

SkinHealthMate app: An AI-powered digital platform for skin disease diagnosis



Amina Aboulmira ^{a,*}, Mohamed Lachgar ^{b,c,d}, Hamid Hrimech ^a, Aboudramane Camara ^{d,e}, Charafeddine Elbahja ^{d,e}, Amine Elmansouri ^e, Yassine Hassini ^e

^a LAMSAD Laboratory, ENSA, Hassan 1er University, Berrechid, Morocco

^b L2IS Laboratory, Faculty of Science and Technology, University Cadi Ayyad, Marrakech, Morocco

^c Higher Normal School, Department of Computer Science, University Cadi Ayyad, Marrakech, Morocco

^d LTI Laboratory, ENSA, Chouaib Doukkali University, Morocco

^e IITE, ENSA, Chouaib Doukkali University, Morocco

ARTICLE INFO

Keywords:

Artificial intelligence
Dermatology ensemble learning
Skin disease classification
Digital health platforms

ABSTRACT

Accurate diagnosis of skin diseases remains a significant challenge due to the inherent limitations of traditional visual and manual examination methods. These conventional approaches, while essential to dermatological practice, are prone to misdiagnoses and delays in treatment, particularly for conditions like skin cancer. To address these gaps, this paper presents the SkinHealth App, an innovative AI-driven solution that enhances the accuracy and efficiency of skin disease diagnosis. The app integrates a robust ensemble learning model, combining the strengths of EfficientNetB1 and EfficientNetB5 architectures. This ensemble model improves disease classification performance through advanced image processing techniques such as noise reduction and data augmentation. The key contributions of this work include the development of a scalable and secure server-side structure that ensures the safe handling of patient data and efficient processing of diagnostic queries. Experimental results on the HAM10000 dataset demonstrate the model's superior performance, achieving an accuracy of 93%, along with high precision and recall scores, thereby reducing false positives and false negatives. These outcomes clearly establish the app's potential to enhance dermatological diagnosis by providing timely and accurate disease identification. Ultimately, this work bridges the gap between traditional diagnostic methods and modern AI-driven technology, offering a transformative tool for improving patient care in dermatology.

1. Introduction

For decades, dermatologists have relied on traditional diagnostic methods to identify and classify skin diseases. These methods, relying solely on visual inspection and manual examination, have provided valuable insights into various skin conditions [1], but sometimes, they might overlook subtle signs, delaying intervention [2], which can allow the disease to reach advanced stages, particularly in the case of aggressive skin cancers. In addition, exclusive reliance on conventional approaches can lead to inaccurate diagnoses when similar pathologies share common visual features. This susceptibility to misinterpretation or delayed diagnostic accuracy then has an impact on therapeutic considerations and overall patient health outcomes.

Biopsy, the removal and examination of tissue samples for diagnostic purposes, is an essential tool in dermatology to confirm or rule out

various skin conditions, including skin cancers. However, even this crucial step is not without its limitations and risks. Dermatologists' subjective selection of biopsy sites can sometimes lead to crucial indicators being missed due to variations in areas of interest, ultimately resulting in a lack of standardization [3]. Traditional biopsies, limited by restricted tissue sampling, may not encompass the entirety of a lesion, leading to inaccurate diagnoses. Subsurface changes that are important for pathologies such as melanoma may go unnoticed on visual inspection, resulting in missed opportunities for early detection. The scope of traditional biopsies, limited to visible areas, can overlook these subsurface alterations, resulting in false negatives.

Artificial intelligence (AI), particularly deep learning, has opened up a new era of innovation and progress in the medical field. Deep learning algorithms rely on neural networks that simulate the interconnected nodes of the human brain, enabling computers to learn from vast

* Corresponding author.

E-mail address: a.aboulmira@uhp.ac.ma (A. Aboulmira).

datasets, recognize patterns and make intelligent decisions [4].

Those Deep learning models have revolutionized the field of medical diagnostics by enabling the automatic identification and classification of complex patterns in medical images. These models, powered by neural networks, can learn intricate features from large datasets, offering unprecedented accuracy and efficiency. However, relying on a single model may not always capture the full spectrum of variability present in real-world data [5,6].

Ensemble learning addresses this limitation by combining multiple individual models to enhance performance and robustness [7]. This approach leverages the strengths and mitigates the weaknesses of each model, resulting in improved accuracy and generalization capabilities. By aggregating the predictions from diverse models, ensemble methods reduce the risk of overfitting and increase the stability of the overall system. In the context of this research, ensemble learning plays a crucial role in refining the classification system, providing more reliable and consistent results across various skin types and dermatological conditions. This methodological framework, when integrated with advanced DeepLearning algorithms and user-friendly interfaces, aims to deliver a comprehensive and efficient solution for timely identification of skin conditions, ultimately bridging the gap between AI and dermatologists.

The transformative impact of the technology of AI on healthcare is profound, offering solutions that improve diagnosis, treatment and overall patient care, by performing complex tasks in fields such as radiology, pathology, and dermatology [8–14]. It can analyze medical images, such as X-rays, MRIs, and histopathological slides, which relies on visual interpretation, to detect subtle anomalies that might elude human eyes, especially in case of cancers. Deep learning not only relies on medical images, but also integrates patient data, genetic profiles and medical history, for personalized diagnosis [15]. However, integrating AI into healthcare is not without challenges. Data privacy concerns, the need for robust validation, and the importance of maintaining a human touch in patient-doctor interactions are critical aspects that must be carefully navigated [16].

In the literature, several algorithms and software solutions have been explored for dermatological diagnostics. Common approaches involve image analysis, pattern recognition, and machine learning. These Dermatological Diagnostic Assistance Platforms leverage advanced machine learning techniques to enhance dermatological diagnostics. By automating the analysis of dermatological images, they aim to overcome the limitations of traditional methods, providing a more efficient and accurate means of detecting and managing skin disorders. The significance lies in improving the precision of predictions, which enables healthcare professionals to make more accurate and timely diagnoses, potentially leading to advancements in understanding skin disorders and their treatment.

In this article, a web and mobile application that simplify the skin disease diagnostic process for dermatologists is proposed. This application offers a user-friendly interface with clear visual feedback for user interactions, and it allows direct queries to the deep learning (DL) model to obtain predicted disease results. The architecture of the model is based on an ensemble learning approach, combining EfficientNetB1 and EfficientNetB5 to leverage their individual strengths for enhanced feature extraction from skin images. This ensemble method ensures more accurate and robust diagnostic predictions. Moreover, the application includes features for prescribing treatments based on diagnoses, as well as tools for managing appointments, making the entire process more efficient and convenient for both dermatologists and patients. The platform not only enhances the diagnostic capabilities but also focuses on improving the overall user experience, making it a valuable tool for dermatological professionals in their daily operations.

This research paper is structured to systematically explore the multifaceted aspects of the application. Section 2 reviews existing literature, highlighting the foundational studies and identifying gaps that the research seeks to address. Section 3 presents the methodology, delving into the specifics of the proposed approach, detailing the

Convolutional Neural Network (CNN) model, the architecture, and the functionalities of the developed application. Section 4 provides the analysis and discussion part, interpreting the results and offering a critical evaluation of the findings in the context of existing research. In Section 5, the functionalities of both the web and mobile versions of the application are presented, along with illustrative examples. Section 6 explores the broader implications of the work, while Section 7 acknowledges the study's limitations and proposes directions for subsequent research. Section 8 outlines the quality assurance processes, including tests conducted on both the web and mobile applications to ensure reliability. Finally, in Section 9, the findings presented in earlier sections are synthesized and their implications discussed.

