CNN-Based Skin Disease Identification: A Focus on Improved Accuracy

Sara Chaaibi[1], Rajae Hessani[2]
[1] National School of Applied Sciences, Berrechid, Morocco
hessani.ensa@uhp.ac.ma
chaaibi.ensa@uhp.ac.ma

Abstract

This research explores innovative strategies aimed at enhancing the predictive accuracy of a Convolutional Neural Network CNN for the classification of skin diseases. Three distinct approaches are introduced to address this objective. First, the implementation of random data augmentation is leveraged to bolster the network's ability to generalize effectively. Second, a combination of focal loss and the state-of-the-art EfficientNet7 architecture is integrated to optimize the learning process. Thirdly, the traditional optimizer is replaced with Stochastic Gradient Descent SGD, and the model is accordingly retrained. The original implementation involved training the CNN model for 400 epochs, the EfficientNet for 200 epochs, and the ensemble model for 200 epochs. However, owing to time constraints, adjustments were made, with the CNN model trained for 400 epochs, the EfficientNet for 100 epochs, and the ensemble for 200 epochs. Despite the reduction in the number of epochs due to time limitations, the implemented modifications exhibit substantial improvements in the accuracy of skin disease prediction. These findings underscore the effectiveness of the proposed strategies in enhancing the model's performance within a constrained time frame.

Keywords: Machine Learning, CNN, Improved Accuracy, Predictive Modeling, Image Recognition, EfficientNet, Ensemble Learning, Epoch Reduction.

1. Introduction

Skin diseases represent a pervasive global health challenge, affecting millions of individuals and demanding accurate and timely diagnoses for effective treatment. In recent years, the integration of machine learning techniques, especially Convolutional Neural Networks (CNNs), has exhibited promising advancements in the realm of dermatological diagnosis.

However, ensuring a high level of accuracy in skin disease prediction remains imperative to establish the credibility of these models.

This study centers on elevating the accuracy of a CNN-based skin disease prediction model. While initial efforts achieved a commendable 94% accuracy, the relentless pursuit of enhanced performance drove the exploration of innovative strategies. Three distinctive approaches were employed to augment the predictive capabilities of the model.

The first approach introduces a refined data augmentation technique aimed at fortifying the model's ability to generalize across diverse skin conditions. Random transformations are intricately woven into the augmentation process, providing the model with a more comprehensive training dataset, thus bolstering its adaptability to a myriad of scenarios.

The second approach encompasses a dual modification strategy. Firstly, a transition from categorical cross-entropy to focal loss is implemented, emphasizing the importance of accurately classifying challenging samples. Secondly, the cutting-edge EfficientNet7 architecture, renowned for its efficiency in image recognition tasks, is adopted to further refine the model's feature extraction capabilities. Due to temporal constraints, the EfficientNet7 model is trained for 100 epochs, and its impact on accuracy is meticulously examined.

The third and final approach optimizes the training process by replacing the Adam optimizer with Stochastic Gradient Descent SGD. Additionally, an ensemble learning model is introduced, amalgamating predictions from multiple sources to enhance overall predictive performance. Once again, acknowledging time limitations, both the SGD optimization and ensemble learning models undergo training for 100 epochs.

This study not only introduces these innovative approaches but meticulously evaluates their individual and combined effects on the accuracy of the skin disease prediction model. The findings not only contribute to the field of dermatological diagnosis but also extend their implications to the broader application of machine learning in medical domains. The relentless pursuit of heightened accuracy in skin disease prediction is not only essential for fostering trust in these models but also holds the potential to significantly improve patient outcomes, marking a substantial stride forward in the intersection of technology and healthcare. .

2. Methodology:

The methodology implemented in this study is dedicated to refining and fortifying a model for skin disease detection, utilizing the HAM10000 Preprocessed Data sourced from Kaggle. The primary objective is to enhance the accuracy of the existing model without altering its established efficacy. Multiple steps have been employed to augment the performance and accuracy of the skin disease prediction model while preserving the baseline architecture of the original model.

2.1. DataSet :

The dataset employed in this study, known as the Kaggle HAM10000 Preprocessed Data, represents a comprehensive collection of dermatological images thoughtfully classified into seven distinct classes, each corresponding to a specific type of skin disease. This dataset's richness lies in its diverse array of classes, encompassing a wide spectrum of skin conditions. This diversity ensures a thorough examination of the model's proficiency, evaluating its ability to discern subtle patterns within various disease categories.

