



# Artificial Intelligence for Melanoma Diagnosis: A Qualitative Synthesis

Lucas Augusto Souza <sup>a++\*</sup>, Eduarda Breancini <sup>a++</sup>,  
Fernanda Staub Zembruski <sup>a#</sup>, Marina de Queiroz <sup>at</sup>  
and Junir Antônio Lutinski <sup>at</sup>

<sup>a</sup> Community University of the Chapecó Region – UNOCHAPECÓ, Rua Servidão Anjo da Guarda, 295-D, Efapi, Chapecó – SC, 89809-900, Brazil.

## *Authors' contributions*

*This work was carried out in collaboration among all authors. Authors LAS, EB, MdQ and JAL conducted the study, did bibliographic review, and contributed to wrote the manuscript. Author FSZ was responsible for the translation and formatting of the manuscript. All authors read and approved the final manuscript.*

## *Article Information*

DOI: <https://doi.org/10.9734/acri/2025/v25i61251>

## *Open Peer Review History:*

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://pr.sdiarticle5.com/review-history/136414>

## **Systematic Review Article**

**Received: 21/03/2025**

**Accepted: 24/05/2025**

**Published: 31/05/2025**

## **ABSTRACT**

**Aims:** To identify what the scientific literature presents regarding the accuracy of Artificial Intelligence (AI) in diagnosing melanoma, to compare it with the accuracy of dermatologists, and to identify possible limitations.

**Study Design:** Review

**Place and Duration of Study:** Brazil, 01/19/2024.

**Methodology:** The methodology used was a literature review, in which the studies included in the analysis were collected from the PubMed, Scielo, Cochrane, and Science Direct databases.

<sup>++</sup> Medical Student;

<sup>#</sup> PhD Student;

<sup>†</sup> Faculty Member;

\*Corresponding author: Email: [lucas\\_szza@unochapeco.edu.br](mailto:lucas_szza@unochapeco.edu.br);

**Cite as:** Souza, Lucas Augusto, Eduarda Breancini, Fernanda Staub Zembruski, Marina de Queiroz, and Junir Antônio Lutinski. 2025. "Artificial Intelligence for Melanoma Diagnosis: A Qualitative Synthesis". *Archives of Current Research International* 25 (6):78-88. <https://doi.org/10.9734/acri/2025/v25i61251>.

Inclusion and exclusion criteria were followed, after which the selected studies were classified by two investigators and analyzed through full-text reading.

**Results:** Eleven studies were included in the review. Most of the target population analyzed in the studies was male (36.4%), the predominant age group was 60 years or older (36.4%), and the predominant race was White. Three main diagnostic methods were identified for training the AIs: dermatoscopy, spectroscopy, and total body mapping. Fourteen different AI algorithms were evaluated by the studies. The highest sensitivity achieved by dermatologists in the studies was 96.5%, while the algorithms reached 97%. On the other hand, in terms of specificity, dermatologists achieved a maximum value of 99%, higher than the 97.1% achieved by the AIs.

**Conclusion:** AI demonstrated sensitivity values similar to those of dermatologists, surpassing them in overall values. On the other hand, the maximum specificity achieved by dermatologists was higher. Studies revealed biases in the databases, a lack of demographic information about the patients, dataset restrictions and a small number of participating specialists. Therefore, AI emerges as a potential complementary method for melanoma diagnosis rather than a replacement for specialists; however, more robust studies are needed for its implementation in clinical practice.

**Keywords:** Dermatologists; dermoscopy; algorithms; machine learning.

## 1. INTRODUCTION

Melanoma is a tumor derived from melanocytes, mostly of cutaneous origin. The four main types of malignant melanoma are: lentigo maligna, acral lentiginous, superficial spreading, and nodular. The lentigo maligna subtype generally presents poorly defined, irregular margins and usually occurs on sun-exposed skin in elderly patients. Acral lentiginous melanoma occurs on the skin of the palms, soles, and subungual regions, and often shows extensive invasion during the vertical growth phase. Superficial spreading melanoma is the most prevalent type, accounting for about 70% of cases, and affects the trunk and extremities more frequently. Nodular melanoma is the second most common type, characterized by rapid growth and early invasiveness, and is therefore considered the type with the worst prognosis (Bologna, 2015).

Epidemiologically, melanoma is more commonly found in light-skinned adults and can appear anywhere on the skin and mucous membranes. In individuals with darker skin, it is more common in lighter areas, such as the palms of the hands and the soles of the feet (Bologna, 2015). In Brazil, the estimated number of new cases in 2023 is 8,980, corresponding to 4.13 cases per 100,000 inhabitants. The state of Santa Catarina has the highest incidence of melanoma skin cancer in the country (INCA, 2022).

