revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

A DEEP-LEARNING BASED TOOL FOR HISTOPATHOLOGICAL DIAGNOSIS OF MELANOMA USING CONVOLUTIONAL NEURAL NETWORKS

DOI: 10.16891/2317-434X.v13.e4.a2026.id2848 *Recebido em: 13.03.2025 | Aceito em: 17.07.2025*

Thiago Magalhães Amaral[a]*, Jefferson Tales Oliva[b], Henrique Takashi Idogava[a]

Federal University of Vale do São Francisco – UNIVASF, Petrolina – PE, Brasil[a]
Federal University of Technology - Paraná – UTFPR, Pato Branco – PR, Brasil[b]
****E-mail: thiago.magalhaes@univasf.edu.br***

ABSTRACT

The application of deep learning to medical image analysis to aid diagnosis in healthcare has emerged as a powerful tool for early disease detection, thus improving treatment prospects and patient recovery. The aim of this study was to develop a convolutional neural network (CNN)-based predictive model to support the histopathological diagnosis of melanoma. To develop the proposed model, a dataset consisting of 411 images was used, including 393 of these in the experimental phase. The dataset was divided into 70% of the images for training and 30% for testing, and a model was constructed using ResNet50 architecture. The results showed that ResNet50 rapidly acquired the ability to distinguish features to accurately perform histopathological melanoma diagnoses. The error rate rapidly converged, achieving accuracy of approximately 90%. This proposed model is able to enhance diagnostic accuracy and support clinical practice in melanoma detection.

Keywords: Artificial Intelligence; Pathology; Melanoma.

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

INTRODUCTION

The use of deep learning in medical image diagnosis plays a crucial role in healthcare and has progressed significantly since 2012. It is applied in various tasks, including medical image classification (SILVA et al., 2020). Deep learning uses convolutional neural networks (CNNs) to analyze images and extract relevant information, promoting accurate diagnosis of several pathologies, including cancer and pneumonia (SILVA et al., 2020; DILDAR et al., 2021; KUMAR and SINHA, 2022).

Specifically for melanoma, deep learning has supported medical decision-making, as highlighted by Vollmer et al. (2025). Moreover, it can recognize patterns, which is effective for early detection of melanoma, thereby contributing to better therapeutic outcomes and patient survival (VOLLMER et al., 2025). Recent studies have reported gains in applying deep learning to identify melanoma (KUMAR and SINHA, 2022; VOLLMER et al., 2025; and ARUK, PACAL, and TOPRAK, 2026).

According to the Brazilian Society of Dermatology (SBD, 2025), skin cancer accounts for approximately 33% of all cancer diagnoses in Brazil, with an estimated 185,000 new cases per year. Researchers generally classify skin cancer into three main types: basal cell carcinoma, squamous cell carcinoma, and melanoma, the last of which is the most lethal (BHATT et al., 2022). Melanoma is an aggressive form of skin cancer that can rapidly metastasize throughout the body if not identified and treated early. Melanoma accounts for about 90% of deaths from skin cancer (SIEGEL et al., 2020) and in Brazil, this represents about 8,400 cases annually (SBD, 2025). Although melanoma is the least frequent type of skin cancer, it has the poorest prognosis (SBD, 2025). However, the chances of cure are over 90% when it is detected early, reducing the risk of severe complications (SBD, 2025).

According to the American Cancer Society (2021), the cure rate for melanoma can reach up to 99% when the disease is confined to the superficial skin layers, but this rate drops to approximately 15% when metastasis to other organs occurs. In this context, dermoscopic and histological evaluations play an essential role in melanoma diagnosis. Dermoscopy is a noninvasive semiological examination that assists dermatologists in evaluating pigmented lesions such as melanoma (SILVEIRA and GOULART, 2021). This tool enables detailed analysis of the skin surface and identifies characteristics such as asymmetry, irregular borders, color variations, and atypical pigment structures. According to the SBD (2025), melanoma generally has the appearance of a mole or mark on the skin, with brownish or blackish tones. For definitive diagnosis, a biopsy followed by histopathological analysis is required. This procedure enables precise evaluation of cellular characteristics and assists differentiation of benign lesions from malignant ones and determination of the stage and type of melanoma (BHATT et al., 2022). However, biopsy is also considered a rigorous, painful, expensive, and time-consuming procedure (BHATT et al., 2022).

Artificial intelligence (AI), in turn, has been used as a collaborative tool to assist dermatologists and has been improving the diagnostic accuracy of melanoma (VOLLMER, 2025). AI has also been used to analyze histopathological images to suggest possible cases of melanoma to pathologists (BHATT et al., 2022). Although these tools cannot replace the essential role of histopathological confirmation, they can assist clinicians in early detection of malignancies (BHATT et al., 2022).

To predict skin cancer, researchers most often use supervised learning methods, employing classification algorithms based on conditional decisions or probabilities. The most widely used methods include decision trees, artificial neural networks (CNN and ViT models), support vector machines (SVM), k-nearest neighbors (KNN), and others (LEE et al., 2018; SHAHIN et al., 2019; SALEM et al., 2019; KHAN et al., 2019; ZHANG et al., 2020; JUTZI et al., 2020; DILDAR et al., 2021; SHALINALI et al., 2021; ARAUJO, ARAÚJO, and SILVA, 2021, ARUK, PACAL, and TOPRAK, 2026).

The literature holds several important studies comparing diagnoses made by specialists with those made by artificial neural networks for the classification of skin lesions (LEITER et al., 2014; BHATT et al., 2022; VOLLMER et al., 2025). Leiter et al. (2014) reported that predictive models achieved sensitivity of up to 95% and specificity of up to 88%, while dermatologists showed similar sensitivity and specificity rates of 95% and 90%, respectively. Vollmer et al. (2025) showed that when dermatologists incorporated predictions provided by deep learning models based on CNN (level III), sensitivity

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

increased to 91.4% (88.3–94.5%, $p < 0.001$) without significantly altering specificity (74.2%) or AUC-ROC (0.954). The authors emphasized that the use of deep learning slightly improved the diagnostic accuracy of dermatologists, increasing sensitivity without loss of specificity, thus highlighting the potential of the tool in melanoma diagnosis.