2. Literature survey

2.1. Related works

Recent advancements in artificial intelligence (AI) have significantly impacted the field of dermatology, particularly in the classification and diagnosis of skin diseases [17]. Various studies have demonstrated the efficacy of AI systems in improving diagnostic accuracy and assisting clinicians in triage. Table 1 presents a summary of related works, in terms of accuracy.

Pathak et al. [18] developed a CNN model using the MobileNet architecture, achieving an accuracy of 85% on the HAM10000 dataset. While this model demonstrates the effective application of deep learning

Table 1
Summary of related works.

Paper Title	Deep learning techniques	Dataset	Accur
Identification of Skin Diseases Using Convolutional Neural Network [18]	Applied MobileNet architecture for skin disease classification	HAM10000 with 7 classes	85%
A Web-Based Skin Cancer Assessment and Classification Using Machine Learning and Mobile Computerized Adaptive Testing in a Rasch Model [19]	Developed a k-nearest neighbors model for early-stage skin cancer risk assessment	Population-based Australian cohort study	91%
An Adaptive Federated Machine Learning-Based Intelligent System for Skin Disease Detection [20]	Proposed a federated learning system for skin disease detection	ISIC2019 with 8 classes	90%
A deep learning system for differential diagnosis of skin diseases [21]	Developed a deep learning system for differential diagnosis of 26 common skin conditions	Teledermatology dataset with 26 classes	66%
Identification and Classification of Skin Diseases using Deep Learning Techniques [24]	Utilized VGG16 and Streamlit for classifying skin diseases	Private collected dataset with 5 classes	86%
Machine Learning on Web: Skin Lesion Classification using CNN [25]	Implemented CNNs with Flask for lesion classification	ISIC with 4 classes	90%
An application for automated diagnosis of facial dermatological diseases [22]	Segmentation and classification Deep learning model using Matlab	Digital photographs provided from public databases	96.01%
Deep Learning-Based Skin Diseases Classification using Smartphones [23]	AN Android mobile App using CNN models implemented with TensorFlow Lite	Seven classes on a combined dataset	74.27%

techniques in dermatology, its accuracy could potentially be enhanced with further refinement of the training process and the use of larger, more diverse datasets. Additionally, the study primarily focuses on diagnostic accuracy and does not explore the integration of this model into practical clinical workflows.

Yang et al. [19] presented a web-based adaptive testing model for skin cancer risk assessment using k-nearest neighbors. Although the model achieved high precision, the use of k-nearest neighbors, a simpler algorithm, may hinder its scalability to more complex, diverse datasets. Additionally, the model is tailored for skin cancer, limiting its applicability to a broader range of dermatological conditions.

Hashmani et al. [20] proposed a federated machine learning system for skin disease detection, tested on the ISIC 2019 dataset. The study's federated learning approach successfully addresses privacy concerns, but the lack of emphasis on computational and communication overheads, limits its practical application. Moreover, it focuses primarily on privacy and less on the holistic patient management process, which is crucial for comprehensive healthcare solutions.

Liu et al. [21] developed a deep learning system for differential diagnosis that performed on par with dermatologists and outperformed primary care physicians. While the model demonstrated robust diagnostic capabilities, the integration of such a system into clinical workflows remains an under-explored aspect. The lack of focus on practical deployment and scalability in real-world clinical environments presents a barrier to its widespread adoption.

Goceri et al. [22] proposed a deep learning-based automated diagnosis system for facial dermatological diseases, using a Matlab-based application. Although the study achieved significant diagnostic accuracy, the reliance on Matlab, a specialized tool, limits the system's accessibility and usability, particularly in resource-limited settings or for non-technical healthcare professionals.

Oztel et al. [23] developed an Android app using TensorFlow Lite to classify seven skin diseases, achieving a classification accuracy of 74.27%. Although the application showcases the potential of mobile-based diagnostics, the relatively lower accuracy compared to desktop-based models suggests that further optimization is needed. Additionally, the scope of the diseases covered is limited, which restricts its utility in broader dermatological diagnostics.

Pai et al. [24] used the VGG16 architecture in a web application to classify skin diseases, achieving an accuracy of 86%. However, the study lacks a focus on user-friendliness and long-term patient care, which are critical for real-world healthcare applications. These studies collectively highlight the integration of machine learning frameworks in developing robust web and mobile applications for skin disease classification, showcasing advancements in accuracy and usability.

2.2. Identified gaps

While the related works demonstrate significant advancements in dermatological diagnostics using machine learning and deep learning techniques, several gaps remain unaddressed. For example, systems proposed by Pathak et al. [18], Liu et al. [21], and Pai et al. [24], focus primarily on achieving high diagnostic accuracy but do not provide comprehensive solutions that integrate continuous patient monitoring and management. Others, like Yang et al. [19] and Hashmani et al. [20], emphasize specific aspects like skin cancer risk assessment or federated learning for privacy but lack a broader application to a variety of skin conditions. Additionally, most of the studies, including those by Goceri et al. [22] and Oztel et al. [23], utilize platforms that may not be as accessible or user-friendly, particularly for non-technical healthcare professionals. These gaps highlight the need for a more holistic solution that not only ensures high diagnostic accuracy but also incorporates continuous patient care, user-friendly interfaces, and comprehensive healthcare management functionalities, which are essential for optimizing dermatological care and enhancing clinical decision-making processes.

2.3. Connection to the proposed work

SkinHealthMate App fills the gaps in previous research by integrating advanced image analysis with dedicated classification models, continuous patient monitoring, and comprehensive healthcare management. This holistic approach emphasizes user-friendly interfaces and continuous patient care, making it a versatile and accessible tool for healthcare professionals. Unlike prior works that focus solely on AI model classification, SkinHealthMate App offers comprehensive management of dermatological care, prioritizing a centralized approach with functionalities such as patient management and appointment scheduling. Additionally, by providing both web and mobile applications, SkinHealthMate App enhances accessibility and usability, positioning itself as a benchmark for future advancements in dermatological care through its holistic and integrated approach.

3. Methodology

In this section, the methodology employed in the research is outlined, detailing various aspects from data preparation to the software implementation of the solution.

3.1. Dataset

HAM10000 dataset, created for the International Skin Imaging Collaboration (ISIC) Challenge, is a comprehensive and widely utilized resource for skin lesion classification [26]. With over 15,000 high-resolution dermoscopic images labeled into eight different classes ('MEL', 'NV', 'BCC', 'AKIEC', 'BKL', 'DF', 'VASC'). Fig. 1 illustrates the distribution of images across different classes in the dataset.

HAM10000 aims to support the development and benchmarking of algorithms for differentiating between various skin conditions, including both benign and malignant lesions. However, it presents challenges such as class imbalance, high variability in image characteristics, and potential label noise due to diagnostic uncertainties.

3.2. Data augmentation

A variety of data augmentation techniques was employed to enhance model robustness and generalization. These techniques help in artificially increasing the diversity of the training data, which allows the model to better generalize to unseen data [27].