Malignant melanoma is one of the most challenging cancers to diagnose, as it requires the expertise of a dermatologist to detect lesions at an early stage, and of a pathologist to interpret the complex architecture of skin biopsies (Cabrera & Recule, 2018). Furthermore, malignant melanomas present various

clinicopathological and cytopathological manifestations. The clinical presentation of melanoma varies by subtype, and the diagnostic suspicion may be raised through physical examination using simplified screening methods, such as the "ABCDE" rule, which uses the initials of five clinical characteristics of skin lesions that may indicate malignancy (Asymmetry, Border irregularity, Color variation, Diameter greater than 6 mm, and Evolution).

The use of tools such as the dermatoscope since the 1980s has allowed the observation of structures and colors of the epidermis, the dermo-epidermal junction, and the superficial dermis, which are not visible to the naked eye, thereby increasing the diagnostic accuracy of malignant lesions. However, histological analysis remains the gold standard for diagnosis and should always be requested. This analysis provides the clinicopathological subtype, Breslow depth, mitotic rate, and the presence of ulceration and microsatellitosis (Garbe et al., 2022).

Currently, artificial intelligence (AI) has been the subject of studies as a new tool for melanoma diagnosis. A pioneering study in this area by Sboner et al. (2003) reported that both specialists and non-specialists often fail to fully follow diagnostic guidelines. This leads to poor medical practice and delayed treatments, contrary to what is expected in skin cancer with high mortality rates. Consequently, computerized system models were created, trained to analyze various diagnostic variables and pattern recognition methods in melanoma diagnosis, and to identify inconsistencies and divergences in diagnoses made by physicians.

Zhang *et al.* (2022) found that convolutional neural network (CNN) projects can achieve an accuracy of over 90% after being trained on dermoscopic image databases. AI systems perform better in several tasks, which could improve patient prognosis and aid in diagnosis.

The most accessible machine learning models used for melanoma diagnosis rely on clinical photographs, either from databases or personal collections. These images are input into a pre-trained network such as ResNet, VGG-16, or Inception V3, or into an untrained CNN (Grossarth *et al.*, 2023). Thus, the International Skin Imaging Collaboration (ISIC) created a dermatological image database to assist in testing machine learning models in studies (Rotemberg *et al.*, 2021).

In this context, this study aimed to identify what the scientific literature reports about the efficiency and accuracy of artificial intelligence in diagnosing melanoma skin cancer, in addition to comparing it with dermatologists' accuracy and identifying possible limitations of AI in effective melanoma diagnosis.

## 2. MATERIALS AND METHODS

This is an integrative literature review. This type of study aims to apply explicit and systematic search methods to select studies capable of

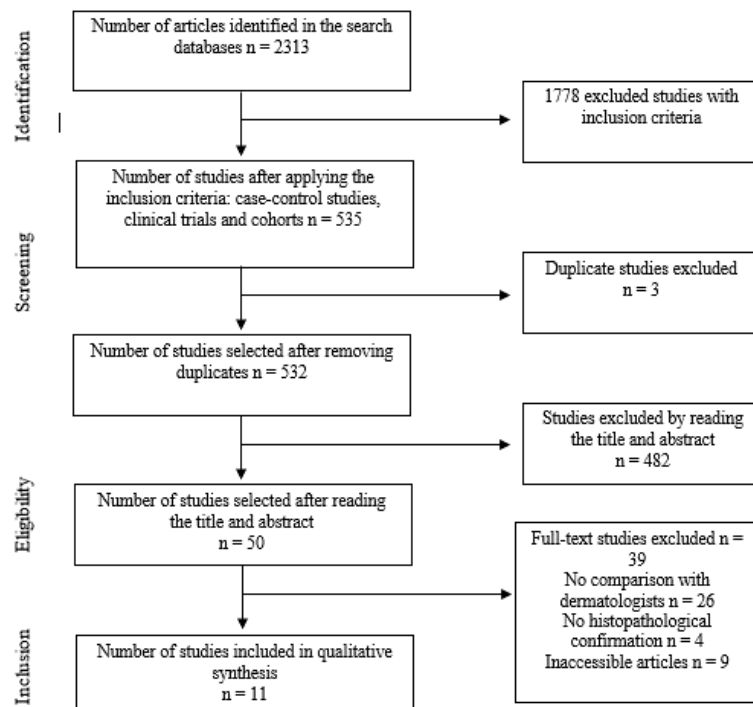
providing scientific evidence and to offer a synthesis of that evidence, aiming to facilitate its implementation in evidence-based practice in the health field (Guanilo *et al.*, 2011).

Studies were searched in the PubMed, Scielo, Cochrane, and Science Direct databases. Only scientific texts published in the form of articles were included. Case-control studies, clinical trials, and cohort studies that analyzed the accuracy of artificial intelligence in the diagnosis of melanoma-type cancer were included. No restrictions were made regarding language or publication date.

Exclusion criteria included the absence of diagnostic accuracy data, lack of histopathological confirmation of melanoma and comparison with dermatologists, animal studies, and systematic reviews.

The search in the databases used the following keywords: melanoma combined with diagnosis or "diagnosis" and artificial intelligence or "Machine Learning" or "artificial intelligence" or "AI". The searches were performed on January 19, 2024.

