

Comparación de algoritmos de visión artificial en la clasificación de enfermedades de la piel utilizando redes neuronales

Comparison of artificial vision algorithms in the classification of skin diseases using neural networks

Erick García Espinosa¹, José Sergio Ruiz Castilla², Farid García Lamont³

^{1,2,3} Universidad Autónoma del Estado de México. Texcoco, Estado de México, México

¹dev.garcia.espinosa.erick@gmail.com, ²jsruizc@uaemex.mx, ³fgarcial@uaemex.mx

Abstract

This study evaluates three tools for the detection of various dermatological diseases, including melanomas, chickenpox, measles, lupus, herpes, scabies, and monkeypox. The results indicate that Orange Data Mining consistently demonstrated high precision (98.2% in training, 99.8% in validation), while Azure Custom Vision achieved moderate precision (88.6% in training, 62.7% in validation), and the CNN showed lower precision (28.14% in validation). The objective of the research is to provide a diagnostic support tool for medical personnel in Level 1 clinics in Mexico, helping them detect sick patients and channel them to Level 2 clinics or specialists, ultimately leading to an improvement in the quality of life for patients who do not have the resources or a Level 2 hospital near their residence.

Resumen

Este estudio evalúa tres herramientas para la detección de diversas enfermedades dermatológicas, que incluyen melanomas, varicela, sarampión, lupus, herpes, sarna y viruela del mono. Los resultados indican que Orange Data Mining demostró consistentemente una alta precisión (98.2% en entrenamiento, 99.8% en validación), mientras que Azure Custom Vision alcanzó una precisión moderada (88.6% en entrenamiento, 62.7% en validación), y la CNN mostró una menor precisión (28.14% en validación). El objetivo de la investigación es proporcionar una herramienta de apoyo al diagnóstico para el personal médico en clínicas de Nivel 1 en México, ayudándoles a detectar pacientes enfermos y canalizarlos hacia clínicas de Nivel 2 o a especialistas, lo que en última instancia conduce a una mejora en la calidad de vida de los pacientes que no cuentan con los recursos o con un hospital de nivel 2 cerca de su domicilio.

Keywords and phrases: Skin disease classification, Artificial vision, Deep Learning, Cloud Learning
2020 Mathematics Subject Classification: 68N01

Reception date: November 14, 2023 / Acceptance date: February 6, 2024

1 Introduction

Detecting dermatological diseases is crucial for timely diagnosis and treatment, which can help prevent serious complications. In this regard, Artificial Intelligence and Computer Vision can assist healthcare specialists in performing this task. We will evaluate the capabilities of various tools, such as Orange Datamining, Microsoft Azure's Custom Vision, and convolutional neural networks, leveraging resources provided by Google Colab. In this article, we will analyze the following diseases: melanomas, chickenpox, measles, lupus, herpes, scabies, and monkeypox.

The objective of this article is to support medical personnel in level 1 clinics in Mexico in the task of referring patients to a specialist or level 2 clinics for treatment. This timely referral can help prevent serious complications and improve the quality of life for patients.

2 Theoretical Framework

2.1 CNN

Convolutional Neural Networks (CNNs) have several layers to process data. The idea for this model started in the 1950s when researchers, Hubel and Wiesel, looked at how animals' visual systems work [1].

A CNN consists of an input layer, one or more hidden layers, and an output layer. CNNs are mainly used for tasks like sorting things and computer vision. They're great at sorting images because they can pick out important patterns on the images, making it easier to recognize and classify them.

2.2 MISH

The MISH activation function is a promising alternative to ReLU in image classification algorithms. One of its main advantages is its ability to tackle the gradient vanishing problem. While ReLU can cause excessively large gradients and neuron saturation, MISH is smoother and helps alleviate these issues. This function is a new, smooth, and non-monotonic neural activation function defined as: $f(x) = x \cdot \tanh(\zeta(x))$ (1) where, $\zeta(x) = \ln(1 + e^x)$ is the softplus activation function [2]

2.3 Azure Custom Vision

Microsoft Azure's Custom Vision is an image recognition service that allows users to create, deploy, and enhance their own image identification models. Unlike the Azure AI Vision service, Custom Vision enables users to define their own labels and train customized models to detect them. [3]

2.4 Orange Datamining

Orange Data Mining [4] is an open-source software used for data visualization, machine learning, data mining, and data analysis. It provides a visual programming interface that allows users to explore qualitative data and interactively visualize it.