Specifically, based on the study of Perez et al. [28], several augmentation techniques were employed: images were rotated by up to 90 degrees, sheared by up to 20 degrees, and scaled within the range of [0.8, 1.2]. Additionally, random cropping was applied to simulate variability in lesion placement, helping the model detect relevant features regardless of their position in the image, thus improving its ability to generalize without relying on specific spatial patterns. Furthermore, images were randomly flipped both horizontally and vertically to provide additional variability. To ensure color robustness, hue shifts by sampling values from a uniform distribution in the range of [-0.1, 0.1] was used to account for variations in skin tones and lighting conditions, which are common in dermatological imaging. Additionally, cutout augmentation was applied, where random patches of the images were masked out, further promoting robustness to occlusions and partial visibility, a common challenge in clinical images where lesions might be partially obscured by hair or other objects. These augmentations collectively contribute to creating a more resilient and adaptable model capable of handling diverse real-world data scenarios. The Fig. 2 illustrates a series of images showcasing the various data augmentation techniques used.

3.3. Noise reduction

The application of median filtering on skin disease images plays a

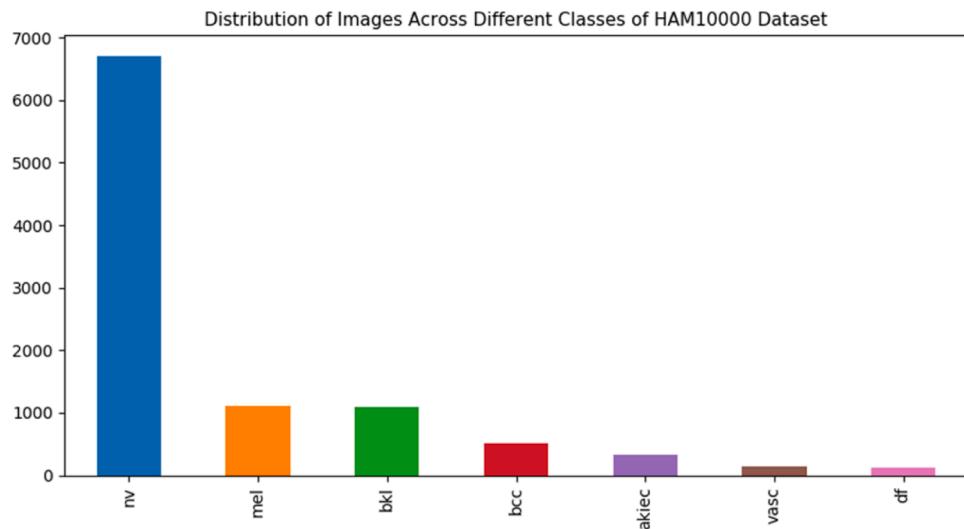


Fig. 1. Distribution of images across different classes in HAM10000.

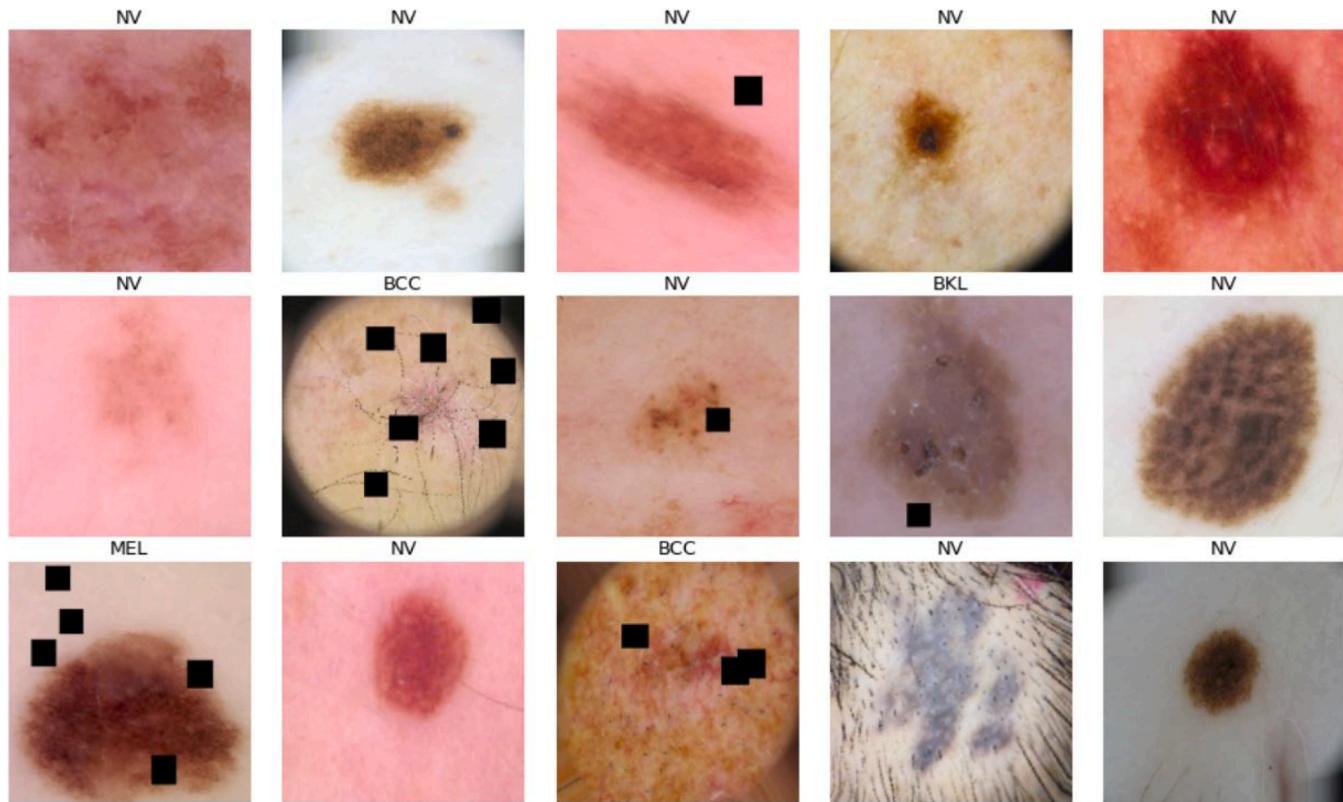


Fig. 2. Examples of data augmentation techniques applied to images.

crucial role in enhancing the accuracy of automated pattern recognition systems [29]. Median filtering is particularly effective in reducing salt-and-pepper noise, which is common in medical imaging. This technique works by replacing each pixel's value with the median value of the pixels in its neighborhood. In the context of skin disease images, where precise texture and boundary information are vital for accurate diagnosis, median filtering helps maintain the integrity of key features such as lesion borders and texture patterns. Median filtering was selected for this study due to its ability to preserve important edges in skin disease images, which is crucial for maintaining the sharp boundaries and textures necessary for accurate diagnosis by deep learning

models. Its effectiveness in removing impulsive noise, such as salt-and-pepper noise, commonly caused by lighting variations or imaging defects, makes it highly suitable for biomedical image processing. This is supported by prior research [30,31], which demonstrated the superiority of median filtering over Gaussian and Wiener filters in segmenting skin tumor images, effectively reducing noise while preserving critical details for accurate lesion detection. Additionally, median filtering is computationally efficient, making it appropriate for processing large datasets like HAM10000 without excessively smoothing important diagnostic features (Fig. 3).