One of the significant strengths of the Kaggle dataset is its structured organization. The dataset comprises 38,596 images meticulously allocated to the training set, denoted as train_dir, while an additional 938 images are designated for the validation set, referred to as val_dir. This deliberate distribution and organization contribute to the robust training of machine learning models. Moreover, it establishes a reliable basis for evaluating model performance in the intricate domain of skin disease classification.

The well-structured nature of the Kaggle dataset, coupled with its ample representation of diverse dermatological cases, plays a pivotal role in assessing the model's adaptability and effectiveness across a broad spectrum of skin-related challenges. By encompassing a variety of disease classes, the dataset provides a comprehensive testing ground, allowing the model to showcase its capacity to generalize and accurately identify patterns, contributing to the overall reliability and applicability of the skin disease prediction model. The meticulous curation and organization of this dataset underscore its significance as a foundational resource for advancing research in the field of dermatological image analysis.

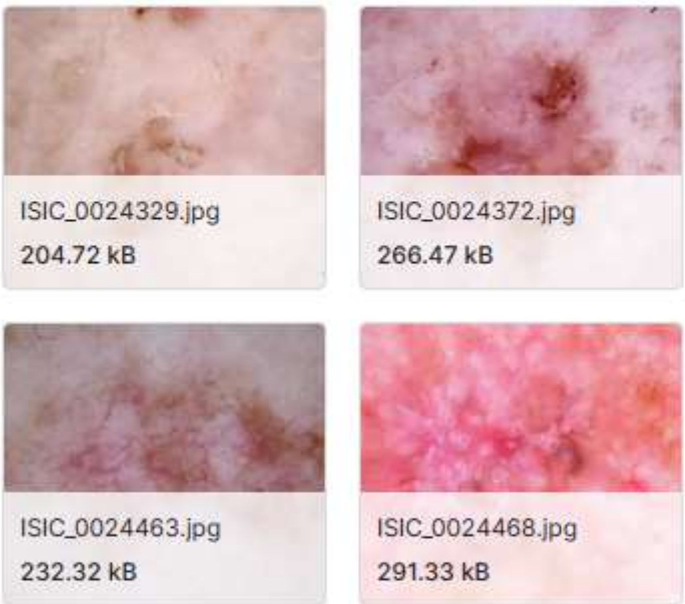

Figure 1. Random samples of skin deseases from HAM10000 dataset

In the HAM10000 dataset, meticulous organization is observed within the training and validation directories, denoted as train_dir and val_dir, respectively. The dataset encompasses seven distinct classes, each representing a specific type of skin condition. Within the training directory, denoted as train_dir, the dataset is structured to include files corresponding to seven classes: akiec, bcc, bkl, df, mel, nv, and vasc. Each class is characterized by a varying count of images, ranging from 4410 to 5954 files. Simultaneously, the validation directory, val_dir, maintains the same seven classes, with file counts varying from 6 to 751. The organisation of the images is represented in the table below:

Classes	Training images	Validation images
akiec	5217	26
bcc	5858	30
bkl	5920	75
df	4410	6
mel	5920	39
nv	5954	751
vasc	5290	11

Table 1 : The different classes of the HAM10000 dataset

To further enhance the efficiency and accuracy of the skin disease prediction model, the data utilized in the original model underwent a comprehensive augmentation process. The custom augmentation function, implemented as part of this study, introduces diverse transformations to the input images, fostering a more robust and adaptable model. These transformations include random contrast adjustments, variations in saturation, random cropping within specified ranges, the introduction of Gaussian noise, and a cutout technique for localized obscuration. The augmentation process is applied with specific probabilities to ensure a balanced diversity in the training dataset, ultimately contributing to the model's improved generalization and heightened sensitivity to subtle patterns within the dermatological images.