Within the field of deep learning and its models, CNNs are a type of algorithm specifically designed to handle multidimensional data and to solve image recognition problems (AREL, ROSE and KARNOWSKI, 2010). According to Lecun, Kavukcuoglu, and Farabet (2010), CNN architectures are inspired by biological systems, and they are able to learn invariant features. Several important studies have applied CNNs to melanoma identification. For example, Kumar et al. (2022) proposed a CNN and a transfer learning architecture to assist in melanoma detection, training and testing the model using the publicly available dataset from the International Skin Imaging Collaboration (ISIC). The model achieved an accuracy of 81.2%, precision of 71%, recall of 97%, and an F1 score of 85%.

In another study, Rashid et al. (2022) employed a transfer deep learning model based on the MobileNetV2 architecture for melanoma classification, also using the 2020 dataset from ISIC, and achieved an impressive accuracy of 98.2%. Sayed, Soliman, and Hassanien (2021) proposed another hybrid approach, combining a CNN with a search optimization algorithm known as "Eagle" for melanoma recognition in skin lesions. This algorithm optimized the parameters for the SqueezeNet architecture, resulting in an accuracy of 98.37% in melanoma classification, with the ISIC 2020 dataset as the basis for the research. Additionally, Kaur et al. (2022) presented a CNN based on the LCNet network for automated melanoma classification. They used datasets from the ISIC database from 2016, 2017, and 2020; and the results showed accuracies of 81.41%, 88.23%, and 90.42%, respectively, demonstrating high performance for the model.

Several studies have applied deep learning at the histopathological level. For example, Hekler et al. (2019) reported a discrepancy of 25-26% in the classification of benign nevi and malignant melanoma among histopathologists. In their study, the authors trained a CNN using 595 histopathological images of melanoma and nevi, achieving mean sensitivity, specificity, and accuracy values of 76%, 60%, and 68%, respectively, across 11 tests. The authors then compared these results with those from 11 pathologists, who achieved mean sensitivity, specificity, and accuracy values of 51.8%, 66.5%, and 59.2%, respectively. The CNN thereby significantly ($p = 0.016$) outperformed the pathologists in image classification. Although AI can assist in diagnostic decision making, a pathologist should make the final diagnosis. It is important to note that only 2,006 pathologists in Brazil have this medical specialization, representing 0.75% of the total number of physicians (CREMESP, 2023).

Alheejawi et al. (2020) presented a state-of-the-art review regarding computer-assisted diagnosis of skin cancer/melanoma using histopathological images. The authors reported that deep learning techniques, such as SegNet and U-Net, have been successfully used to analyze histopathological images. Although a large number of studies have applied artificial intelligence techniques for detection and identification of melanoma using clinical photographs of skin or dermoscopic images (Fig. 1), few studies have been conducted at the histological level, focusing on identification of tumor cells after a biopsy (Fig. 2).

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

Figure 1. Images of skin lesions, a) melanoma and b) normal patient. Source: Alenezi, Armghan and Polat (2023).

