs, with adaptative batch size, anticipated stop and continuous internal validation during training which allows the optimization of the model performance. This methodological rigor aligned with the best practices for training and use of classification models[24] ensures that the model presents relevant patterns and minimizes bias in identifying lesions.

The validation of AI models was conducted in two complementary stages[25]:

1) Internal validation utilizing a stratified random sampling of 90% of the images to train the model and 10% for testing. This technique consists in evaluating the controlled performance of the model utilizing the data of the original set, allowing to estimate the predictive capacity of the model and ensuring that all classes of skin lesions were presented proportionally in the sets of training and test.

2) External validation with unreleased images (not utilized before training captured in clinical environment). These images were processed with the model trained, allowing to verify whether the model holds its accuracy in practical situations out of the initial test environment.

The following performance metrics were evaluated in the internal validation: mean Average Precision – mAP, mean sensitiveness, mean precision and table of diagnostic performance.

The external validation sample consisted in 58 dermatoscopy images obtained at a dermatology outpatient unit of a private health service. The images were captured with "*Dermatoscópio Dermlite DL4*", with polarized light and non-polarized 24 white LEDS and Pigment Boost Plus (orange light), following the routine of clinical triage of suspicious lesions. Images with important artifacts (dressing, blood, thick hair) have been excluded.

There was no balancing of classes in the external validation, therefore, the distribution of the lesions reflects the prevalence observed in clinical practice (more presence of benign lesions as nevus and keratosis). Non-balancing can favor the performance of the model for the most frequent classes and damage the evaluation of least represented classes as melanoma and SCC, resulting in high likelihood of false-negatives in these categories. This limitation will be treated in future studies with larger clinical samples and balancing strategies.

The training was conducted on high performance environment Saturn[26], equipped with eight GPUs NVIDIA A100 and 2Tb of memory. The system was implemented in Python 3.10, utilizing libraries PyTorch, Ultralytics, Pandas, NumPy, OpenCV and Matplotlib, among others.

Submission and review by an ethics committee was waived because only secondary, public and deidentified data have been utilized in compliance with Directive 466/2012[27] of the National Health Council for studies with human beings and specifically, Directive 510/2016[28], that addresses rules applicable to social and human sciences trials utilizing public data.

RESULTS

Preliminary tests were performed with different architectures of the family YOLOv11 (nano, small, medium and large) to evaluate the detection of skin lesions in dermatoscopy images. Of these variables, the architecture YOLOv11s (small) presented the best results reported herein. Figure 2 (diagnostic performance table) shows the hits and errors per class, x axis for actual classes and y axis for predicted classes. The analysis of this matrix allowed to identify patterns of hits and errors of the model in different types of skin lesions.

The performance varied according to the class evaluated. For vascular lesions, the hit rate was 93% with 7% of failure to recognize, showing elevated specificity but limited sensitivity in marginal cases. The hit rate was 90% for actinic keratosis, however, 10% were erroneously classified as melanoma, indicating relevant clinical risk due to the aggressiveness of this neoplasm.

83% of melanocytic nevus were correctly classified while 16% were erroneously recognized suggesting difficulty of the model to differentiate benign and malignant lesions. The hit rate was 78% for dermatofibroma, and 6% of errors as melanoma and 6% as nevus.

75% of the melanomas were identified correctly but 12% were confounded with nevus and 7% with other categories, revealing the necessity of improved accuracy in this class. Heterogenous performance was observed in carcinomas: 67% of accuracy for SCC and only 45% for BCC. Benign keratosis presented the lowest hit rate (33%) possibly because of scarcity of examples in the training set and morphological similarity with other lesions.

Cases where the model attributed the background class were false-negatives, situations where the actual lesions were not recognized. This type of error is clinically relevant because it can delay the diagnosis and compromise the therapeutic conduct.

In addition to the quantitative evaluation, a qualitative analysis of the visual predictions was performed in dermoscopy images. Figure 3 shows examples, highlighting the correct location and classification of different types of lesions and practicality of the model.

To facilitate the visualization of the discriminative performance of the models, the main metrics per class were consolidated in Table 2 where differences of sensitivity, accuracy and profiles of errors of the types of lesions can be seen.

