iMedPub Journals www.imedpub.com

Vol.12 No.1:001

Methodological Approaches to Machine Learning based on a Small Sample Size for Contributing Assessment to the Decision Support System in Dermatology

Andriambololoniaina Faly Herizo^{1,2,3*}, Randriamanalina Rojo Nofidiantsoa¹, Gottfrois Philippe⁵, Rakotoarisaona Mendrika Fifaliana², Fenohasina Georges Cheny², Ramily Leophonte², Rakotomanana Kiady Armando², Ratovohery Andry Nirilalaina^{1,3}, Randriamboavonjy Rado Lalao^{1,3}, Bagayoko Cheick Oumar⁴, Rakotomalala Mamy Alain¹, Rapelanoro Rabenja Fahafahantsoa^{2,3}, Hsu Christophe⁵, Navarrini Alexander⁵, Ravonimanantsoa Ndaohialy Manda-Vy¹

¹Department of Science and Engineering Technology, University of Antananarivo, Antananarivo, Madagascar

²Department of Dermatology, Joseph Raseta Befelatanana University Hospital, Antananarivo, Madagascar

³Department of Public Health, Research Laboratory on Information Technology in Construction (LaRTIC)

⁴Center for Innovation and Digital Health, University of Science, Technology and Technology of Bamako, Mali, West Africa

⁵Department of Dermatology, University Hospital of Basel, Basel, Switzerland

*Corresponding author: Andriambololoniaina Faly Herizo, Department of Science and Engineering Technology, University of Antananarivo,

Antananarivo, Madagascar; Email: herizo. andriambololoniaina@univ- Antananarivo.mg

Received: October 15, 2024 Manuscript No. IPACSIT-25-19816; **Editor assigned:** October 18, 2024, PreQC No. IPACSIT-25-19816 (PQ); **Reviewed:** November 01, 2024, QC No. IPACSIT-25-19816; **Revised:** January 31, 2025, Manuscript No. IPACSIT-25-19816 (R); **Published:** February 28, 2025, DOI: 10.36648/2349-3917.12.1.1

Citation: Herizo AF, et al. (2025) Methodological Approaches to Machine Learning based on a Small Sample Size for Contributing Assessment to the Decision Support System in Dermatology. Am J Compt Sci Inform Technol Vol.12 No.01

Abstract

In the field of dermatology, the diagnostic process is often difficult due to a wide variety of skin conditions requiring the expertise of a dermatologist, who are currently in very insufficient numbers. Since 2003, tele dermatology initiatives have been implemented in Madagascar, but local experience has shown unsatisfactory results due to the unavailability of specialists to respond to requests posted on the platforms. Over the past five years, the literature is full of several studies proposing Artificial Intelligence (AI) algorithms to aid in diagnosis. Thus, our study aims to seek a better approach to diagnostic support by conducting a meta-analysis of experiences combining telemedicine and artificial intelligence in dermatology and to propose a model for a diagnostic support system adapted to our context. The first draft of model training results with U-Net for 458 annotated images presented in this article shows training parameters better suited to a small sample of images. The result gave us a stable accuracy on average at 97.83% from the 5th epoch of training. With the data collected as part of the passion dermatology project, which contains 1960 photos of skin lesions, our solution focuses on (a) a semantic segmentation model proposing a tool for evaluating knowledge about skin lesions and (b) a diagnostic orientation mechanism. The model thus obtained will become a self-assessment tool for students and physicians to visually familiarize themselves with skin pathologies in the absence of a dermatologist. The pre-trained machine will thus become a medical training tool to improve the diagnostic approach.

Keywords: Self-assessment; Dermatology; Artificial intelligence; Diagnostic support system; Telemedicine

Introduction

In the current era, dermatology is fully leveraging technological advancements, particularly tele dermatology and Artificial Intelligence (AI), which offer innovative solutions for the identification and classification of skin conditions. However, the success of an AI model undeniably depends on the quality of the data used for its training. The management of datasets plays a crucial role in this process, as it influences the model's ability to accurately differentiate between various dermatological conditions based on images and clinical elements. Given the diversity and complexity of skin pathologies, a rigorous approach is necessary to create a representative and coherent dataset. The effectiveness of a model relies on its ability to generalize from a wide range of data, including different types and colors of skin pigmentation, as well as the stages of dermatological conditions and the lighting conditions of the captured photos. Therefore, data management involves not only the comprehensive collection of dermatological images but also the use of appropriate acquisition techniques and precise annotation to facilitate the proper training of the artificial intelligence models used [1].