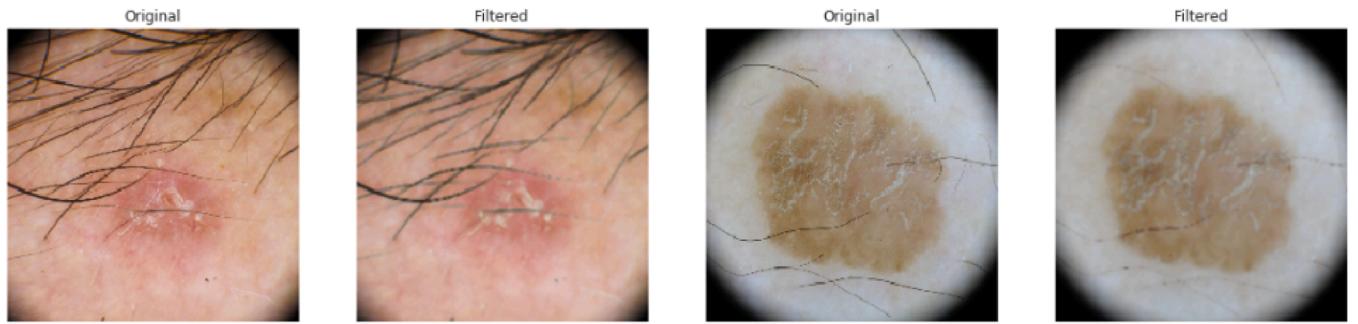


Fig. 3. Effect of applying median filter to skin diseases images.

3.4. Workflow

The architecture of the model used is based on an ensemble learning approach, based on EfficientNetB1 and EfficientNetB5 [32], to leverage their individual strengths for enhanced feature extraction from skin images. Fig. 4 illustrates the sequential steps from data preprocessing to model training and ensembling, highlighting the comprehensive approach undertaken to develop a robust and accurate classification system. First, the dataset of skin lesion images is loaded, preprocessed, and augmented with various techniques such as resizing, normalization, and transformations to improve model generalization and reduce overfitting. The normalization is represented mathematically as shown in this Eq. 1:

$$X' = \frac{X - \mu}{\sigma} \quad (1)$$

where X' is the normalized output, X is the input image pixel value, μ is the mean of the pixel values, and σ is the standard deviation, ensures that the pixel values are centered and scaled.

Two variations of EfficientNet, EfficientNetB1 and EfficientNetB5, are defined and compiled with a classification head composed of Global

Average Pooling, Dense, Dropout, and Softmax layers. The models are trained individually using the Adam optimizer and CategoricalCrossEntropy loss function. The categorical crossentropy loss is calculated as show in Eq. 2 :

$$L(y, \hat{y}) = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (2)$$

where L is the loss, y is the true distribution, \hat{y} is the predicted probability distribution, and C is the number of classes. This function quantifies the difference between the true labels and predicted probabilities.

Once trained, the ensemble learning strategy is applied by averaging the predictions from both EfficientNetB1 and EfficientNetB5, leveraging the strengths of each model to improve overall classification performance. The final prediction P is obtained by:

$$P = \frac{1}{N} \sum_{n=1}^N \hat{y}_n \quad (3)$$

where N is the number of models (in this case, 2), and \hat{y}_n is the predicted output from each model. This approach leverages the strengths of each model to improve overall classification performance, as described in Eq.

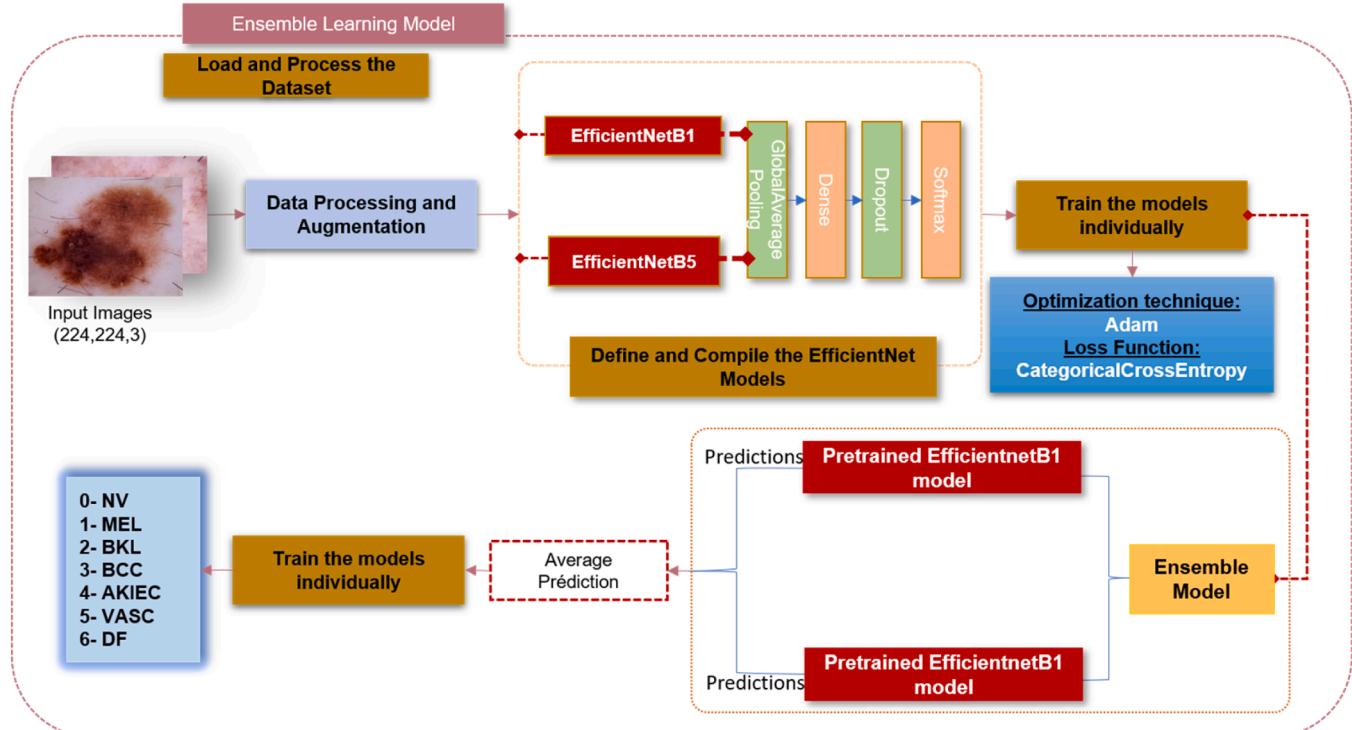


Fig. 4. Training process overview for the skin disease classification model.

3:

The ensemble then generates the final prediction, classifying the skin lesions into one of seven categories: NV, MEL, BKL, BCC, AKIEC, VASC, or DF.

The decision to use EfficientNetB1 and EfficientNetB5 for the ensemble learning approach was based on their complementary strengths. EfficientNetB1 strikes a balance between model complexity and efficiency, making it suitable for extracting features with lower computational cost, while EfficientNetB5 excels in capturing higher-level, complex features due to its increased depth and parameters. By averaging their outputs, the ensemble leverages both the efficiency of EfficientNetB1 and the advanced feature extraction capabilities of EfficientNetB5, resulting in a diverse set of features and improved classification accuracy. These models were pre-trained on large datasets using the 'Noisy Student' weights, providing a strong foundation for generalization. Both models are initialized with these weights and expect input images of size 224×224 pixels, allowing the ensemble to handle the diverse characteristics of skin disease images more effectively. Fine-tuning was used to adapt these pre-trained models to the skin disease classification task by initially freezing most layers and training only the new top layers, then gradually unfreezing and training more layers with a lower learning rate to avoid overfitting. Specifically, the last 20 layers of the model were selectively unfrozen, excluding batch normalization layers to ensure training stability. This process can be described as:

$$\theta' = \theta - \eta \nabla L(\theta) \quad (4)$$

where θ' represents the updated weights, θ is the current weight, η is the learning rate, and $\nabla L(\theta)$ is the gradient of the loss function with respect to the weights. This allowed the model to fine-tune deeper layers relevant to skin disease classification, while preserving the robust general features learned from the "noisy-student" training process, as illustrated in Eq. 4.