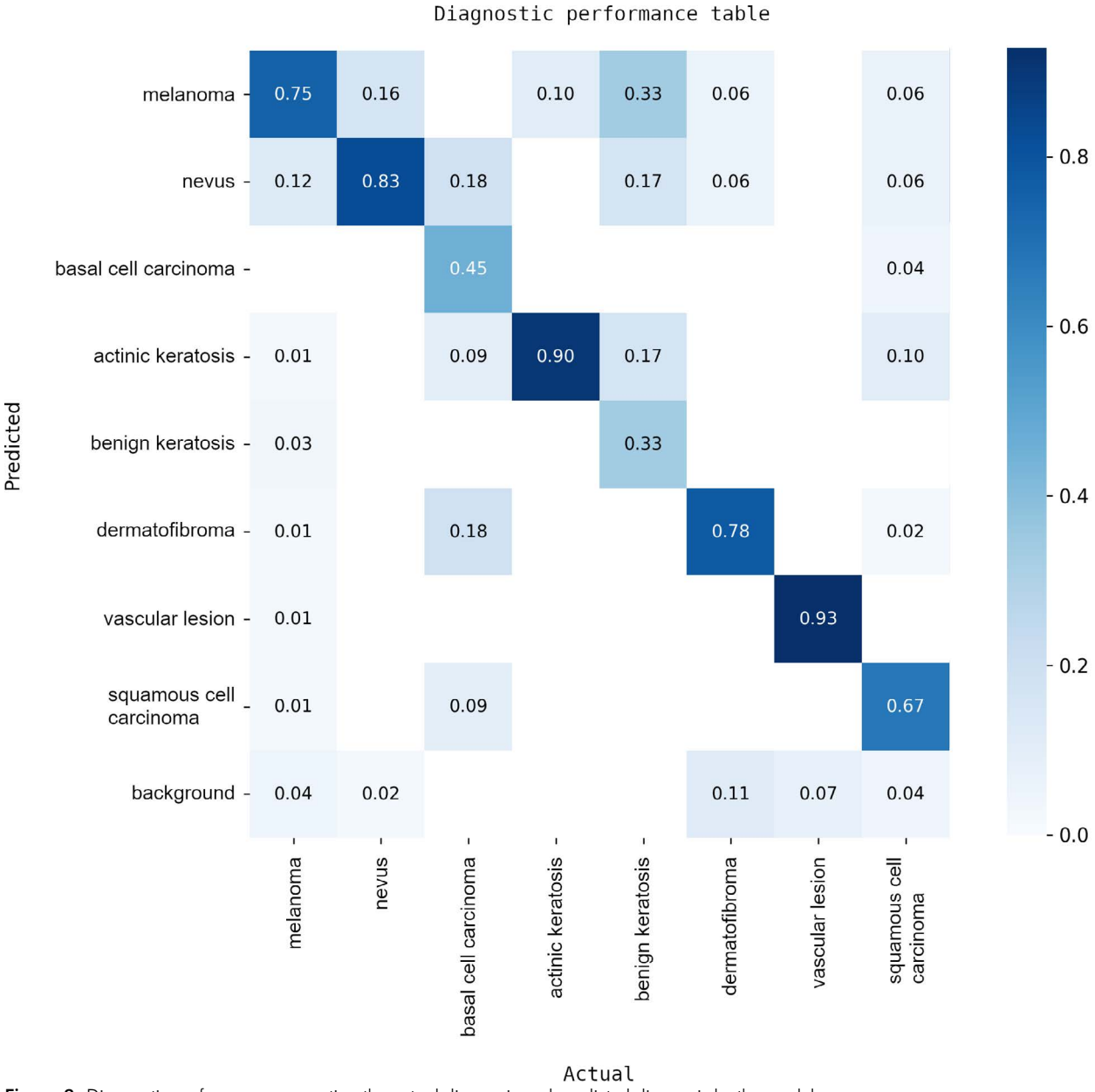

Figure 2. Diagnostic performance presenting the actual diagnosis and predicted diagnosis by the model
Note: The actual classes of the lesions are shown in the x axis and the predictions of the model in the y axis. The main diagonal represents the hits. When the model classifies a malignant as benign lesion, a false-negative occurs, a situation of major clinical risk because it delays the diagnosis. When a benign lesion is classified as malignant, a false-positive occurs, potentially leading to biopsies or unnecessary referrals.

DISCUSSION

The results show the potential of the model proposed as supporting tool to the medical diagnosis in cutaneous oncology, most of all in screening and support to clinical decision. Notwithstanding the advances achieved, the heterogeneous performance among the classes of lesions indicates relevant limitations. Classes with similar morphologic patterns as nevus and melanomas are still subject to significant confusion, indicating that there is still space to improve the discriminative capacity of the system.

The low performance in specific categories as BCC and benign keratosis suggests impact either due to the sample disproportion or intrinsic complexity of these lesions. Strategies of expansion and diversification of the databases emerge as priority pathways to increase the model robustness.

Clinically, the occurrence of false-negatives is quite challenging because failing to detect a malignant lesion may lead to severe repercussions for the patient. The reduction of this type of error should be prioritized in future training to attempt to elevate the global accuracy above 95%.