Material and Methods

General view of methodological approaches

This study examines the challenges and best practices related to data management for combining two high-performing models in the recognition of dermatological lesions and diagnostic prediction based on associated clinical elements. The goal is to improve the efficiency of the diagnostic support system in dermatology by using an artificial intelligence engine (Figure 1).





This conceptual diagram summarizes the process-oriented approach adopted in this study. It highlights the sequence of the two combined learning systems, which aim to establish an artificial intelligence engine that effectively contributes to the diagnostic support system in dermatology. This approach is focused on a pedagogical method, rather than simply facilitating diagnostic procedures and reducing the effort required by the practitioner, in order to minimize any risk of diagnostic error related to a lack of human expertise [2].

Methodology related to data collection

Tele dermatology is used for data collection. As an emerging field of telemedicine, it has significantly evolved over the past ten years largely due to the integration of innovative technologies. One of the essential components of tele dermatology is the gathering of information, including photographs of dermatological lesions and clinical data, which are crucial for diagnosing skin conditions. This study explores current techniques for collecting this data through a digital application called My CRF ONE, a tele dermatology platform, highlighting technological advancements, challenges encountered and clinical implications. The data collection medical team typically uses a smartphone to capture photos. The average resolution of the images is 2 megapixels. For each case, one photo is taken at a focal distance of 25 cm to obtain an overview of the lesion and another at 10 cm for a detailed close-up. A recently painted light green room is dedicated to tele dermatology activities, including photo capture, within the dermatology and venereology department of CHU Joseph Raseta Befelatanana. The collected cases primarily come from CHU Joseph Raseta Befelatanana, ESSPA Analakely in Antananarivo, CHRR in Antsirabe, Toamasina and a hospital center in Mahajanga. The initial goal is to collect 10,000 cases per

year from these various sites. A secure platform is specifically dedicated to this collection. Data entry occurs in three steps: The first involves entering age, sex and phototype, followed by uploading photos of the dermatological lesions according to the previously detailed photography guidelines. Then, the topography of the lesions and clinical elements are recorded in the subsequent steps. With an average connection speed of 20 to 40 Mbps, it takes about 5 to 10 minutes to record a case [3].

The collection of dermatological image data and clinical data from tele dermatology platforms holds considerable potential for ensuring the reliability, speed of transmission and quality of the data collected. However, it is crucial to address the challenges related to data security and accessibility to ensure an ethical and equitable implementation of these technologies. The success of this system therefore depends on the collaboration between the physicians and interns who collect the data, the technical team of artificial intelligence experts, the application developers and the dermatologists who validate each annotation of lesions in the collected cases.

During data collection, patients' personal information is properly anonymized and secured. Medical images are directly stored securely, with robust encryption in an S3 bucket of Amazon Web Services (AWS) cloud storage, thus ensuring data performance and durability while preventing unauthorized access. Transmission channels are secured using an SSH tunnel to avoid any risk of data breaches during the transfer of sensitive medical information. Regarding patient consent, they are informed about how their data will be used and must provide consent before the collection of information informed pertaining to them (photos of lesions and clinical elements). This also includes understanding the risks associated with tele dermatology. Compliance with medical data collection standards through tele dermatology adheres to regulations regarding the protection of personal data. The research protocol for data collection within the passion dermatology project has received approval from the ethics committee of the ministry of public health [4].

Methodology related to data preprocessing

The preprocessing of data for the classification of skin lesions is a preliminary step before implementing the lesion recognition algorithm using an object detection model. This phase includes essential actions: (a) Image normalization, which several standardizes the dimensions, resolution and colors of images to ensure consistency within the dataset, helping the model to generalize better. (b) Data augmentation, which enriches the dataset by applying various transformations to the images, such rotation, flipping, cropping or adding noise, thereby as increasing the diversity of examples and improving the model's robustness. We also use annotated images from the SCIN dataset to supplement our data for segmentation. (c) Labeling, which involves assigning labels or categories to the images while ensuring that they are accurate and consistent. (d) Data cleaning, which involves removing unnecessary, duplicate or low-quality images to optimize the efficiency of the model. (e) Class balancing through resampling, which ensures that each

class has a sufficient number of occurrences to avoid class bias. (f) The creation of training and testing sets, which involves dividing the dataset into two parts to independently evaluate the model's performance.