Furthermore, it includes the Inception v4 model as a pre-trained CNN model for image classification tasks. This deep learning model, developed by Google, is renowned for its high accuracy in image

classification. It's a variation of the Inception architecture that uses a combination of convolutional layers, pooling layers, and fully connected layers to learn features from input images. [5]

3 Related Work

In the work of In Hoang et al (see Table 1), they present a deep learning model that employs a segmentation approach and Wide-ShuffleNet for the classification of skin lesions. The method was evaluated on two well-known datasets: HAM10000 and ISIC2019. They achieved an accuracy of 86.33% in their second experiment, which primarily focused on the HAM10000 dataset. Additionally, they obtained an accuracy of 82.56% with the ISIC dataset [6].

In the work by Aldhyani et al., they present a lightweight and efficient model designed for accurate skin lesion classification. This model integrates dynamically sized kernels in its layers and employs ReLU and leaky ReLU activation functions. The model achieved an impressive overall accuracy of 97.85% when tested on the HAM10000 dataset [7]

The work by Ozkan et al. aims to classify skin lesions into three groups using four different machine learning methods (ANN, SVM, KNN, and Decision Tree), resulting in a classification rate of 90.45% [8].

The work presented by Barros et al demonstrates the construction of a classification model for 12 types of lesions, generating an extensive dataset with augmentation algorithms, and achieving an accuracy of 94.50% using a model based on the ResNet-152 architecture [9].

In the work by Hammed and others, they introduce a Multi-Class Multi-Level (MCML) classification algorithm. The proposed algorithm was evaluated on 3672 images from various sources and achieved an accuracy of 96.47% [10].

Table 1. Results of similar works focused on lesion detection.

Method	Dataset	Number of Images	Tags	Model	Accuracy
Hoang et al [15]	ISIC 2019	25,331	8	EW-FCM + EfficientNet-B0	84.66%
Aldhyani et al [16]	HAM10000	10,015	7	DKCNN model	98.00%
Ozkan et al [17]	PH2	200	3	Neuronal Network	92.50%
Barros et al [18]	MED-NODE, Edinburgh	111,069	12	ResNet-152	94.50%
Hammed et al [19]	11K, Dataset, Dermls, DermQuest, DermNZ, PH2, ISIC	3672	4	AlexNet	96.47%

4 Our Method

4.1 Dataset

In this study, we collected a dataset of skin disease images from various sources. These sources included the HAM10000 challenge 2019, Kaggle, Google Images, Dermnet NZ, Bing Images, Yandex, Hellenic Atlas, and Dermatological Atlas. To compile a wide range of representative images [11-18], we employed web scraping techniques using the Selenium library. This technique facilitated the systematic extraction of relevant images from the mentioned sources, ensuring the necessary diversity and quantity to build the dataset. This dataset is available on Kaggle [19-20] in two different sets: the "Raw dataset" and the "Skin Lesion Dataset Using Segmentation."

Subsequently, the dataset was divided into seven categories representing different skin diseases. These categories included smallpox, chickenpox, monkeypox, herpes, lupus, melanoma, measles, and scabies (see Table 2). To ensure an accurate evaluation of the model, the data was strategically split in 80/20. This meant that 80% of the images were assigned to the training set, while the remaining 20% were reserved for the test set. This division provided an adequate amount of data to train the model and, at the same time, a representative sample to assess its performance on previously unseen data.

Table 2. The structure of the dataset, total of images: 84794 images

Name	ID	Number of Images	Resolution	Format
Chicken Pox	CH	12204	224x224	JPG, PNG
Herpes	HZ	11918	224x224	JPG, PNG
Lupus	LP	12096	224x224	JPG, PNG
Melanoma	Mel	12738	224x224	JPG, PNG
Scabies	SC	11901	224x224	JPG, PNG
Monkey Pox	MP	12040	224x224	JPG, PNG
Measles	MS	11901	224x224	JPEG, PNG

4.2 Preprocessing

4.2.1 Segmentation

To segment the images, we followed the next steps (see Image 1), we start by converting them to the HSV color space and extracting their hue (H), saturation (S), and value (V) channels. Then, we apply color segmentation using defined thresholds in the hue channel (H) within the range [0, 60], a saturation (S) greater than 90, and a value (V) between 10 and 200, which allows us to identify the regions of interest. After this step, the image is converted to grayscale, and adaptive thresholding is applied with a block size of 11 and a constant of 2 to highlight the edges. Subsequently, pixel values in the thresholded image are inverted, followed by erosion operations with a 4x4 kernel and then dilation operations with a 3x3 kernel to refine the segmentation.