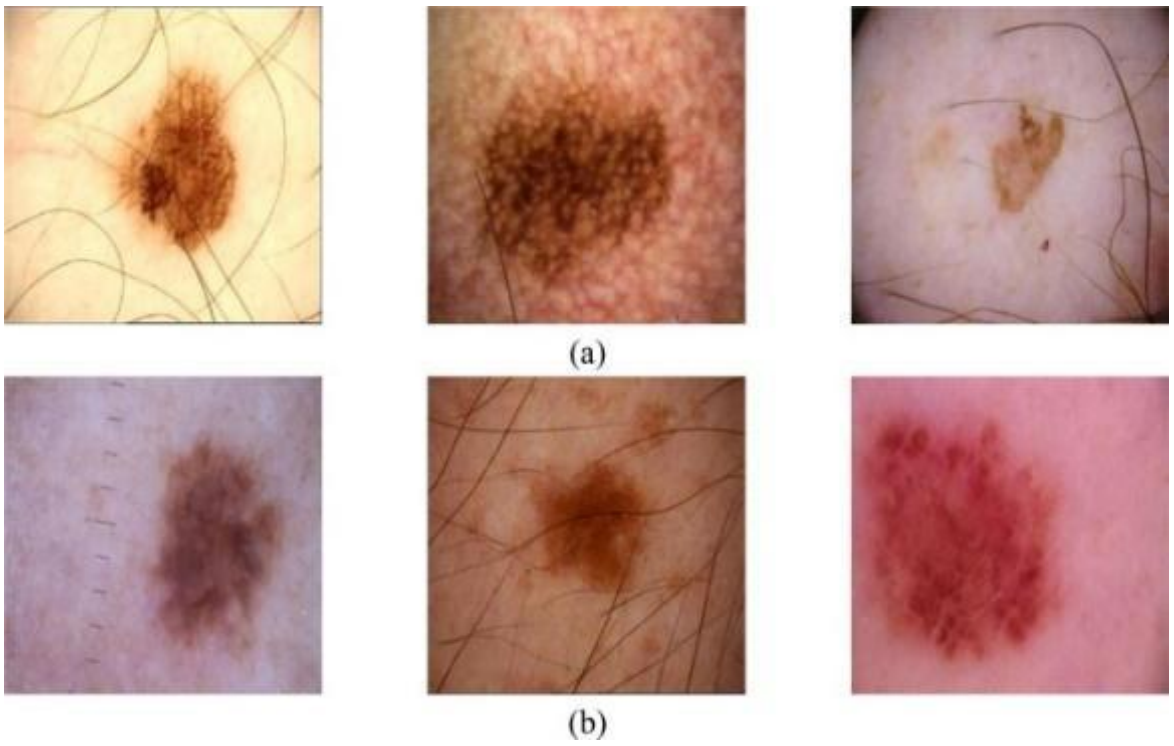
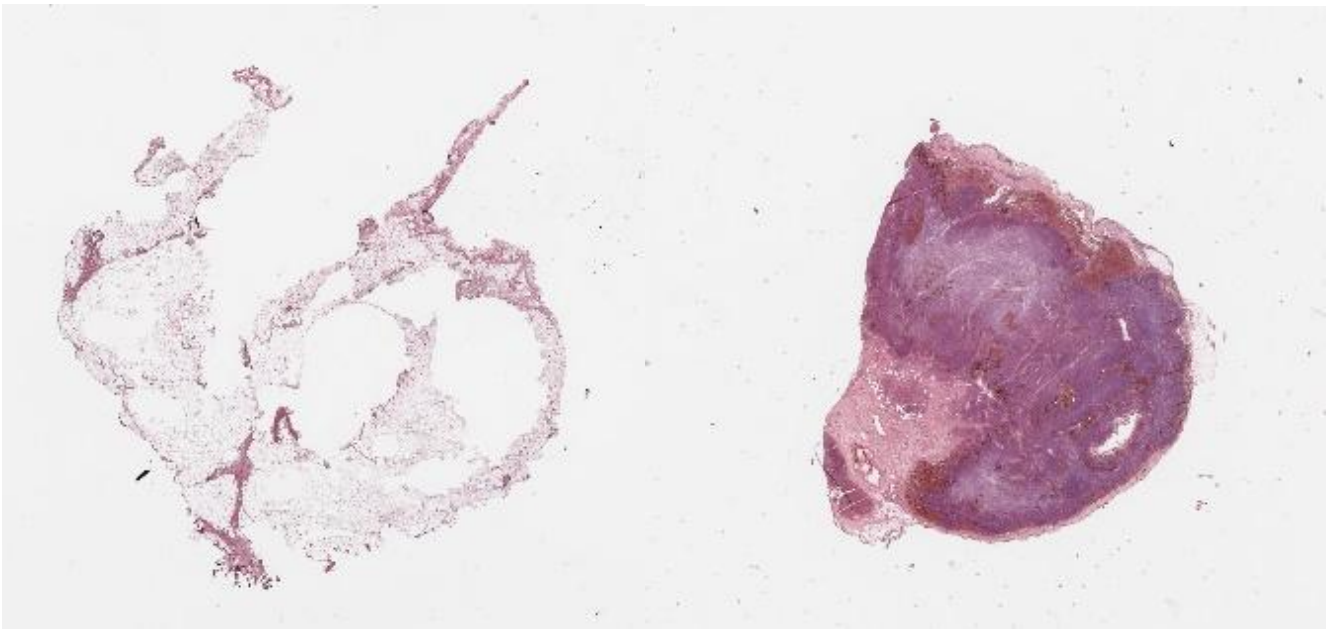


Figure 2. Histopathological images of a normal skin (left) and melanoma (right). Source: Nolan (2028).



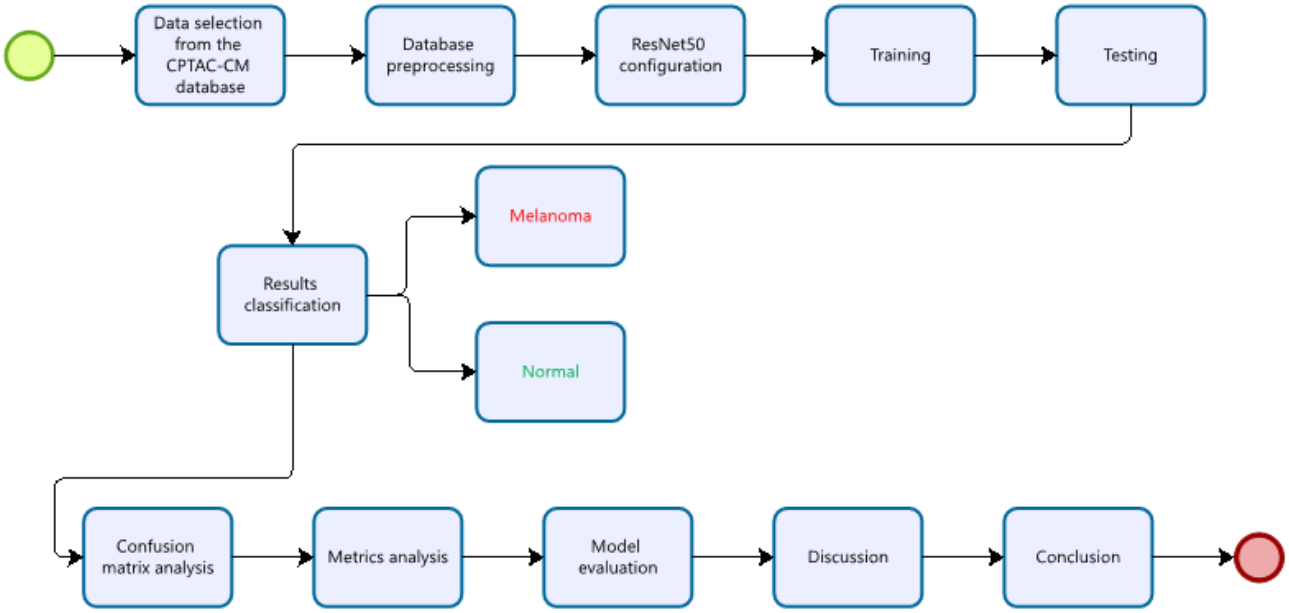
revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

Therefore, this aims of this study was to develop a CNN-based predictive model to support the histopathological diagnosis of melanoma.

MATERIALS AND METHODS

This research follows 11 steps as shown in Fig. 3, namely: Data selection from the CPTAC-CM database, Database preprocessing, ResNet50 configuration, Training, Testing, Results classification, Confusion matrix analysis, Metrics analysis, Model evaluation, Discussion, and Conclusion.