4. Experimental results and discussion

The ensemble model utilizes a learning rate of 0.0001, with Adam optimizer, with training controlled by an early stopping mechanism. The learning rate's modest value allows precise weight adjustments, preventing overfitting during training, while the Adam optimizer efficiently updates model weights. These hyperparameters reflect a balanced approach to ensure effective learning and generalization, optimizing the ensemble model's performance for skin disease classification.

The performance of the developed skin disease classification model was evaluated using various metrics, including accuracy, precision, recall, F1 score, and area under the curve (AUC). The model exhibited high accuracy, reaching 93%, showcasing its ability to correctly classify skin diseases. Table 2 presents the precision, recall, and F1 score, highlighting the model's effectiveness in minimizing false positives and false negatives.

The classification report reveals a detailed performance analysis of the model across various skin lesion categories. Overall, the model achieves an impressive accuracy of 93%, indicating its robust performance in identifying and classifying skin lesions accurately. The

weighted average precision, recall, and F1-score are all 0.93, further supporting the model's effectiveness in this task.

Breaking down the results for individual categories, the model performs exceptionally well in identifying nevus lesions ('nv'), with a precision of 0.97, recall of 0.99, and F1-score of 0.98. This high performance can be attributed to the large number of samples (748), which likely provided ample data for the model to learn from. Similarly, vascular lesions ('vasc') also show excellent results, with a precision and recall of 0.92 and 1.00 respectively, and an F1-score of 0.96, despite having only 11 samples. This indicates that the model can generalize well even with fewer samples in certain categories.

However, the model's performance on melanoma ('mel') is notably lower, with a precision of 0.63, recall of 0.49, and an F1-score of 0.55. This suggests that the model has more difficulty correctly identifying melanoma lesions, which could be due to their similarity to other lesion types or insufficient representation in the training data. The category Actinic keratoses and intraepithelial carcinoma ('akiec') also shows a lower recall of 0.54, although it has a higher precision of 0.88, resulting in an F1-score of 0.67. This indicates that while the model can accurately identify 'akiec' when it predicts it, it misses a substantial number of true 'akiec' cases.

For basal cell carcinoma ('bcc') and benign keratosis-like lesions ('bkl'), the model shows balanced performance with F1-scores of 0.84 and 0.76, respectively. The precision and recall for these categories are relatively high, indicating that the model is both accurate and consistent in identifying these types of lesions.

The macro average, which treats all classes equally regardless of their support, shows an F1-score of 0.81. This lower macro average compared to the weighted average suggests that while the model performs well overall, its performance is less consistent across all classes, particularly those with fewer samples. The Fig. 5 illustrate the training and validation performance of the model in terms of both accuracy and loss over 30 epochs. The first figure shows the model's loss, where both training and validation losses decrease consistently and converge, indicating the absence of overfitting. The second figure presents the accuracy metrics, demonstrating steady improvement in both training and validation accuracy, further confirming that the model generalizes well without overfitting to the training data. Together, these figures highlight the model's robustness and effective learning process.

The confusion matrix provides a detailed and insightful overview of the performance of classification models. Fig. 6 visually presents the confusion matrix of the system, offering a clear representation of the model's predictive accuracy and errors.

An essential component of the evaluation of a Deep Learning models is the ROC curve, which offers a thorough representation of the model's effectiveness across various decision thresholds. Fig. 7 displays the ROC curve of the ensemble learning model.

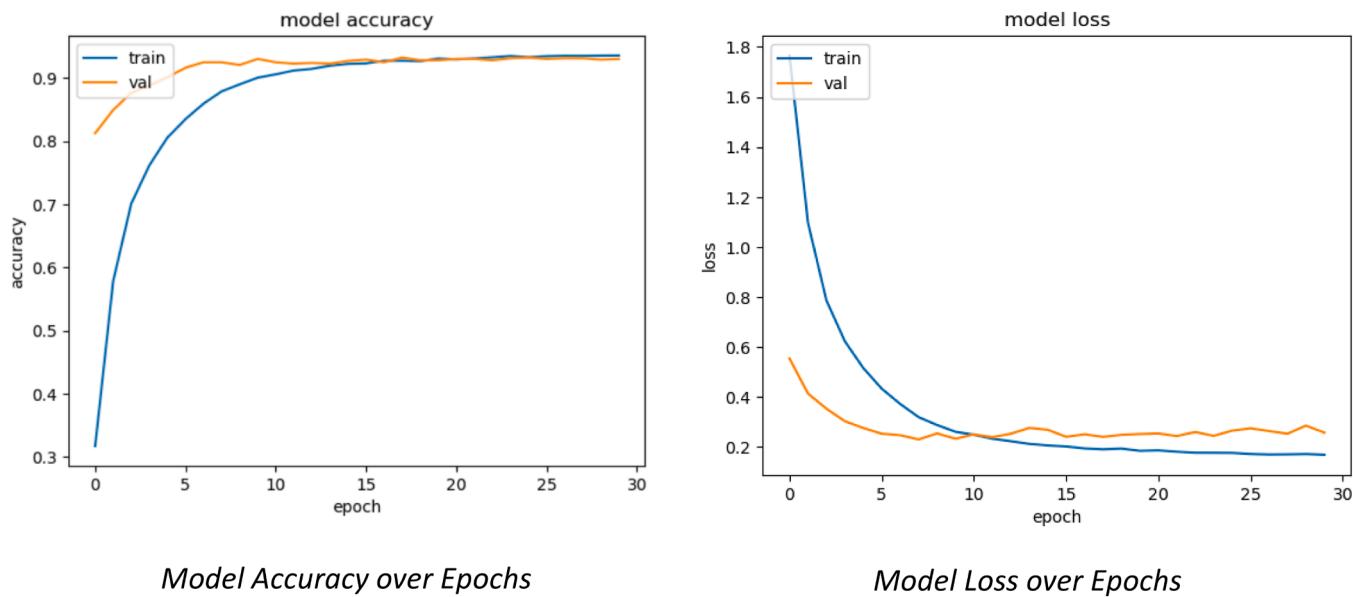
To illustrate the performance of the proposed model in classifying skin diseases, Gradient-weighted Class Activation Mapping (Grad-CAM) was utilized.

Grad-CAM provides visual explanations by highlighting the regions of the input image that are most influential in the model's decision-making process. The Grad-CAM presented in Fig. 8 showcases the model's focus areas, when predicting various skin disease categories. These visualizations help in understanding how the model interprets and distinguishes different skin conditions, thereby validating its robustness and accuracy. For dermatologists, these visualizations offer a layer of transparency, allowing them to assess whether the model is focusing on clinically relevant areas, such as lesion borders or irregularities, rather than irrelevant features like background or artifacts. This ensures the model's decisions align with expert knowledge and diagnostic criteria.