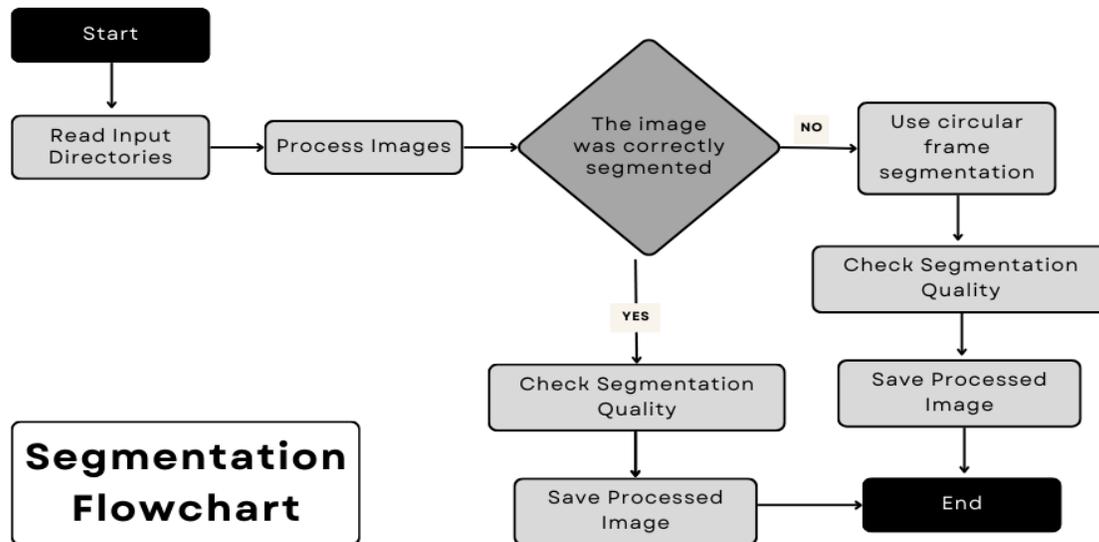


Figure 1. Segmentation Flowchart

It's worth noting that the algorithm doesn't correctly segment all the images in the dataset due to factors like the proximity of the lesions, ambient lighting, etc. To address this, a decision was made to create a circular frame around the image borders in order to improve highlight the lesions (see Figure 2).

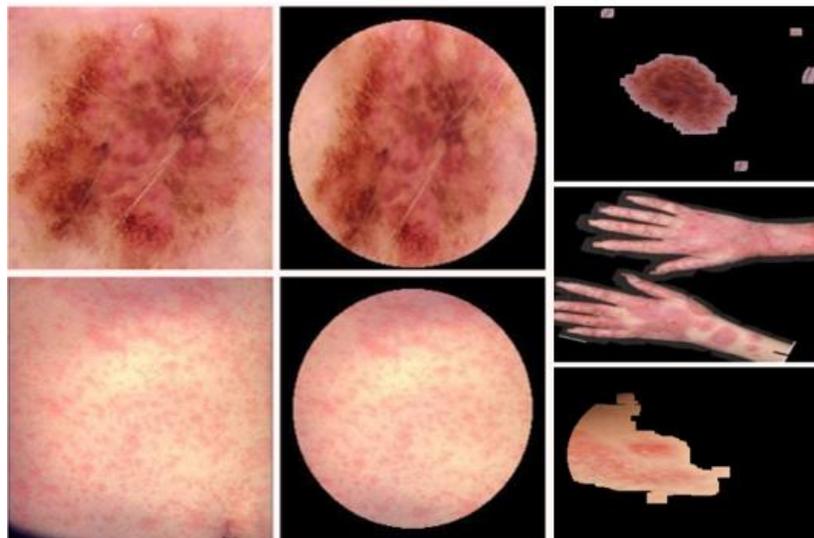


Figure 2. Regular segmentation and circular segmentation

4.2.2 Augmentation

Later, we perform augmentation to make the dataset more robust (see Figure 3). For this purpose, three augmentation functions are defined: horizontal flip, random angle rotation, and random scale factor resizing. Next, we process the images from the source directory and create copies of each image in the destination directory. Subsequently, for each copy, a random combination of these filters is applied, generating multiple augmented versions of each image.