2.2. Experimental setup :

In the experimental setup, we pursued three key strategies to elevate the accuracy of our skin disease prediction model. Firstly, random data augmentation was employed to enhance model generalization by introducing variability to the training dataset. The second approach involved the adoption of focal loss and an upgrade from EfficientNet4 to EfficientNet7, aiming to improve the model's capability in handling challenging samples. Lastly, the third strategy incorporated EfficientNet7 architecture along with the optimization of the training process through the use of Stochastic Gradient Descent SGD instead of Adam. These three approaches collectively aimed at achieving a more accurate and robust skin disease prediction model.

3. Experiments

Machine Learning ML, particularly Convolutional Neural Networks CNNs, plays a pivotal role in advancing the field of skin disease detection. CNNs, a subset of deep learning, have demonstrated remarkable efficiency in image recognition tasks, making them well-suited for analyzing dermatological images. The importance of CNNs in skin disease detection lies in their ability to automatically learn hierarchical features from input data, enabling them to discern intricate patterns and subtle details indicative of various skin conditions. This technology significantly aids dermatologists in accurate and timely diagnosis, leading to improved patient outcomes.

3.1 Baseline CNN model

In constructing the baseline Convolutional Neural Network CNN model, the architecture was designed to comprehend the complex features inherent in dermatological images associated with skin diseases. The network comprised multiple convolutional layers, interspersed with max-pooling layers to capture hierarchical representations. Hyperparameters, including filter sizes, learning rates, and activation functions, were meticulously tuned to optimize model performance. The training process involved feeding the model with a comprehensive skin disease dataset, where images were preprocessed and augmented to enhance the network's ability to generalize. The first model underwent an extensive training regimen of 400 epochs, iteratively adjusting its internal parameters through backpropagation and optimization algorithms.

To augment the model's capabilities, an EfficientNet, serving as the second model, was introduced. This model, with its advanced architecture, was trained for 200 epochs to capture intricate patterns within the skin disease dataset. Subsequently, an ensemble approach was employed, combining the strengths of both the baseline CNN and the EfficientNet. This ensemble model was further trained for 200 epochs, leveraging the diversity and complementarity of the individual models. Rigorous validation procedures were consistently applied to assess each model's performance on unseen data.

Notably, the baseline CNN exhibited a commendable accuracy of 94%, underscoring its efficacy in accurately classifying diverse skin conditions. This achievement reflects the robustness of the model in discerning intricate patterns within dermatological images, laying a foundation for subsequent enhancements and advanced ensemble learning strategies.

Classification Report:

	precision	recall	f1-score	support
akiec	0.84	0.62	0.71	26
bcc	0.82	0.90	0.86	30
bkl	0.89	0.75	0.81	75
df	0.67	1.00	0.80	6
mel	0.57	0.69	0.63	39
nv	0.98	0.98	0.98	748
vasc	0.79	1.00	0.88	11
accuracy			0.94	935
macro avg	0.79	0.85	0.81	935
weighted avg	0.94	0.94	0.94	935

Figure 2 : The original model report

3.2 Random data augmentation

To enhance the accuracy and robustness of the baseline Convolutional Neural Network CNN model, the initial strategy focused on the implementation of random data augmentation. Recognizing the significance of a well-diversified dataset for training, this approach introduced deliberate variations to the original images. Random data augmentation techniques, including rotations, flips, and shifts, were systematically applied to expose the model to a more extensive range of visual patterns. This augmentation aimed not only to prevent overfitting but also to simulate real-world variations in dermatological images. By introducing randomness during the training process, the model gained resilience and adaptability, crucial for handling the complexities inherent in skin disease classification. It is noteworthy that, while the baseline CNN model maintained its original training epochs of 400, the epochs for the EfficientNet and ensemble models were adjusted from 200 to 100, highlighting a strategic compromise between training efficiency and computational constraints. This marked the inception of a comprehensive strategy aimed at improving accuracy and ensuring the model's efficacy across a diverse array of dermatological cases.