By overlaying the heatmaps on the original images, it can be observed that the model correctly identifies and emphasizes the critical regions associated with the diseases, demonstrating its capability to make informed and reliable predictions. This visualization technique not only aids in model interpretability but also provides insights into

Table 2
Ensemble model performance of HAM10000 dataset.

Class	Precision	Recall	F1-Score	Support
akiec	0.88	0.54	0.67	26
bcc	0.74	0.97	0.84	30
bkl	0.77	0.75	0.76	75
df	1.00	0.83	0.91	6
mel	0.63	0.49	0.55	39
nv	0.97	0.99	0.98	748
vasc	0.92	1.00	0.96	11
accuracy	0.93			



Model Accuracy over Epochs

Model Loss over Epochs

Fig. 5. Training and validation loss and accuracy.

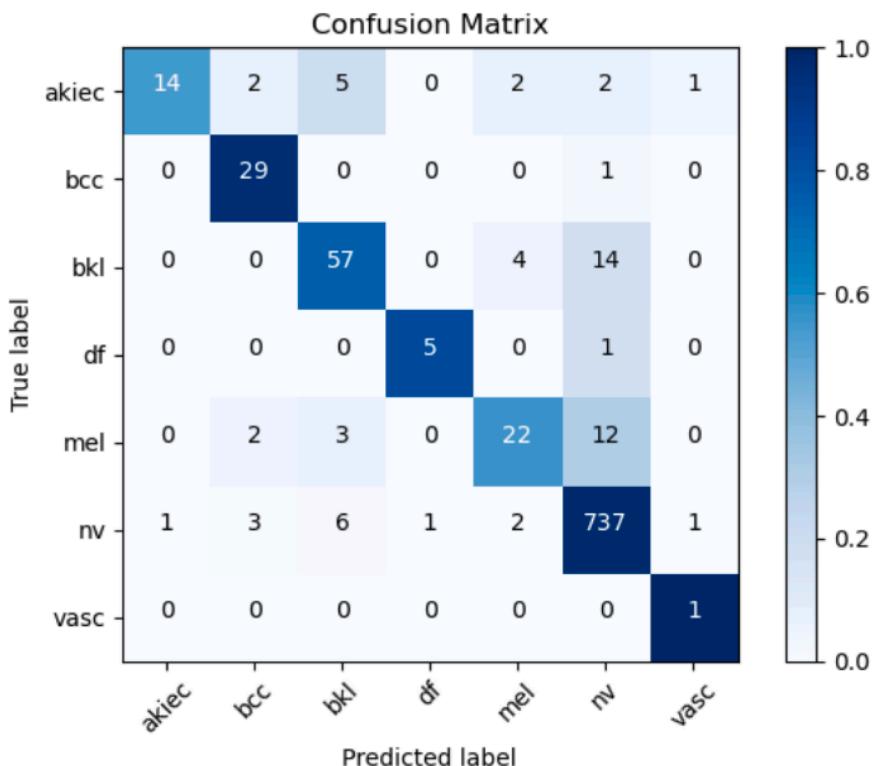


Fig. 6. Confusion matrix of the proposed model.

potential areas for further improvement and refinement of the model.

5. Software description

5.1. Software architecture

The diagnosis software is thoughtfully designed, offering a robust server-side component and versatile client-side applications. The overall architecture is presented in Fig. 9.

The Dermatological Diagnostic Assistance Platform is built on a modern and scalable architecture using the JHipster framework. The

architecture consists of three main components: the Spring Boot back-end, the React web application, and the React Native mobile application. The Spring Boot Backend forms the foundation, providing robust and scalable support for data processing and managing interactions with the database. It adopts a monolithic architecture and incorporates RESTful APIs to enable seamless communication with the frontend components. In parallel, the Web front-end is constructed using React, delivering a responsive and user-friendly interface tailored for dermatologists, secretaries, patients, and administrative users. This component introduces essential features like patient management, appointment scheduling, and diagnostic report generation. State management is handled

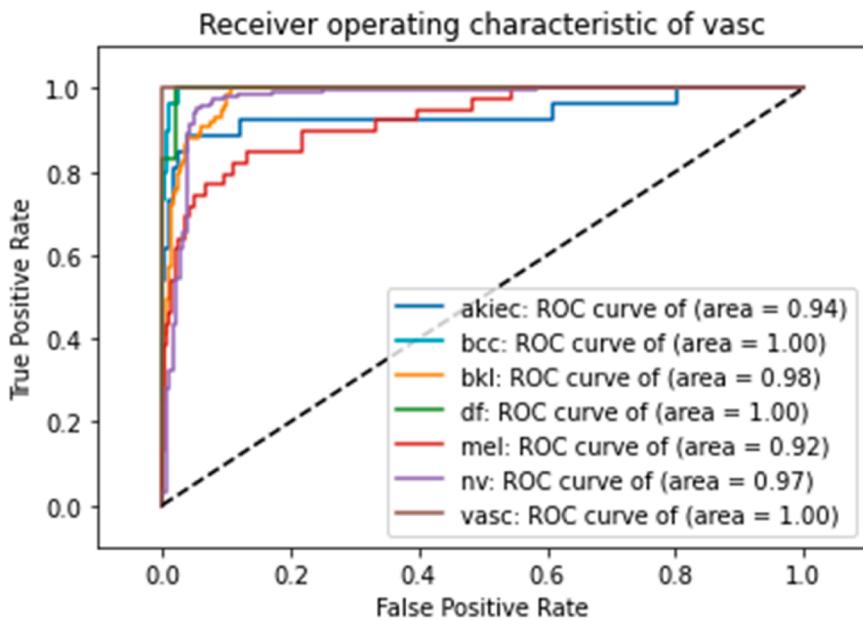


Fig. 7. ROC curve graph.

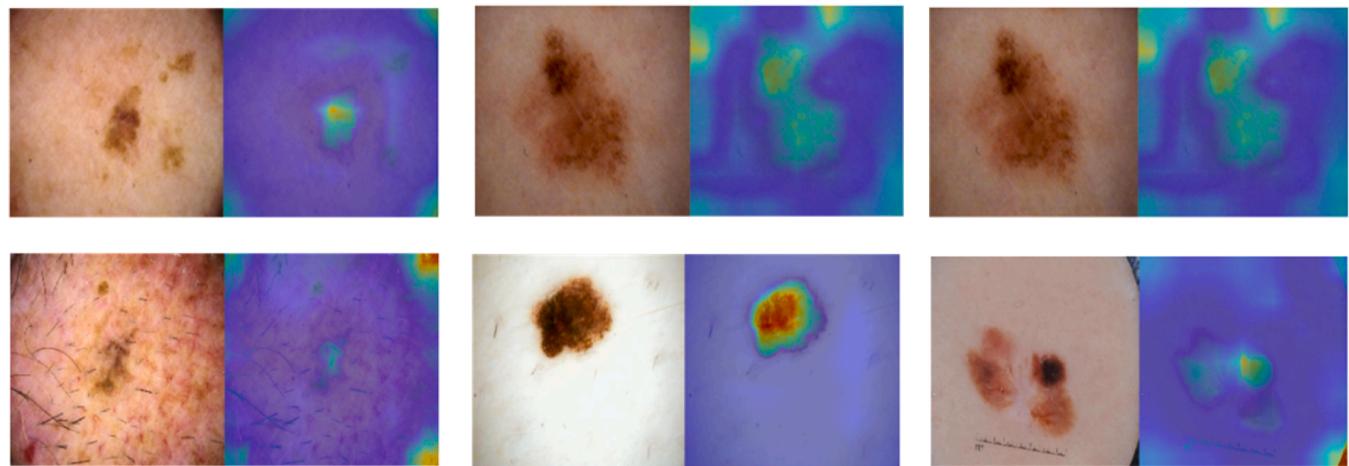


Fig. 8. Grad-CAM visualization of skin disease classification.

efficiently with React hooks, ensuring smooth communication with the backend through API calls. Complementing the web application, the Mobile Application extends the platform's capabilities to dermatologists. Developed with React Native, it enables secure access to medical records, viewing upcoming appointments, and receiving diagnostic results. This mobile component significantly enhances user accessibility, empowering dermatologists to actively engage in their healthcare journey.