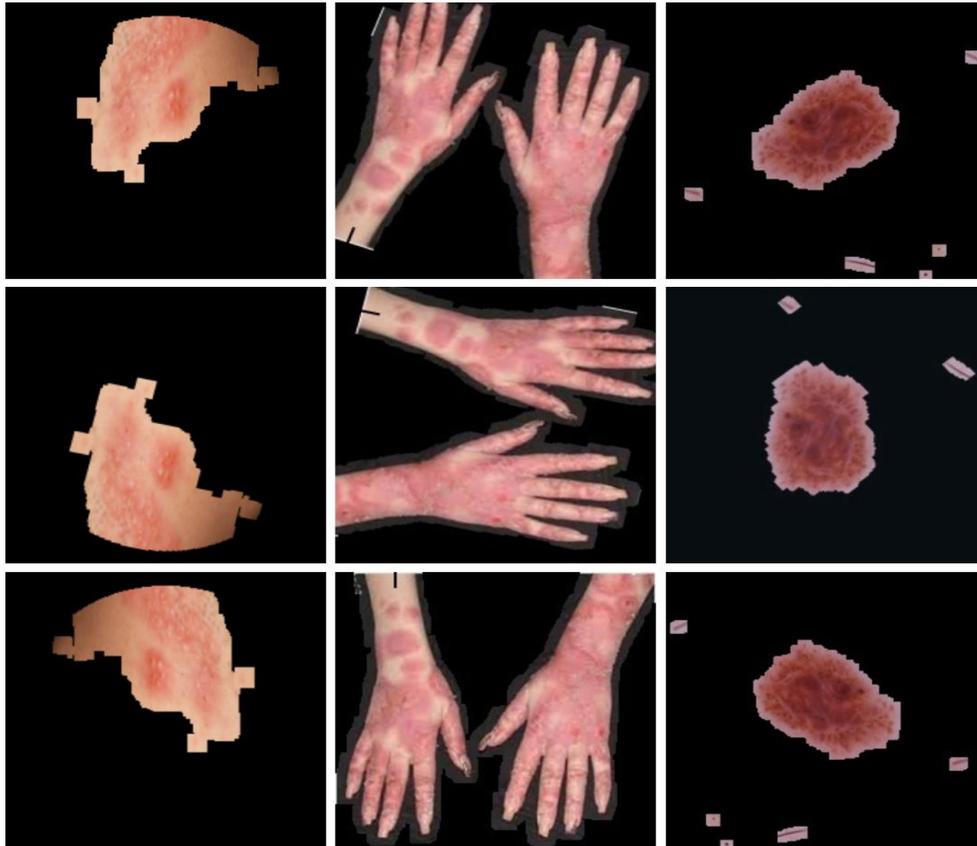


Figure 3. Augmentation

5 Results

Azure Custom Vision achieved precision rates of 88.6% on the training set and 62.7% on the validation set using the following parameters: Probability Threshold of 60% and the General (compact) model. It's worth noting that the model was trained with only 52,576 images because the model determined that some images were similar (see Table 3).

Table 3. Azure Results in both tests

Disease	Training	Number of images	Validation	Number of Images
Melanoma	99.3%	8334	100.0%	100
Measles	90.3%	6863	38.0%	100
MonkeyPox	89.7%	7105	49.0%	100
Scabies	86.3%	6918	78.0%	100
Herpes	84.5%	7069	70.0%	100
Chickenpox	83.8%	6851	73.0%	100
Lupus	83.6%	7436	95.0%	100

On the other hand, Orange Data Mining stands out as the most effective method based on the results obtained, as shown in Table 4. It achieves high AUC and precision values in both the training and validation sets. Additionally, it is easy to implement and visualize, as indicated in the figures 4-5. However, it can be a bit challenging to understand the program's inner workings when analyzing the library code step by step.