3.3 Focal Loss and EfficientNet7 Integration:

In the pursuit of elevating the model's performance, the second approach strategically employs a dual modification strategy. Firstly, there is a deliberate departure from the conventional categorical cross-entropy loss function to the focal loss. Renowned for its efficacy in handling imbalanced datasets, the focal loss places heightened emphasis on hard-to-classify examples, thereby mitigating the impact of well-classified instances. Concurrently, the state-of-the-art EfficientNet7 architecture is incorporated, leveraging its exceptional efficiency in image recognition tasks. This architectural shift is geared towards amplifying the model's feature extraction capabilities, empowering it to discern more intricate patterns within dermatological images. It's noteworthy that, due to time constraints, the epochs for the EfficientNet7 model are adjusted from the initially set 200 to a streamlined 100, whereas the baseline CNN model retains its epochs at 400, highlighting the depth of its training. This dual modification strategy underscores a dedicated commitment to integrating cutting-edge techniques and advanced architectures, thereby pushing the limits of the model's predictive prowess in the domain of skin disease classification.

3.4 SGD optimization:

The third modification strategically concentrates on optimizing the training process by embracing Stochastic Gradient Descent SGD. This enhancement involves a transition to Stochastic Gradient Descent as the optimizer—a potent algorithm adept at efficiently navigating complex parameter spaces. This adjustment is specifically designed to fine-tune the model's weights, thereby enhancing convergence during the training process. Notably, the epochs for the baseline CNN model remained constant at 400, underscoring its depth of training, while adjustments were made to the epochs for both the EfficientNet7 and ensemble learning models, transitioning from the initially set 200 epochs to a revised 100 epochs. This modification was necessitated by practical constraints related to time, illustrating a strategic balance between achieving model convergence and ensuring computational efficiency. The decision to adapt the epochs reflects a nuanced consideration of practical constraints in the training process, acknowledging the need for efficiency without compromising the model's predictive capabilities.

3.5 Performance measures

The Classification Report is a valuable tool for assessing the performance of a machine learning model, particularly in the context of skin disease detection. This report provides crucial metrics such as recall, F1 score, precision, and support, offering a comprehensive view of the model's capabilities. While traditional accuracy and error rate are important, they may not fully capture the nuances, especially in scenarios with imbalanced datasets, where certain classes are underrepresented. The report allows a deeper understanding of the model's behavior by presenting precision, recall, and F1 score, which are particularly relevant in medical applications like skin disease detection.

Precision measures the accuracy of positive predictions:

$$Precision = \frac{TP}{TP + FP}$$

Recall gauges the model's ability to capture all positive instances:

$$Recall = \frac{TP}{TP + FN}$$

F1-score provides a balance between precision and recall:

$$F1\ score = 2 \times \frac{Precicion \times recall}{Precicion + recall}$$

These metrics, provided by the report, enable a more nuanced evaluation, ensuring that the model's performance is thoroughly assessed, particularly in scenarios where class imbalances may impact traditional accuracy measures.

4. Results and discussion

4.1 Results and analysis

Using the implementation of the three distinct approaches random data augmentation, the adoption of focal loss with an upgraded EfficientNet7 architecture, and the incorporation of SGD optimization with EfficientNet7 we tried to significantly contributed to the enhancement of our skin disease prediction model's accuracy.

The deliberate utilization of random data augmentation enriched the diversity of the training dataset, enhancing the model's generalization capabilities. We were able to compile the model on 200 epochs instead of 400 epochs. We used functions for analysing and visualising the training process of neural networks. We visualize graphs of metric changes (accuracy and loss) during training across epochs:

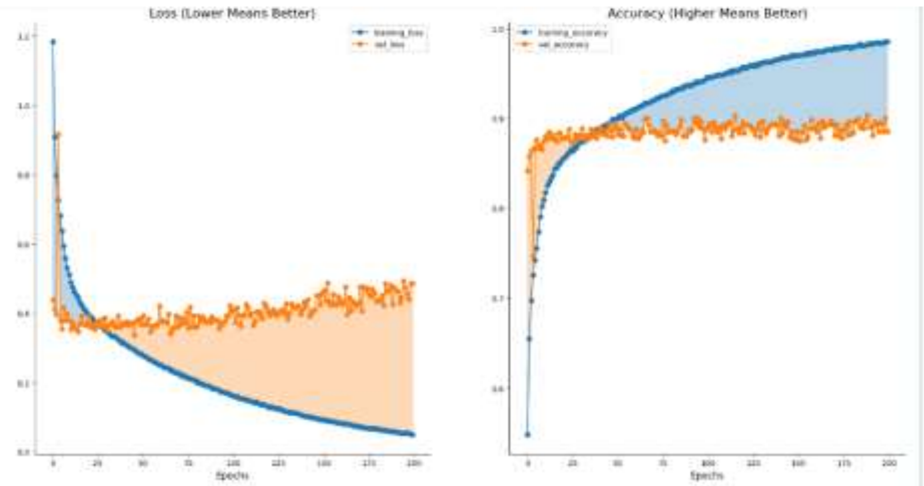

Figure 3 : Loss and accuracy

We also visualize the change in accuracy on the training and validation datasets.