Figure 3. Research steps.



The dataset was obtained from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium - Cutaneous Melanoma (CPTAC-CM) platform, which aims to advance understanding of the molecular mechanisms of cancer through extensive proteomic and genomic analyses. The Cancer Imaging Archive systematically collects radiology and pathology images of CPTAC patients and makes these images available to the scientific community, facilitating investigations of cancer phenotypes that can be correlated with corresponding proteomic, genomic, and clinical data. The dataset comprises 411 images, 393 of which were used in the experiment. Among these, 282 presented images of melanoma in patients, while 111 were images of normal skin structures without melanoma.

The ResNet50 network was selected in this study due to its extensive parameter count, exceeding 23 million, and its validation in previous research with medical images (SILVA et al., 2020). According to Silva et al. (2020), ResNet is a well-established CNN model that serves as a fundamental framework for various computer vision tasks. This architecture has demonstrated remarkable effectiveness in several applications, including object recognition, object detection, and semantic segmentation. A distinctive feature of this network is its depth, consisting of 50 layers, which explains its name.

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

Silva et al. (2020) present further details on the ResNet architecture.

The widely used confusion matrix was adopted in this study to evaluate classification performance and to facilitate identification of error patterns and make adjustments to improve predictive accuracy. The confusion matrix provides a table with the frequencies of correct and incorrect classifications made by a machine learning model in relation to the true classes within a dataset. In binary classification, the confusion matrix comprises four main parameters (SILVA, 2018): true positive (TP), defined as the number of samples correctly classified as positive; true negative (TN), the number of samples correctly classified as negative; false positive (FP), the number of samples incorrectly classified as positive; and false negative (FN), the number of samples incorrectly classified as negative. In this research, accuracy, precision, recall, and the F1-Score were used. According to Silva et al. (2020), these metrics are defined and computed as follows:

Accuracy refers to the proportion of the total number of correct classifications (TP and TN) within the total number of classifications, including false positives and false negatives (FP and FN). Accuracy = $\frac{(TP + TN)}{(TP + FP + FN + TN)}$ (Eq. 1). Precision refers to the proportion of TP classifications among all positive classifications. Precision = $\frac{TP}{(TP + FP)}$ (Eq. 2). Recall is the proportion of TP and the number of TP plus FN. Recall = $\frac{TP}{(TP + FN)}$ (Eq. 3). The F1-Score is a weighted average of precision and recall, with values ranging from 0 (worst) to 1 (best) (SILVA, 2018). F1-Score = $\frac{2\ (Precision\ x\ Recall)}{(Precision + Recall)}$ (Eq. 4).

We organized the dataset, mainly because some images did not have a corresponding diagnosis (label) and had different dimensions. We manually removed 18 images from the original dataset containing 411 images. In this preprocessing step, the “ImageDataGenerator” object was used with the parameter “rescale=1/255” to normalize the pixel values of the test images, converting them from the original range of 0-255 to a range of 0-1, which improved the neural network performance. The “flow_from_directory” method was then applied to create a data generator that automatically reads images from a specific directory, resizing them all to 256 × 256 pixels, as required by the ResNet50 model architecture. The images were organized into batches of three and shuffled to avoid bias in the evaluation order. In addition, class categorization was defined as 'categorical', meaning that the labels were provided in a one-hot encoded format, suitable for binary classification problems such as melanoma detection. Thus, the test images were efficiently loaded, resized, normalized, and organized for evaluation by the trained model.

In this study, the main hyperparameters included batch size, which was initially set to 10 for the train generator function and later adjusted to 3 before model training. Model training was conducted for 30 epochs, where one epoch represents a complete pass through the entire training dataset. The number of epochs determines the length of the model learning period. Too many epochs may lead to overfitting, while too few may result in underfitting.

The categorical_crossentropy function was used as a loss function to quantify model error during training. In addition, the sigmoid activation function was used in the output layer. For binary classification, the sigmoid function is appropriate because it produces an output from 0-1, which can be interpreted as the probability of the positive class. Each of the two additional hidden dense layers including definition of 256 neurons were added to the base model. A dropout rate of 0.2 was used for the dense layers. Finally, the pooling size of the average pooling layer (AveragePooling2D) was initially set to (4, 4) and subsequently adjusted to (8, 8) during cross-validation.

The dataset was divided into 70% for training and 30% for testing. Figure 2 illustrates two histopathological images from this dataset, which were converted to grayscale. Python programming language was used to develop the CNN architectures. Google Colab was chosen to run the ResNet50 simulations, as it provides pre-configured Python 2 and 3 runtime environments and comes equipped with essential artificial intelligence libraries essential for our study (TensorFlow, Keras, NumPy, and Matplotlib). In addition, five simulations were performed to evaluate different model parameters, with the goal of achieving optimal accuracy.

The K-Fold cross-validation technique was applied to evaluate model performance since the dataset

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

was limited to 393 images. Instead of splitting the data into a single fixed training and test set, cross-validation divides the dataset into K subsets (folds) of approximately equal size. K-Fold cross-validation in this study was n_splits = 5. The process was performed five times. Upon concluding the five iterations, five sets of evaluation metrics were obtained, one for each fold.

RESULTS AND DISCUSSION

Fig. 4 illustrates the error and success rate during the training process. These results indicate that ResNet50 rapidly learned to distinguish melanoma at the histopathological level, as confirmed by experimental validation. The error rate quickly converged, ultimately achieving an accuracy of approximately 89.83%. The error rate stabilized between 0.2 and 0.3 toward the end of the training process (around epochs 25-30).

Figure 4. Error and success rate during training.