After the initial collection in all databases, inclusion criteria were applied for the next analysis phase, in which two investigators evaluated and extracted data using a standard organization tool (EndNote).



**Fig. 1. Methodological flowchart for acquiring studies in the systematic review**

Source: The authors (2024)

Initially, articles were imported into the platform and duplicates were removed. Each investigator evaluated the abstracts of the studies and organized them into three subgroups: included, excluded, and uncertain, according to the eligibility criteria. Titles and abstracts were reviewed, and studies were placed into the "included" group. A third reviewer evaluated articles categorized as "uncertain" to determine their inclusion. The flowchart is presented in Fig. 1.

After article selection, full-text reading was conducted, and data were entered into a database created in Excel (Microsoft Corporation, 2016). The table included the following participant profile information: authors, journal, publication year, study title, country where the study was conducted, study objective, types of melanoma presentation, diagnostic methods used, AI applied, accuracy of the AI,

accuracy of the dermatologists, dermatologists' method of analysis, and intrinsic and extrinsic limiting factors.

This study ensures ethical considerations, guaranteeing the authorship of the ideas, concepts, and definitions of the analyzed articles, using ABNT norms for citations and references. Since this is an integrative literature review, it does not require submission to or approval by the Research Ethics Committee (CEP).

### 3. RESULTS AND DISCUSSION

A total of 11 articles were included in this review. It was found that 90% of the articles were published in the last five years, during which 54.5% of the studies focused exclusively on melanoma-type skin cancer. Each study proposed its own artificial intelligence algorithm using different technologies (Table 1).

**Table 1. Authors, journal, and objectives proposed by the studies included in the integrative review, 2024**

Authors and Year	Journal	Objectives
Sboner <i>et al.</i> , 2003	Artificial Intelligence in Medicine	Improve the performance of single-system algorithms for early-stage melanoma detection and compare with dermatologists
Brinker <i>et al.</i> , 2019a	European Journal of Cancer	Train a convolutional neural network algorithm to classify melanoma images
Brinker <i>et al.</i> , 2019b	European Journal of Cancer	Compare the performance of a convolutional neural network trained with dermoscopic images for identifying melanoma in clinical photographs versus manual classification by dermatologists
Maron <i>et al.</i> , 2019	European Journal of Cancer	Compare the sensitivity and specificity of 112 German dermatologists with a single AI for lesion detection
Phillips <i>et al.</i> , 2019	JAMA Netw Open	Compare diagnostic accuracy between an AI and dermatologists using photos of potential melanomas
MacLellan <i>et al.</i> , 2021	Journal of the American Academy of Dermatology	Compare the diagnostic accuracy of bedside clinical examination by dermatologists, teledermatology, and non-invasive imaging techniques for melanoma (FotoFinder®, Melafind®, Verisante AuraTM)
Combalia <i>et al.</i> , 2022	The Lancet Digital Health	Simulate various AIs in real-world scenarios and compare with dermatologists
Lucieri <i>et al.</i> , 2022	Computer Methods and Programs in Biomedicine	Evaluate a new explainable AI framework for biomedical image analysis, providing multimodal concept-based explanations supported by visual maps to justify predictions
Aishwarya <i>et al.</i> , 2023	Procedia Computer Science	Train a convolutional neural network algorithm to classify images and compare results with dermatologists and residents
Menzies <i>et al.</i> , 2023	The Lancet Digital Health	Compare diagnostic decisions made by an algorithm versus healthcare professionals in clinical practice
Manolakos <i>et al.</i> , 2024	JAAD International	Evaluate the safety and effectiveness of an elastic-scattering spectroscopy (ESS) device in assessing melanocytic lesions suggestive of skin cancer

Source: The authors (2024)

**Table 2. Participant profile reported by the included studies by sex, age, and race, 2024**

Variable	Number of Articles	Percentage (%)
Age Group		
≥ 60 years	4	36.4
< 60 years	3	27.3
Not reported	4	36.4
Sex		
Male	4	36.4
Female	1	9.1
Not reported	6	54.5
Race		
White	3	27.3
Not reported	8	72.7

Source: The authors (2024)

**Table 3. Diagnostic methods used in the studies with their respective Artificial Intelligences, dermatologists' analysis methods, and authors, 2024**