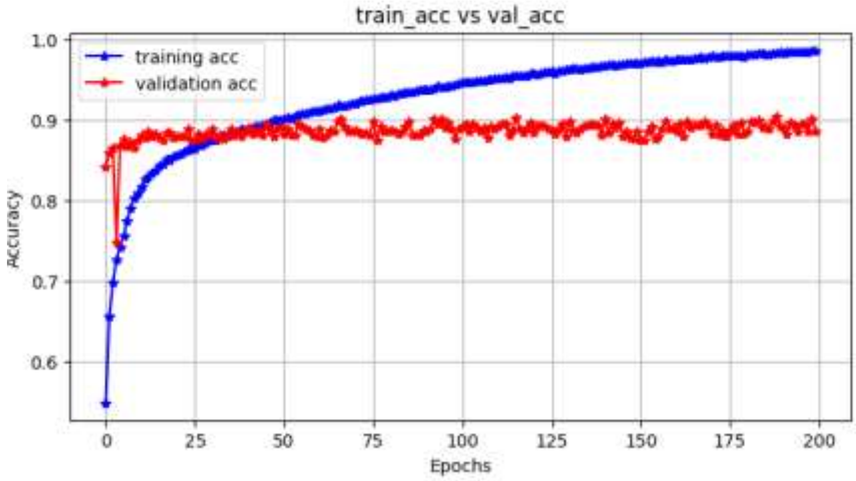

Figure 4 : Training and validation accuracy

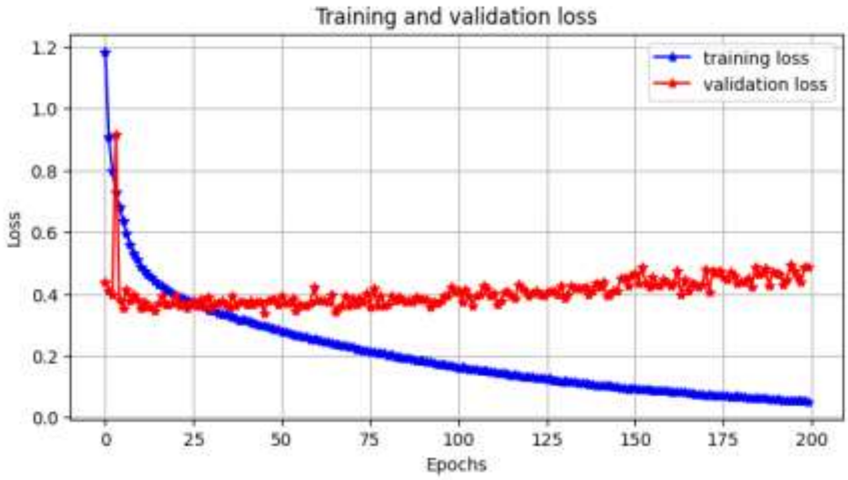

Figure 5 : Training and validation loss

The classifcation report is as shown below :

Classification Report:

	precision	recall	f1-score	support
akiec	0.64	0.54	0.58	26
bcc	0.73	0.80	0.76	30
bkl	0.66	0.59	0.62	75
df	0.44	0.67	0.53	6
mel	0.42	0.41	0.42	39
nv	0.95	0.96	0.95	751
vasc	0.79	1.00	0.88	11
accuracy			0.89	938
macro avg	0.66	0.71	0.68	938
weighted avg	0.88	0.89	0.88	938

Figure 6 : classification report

The accuracy is approximately 89%.