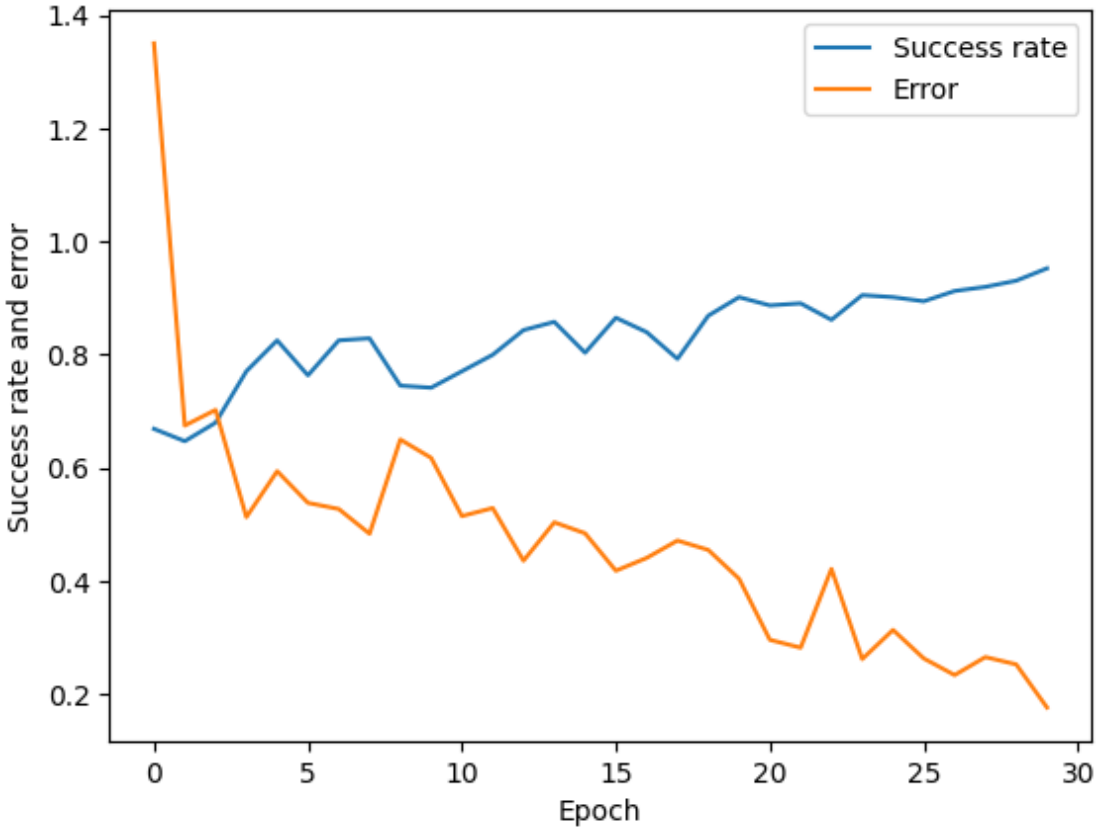


Fig. 5 presents the confusion matrix, with the results for True Positives (76), True Negatives (31), False Positives (2), and False Negatives (9).

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

Figure 5. Confusion Matrix.

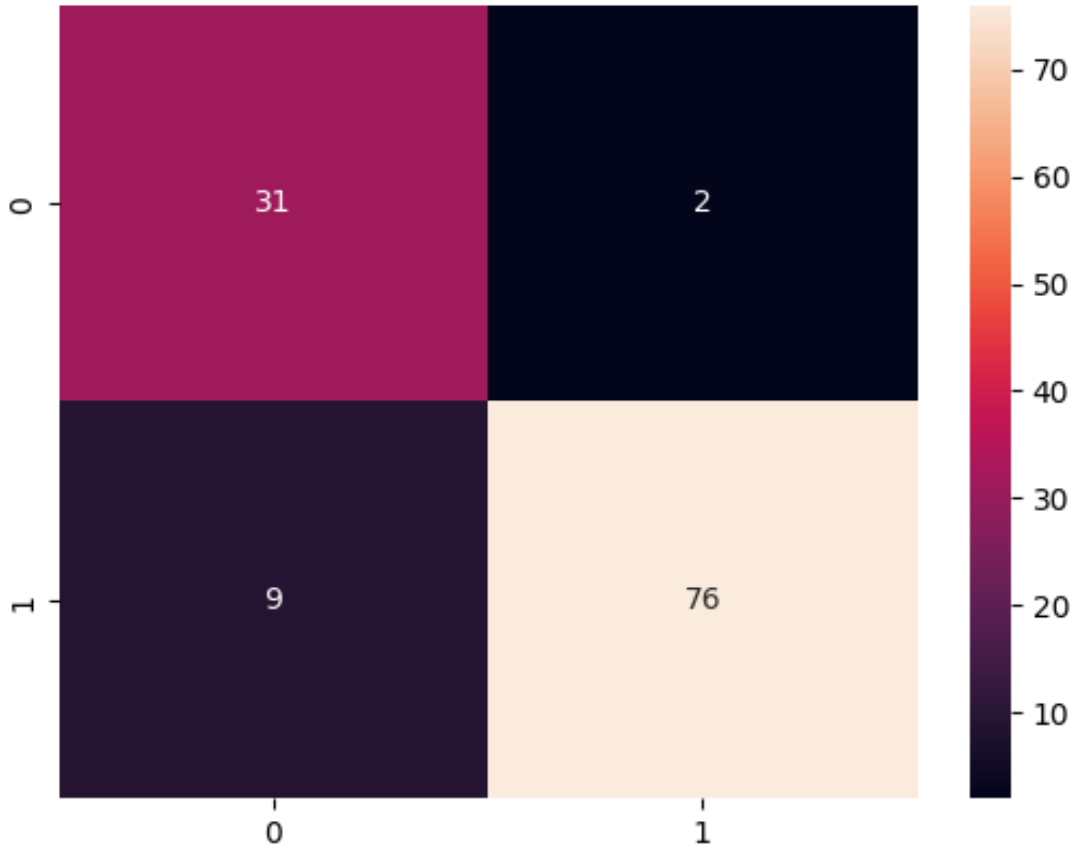


Table 1 presents the metrics obtained, including precision, recall, and F1-Score. Notably, the accuracy in identifying melanoma (97%) was higher than that for normal images (78%). This difference highlights the classifier's ability to avoid misclassifying negative samples as positive. Recall exhibited a higher value for normal images (94%) compared to melanoma images (89%). Finally, the F1-Score illustrates that the classifier performed well in identifying melanoma (93%), as the result approached a value closer to 1.