Diagnostic Methods	AI Used	Dermatologist Analysis Method	Authors
Dermoscopy	MEDS architecture	Dermoscopy image database	Sboner <i>et al.</i> , 2003
	ResNet50 CNN 1	Dermoscopy image database	Brinker <i>et al.</i> , 2019a
	ResNet50 CNN 2	Dermoscopy image database	Brinker <i>et al.</i> , 2019b
	ResNet50 CNN 3	Dermoscopy image database and clinical analysis via questionnaire	Maron <i>et al.</i> , 2019
	Deep Ensemble for Recognition of Malignancy Algorithm (DERMA)	Dermoscopy image database	Phillips <i>et al.</i> , 2019
	MelaFind	Teledermoscopy and clinical analysis	MacLellan <i>et al.</i> , 2021
	FotoFinder Moleanalyzer Pro	Teledermoscopy and clinical analysis	MacLellan <i>et al.</i> , 2021
	EfficientNet	Dermoscopy image database	Combaila <i>et al.</i> , 2022
	XAI	-	Lucieri <i>et al.</i> , 2022
	YOLO V3 e V4	-	Aishwarya <i>et al.</i> , 2023
Spectroscopy	Elastic-scattering spectroscopy	Dermoscopy and clinical analysis	Manolakos <i>et al.</i> , 2024
	Verisante Aura	Teledermoscopy and clinical analysis	MacLellan <i>et al.</i> , 2021
Total body mapping	7-class-1 diagnosis AI	Dermoscopy and clinical analysis	Menzies <i>et al.</i> , 2023
	ISIC algorithm	Dermoscopy and clinical analysis	Menzies <i>et al.</i> , 2023

Source: The authors (2024)

The majority of the study populations were male (36.4%), female (9.1%), and 54.5% did not report gender. The predominant age group was 60 years or older (36.4%), and the predominant race was white. More than half of the articles did not report gender or race (Table 2).

A total of three main diagnostic methods were identified in the studies for training the algorithms: dermoscopy, spectroscopy, and total body mapping. For dermoscopy, different algorithms were used, with three based on the same ResNet50 CNN architecture. For both spectroscopy and total body mapping, two

different algorithms were identified for each method. The most used method by dermatologists was dermoscopic image banks; only three studies performed dermoscopy and in vivo clinical analysis (Table 3).

The main parameters evaluated for AI model accuracy and comparison with dermatologists were sensitivity and specificity. The highest sensitivity achieved by dermatologists was 96.5%, while the algorithms reached up to 97%. However, dermatologists reached a maximum specificity of 99%, higher than the 97.1% achieved by AIs. The lowest specificity among

algorithms was observed in spectroscopy as the standard diagnostic method (26.2%), while the lowest sensitivity occurred in total body mapping (55.6%). Not all articles reported complete sensitivity and specificity data (Table 4).

In total, twelve AI models presented comparative information on specificity (Fig. 2) and eight on sensitivity (Fig. 3), both from dermatologists and the algorithm, allowing for an overall comparison.

Artificial intelligence plays an important role in melanoma management, presenting both opportunities and challenges in its application (Zhang *et al.*, 2022). For training the algorithms included in this review, a database with images of benign and malignant lesions was required. Most of the studies (85%) used dermoscopy and total body mapping, as these are the most available and widespread diagnostic methods. Confirmation to guide the comparison with dermatologists was performed through histopathology of the lesion, which is considered the gold standard for melanoma diagnosis (Bologna, 2015).

An analysis of demographic data reveals a lack of information, particularly concerning race—an important variable when dealing with skin lesions. This is likely because the majority of studies (54%) used pre-existing dermatological image databases from the internet that had been previously used in other research to train the AI. In the studies that reported demographic data, white patients were predominant (Phillips *et al.*, 2019; Menzies *et al.*, 2023; Manolagos *et al.*, 2024), highlighting the need for further studies involving populations with different skin phototypes to determine whether the accuracy would remain consistent. Similarly, Jones *et al.* (2022) found that commonly used image databases often fail to consistently report inclusion criteria and patient demographic data (including Fitzpatrick skin type), raising concerns about the quality and generalizability of the data. Moreover, MacLellan *et al.* (2020) excluded patients with Fitzpatrick skin types III and above due to limitations of the devices used in melanoma diagnosis for darker skin types.

Each algorithm used in the studies is unique, employing different programming techniques and trained with a specific number of images to enhance its accuracy. Each algorithm was named individually by the authors. Aishwarya *et al.* (2023) described the You Only Look Once (YOLO) algorithm, which uses a single neural network to handle all categorization and

prediction processing, potentially making it faster than models using multiple networks. On the other hand, Sboner *et al.* (2003) used three different classification systems combined to improve performance in recognizing malignant lesions. The Explainable AI (XAI) model employed two different classifiers—one to detect dermatological concepts in an image, and another to classify the findings in skin lesions—thus producing a diagnosis (Lucieri *et al.*, 2022).

Among studies that used dermoscopy as the diagnostic method, various image sources were used for training and testing the algorithms. Phillips *et al.* (2019) obtained images using smartphone cameras (iPhone 6s and Galaxy S6), while MacLellan *et al.* (2020) used the DermLite Cam, an electronic dermatoscope attached to a cellphone or camera. Other authors used online image databases. In spectroscopy, Manolagos *et al.* (2024) developed a wireless, battery-powered portable device that emits light pulses at various wavelengths (ranging from 360 to 810 nm), illuminating the skin lesion and capturing optical backscatter reflectance from the tissue. This heterogeneity in diagnostic methods across studies demonstrates the wide range of applications for AI in melanoma diagnosis—possibly even combining different algorithm types to enhance accuracy.