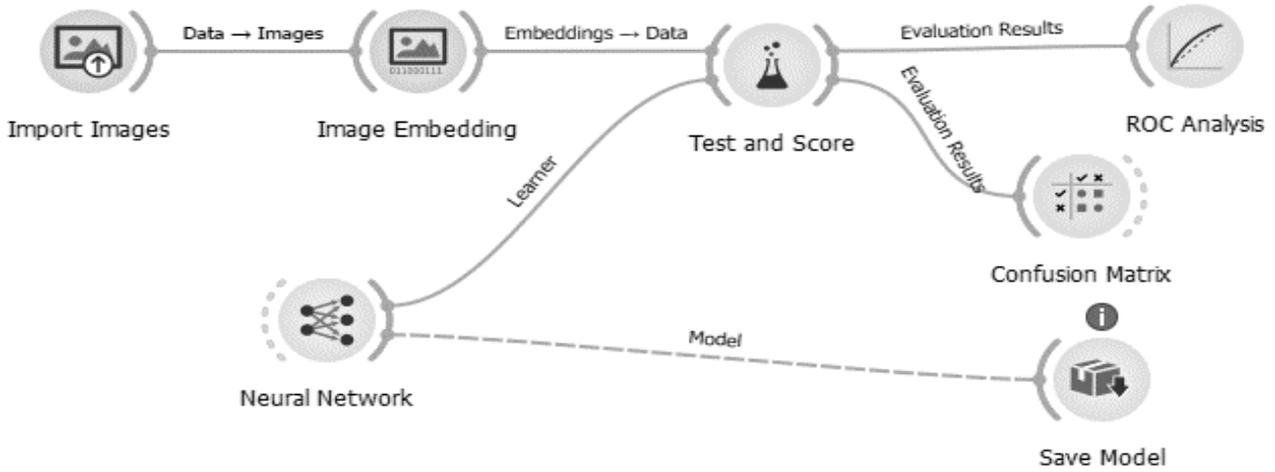


Figure 4. Training Flow

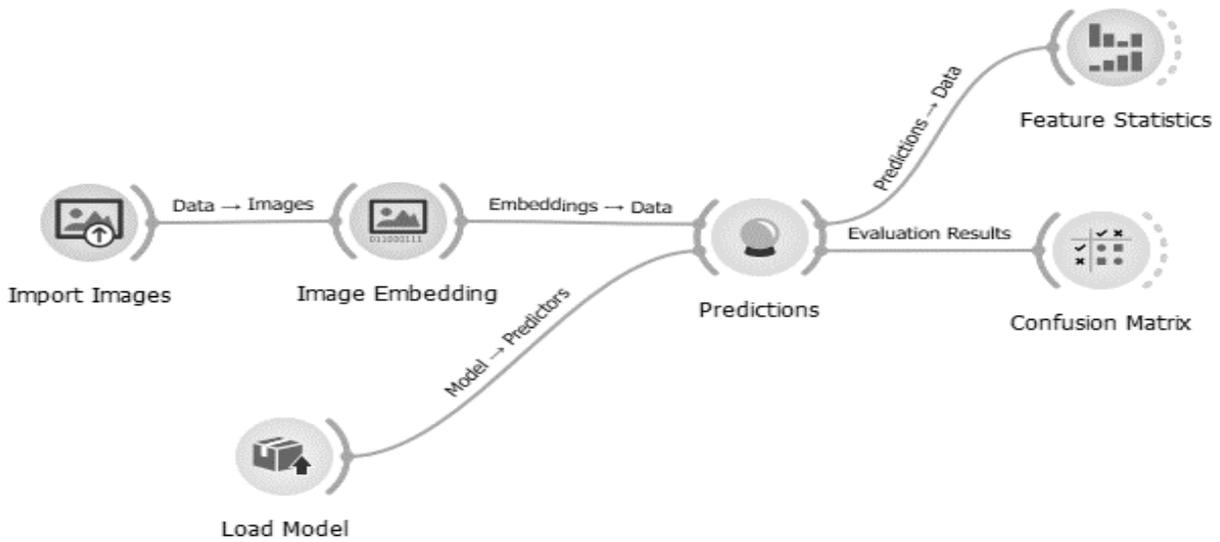


Figure 5. Validation Flow

Table 4. Orange Datamining Results in both tests

Disease	Training	Number of images	Validation	Number of Images
Melanoma	99.6%	12738	100.0%	100
Measles	99.1%	11901	100.0%	100
MonkeyPox	98.4%	12040	97.0%	100
Scabies	97.7%	11901	99.0%	100
Herpes	97.5%	11918	100.0%	100
Chickenpox	97.9%	12204	99.0%	100
Lupus	97.5%	12096	100.0%	100

For feature extraction (see Table 5), Orange primarily displays size(a), width(b), and height(c), along with a variety of features that are not explicitly(d) specified in terms of what they extract from the image, as seen in Figure 6. However, you can use a heatmap to identify which parts of the image are relevant for the InceptionV3 model, allowing you to understand what it extracts from each image, as shown in Figure 7. This helps provide insights into the model's feature selection and what it finds significant in the images.