Table 1. Metrics referring to the proposed model.

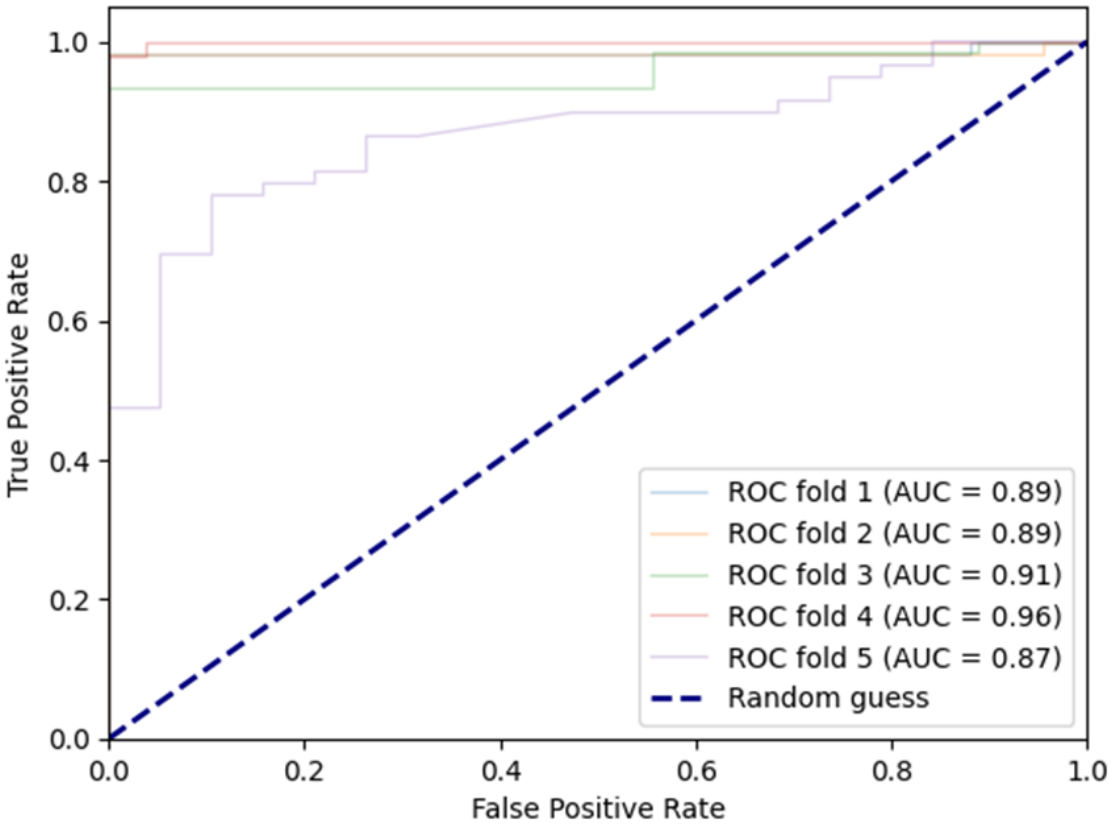
	Precision	Recall	F1-Score
Normal	0.78	0.94	0.85
Melanoma	0.97	0.89	0.93
Accuracy			0.91

Figure 6 shows the Receiver Operating Characteristic (ROC) curve used to evaluate the performance of the classification model. The ROC curve is a widely used tool for assessing the performance of binary classification models (normal vs. melanoma). The x-axis (false positive rate - FPR) is known as specificity, representing the proportion of negative cases that the model incorrectly classified as positive. The Y-axis (true positive rate - TPR) is known as sensitivity or recall, representing the proportion of positive cases correctly

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

classified by the model. The Area Under the Curve (AUC) refers to the area under the ROC curve, which is an aggregated measure that summarizes the overall performance of the classifier across all possible classification thresholds. The AUC value ranges from 0 to 1.

Figure 6. ROC Curves for each fold.



In this study, the ROC curve was well above the random dashed line, and the mean AUC value was 0.90. This indicates that the model performed well in distinguishing the "Normal" from the "Melanoma" classes. An AUC of 90% suggests a high probability that the model assigns a higher score to a randomly chosen positive case (melanoma) than to a randomly chosen negative case (normal). The consistent position of the curves above the random line and the AUC values consistently above 85% across all folds indicate that the model is reasonably robust. The mean AUC (90.19%) from cross-validation was very close to the accuracy of the model on the test set (89.83%). Fold 4 achieved the best ROC, with AUC = 0.96. Analysis of the ROC curve and the AUC is important for understanding the trade-off between sensitivity and specificity at different decision thresholds and for comparing the performance of different models.

The results demonstrate that the ResNet50-based model can effectively support medical decision-making. Our research findings are consistent with the existing literature. For example, the accuracy of our model is comparable to that of previous studies on melanoma identification, such as Lee et al. (2018) with 80.91%, Barata et al. (2015) with 84.30%, and Kaur et al. (2015) with 87.90%. Additionally, the work of Kumar and Sinha (2022), which employed CNNs and transfer learning, achieved an accuracy of 81.2%, precision of 71%, recall

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

of 97%, and an F1-Score of 85%, values closely mirroring those attained by our model, albeit with distinct datasets. It is noteworthy that all the aforementioned studies were conducted using dermoscopic images rather than histopathological ones.