Of the 14 AI algorithms included in the reviewed studies, 8 demonstrated similar or superior accuracy compared to dermatologists. The study that showed the greatest difference in favor of the AI was that using ResNet50 CNN1, with a 26.5% higher specificity than dermatologists (Brinker *et al.*, 2019). Similarly, a review by Haggemüller *et al.* (2021) comparing AI and medical experts in dermoscopic image classification found that five of six studies favored the algorithms, with only one showing superior performance by dermatologists.

Menzies *et al.* (2023) reported that algorithms had lower accuracy than dermatologists but outperformed less experienced physicians. A group of dermatologists showed significantly better performance compared to algorithms using only one classification (Sboner *et al.*, 2003). A systematic review found that smartphone apps using AI-based analysis are not yet sufficiently promising in terms of accuracy and are associated with a high probability of failing to detect melanomas (Chuchu *et al.*, 2018).

In spectroscopy, dermatologists showed significantly higher specificity, with a 29.9%

advantage; however, in that study, all non-biopsied lesions recorded by dermatologists were presumed to be benign. Despite this, the overall sensitivity of the device was 96.67% for melanoma, statistically similar to that of dermatologists. Since the elastic-scattering spectroscopy (ESS) device is highly sensitive, non-invasive, safe, and portable, it could assist primary care clinicians in assessing suspicious skin lesions (Manolakos *et al.*, 2024).

Statistical significance (p-values) could be evaluated in only four studies, all addressing specificity. Two algorithms showed no significant difference compared to dermatologists, while one study favored dermatologists and another favored AI. Further statistical analysis was not possible due to the large heterogeneity among the studies.

After comparisons, the hypothesis is raised that collaboration between AI and dermatologists may provide the best outcomes for patients, integrating this technology into clinical practice. MacLellan *et al.* (2020) concluded that AI could serve as a useful aid to physicians, but not as a replacement for clinical decision-making. The development of accessible screening methods, such as algorithms for early melanoma detection, could increase healthcare efficiency, reduce diagnostic delays, and transform clinical workflows (Phillips *et al.*, 2019).

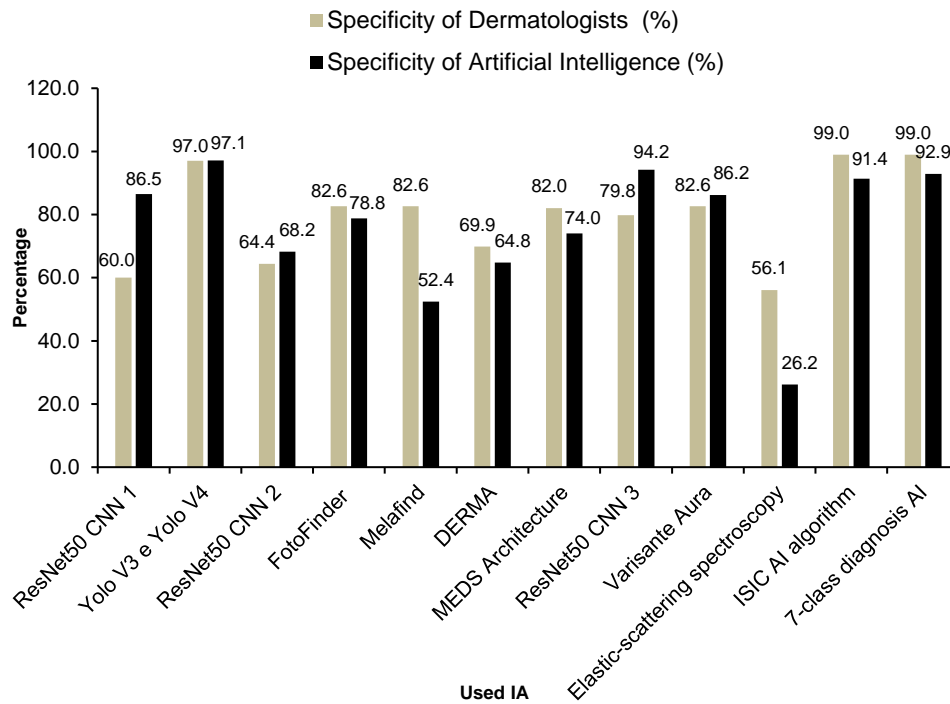
Some limitations of AI were identified in the studies included in this review. Brinker *et al.* (2019) noted that real clinical encounters provide more information than those available from database images, since dermoscopic images not only have higher resolution but also show greater visibility of the skin's underlying layers due to magnification and direct contact with the patient's skin. MacLellan *et al.* (2021) reported that low specificity and diagnostic precision suggest that some machines cannot replace the clinical experience of a dermatologist in selectively choosing which lesions to remove.