The data preprocessing step allows us to proceed with training in two stages: (a) Training a convolutional neural network model for skin lesion recognition and (b) training a regression model. The latter, after exporting the features into a concatenated NumPy array with a dataset containing the topography of the lesions and clinical elements, determines the diagnostic prediction. At each stage, a comparative analysis of the models is conducted to determine the appropriate methodological choice: For stage 1, between the U-Net semantic segmentation model and the fast R-CNN model and for stage 2, between the multivariate logistic regression model and the polynomial regression model, depending on the interaction behavior among the studied variables [5].

Methodology related to model training

The choice between the U-Net semantic segmentation model and the Fast R-CNN model at stage 1 allows for a comparison of their performances, given the limited number of object classes and occurrences per class currently available to us. At this stage, we can determine which model is most effective for training the machine to recognize elementary lesions. U-Net facilitates the localization and distinction of the boundaries of the elements

Table 1: Mathematical approach for U-Net.

representing each elementary lesion by classifying each pixel in the annotated images. This step can be represented by convolution operations followed by nonlinearities (ReLU) and down sampling (pooling) to produce attribute maps. Mathematically, this can be represented by:

 $X^{(I)}=f(X^{(I-1)*}W^{(I)}+b^{(I)})$ (1)

 $X^{(I)}$: The activations of layer (I) are the result of applying the activation function (f) to the weighted sum of the activations from the previous layer (I-1).

 $X^{(l-1)}$: The activations of the previous layer (l-1) are the outputs of layer (l-1) that serve as inputs to layer (l).

 $W^{(l)}{:}\ The weights of layer (l) are a weight matrix applied to the activations of the previous layer (l-1) to produce the weighted sum.$

 $b^{(l)}$: The biases of layer (I) are a bias vector added after the multiplication by the weights to adjust the weighted sum.

f: The activation function, which is a non-linear function applied to the weighted sum of the activations and biases to introduce non-linearity into the model. ReLU is used for the first three convolutional layers and the sigmoid function is used for the final output layer to produce probabilities for each pixel belonging to a class (Table 1).

Convolution layer	Conv-1, Conv-2, Conv-3	Conv-4
Activation function	X ^(l) =ReLU(X ^(l-1) *W ^(l) +b ^(l))	$X^{(l)}=\sigma(x) (X^{(l-1)*}W^{(l)}+b^{(l)})$

For the sigmoid function, the base of the natural logarithm is approximately 2.71828.

In our case, the classes studied primarily concern elementary lesions: The primary class includes papules, macules, pustules, nodules and plaques, while the secondary class includes the underlying elements of each primary class, such as erythematous papules, pigmented papules, erythematous plaques, pigmented plaques, etc. This mechanism is illustrated by the conceptual diagram below (Figure 2).



Figure 2: Elementary lesions classification process.

The characteristics extracted from the annotations of these images are exported to a NumPy table. This table will be used in step 2 as values for an explanatory variable in a dataset that includes the topography of these lesions and clinical elements to predict diagnosis. For this extraction, we utilize the pre-trained VGG19 model with image net weights. In step 2, the NumPy table containing the characteristics extracted from the photo annotations is concatenated with other data collection variables to form a data frame, which will be used for training the diagnostic prediction model. The choice between the multivariate logistic regression model and the polynomial model depends on the interaction behavior of these explanatory variables [6].

Two model choices are proposed for the training: (a) the multivariate logistic regression model, which essentially includes k variables. These variables include the characteristics of the annotated elementary lesions and their topographical features, such as location, extent and edge. In addition, clinical signs are characterized by the presence or absence of fever, pruritus at night or during the day and pain expressed as binary and ordinal numerical values. In our case, the value of k is 6. Fever is represented by values 0 for no fever, 1 for moderate fever and 2 for severe fever. Similarly, pain and pruritus are coded as 0 for no pruritus, 1 for day pruritus and 2 for night pruritus. This model is represented by the equation:

$$z = \log\left(\frac{p}{1-p}\right) \sum_{i=1}^{k} \beta 0 + \beta i Xi$$
⁽²⁾

American Journal of Computer Science and Information Technology

Where p is the probability that z takes the value 1, with β 0 as the logit estimate of the dependent variable and β i as the logistic regression coefficient (parameter estimates), while k represents the number of independent variables Xi.