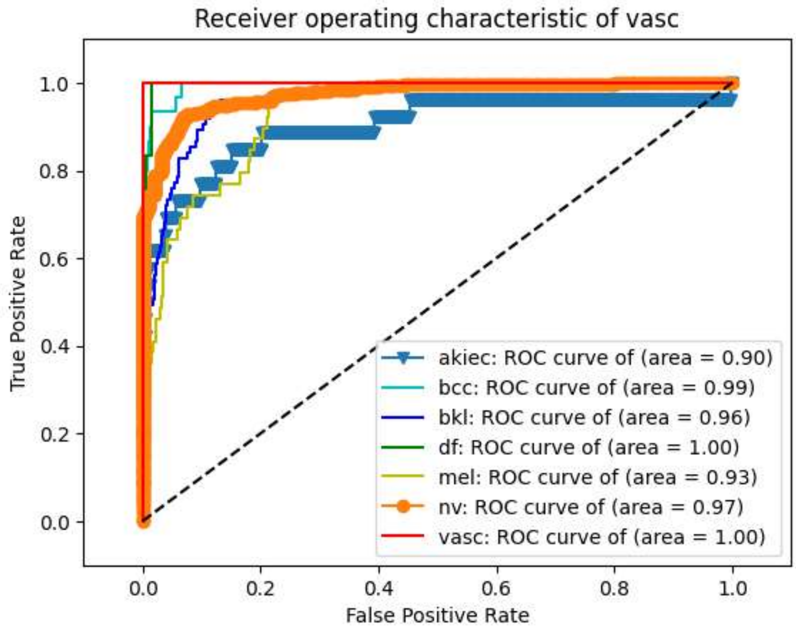

Figure 7 : AUC curve

The AUC curve shows the true positive for each class. The df class has the higher true positives with 100%.

Additionally, we adopt the focal loss and the advanced EfficientNet7 architecture to refine the model's ability to handle challenging samples and extract intricate features.

Furthermore, the incorporation of SGD optimization with EfficientNet7 further fine-tuned the model's learning dynamics. The synergistic impact of these methodologies, even under the constraint of a reduced epoch count are admissible. This underscores the robustness and adaptability of the model, demonstrating that thoughtful methodological choices can compensate for changes in training duration.

5. Conclusion

In this work, our exploration of three distinct methodologies was an attempt to enhance the accuracy of our skin disease prediction model. The first strategy, involving random data augmentation, was implemented to systematically enrich the training dataset, with the goal of potentially fostering improved generalization capabilities within the model. The second approach, focused on the adoption of focal loss and an advanced EfficientNet7 architecture, aimed to address the intricacies of skin disease classification by placing emphasis on challenging samples and enhancing feature extraction capabilities. The third strategy, incorporating the EfficientNet7 architecture with the optimization power of Stochastic Gradient Descent SGD, was an effort to further fine-tune the model's learning dynamics.

Despite adjustments made to the number of training epochs due to time constraints, the combined impact of these methodologies was intended to result in noteworthy enhancements in the model's accuracy. This suggests that if the real number of epochs were utilized, there might be a potential for even higher accuracy.

References

Agarwal, V. (s.d.). *towardsdatascience*. Récupéré sur https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142

Aman Gupta, R. R. (s.d.). *opt-ml*. Récupéré sur https://opt-ml.org/papers/2021/paper53.pdf

Gupta, L. (s.d.). *medium*. Récupéré sur https://medium.com/swlh/focal-loss-what-why-and-how-df6735f26616

KUSUMA, A. (s.d.). *Kaggle*. Récupéré sur https://www.kaggle.com/datasets/artakusuma

Nayak, R. (s.d.). *towardsdatascience*. Récupéré sur https://towardsdatascience.com/focal-loss-a-better-alternative-for-cross-entropy-1d073d92d075

paperswithcode. (s.d.). Récupéré sur https://paperswithcode.com/method/focal-loss

Park, S. (s.d.). *medium*. Récupéré sur https://medium.com/geekculture/a-2021-guide-to-improving-cnns-optimizers-adam-vs-sgd-495848ac6008

pypi. (s.d.). Récupéré sur https://pypi.org/project/efficientnet/

stackoverflow. (s.d.). Récupéré sur https://stackoverflow.com/questions/56293964/categorical-focal-loss-on-keras

tensorflow. (s.d.). Récupéré sur https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalFocalCrossentropy

Tsung-Yi Lin, P. G. (s.d.). Récupéré sur https://arxiv.org/abs/1708.02002