Likewise, methodological limitations and biases were identified. One major challenge in reducing bias is related to the image databases used to train the algorithms. Lucieri *et al.* (2022) inferred that current public datasets often suffer from low sample quality, due to lack of process standardization and lack of histological confirmation. Additionally, the limited number of images available—especially those with detailed conceptual annotations—leads to significant shifts in data distributions across datasets. Phillips *et al.* (2019) also reported potential bias in datasets due to image quality variations across different camera types. Therefore, a broader, more consistent, reliable, detailed, and high-resolution image database must be developed before real-world clinical implementation becomes viable.

**Table 4. Comparison of sensitivity and specificity between dermatologists and artificial intelligences (percentage), stratified by AI used, standard diagnostic method, and whether there was a significant difference (p-value) in specificity in the studies, 2024**

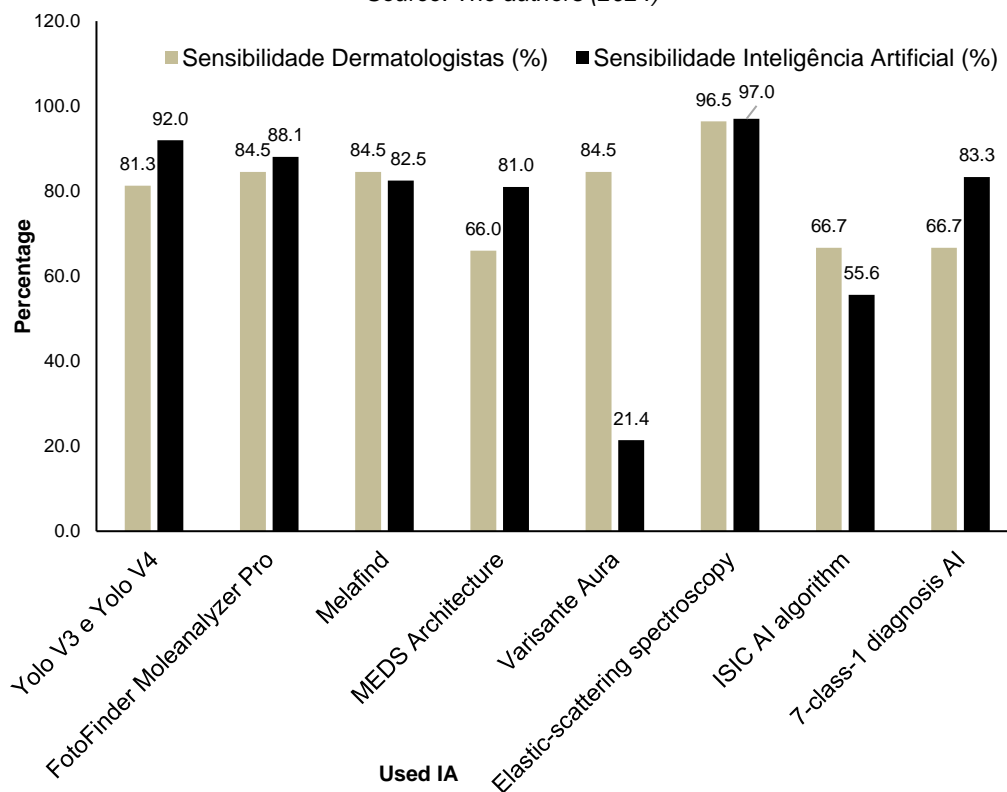
Standard Method	AI Used	Dermatologists		Inteligência Artificial		p-value (Specificity)
		Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	
Dermoscopy	ResNet50 CNN 1	74.1	60.0	74.1	86.5	p = 0.31
	Yolo V3 e Yolo V4	81.3	97.0	92.0	97.1	-
	ResNet50 CNN 2	89.4	64.4	89.4	68.2	-
	FotoFinder	84.5	82.6	88.1	78.8	-
	Moleanalyzer Pro					
	Melafind	84.5	82.6	82.5	52.4	-
	DERMA	-	69.9	-	64.8	-
	MEDS	66.0	82.0	81.0	74.0	p = 0.185
	architecture					
	ResNet50 CNN 3	63.5	79.8	63.5	94.2	p < 0.001
Spectroscopy	EfficientNet	-	-	-	-	-
	XAI	-	-	-	-	-
	Varisante Aura	84.5	82.6	21.4	86.2	-
Total Body Mapping	Elastic-scattering spectroscopy	96.5	56.1	97.0	26.2	p < 0.0001
	ISIC AI algorithm	66.7	99.0	55.6	91.4	-
	7-class-1 diagnosis AI	66.7	99.0	83.3	92.9	-

Source: The authors (2024)



**Fig. 2. Comparison of the specificity of melanoma diagnosis performed by dermatologists and artificial intelligences from each architecture used, in percentage, 2024**

Source: The authors (2024)



**Fig. 3. Comparison of the sensitivity of melanoma diagnosis performed by dermatologists and artificial intelligences of each architecture used, in percentage, 2024**

Source: The authors (2024)



Menzies *et al.* (2023) reported that their study was restricted to pigmented skin cancer and excluded low-quality images, which would not be feasible in broader sampling. Furthermore, testing was conducted in two academic centers with seven specialists and eighteen junior physicians, limiting generalizability to wider clinical contexts. Additionally, AI systems are unable to communicate what they do not know. For example, when presented with an image of a condition not represented in the training data, the system cannot label it as unknown and instead classifies it as one of the trained conditions (Combalia *et al.*, 2022).