The probability that z takes the value 1 is expressed using the logistic function:

$$p = \frac{1}{1 + e^{-z}} | p = \frac{1}{1 + e^{-\sum_{i=1}^{k} \beta_{0} + \beta_{i} X_{i}}}$$
(3)

(b) the Polynomial regression model, with Y as the explained variable corresponding to Diagnostic and the explanatory variables X1 to X6. The polynomial regression of order (n) is represented by the equation:

$$\left[Y = \beta_0 + \beta_{1X_1} + \beta_{2X_2} + \dots + \beta_{6X_6} + \sum_{\{i=1\}}^{\{n\}} \beta_{\{i,j\}} X_1^i X_2^j\right]$$
(4)

With β 1... β 6 as the coefficient for the linear term, β {i,j} for the polynomial terms and n represents the degree of polynomial. The equation presented here corresponds to polynomial terms of order 2 to {n}.

Methodology related to the evaluation of models

For the segmentation task, the model evaluation method is based on accuracy which gives the total number of correct predictions, for all classes divided by the total number of samples. This metric is translated into the formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
(5)

The training uses Adam as an optimizer and balanced crossentropy as a loss metric that aims to compensate for class imbalance. The Balanced Cross-Entropy (BCE) formula is defined as follows.

$$\left[BCE = -\frac{1}{N} \sum_{i=1}^{N} \left(w_{pos} \cdot y \cdot \log(p) + w_{neg} \cdot (1-y) \cdot \log(1-p) \right) \right]$$
(6)

N is the total number of samples, w_{pos} a weight associated with the positive class, w_{neg} a weight associated with the negative class, y the true label (0 for the negative class, 1 for the positive class) and p the prediction of the positive class [7].

For regression problems, precision, recall and F1 score. Precision is calculated using the fraction of True Positives (TP) on all True Positives (TP) and False Positives (FP). It provides a perspective on the accuracy of the model's positive predictions. This metric is translated into the formula:

$$Precision = \frac{TP}{TP + FP}$$
(7)

The recall, which is the relationship between True Positives (TP) on the set of True Positives (TP) and False Negatives (FN), offers a perspective on the model's ability to capture all positive results. This metric is translated into the formula:

$$Recall = \frac{TP}{TP + FN}$$
(8)

The F1 Score is useful to have a global view of the balance between accuracy and recall, especially when class imbalance is present as in our case. This metric is translated into the formula:

$$F1 Score = 2 X \frac{Precision X Recall}{Precision + Recall}$$
(9)

Methodology related to the application

The AI engine resulting from the combination of the two models described in the learning methodology approach is implemented in an application that involves asking the user to enter his own data (images of the lesion, its topography and clinical elements). The choice of target audience is justified by the need for learning and continuing medical training using as reference real patient records: (a) medical students in 6th and 7th year of transition to the department of dermatology at CHU JRB Befelatanana, at CHU Morafeno Toamasina and CHU Antsiranana; (b) The application is also offered to general practitioners in areas outside the city that express needs for medical assistance in dermatological consultations. For students, the application is integrated into their digital work environment of the faculty of medicine of Antananarivo. For doctors, through a downloadable mobile application. The application's operating mechanism is based on medical data input and ask the user about the characteristics of the lesions identified on the proposed images and the diagnosis he is considering in relation to the clinical data entered. In this way, he becomes familiar with skin pathologies without the direct help of a dermatologist. The previously trained machine will become a training tool for humans to strengthen their skill and not to replace their skill. For security reasons, the application is strictly professional and should not be made available to the public. Each user who requires its use is subject to a signed commitment for the guarantor of security.