In studies specifically involving histopathological images, our model achieved higher accuracy than that of Hekler et al. (2019), who reported 68%, although it fell short of the performance achieved by Xie et al. (2020), who obtained 98.5%. Hekler et al. (2019) used a dataset comprising 695 lesions classified by 11 expert histopathologists in accordance with current guidelines (350 nevi and 345 melanomas, resulting in 695 images). Similar to our approach, the authors employed ResNet-50 as the convolutional neural network (CNN). The ResNet-50 model achieved mean sensitivity of 76%, specificity of 60%, and accuracy of 68% across 11 test runs. In comparison, the 11 pathologists achieved mean sensitivity of 51.8%, specificity of 66.5%, and accuracy of 59.2%. Therefore, the authors concluded that the CNN was significantly superior ($p = 0.016$) in classifying histopathological images.

In contrast, Xie et al. (2020) developed a deep learning pipeline to enhance melanoma diagnosis using CNNs and to localize potential malignant melanoma tissue through gradient-weighted class activation mapping. Their dataset included 1,806 melanoma and 435 nevus images, a dataset approximately six times the size of ours. The CNN model achieved an AUC of 0.985, sensitivity of 93.6%, and specificity of 95.9%. The accuracy of their model (93.3%) was also compared with the accuracy of 20 pathologists (73.2%), with the model showing a clear performance advantage. Similar to Hekler et al. (2019), the authors compared the neural network results with expert pathologist assessments.

LIMITATIONS

A limitation of the present study is the relatively small sample size, comprising 411 images, of which 393 were used in the experiment. Future research should expand the dataset to enhance the accuracy and robustness of the model. A possible approach to increase the number of images is to add the dataset provided by Thomas et al. (2021). Data augmentation techniques were not applied in our study, due to the size and nature of the available dataset. Although data augmentation methods can help increase dataset diversity and potentially improve the generalization ability of the model, the relatively limited number of images and the clinical specificity of the melanoma images required a more conservative approach, to avoid introducing artificial variations that might not accurately reflect the true characteristics of the lesions. In other words, the primary focus of our work was to evaluate model performance using original data to ensure that the results were directly representative of clinical images without additional manipulations. Another limitation of our study was that a team of pathologists did not validate the model, unlike the approaches adopted by Hekler et al. (2019) and Xie et al. (2020). This validation could have confirmed and strengthened the potential applicability of our results.

CONCLUSION

In this study, we developed a CNN model for melanoma detection in histopathological evaluations. The proposed classification model achieved an accuracy of approximately 90% over 30 epochs, as indicated by the evaluation metrics. Although the CNN model demonstrated promising results in melanoma classification, including high accuracy metrics, it is important to emphasize that it should be considered a complementary tool in clinical assessment. A qualified pathologist must always make the final diagnosis. In future research, we intend to compare our results with alternative neural network architectures, such as DenseNet, Xception, MobileNetV2, and LCNet. We also intend to expand the dataset of pathological images, as we recognize that the dataset used here is still relatively small given the importance of the topic. Furthermore, we aim to implement the model within a healthcare setting to further evaluate its potential in supporting medical decision-making in real-world scenarios. With additional validation, we believe the model may help improve diagnostic accuracy and contribute to early detection efforts, ultimately enhancing clinical practice in management of melanoma.

revista interfaces
Artigos Originais
V. 13, N. 4 (2025) | ISSN 2317-434X

REFERENCES

ABADI, M.; BARHAM, P.; CHEN, J.; *et al*. TensorFlow: A System for Large-Scale Machine Learning. **12th USENIX Symp Oper Syst Des Implement (OSDI).** 2016.

ALENEZI, F.; ARMGHAN, A.; POLAT, K. A multi-stage melanoma recognition framework with deep residual neural network and hyperparameter optimization-based decision support in dermoscopy images. **Expert Systems with Applications,** v. 215, 119352, 2023. DOI: https://doi.org/10.1016/j.eswa.2022.119352.

ALHEEJAWI, S.; MANDAL, M.; XU, H, *et al*. Deep learning-based histopathological image analysis for automated detection and staging of melanoma. **Deep Learn Biomed Health Inform,** p. 237–265, 2020. https://doi.org/10.1016/B978-0-12-819061-6.00010-0.

AMERICAN CANCER SOCIETY. **Melanoma skin cancer.** https://www.cancer.org/cancer/melanoma-skin-cancer/detection-diagnosis-staging/survival-rates-for-melanoma-skin-cancer-by-stage.html; 2021 [accessed 4 February 2023].

ARAÚJO, R.; ARAÚJO, F.; SILVA, R. Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning. **Multimed Syst.** v. 28, n. 4, p.1239–1250, 2021. DOI: https://doi.org/10.1007/s00530-021-00840-3.

AREL, I.; ROSE, D.; KARNOWSKI, T. Deep machine learning – a new frontier in artificial intelligence research. **IEEE Comput Intell Mag,** v. 5, n. 4, p. 13–18, 2010.

ARUK, I.; PACAL, I., TOPRAK, A. N., A comprehensive comparison of convolutional neural network and visual transformer models on skin cancer classification, **Computational Biology and Chemistry,** v. 120, 2026. DOI: https://doi.org/10.1016/j.compbiolchem.2025.108713.

BARATA, C.; CELEBI, M.; MARQUES, J. Improving dermoscopy image classification using color constancy. **IEEE J Biomed Health Inform.** 2015. v. 19, n. 3, p. 1146–1152, 2015. DOI:. http://dx.doi.org/10.1109/JBHI.2014.2336473.

BHATT, H.; SHAH, V.; SHAH, K.; SHAH, R.; SHAH, M. State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review. **Intell Med.** 2022.

CREMESP. **Conselho Regional de Medicina do Estado de São Paulo. Sociedade alerta sobre falta de patologistas no Brasil.** https://www.cremesp.org.br/?siteAcao=Imprensa&acao=crm_midia&id=706 [accessed 4 February 2023].