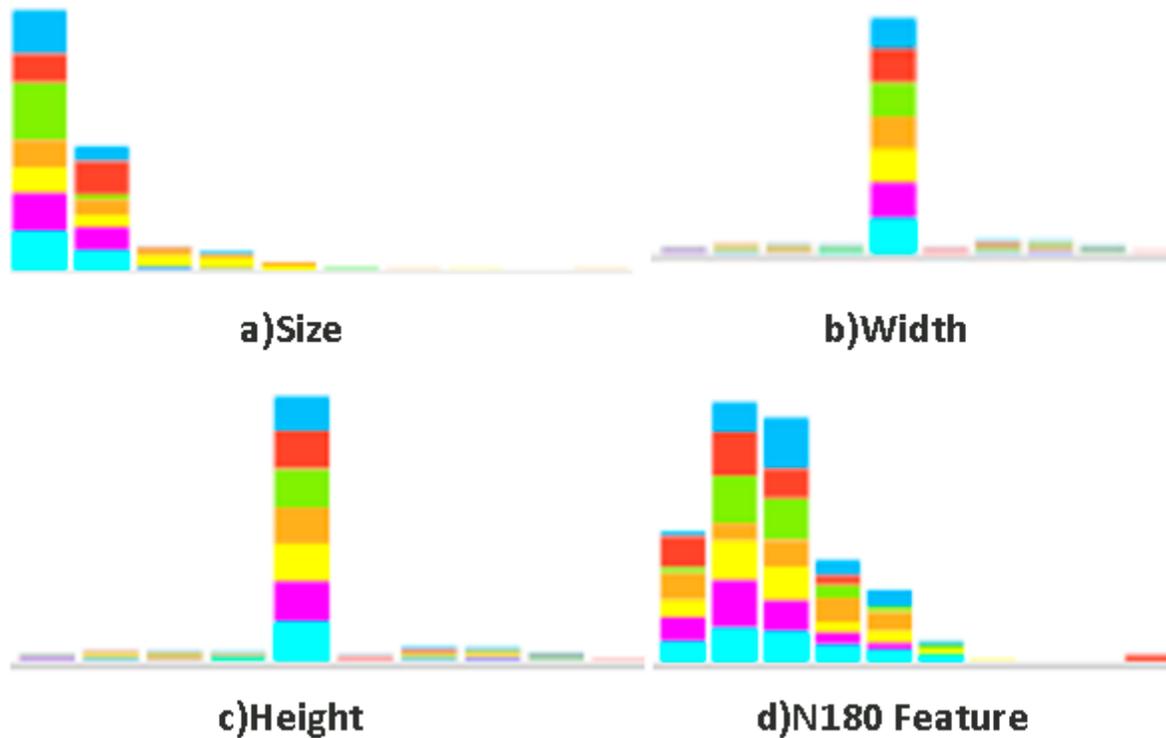


Figure 6. Features extracted by Orange.

Table 5. Orange Datamining feature extraction

Feature	Mean	Mode	Median	Dispersion	Min	Max	Missing
Size	21144.53	1411	16347	0.86	1203	159351	0%
Width	225.85	224	224	0.14	113	334	0%
Height	225.85	224	224	0.14	113	334	0%
N180	1.43446	5.435	1.32143	0.5641	0.081473	5.435	0%