Results and Discussion

The result presented in this article is focused on the recognition of elementary lesions that determines the first part

of the process approach. From all the photos collected with clinical data through the tele dermatology platform My.crf.one we obtained a total of 102 classes representing all the possibilities of elementary lesions in dermatology. These classes are divided into several levels: Level 1 corresponds to the basic elementary lesions such as papule, macula, pustule, nodule, plaque, scar, squame, crust, vesicle, etc... Level 2 corresponds to the first subclass as erythematous papules, pigmented papules, erythematous macula, erythematous plaque, etc... Level 3 corresponds to the second subclass as atrophied pigmented macula, squamous erythematous plaque (or erythematos quameous plate), hypochromic macular plaque, etc... [8].

With all these classes, compared to the small sample quantity for 458 images and 458 masks, we have a strong class imbalance. Thus, we filtered data only to classes with at least 100 occurrences for training with U-Net. We have thus obtained 8 classes represented graphically in the Figure 3 below:

Table 2: Description of the combination of hyper parameters.



The parameters used are summarized in the table below (Table 2).

Gradient batch	Optimizer	Loss function	Learning rates
8	Adam	Binary cross entropy	0.001
8	Adam	Binary cross entropy	0.01
16	Adam	Binary cross entropy	0.01

American Journal of Computer Science and Information Technology

In this series, we have between 124 and 714 occurrences per class. The U-Net training described in the methodology above is conducted in two stages [8]. For the first step of the drive, 256 × 256 size input images were analyzed in 4 convolutional layers with ReLU as activation function for the first 3 layers and sigmoid for the output layer. Training was conducted in 10 epochs to obtain the following result (Figure 4).



Convolution and pooling layers, followed by dense layers and the ReLU activation function lead to a form output (None, 256, 256, 1), with a total of 407.809 parameters in the model. For this first step, the model was trained over 10 epochs. This curve shows that there is an instability in the performance of the model but a tendency to convergence towards higher accuracy at the last time (Accuracy at 96.80%).

A second stage of training was carried out in 20 epochs to obtain the following result (Figure 5).



The results appear initially low in the first 2 epochs, with an accuracy close to zero and an unusual negative loss. Then, from epoch 3 to 7 we observe an improvement in accuracy and a reduction in loss. However, the validation accuracy remains quite low. From time 8 to 20 the accuracy stabilizes at high values between 97.42% and 98.24%, indicating a significant improvement of the model. The validation accuracy remains stable at 99.94%, which suggests that the model generalizes well.

Overall, there is a significant improvement in the performance of the model on later periods, with high validation accuracy and low loss. This indicates that the model has learned to generalize well on validation data and will continue to improve as more epochs are added and annotated data is added [9].

We initially have a small sample size, 700 images were annotated but after the filter of 200 occurrences per class we could retain only 458 images with 5 classes of the 102 obtained. This small amount of data usually makes deep learning training difficult, but this article shows that it is possible to adapt the model in relation to the size of the available sample and it is possible to improve it by increasing the annotated data, the number of epochs and iterations.

Trained at 30 epochs, with the same parameters but this time with 5 classes having at least 200 occurrences, the accuracy remains stable between 97.46% and 98.56% from the 5th epoch (Figure 6).



Our perspective is to test it with the fast R-CNN model once annotated images reach 1000 and at least 200 occurrences per class. Hyper parameters will be evaluated with other techniques as presented in the study conducted by Debarpan D et al. which shares experiences of combining hyper parameters summarized in the Table 3 below.

Grandient accumulation batch	Optimizer	Loss funtion	Learning rate
32	Adam	Cross entropy	0.01
32	Adam	Asymetric single label	0.1
32	Adam	Cross entropy	0.1
32	Adam	Focal	0.01
32	Adam	Focal	0.001

Table 3: Hyper parameters combination skills.

The number of eras is also among the criteria that improve learning performance, as presented in the study by Shanthi et al. who used 174 images for model training with convolutional neural networks (AlexNet) with 1000 epochs (100,000 iterations) achieved in 3 days to obtain an accuracy of 96.32%. But deep learning, especially with models such as U-net, presents various challenges related to the number of epochs and the choice of hyper parameters. These challenges can have a significant impact on the performance and effectiveness of learning models. When a deep learning model is trained over too many eras, there is a risk of overlearning, where the model stores training data instead of learning generalizable patterns. Conversely, insufficient number of eras can lead to under learning, where the model fails to capture patterns in data exhaustively. In addition, a high number of epochs require increased computing resources, including computational time and computing power. This is a constraint in our case, due to limited resources. In terms of data sizing, larger or more complex datasets may require a higher number of epochs and a finer optimization of hyper parameters. Our goal is to work on 1760 photos and in a long-term perspective with 10000 photos. So, at this time, a choice between 30 and 50 times would be ideal for our study [10].