The review not only revealed limitations but also highlighted future prospects for AI application. Brinker *et al.* (2019) reported that AI offers many advantages in diagnosing melanocytic lesions, including consistent interpretation, as AI assigns a distinct classification to each image at any given time, and potentially more accurate diagnoses than human experts of all training levels. Their findings suggest that AI algorithms could successfully assist dermatologists in detecting melanoma in clinical practice, though careful evaluation through prospective trials is necessary.

Aishwarya *et al.* (2023) suggested that their proposed model could be improved in the future by incorporating larger datasets for each type of skin lesion, aiding in more accurate investigation. Furthermore, newer versions of the algorithm could be tested to achieve better classification results. Lucieri *et al.* (2022) stated that their diagnostic modality aims to mediate subjectivity by offering a second opinion to the experienced physician, enhancing clinical reasoning and avoiding diagnostic oversight.

Authors also proposed the use of diagnostic algorithms as potential allies in primary care and remote areas. MacLellan *et al.* (2021) showed that computational analysis of dermoscopic images enhanced clinical diagnosis, reducing melanoma misdiagnoses. Additionally, advancements in AI tools could assist less experienced physicians in evaluating pigmented lesions. A low-cost AI-based screening tool could be employed in primary care settings to quickly assess concerning lesions. Phillips *et al.* (2019) echoed this, noting that AI-based services could transform patient diagnostic pathways, improving healthcare system efficiency.

Sboner *et al.* (2003) suggested future plans to integrate the algorithm into an electronic patient record, allowing clinical data to be combined with extracted image features, thereby expanding clinical knowledge and improving performance. Therefore, there are several promising prospects for using AI in early melanoma diagnosis, particularly in primary care. Despite the limitations, these findings point to opportunities for better leveraging this technology in healthcare.

#### 4. CONCLUSION

Artificial intelligence demonstrated sensitivity values comparable to those of dermatologists, even surpassing them in total values. On the other hand, dermatologists showed higher maximum specificity. Due to many heterogeneous characteristics, it was not possible to perform pooled parameter analysis; thus, p-values were not available for the majority of studies included in this review.

Studies revealed biases in the databases, a lack of demographic information about patients, and a small number of participating specialists. Moreover, a clinical consultation with a dermatologist extracts more information than a simple image analysis, and the use of a dermatoscope is indispensable in some cases, given its higher image resolution. In addition, AI is incapable of communicating what it has not been trained on, which can lead to misclassification.

It is inferred that AI is a promising tool to assist physicians—both non-specialists and dermatologists—in the early diagnosis of melanoma-type skin cancer. In this context, its role may be useful in areas with limited access to dermatologists, such as primary care units, serving as a tool to assist healthcare professionals in identifying suspicious lesions that require referral for specialized care.

Well-trained algorithms can achieve accuracy comparable to that of dermatologists using dermoscopy, with various possible implementations including smartphones, spectroscopy devices, electronic dermatoscopes, and professional cameras. Therefore, AI can be seen as a complementary method in melanoma diagnosis, rather than a replacement for specialists. However, more robust studies are necessary for the implementation of AI in clinical practice to become a reality.

## DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

## ACKNOWLEDGEMENTS

A brief acknowledgement section may be given after the conclusion section just before the references. The acknowledgments of people who provided assistance in manuscript preparation, funding for research, etc. should be listed in this section. All sources of funding should be declared as an acknowledgement. Authors should declare the role of funding agency, if any, in the study design, collection, analysis and interpretation of data; in the writing of the manuscript. If the study sponsors had no such involvement, the authors should so state.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES

- Aishwarya, N., et al. (2023). Skin cancer diagnosis with Yolo Deep Neural Network. *Procedia Computer Science*, 220, 651–658. <https://doi.org/10.1016/j.procs.2023.03.083>
- Bolognia, J. (2015). *Dermatologia* (3<sup>a</sup> ed.). Barueri-SP: Grupo GEN.
- Brinker, T. J., Hekler, A., Enk, A. H., Berking, C., Haferkamp, S., Hauschild, A., von Kalle, C., et al. (2019). A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. *European Journal of Cancer*, 111, 148–154. <https://doi.org/10.1016/j.ejca.2019.02.005>
- Cabrera, R., & Recule, F. (2018). Unusual clinical presentations of malignant melanoma: A review of clinical and histologic features with special emphasis on dermoscopic findings. *American Journal of Clinical Dermatology*, 19(S1), 15–23. <https://doi.org/10.1007/s40257-018-0373-6>
- Chuchu, N., et al. (2018). Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma. *Cochrane Database of Systematic Reviews*, 2018(12). <https://doi.org/10.1002/14651858.CD013192>
- Combaila, M., Codella, N., Rotemberg, V., Helba, B., Vilaplana, V., Reiter, O., Puig, S., et al. (2022). Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: The 2019 International Skin Imaging Collaboration Grand Challenge. *The Lancet Digital Health*, 4(5), e330–e339. [https://doi.org/10.1016/S2589-7500\(22\)00021-8](https://doi.org/10.1016/S2589-7500(22)00021-8)
- Garbe, C., et al. (2022). European consensus-based interdisciplinary guideline for melanoma. Part 1: Diagnostics – Update 2022. *European Journal of Cancer*, 170, 256–284. <https://doi.org/10.1016/j.ejca.2022.04.018>
- Grossarth, S., Goepfert, A., Rautenberg, J., & Schadendorf, D. (2023). Recent advances in melanoma diagnosis and prognosis using machine learning methods. *Current Oncology Reports*, 25(6), 635–645. <https://doi.org/10.1007/s11912-023-01407-3>
- Guanilo, M. C., Takahashi, R. F., & Bertolozzi, M. R. (2011). Revisão sistemática: noções gerais. *Revista da Escola de Enfermagem da USP*, 45(5), 1260–1266. <https://doi.org/10.1590/S0080-62342011000500033>
- Haggenmüller, S., Heppt, M. V., Reinholz, M., Berking, C., & Haferkamp, S. (2021). Skin cancer classification via convolutional neural networks: Systematic review of studies involving human experts. *European Journal of Cancer*, 156, 202–216. <https://doi.org/10.1016/j.ejca.2021.06.049>
- INCA - Instituto Nacional de Câncer José Alencar Gomes da Silva. (2019). *Estimativa 2023: incidência de câncer no Brasil*. <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/estimativa-2023.pdf>
- Jones, O. T., Robjant, S., Werth, B. P., Goodman, A., Greenhalgh, J., Hall, P. N., Emery, J., et al. (2022). Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: A systematic review. *The Lancet Digital Health*, 4(6), e466–e476.

- [https://doi.org/10.1016/S2589-7500\(22\)00023-1](https://doi.org/10.1016/S2589-7500(22)00023-1)
- Maclellan, A. N., Matin, R. N., Affleck, A., Barlow, R. J., Bataille, V., Bayne, A. P., Walker, C., et al. (2021). The use of noninvasive imaging techniques in the diagnosis of melanoma: A prospective diagnostic accuracy study. *Journal of the American Academy of Dermatology*, 85(2), 353–359.  
<https://doi.org/10.1016/j.jaad.2020.04.019>
- Manolakos, D., et al. (2023). Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: A comparative effectiveness study. *JAAD International*, 12, 120–126.
- Maron, R. C., Weichenthal, M., Utikal, J. S., Hekler, A., Berking, C., Hauschild, A., von Kalle, C., et al. (2019). Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. *European Journal of Cancer*, 119, 57–65.  
<https://doi.org/10.1016/j.ejca.2019.06.013>
- Menzies, S. W., Emery, J., Staples, M. P., Green, A. C., Buettner, P. G., Lo, S. N., Mann, G. J., et al. (2023). Comparison of humans versus mobile phone-powered artificial intelligence for the diagnosis and management of pigmented skin cancer in secondary care: A multicentre, prospective, diagnostic, clinical trial. *The Lancet Digital Health*, 5(10), e679–e691.  
[https://doi.org/10.1016/S2589-7500\(23\)00130-9](https://doi.org/10.1016/S2589-7500(23)00130-9)
- Phillips, M., Marsden, H., Jaffe, W., Matin, R. N., Greenhalgh, J., & Abbott, R. A. (2019). Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. *JAMA Network Open*, 2(10), e1913436.  
<https://doi.org/10.1001/jamanetworkopen.2019.13436>
- Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chou, S., Codella, N., Halpern, A., et al. (2021). A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific Data*, 8(1), 34.  
<https://doi.org/10.1038/s41597-021-00815-z>
- Sboner, A., et al. (2003). A multiple classifier system for early melanoma diagnosis. *Artificial Intelligence in Medicine*, 27(1), 29–44.  
[https://doi.org/10.1016/s0933-3657\(02\)00087-8](https://doi.org/10.1016/s0933-3657(02)00087-8)
- Zhang, S., Guo, L., Han, Z., Ma, J., Wang, W., & Zhao, Y. (2022). Artificial intelligence in melanoma: A systematic review. *Journal of Cosmetic Dermatology*, 21(11), 5993–6004.  
<https://doi.org/10.1111/jocd.15323>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the publisher and/or the editor(s). This publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright (2025): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Peer-review history:**

The peer review history for this paper can be accessed here:

<https://pr.sdiarticle5.com/review-history/136414>