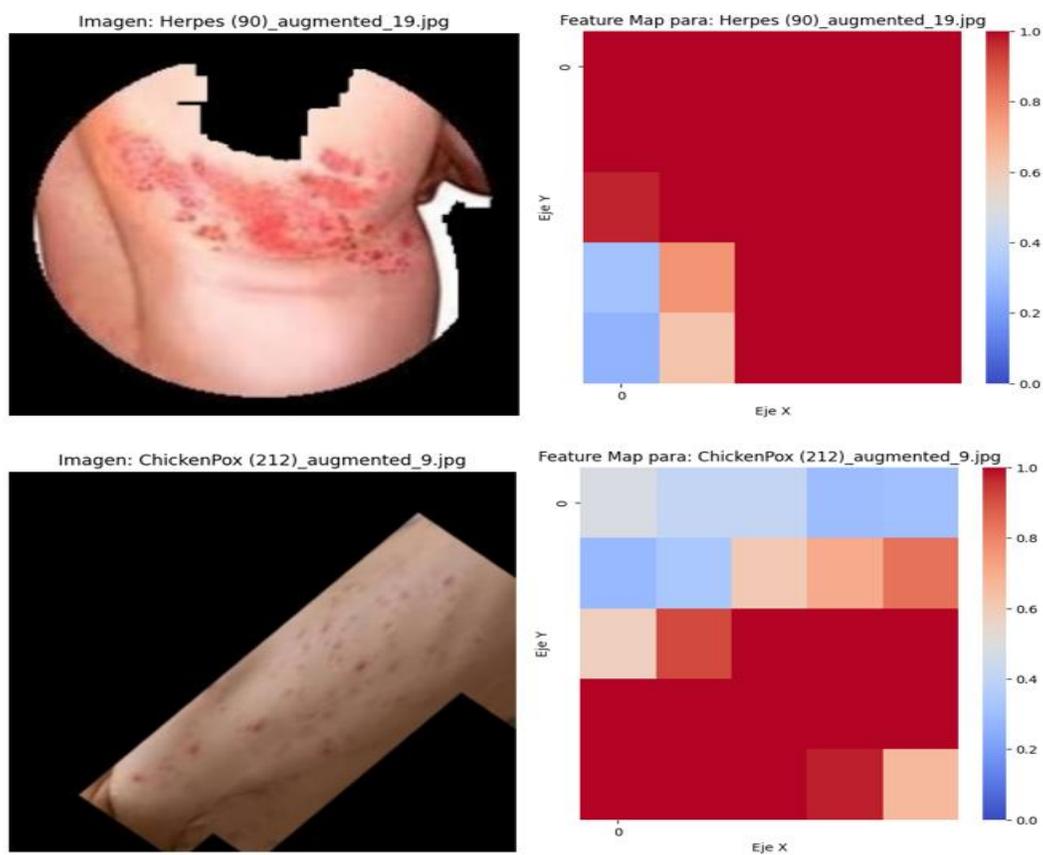


Figure 7. Inceptionv4's heatmaps

The CNN created in Colab achieved a decent AUC score of 0.9742. However, its precision rates were comparatively lower, with 86.4% on the training set and only 28% on the validation set, as can be seen in Table 6. These results suggest that the performance of the CNN created in Colab might be suboptimal for the given task, especially in terms of precision.

Disease	Training	Number of images	Validation	Number of Images
Melanoma	99.6%	12738	100.0%	100
Measles	99.1%	11901	100.0%	100
MonkeyPox	98.4%	12040	97.0%	100
Scabies	97.7%	11901	99.0%	100
Herpes	97.5%	11918	100.0%	100
Chickenpox	97.9%	12204	99.0%	100
Lupus	97.5%	12096	100.0%	100

Table 6. CNN Results

In a nutshell, Azure Custom Vision demonstrated moderate precision, achieving 88.6% on the training set and 62.7% on the validation set. In contrast, Orange Data Mining consistently exhibited high precision in both the training and validation sets, with some diseases showing notably strong results, the CNN displayed lower precision, especially on the validation set, where it only reached 28% (see Table 7).

Table 7. Summary of Performance and Parameters

Name	Parameters	AUC	Precision on Train	Precision on Val
Orange Datamining	Function: Relu Train:90%	0.982	98.2%	99.8%
	Test:10% Epoch:20 Folds:10			
Azure Custom Vision	General Compact S1	-	88.6%	62.7%
CNN made in Colab	Función: MISH Train:80%	0.9742	86.4%	28.14%
	Test:20% Épocas:20 Batch size:128			

6 Conclusion and Future Directions

This work examined platforms like Orange, Azure Custom Vision, and Google Colab in the task of skin lesion identification with the purpose of comparing these tools and gaining a general understanding of them. The long-term goal is the potential implementation of a physical device that can provide support to specialists in triaging individuals who suffer from dermatological conditions.

It's important to highlight the challenges that arose in this research, such as creating an appropriate dataset and lesion segmentation. Additionally, it's relevant to mention the limitations imposed by using a model developed in software that doesn't allow for complete control over the model and its architecture, as would be the case with a model programmed line by line. Therefore, it's essential to continue the development of this research and explore ways to recreate the same model using resources that enable greater control, adaptation, and a more in-depth study of it.