Conclusion

Based on this methodological approach, several recommendations are proposed to improve the accuracy and effectiveness of the decision support system for dermatological diagnosis. (a) Robust quality assurance processes are essential to ensure the accuracy and reliability of data collected by tele dermatology. This includes comprehensive documentation of clinical elements, as well as data validation and image quality assessment protocols. Specific recommendations for photo capture have been developed so that the lesions present in the images can be exploited. These recommendations include the use of a light green background to avoid glare, taking pictures without flash to prevent over-glare and capturing from multiple angles. (b) The inclusion of a diverse and representative data set, covering a wide range of skin conditions, as well as skin types and colours from photo types 1 to 6, is essential. This diversity contributes to the development of more inclusive and accurate AI models for dermatological diagnosis. (c) Establishing collaborative networks and partnerships with dermatologists, medical institutions and technology experts can leverage collective expertise to refine methodology and effectively deploy advanced technologies for federated learning.

In the PASSION dermatology project in which we participate, member countries provide data and train the model using the algorithm developed by the technical team. (d) The continuous validation and improvement of artificial intelligence models, through feedback mechanisms, regular updates and iterative refinement based on real clinical results, are essential to ensure the long-term accuracy and relevance of the decision support system. In conclusion, the methodological approach presented in this article has great potential to contribute to the advancement 2. of dermatological diagnosis through the integration of artificial intelligence through the combination of two models. Implementation of the above recommendations will not only refine the methodology, but also improve the accuracy, inclusiveness and ethical integrity of the decision support system. By adopting these recommendations, the dermatology field will be able to take greater advantage of advanced technologies to provide comprehensive and effective diagnostic support, improving patient outcomes and guality of care.

Conflict of Interest

None.

Acknowledgement

The results of this innovative research were made possible by the support of:

Almighty god, who has graciously granted us his will and blessing to carry out this study. The Botnar foundation, which finances the passion dermatology project.

The malagasy dermatology society (SOMADER), for the time devoted to image annotations and medical and scientific validation.

The public health reception and research laboratory specializing in ICT (LaRTIC) and the laboratory network, artificial intelligence, multimedia, security and information systems (RIMSS) EDSTII doctoral school, Antananarivo polytechnic higher school, for their technical expertise in improving the current model.

And all those who have contributed directly or indirectly to the study, both medically and technically.

References

American Journal of Computer Science and Information Technology

- 1. Lorette G (2022) Teledermatology: From tele-expertise to connected dermatology. La Presse Medicale Formation 3: 459-466
- Jeong HK, Park C, Henao R, Kheterpal M (2022) Deep learning in dermatology: A systematic review of current approaches, outcomes and limitations. JID Innov 3: 100150.
- Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N (2020) Machine learning in dermatology: Current applications, opportunities and limitations. Dermatol Ther 10: 365-386.
- Singh KN, Mantri JK (2024) A clinical decision support system using rough set theory and machine learning for disease prediction. Intelligent Med.
- Ridard E, Secember H, Carvalho-Lallement P, Schuers M (2020) Indicators in teledermatology: A literature review. Ann Dermatol Venereol 147: 602-617.
- Rabenja FR (2022) PASSION project: Data collection in Madagascar and Guinea. Iproceedings 8 :e36903.
- Monnier J, L'Orphelin JM, Bataille M (2024) Artificial intelligence in dermatology: Practical implications. Ann Dermatol Venereol 4: 203-207.
- Demoulins E, Rat C, Martin L, Mamzer MF (2021) Teledermatology practices: Benefits, limitations and perspectives. Qualitative interview-based study with dermatologists. Ethics Med Public Health 16: 100631.
- 9. Phillips CM, Balch D, Schanz S, Branigan A (2002) Teledermatology: Issues in remote diagnosis and management of cutaneous disease. Curr Probl Dermatol 14: 7-38.
- Ward A, Li J, Wang J, Lakshminarasimhan S, Carrick A, et al. (2024) Crowdsourcing dermatology images with google search Ads: Creating a real-world skin condition dataset. arXiv preprint.