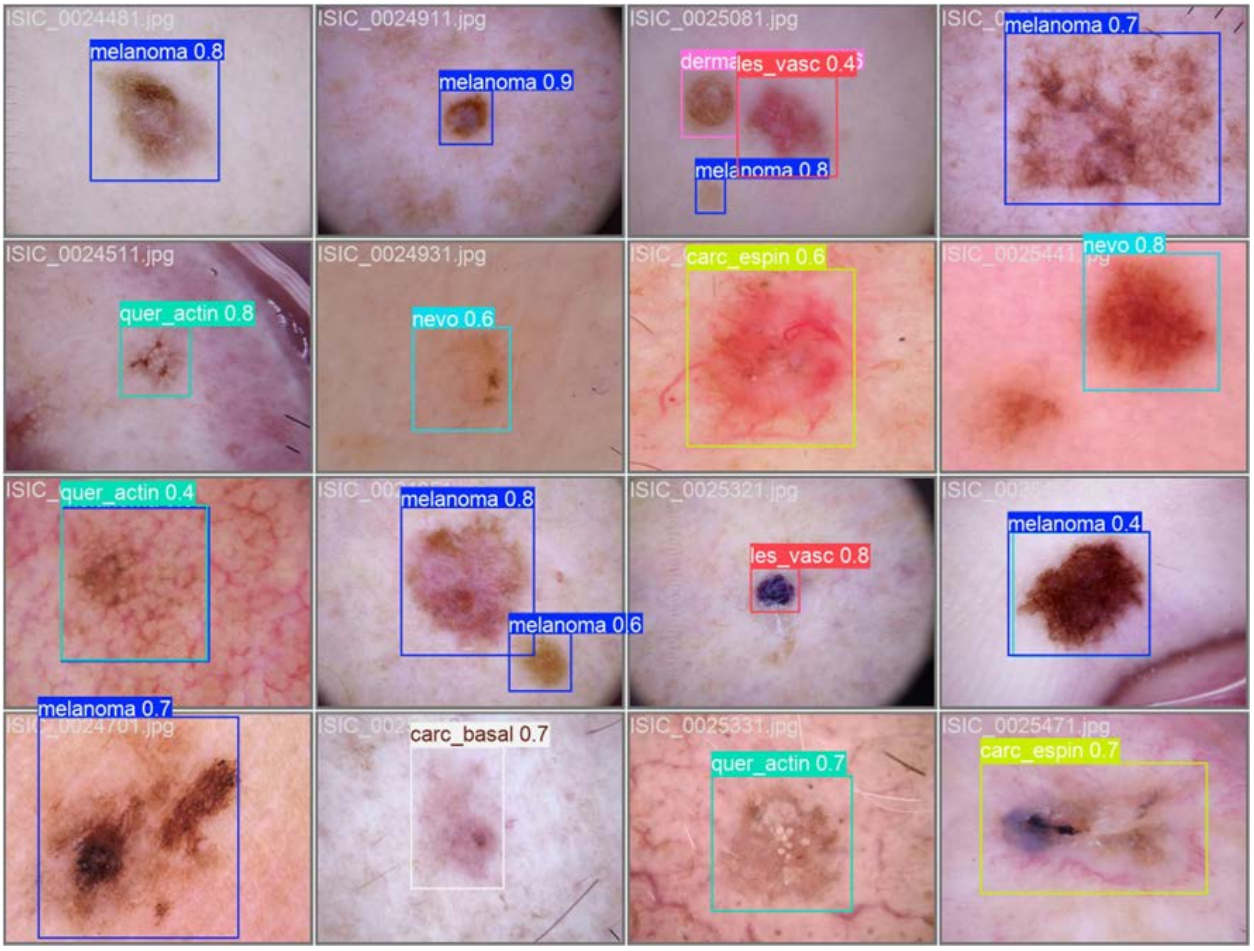

Figure 3. Examples of automated detection of skin lesions through the model developed together with the respective levels of confidence assigned by the planned diagnostic system

Note: Each image portrays the expected class and level of confidence (likelihood) of the prediction. (a) True-positive: the model identifies correctly the lesion and determines its region. (b) False-positive: the model classifies a benign lesion as malignant which can lead to unnecessary investigation. (c) False-negative: the model fails in identifying a malignant lesion which is a clinical risk due to possible diagnostic delay. These examples illustrate how the model responds to different dermoscopy patterns as irregular margins, asymmetry of pigmentation and presence of blood vessels.