References

- [1] C. B. Carrión, "REDES CONVOLUCIONALES.", pp. 15, <https://n9.cl/carmelobonilla>
- [2] Liu, Qian, and Steve Furber. "Noisy Softplus: a biology inspired activation function." In International Conference on Neural Information Processing, pp. 405-412. Springer, Cham, 2016. <https://n9.cl/liu-quian>
- [3] "Azure AI Custom Vision," Microsoft Azure, 2023. [Online]. Available: <https://n9.cl/msazure> . [Accessed: Nov. 9, 2023].
- [4] "Orange Visual Programming," Orange3, 2023. [Online]. Available: <https://n9.cl/orange3> . [Accessed: Nov. 9, 2023].
- [5] S. Sharma and S. Lohchab, "Personal Authentication Using Finger Vein Biometric Technology with Implementation of Transfer Learning CNN Model, pp. 4-5," SSRN Papers, Dec. 28, 2021. [Online]. Available: <https://n9.cl/sharma> . [Accessed: Nov. 9, 2023].
- [6] L. Hoang, S.-H. Lee, E.-J. Lee, and K.-R. Kwon, "Multiclass Skin Lesion Classification Using a Novel Lightweight Deep Learning Framework for Smart Healthcare," Appl. Sci., vol. 12, no. 5, pp. 2677, Mar. 2022, Available: <https://n9.cl/hoang> , doi: 10.3390/app12052677.
- [7] Aldhyani THH, Verma A, Al-Adhaileh MH, Koundal D. Multi-Class Skin Lesion Classification Using a Lightweight Dynamic Kernel Deep-Learning-Based Convolutional Neural Network. Diagnostics (Basel), Available: <https://n9.cl/theyazn> , 2022 Aug 24;12(9):2048. doi: 10.3390/diagnostics12092048. PMID: 36140447; PMCID: PMC9497471.
- [8] I. A. Ozkan and M. Koklu, "Skin Lesion Classification using Machine Learning Algorithms," Int. J. Intell. Syst. Appl. Eng., vol. 5, no. 4, pp. 285-289, Dec. 2017, Available: <https://n9.cl/ozkan> , doi: 10.18201/ijisae.2017534420.
- [9] D. B. Barros and N. C. D. Silva, "Skin Lesions Classification Using Convolutional Neural Networks in Clinical Images," arXiv preprint arXiv:1812.02316, Dec. 2018. [Online]. Available: <https://n9.cl/barros> . [Accessed: Nov. 10, 2023].
- [10] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, "Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques," Expert Syst. Appl., vol. 141, Available: <https://n9.cl/hameed> , pp. 112961, Mar. 2020, doi: 10.1016/j.eswa.2019.112961.
- [11] "Dermnet," Kaggle, Jun. 24, 2020. [Online]. Available: <https://n9.cl/dermnet>
- [12] "Monkeypox Skin Lesion Dataset," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/monkeypox> . [Accessed: Nov. 10, 2023].
- [13] "POX DATASET," Kaggle, Aug. 31, 2022. [Online]. Available: <https://n9.cl/balasubramaniamv> . [Accessed: Nov. 10, 2023].

- [14] D. Bala and M. S. Hossain, "Monkeypox Skin Images Dataset (MSID)," Mendeley Data, Feb. 23, 2023. [Online]. Available: <https://n9.cl/mendeley-pox> . [Accessed: Nov. 10, 2023].
- [15] U. Shahab, "Dermnet Skin Disease Images," Kaggle, May 21, 2021. [Online]. Available: <https://n9.cl/dermnet2> . [Accessed: Nov. 10, 2023].
- [16] "ISIC Challenge Datasets," ISIC Archive, Nov. 8, 2023. [Online]. Available: <https://n9.cl/isic-challenge> . [Accessed: Nov. 10, 2023].
- [17] R. P. Usatine and B. D. Madden, "Interactive Dermatology Atlas," Dermatlas.net. [Online]. Available: <https://n9.cl/dermatlas> . [Accessed: Nov. 10, 2023].
- [18] Hellenic Dermatological Atlas, "Home | Hellenic Dermatological Atlas," <https://n9.cl/hellenic> . [Accessed: Nov. 10, 2023].
- [19] "Skin Lesion Dataset Using Segmentation," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/bigdataset> . [Accessed: Nov. 10, 2023].
- [20] "Skin Lesion Dataset Using Segmentation," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/rawdataset> . [Accessed: Nov. 10, 2023].