DILDAR, M.; AKRAM, S.; IRFAN, M., *et al*. Skin cancer detection: a review using deep learning techniques. **Int J Environ Res Public Health,** v. 18, n. 10, 5479, 2021. DOI: https://doi.org/10.3390/ijerph18105479.

HEKLER, A.; UTIKAL, J.; ENK, A, *et al*. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. **Eur J Cancer,** v. 118, p. 91–96, 2019. DOI: https://doi.org/10.1016/j.ejca.2019.06.012.

JUTZI, T.; KRIEGHOFF-HENNING, E.; HOLLAND-LETZ, T.; *et al*. Artificial intelligence in skin cancer diagnostics: the patients' perspective. **Front Med (Lausanne).** v. 7, 2020. DOI: https://doi.org/10.3389/fmed.2020.00233.

KETKAR, N. Introduction to keras. **Deep Learn Python.** 2017.

KHAN, M.; HUSSAIN, A.; REHMAN, S.; KHAN, U.; MASQSOOD, M.; MEHMOO, K. Classification of melanoma and nevus in digital images for diagnosis of skin cancer. **IEEE Access,** v. 7, p. 90132–90144, 2019. DOI: https://doi.org/10.1109/ACCESS.2019.2926837.

KUMAR, V.; SINHA, B. Skin cancer classification for dermoscopy images using model based on deep learning and transfer learning. **Comput Intell Data Anal,** v. 142,

p. 257–271, 2022. https://doi.org/10.1007/978-981-19-3391-2_19.

LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: Circuits and Systems (ISCAS), **Proceedings of 2010 IEEE International Symposium** on. [S.l.: s.n.], p. 253–256, 2010.

LEITER, U.; EIGENTLER T.; GARBE, C. Epidemiology of skin cancer. **Adv Exp Med Biol.** p. 120–140, 2014. DOI: https://doi.org/10.1007/978-1-4939-0437-2_7.

LEE, H.; MENDES, A.; SPOLAÔR, N.; OLIVA, J., *et al*. Dermoscopic assisted diagnosis in melanoma: reviewing results, optimizing methodologies and quantifying empirical guidelines. **Knowl-Based Syst,** v. 158, p. 9–24, 2018. DOI: https://doi.org/10.1016/j.knosys.2018.05.016.

NOLAN, T. The Clinical Proteomic Tumor Analysis Consortium Cutaneous Melanoma Collection (CPTAC-CM) (Version 10). **Cancer Imaging Arch**. 2018. DOI: https://doi.org/10.7937/K9/TCIA.2018.ODU24GZE.

SALEM, C.; AZAR, D.; TOKAJIAN, S. An image processing and genetic algorithm-based approach for the detection of melanoma in patients. **Methods Inf Med,** v. 57, n. 1, p. 74–80, 2018. https://doi.org/10.3412/ME17-01-0061.

SAYED, G.; SOLIMAN, M.; HASSANIEN, A. A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. **Comput Biol Med,** v. 136, 2021.

SBD. SOCIEDADE BRASILEIRA DE DERMATOLOGIA. Câncer da pele. https://www.sbd.org.br/doencas/cancer-da-pele/ Acessed 13 December 2025.

SHAHINALI, S.; SIPONMIAH, M.; HAQUE, J., *et al*. An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. **Mach Learn Appl,** v. 5, 2021. DOI: https://doi.org/10.1016/j.mlwa.2021.100036.

SHAHIN, A.; KAMAL, A.; ELATTAR, M. Proceedings of 9th Cairo International Biomedical Engineering Conference (CIBEC). **IEEE.** 2019.

SIEGEL, R.; MILLER, K.; JEMAL, A. **Cancer statistics. CA Cancer J Clin**, v. 70, n. 1, p. 7–30, 2020.

SILVA, Rodrigo Emerson Valentim da. **Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens**. 2018. 51 f. TCC (Graduação em Sistemas de Informação) Universidade Federal do Ceará, Campus de Quixadá, Quixadá, 2018.

SILVA, L.; ARAÚJO, L.; SOUZA, V.; SANTOS, A.; NETO, R. Redes neurais convolucionais aplicadas na detecção de pneumonia através de imagens de raio-x. **Computer on the Beach,** v. 11, n. 1, 2020. https://www.researchgate.net/publication/343629637_Redes_Neurais_Convolucionais_Aplicadas_na_Deteccao_de_Pneumonia_Atraves_de_Imagens_de_Raio-X Acessed 13 December 2025.

SILVEIRA, S. J. S.; GOULART, M. J. B. DERMATOSCOPIA, UMA FERRAMENTA QUE PODE BAIXAR CUSTOS NO TRATAMENTO DO MELANOMA. Facit Business and Technology Journal. V1, n. 29. 2021. https://revistas.faculdadefacit.edu.br/index.php/JNT/article/view/1190/0 Acessed 13 December 2025.

THOMAS, S.; HAMILTON, N.; THOMAS, S. **Histopathology Non-Melanoma Skin Cancer Segmentation Dataset.** Univ Queensland Data Collect. 2021.

VOLLMER, A.S.. WINKLER, J.K KOMMOSS, K.S. *et al*. Identifying melanoma among benign simulators – Is there a role for deep learning convolutional neural networks? (MelSim Study), **European Journal of Cancer,** v. 227, 2025. DOI: https://doi.org/10.1016/j.ejca.2025.115706.

XIE, P.; LI, F.; ZHAO, S.; LIU, J.; LU, K.; ZHANG, Y.; CHEN, X. Interpretable pathologist-level classification of melanoma disease pathologies with a convolutional neural

network pipeline. **Journal of Investigative Dermatology.** v. 140, 2020. https://doi.org/10.1016/j.jid.2020.05.046.

ZHANG, N.; CAI, Y.; WANG, Y., *et al*. Skin cancer diagnosis based on optimized convolutional neural network. **Artif Intell Med,** v. 102, 2020. DOI: https://doi.org/10.1016/j.artmed.2019.101756.