Table 2. Performance of the model per class of lesion

Class of skin lesion	Sensitivity (%)	Accuracy (%)	Main errors observed
Melanoma	75	78	False-negative: confounded as nevus
Nevus	83	81	False-negative: confounded as melanoma
Basal cell carcinoma	45	72	Undetected in small or hypopigmented lesions
Actinic keratosis	90	84	Confounded with melanoma in irregular margins
Benign keratosis	33	68	Confounded with other keratosis
Dermatofibroma	78	70	Confounded with nevus or melanoma
Vascular lesion	93	89	Low error rate: easily identified pattern
Squamous cell carcinoma	67	74	Confounded with actinic keratosis

In view of international studies, the performance achieved in the present study is compatible with recent researches that utilized deep neural network (DNN) to detect melanomas. Studies conducted with DNN or adapted YOLO models report sensitivity between 72% and 89% and mean accuracy between 78% and 91% for identification of melanoma and BCC[11,18]. The sensitivity for melanoma in the present study was 75% and overall mean accuracy was 80.53%, within the range observed in the literature, which reinforces the technical feasibility of the model, even based on a reduced set of images annotated manually.

In addition, studies of SkinSage XAI by Munjal et al.[12] and the multicenter prospective study of Heinlein et al.[18] showed that AI-based systems can achieve similar performance or better than dermatologists in screening scenarios, strengthening the application of computational approach that can reduce costs and speed up clinical diagnoses. The performance of this study matches the global tendency, underpinning its scientific relevance.

The perspective of implementation of the model in mobile devices or web platforms is a strategic advance because it expands significantly the possibility of external validation and practical application. This portability favors collaboration tests across different country regions that systematically evaluate the tool. This approach potentializes evidences of robustness and generalization of the model, in addition to contributing to consolidate its applicability in scenarios of telemedicine and precision oncology.

Thus, the study can be seen as a proof of concept, demonstrating the technical and scientific feasibility of the use of the model proposed as supporting tool to dermatologic diagnosis. The creation of a complete pipeline (annotation → training → inference) and the validation with actual images of the clinical environment strengthen the potential of technological transference for SUS, particularly in areas with low availability of dermatologists and limited access to dermoscopy.

Regardless of the limitations, poor sensitivity yet in some classes of lesions and low representation of darker skin tones, the results reveal clinical potential as screening tool in health primary attention.

The model developed, in short, portrays its main potentialities:

- proven technical feasibility for automated detection of multiple types of lesions;
- performance compatible with international studies that utilize deep learning in dermoscopy;
- potential of application in clinical screening mostly in SUS;
- creation of bases for future studies with large diversity of data, multicenter balancing and validation. This activity has a strong potential through expansion and strengthening based on strategic institutional partnerships, mainly with INCA/MS. The support of these entities could expand its reach, integration with national databases and methodological improvement, contributing for the consolidation of a collaborative network focused to oncology monitoring and research in the country.

CONCLUSION

Important advances in the application of AI for assisted diagnosis of skin cancer have been demonstrated, most of all with the utilization of the architecture YOLOv11. The construction of a functional pipeline encompassing pre-processing of images up to model testing with public data manually annotated corroborated the technical feasibility of the proposal.

Among the most significant contributions, the following stand out: implementation of model of automated detection of multiple classes of skin lesions from dermatoscopy images, expanding its potential of clinical application; the development of a dataset with manual annotations, prioritizing relevant visual patterns as vessels and margins with direct impact on the accuracy of the model; integration of computational-based technologies and AI in a reproducible and scalable working flow.

Despite the advances, there are some important limitations: the accuracy of the model, although promising at 80.53% still does not reach the required level to be utilized in clinical practice. This limitation will be addressed in future training cycles that will be performed in the next six months with expanded databases and more diversified; the set of images that were utilized had low representativeness for darker skins, which compromises the generalization of the model for the Brazilian population. This gap exposes the necessity of national initiatives focused to systematic collection of dermatologic images with racial-ethnicity diversity; the external validation, although initiated with unreleased images captured in clinical environment, requires further details on sampling, inclusion criteria and clinical parameters evaluated to ensure the extrapolation of the results to health services.

These considerations show the transformative potential of the model concurrently with the overcoming of the challenges in addition to opening pathways for more robust results and applicable in clinical practice.

ACKNOWLEDGMENT

To NVIDIA for computational support which was critical to perform the experiments and analysis.

CONTRIBUTIONS

All the authors contributed substantially to the conception and design of the study, acquisition, analysis and interpretation of the data, writing and critical review. They approved the final version for publication.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interests to declare.

DATA AVAILABILITY STATEMENT

Dataset generated and analyzed while the study was being developed are available at the repository https://github.com/Sandrocamargo/publications/tree/main/rbc25a.

FUNDING SOURCES

None.

REFERENCES

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, Zalaudeck 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 loca]: 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

Scientific-editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777