5.2. Software functionalities

HealthMate Application boasts a comprehensive set of functionalities designed to revolutionize the field of dermatological diagnostics. Built on a modern and scalable architecture using the JHipster framework, the system accommodates three primary user roles: Doctors, Administrators and Secretaries.

5.2.1. Web application (ReactJS)

The web application is developed using ReactJS and caters to four user profiles: Admin, Secretaries, Doctors, and Patients. It encompasses

the following functionalities:

- User Management:** Administrators wield the power to create and oversee user accounts, encompassing the management of personal data, profile images, and contact particulars.
- Scheduling and Appointment validation:** Secretaries hold the reins in the creation of new patient profiles and possess the capability to schedule appointments efficiently. They can do this by selecting available dates and times from a dynamic calendar, ensuring a seamless appointment management system. Moreover, it plays a pivotal role in validating appointments postpatient visits, thereby granting access to consultations.
- Access to Diagnostic Records:** The platform provides a dynamic dashboard offering a real-time snapshot of daily appointments and consultations, empowering efficient schedule management. The Patient Management feature enables doctors to access, update, and diagnose patients, leveraging machine learning algorithms for preliminary analysis. Dermatologists benefit also from comprehensive access to their patients' complete medical histories, comprising prior consultations and diagnoses. This invaluable feature empowers

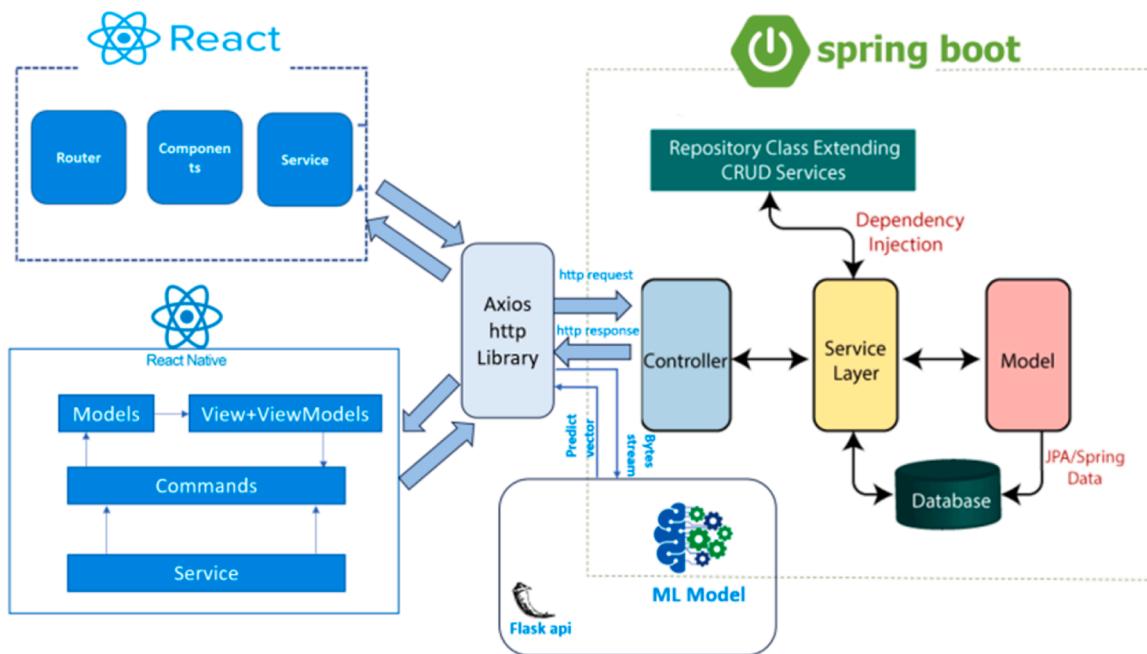


Fig. 9. The project architecture.

dermatologists to monitor the progression of skin conditions over time.

- **Access to appointments and medical records:** Patients, as end-users, also benefit from accessing to HealthMateApp. They have access to their comprehensive medical records, diagnoses, treatment recommendations, upcoming appointments, and personal profiles.
- **Diagnosis:** Dermatologists can delve into their patients' diagnostic histories, input the lesion symptoms, and upload images for diagnosis, whether from their local computers or images sent via the mobile application. The deep learning model meticulously analyzes the provided images, furnishing an initial diagnosis complete with a confidence rating, reflecting the model's confidence level in the proposed diagnosis. To ensure utmost accuracy, the web application bestows upon dermatologists the authority to either validate or

modify the diagnosis, issue prescriptions for patient treatment and make additional comments.

- **Detailed Information Presentation:** Once the diagnosis is validated, medical images correlating with the diagnosis are securely stored within the system. Dermatologists can seamlessly navigate this database, enabling them to compare images of previously validated diagnoses with current patient cases, providing insights into specific lesion characteristics.
- **Secure Communication:** The web application guarantees secure communication with the backend server, safeguarding sensitive information and user medical data.

5.2.2. Mobile application (React Native)

The mobile application, built on React Native, is designed for Doctors to complement the web part. It incorporates the following

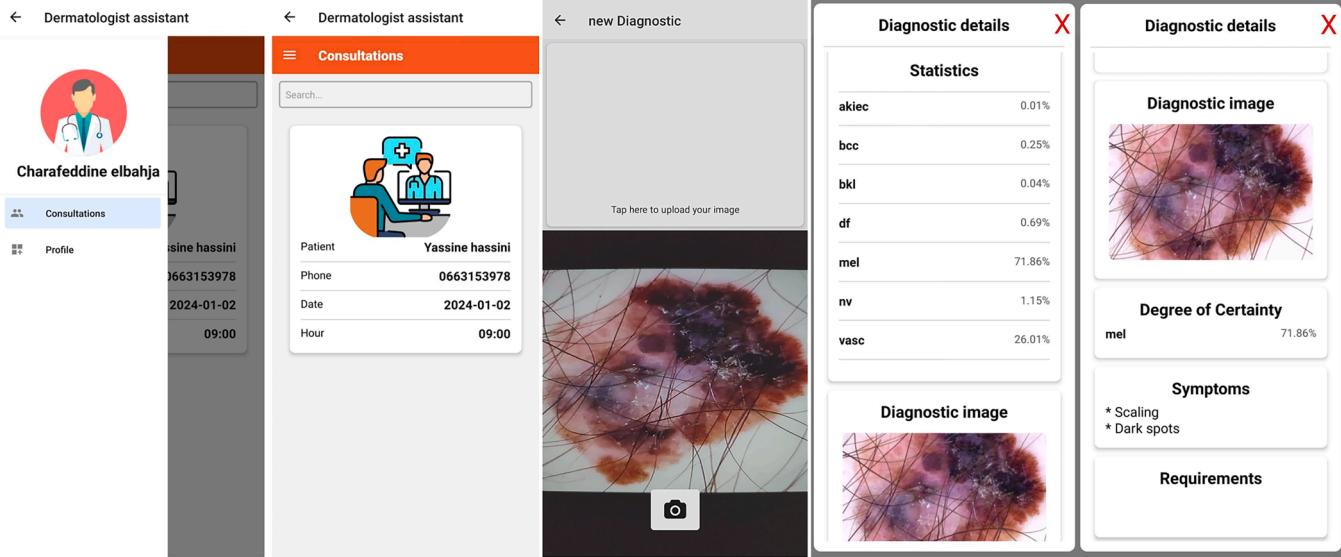


Fig. 10. Mobile App functionalities: (a. Sidebar Menu, b. Today's Consultations, c. Capture/Upload Skin Image, d. Predictions with Probabilities, e. Predicted Diseases with Certainty).

functionalities:

- **Viewing Consultations (Dermatologists):** Dermatologists can access their day's consultations on the mobile app. This feature allows doctors to stay up-to-date with their daily visits.
- **Symptom Input and Diagnosis (Dermatologists):** Doctors can input patient symptoms and submit skin images for diagnosis. The app communicates with the classification model to generate a preliminary diagnosis. Fig. 10 illustrates the user interface components of the Mobile Application.

Images related to diagnoses are securely stored within the app, ensuring patient privacy and data security. The images stored within the database are meticulously managed to uphold the strictest standards of patient confidentiality and ethical considerations. No patient-specific information is ever disclosed or made accessible. Moreover, to ensure the highest quality images, smartphones can be paired with a dermatoscope designed for smartphones, guaranteeing optimal image quality and diagnostic accuracy during the examination process. This integration enhances the overall diagnostic capabilities of the system while maintaining the utmost respect for patient privacy.

5.3. Illustrative examples

5.3.1. Case 1 : Diagnostic creation workflow

The Fig. 11 below illustrates the diagnostic creation form utilized by doctors within the Dermatological Diagnostic Assistance Platform. In this interactive interface, doctors engage in a comprehensive diagnostic process by selecting relevant patient symptoms and uploading an image of the affected area. This critical step initiates the system's response, triggering the underlying machine learning model to analyze the uploaded image and generate a preliminary diagnosis.

The machine learning model, trained to recognize patterns and characteristics indicative of various skin conditions, plays a pivotal role in providing valuable insights into the potential ailment. Subsequently, the diagnostic form acts as the conduit through which doctors receive the model's response, revealing the identified skin disease based on the data inputs. This seamless integration of user inputs, sophisticated machine learning algorithms, and diagnostic outcome not only facilitates an efficient workflow for healthcare professionals but also underscores

the platform's commitment to harnessing cutting-edge technology for accurate and timely dermatological diagnoses.

5.3.2. Case 2 : Detailed diagnostic review

The Fig. 12 bellow illustrates the list of diagnoses associated with a specific patient visit within the Dermatological Diagnostic Assistance Platform.

At this stage, doctors have the capability to enhance the diagnostic information by appending prescriptions and detailed descriptions to a specific diagnosis. This multifaceted functionality empowers healthcare professionals to provide a more comprehensive and personalized assessment of the patient's dermatological condition. Moreover, doctors can delve into the specifics of each diagnosis, reviewing intricate details and gaining a nuanced understanding of the patient's medical situation. As a crucial component of the diagnostic workflow, the platform affords doctors the ability to validate the diagnosis, affirming its accuracy or making necessary adjustments based on their clinical expertise.

6. Impact

Recent research highlights the transformative potential of such integrated healthcare solutions. Studies have shown that the use of advanced image analysis and machine learning models significantly enhances diagnostic accuracy and efficiency in various medical fields [33,34]. These advancements are reflected in the capabilities of the SkinHealthMate Application, which leverages a CNN model to provide preliminary diagnoses with high accuracy, minimizing the need for invasive procedures [35]. The integration of comprehensive patient records, including medical history, symptom descriptions, clinical images, and dermatologist notes, ensures informed decision-making and effective interventions. The application's confidence rate feature allows for validation and adjustment of initial diagnoses, facilitating the collection of correctly labeled data for ongoing model improvement.

Additionally, the dual availability of the application in web and mobile versions enhances accessibility, enabling healthcare professionals to access and utilize the platform's services efficiently. This ease of access significantly reduces diagnosis times and optimizes the expertise of healthcare providers in managing complex cases. According to [36], machine learning applications in dermatology have the potential to enhance the dermatologist's practice from diagnosis to

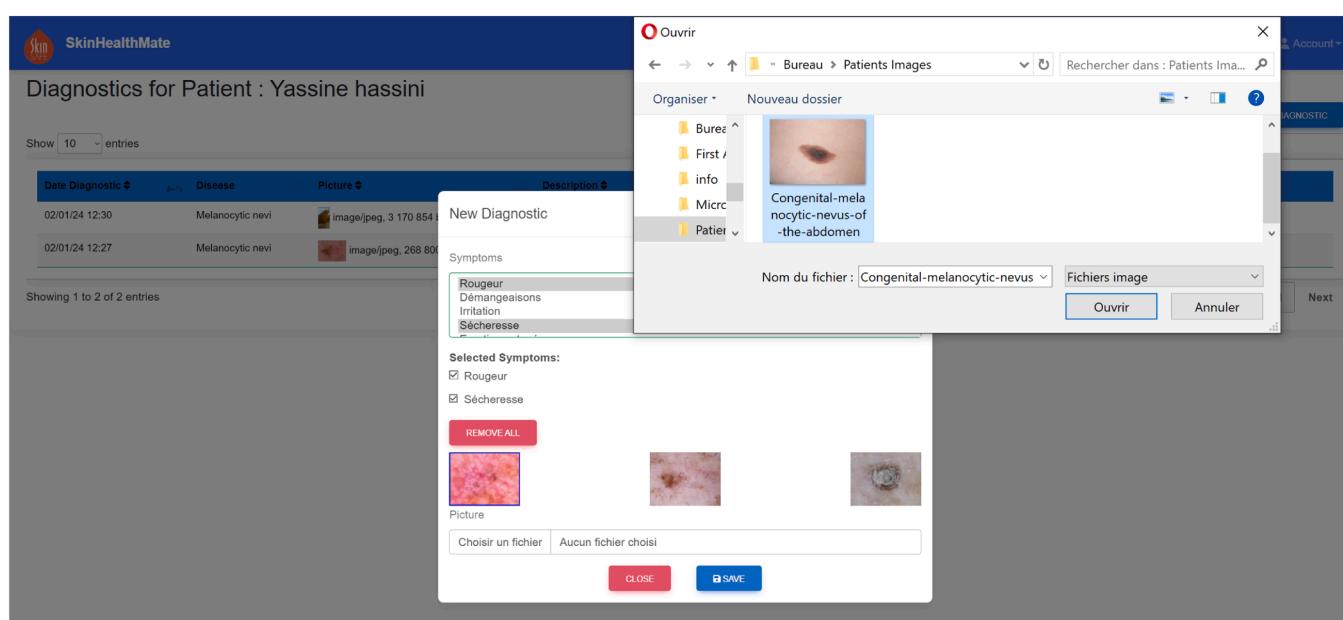


Fig. 11. Form for creating a diagnosis.

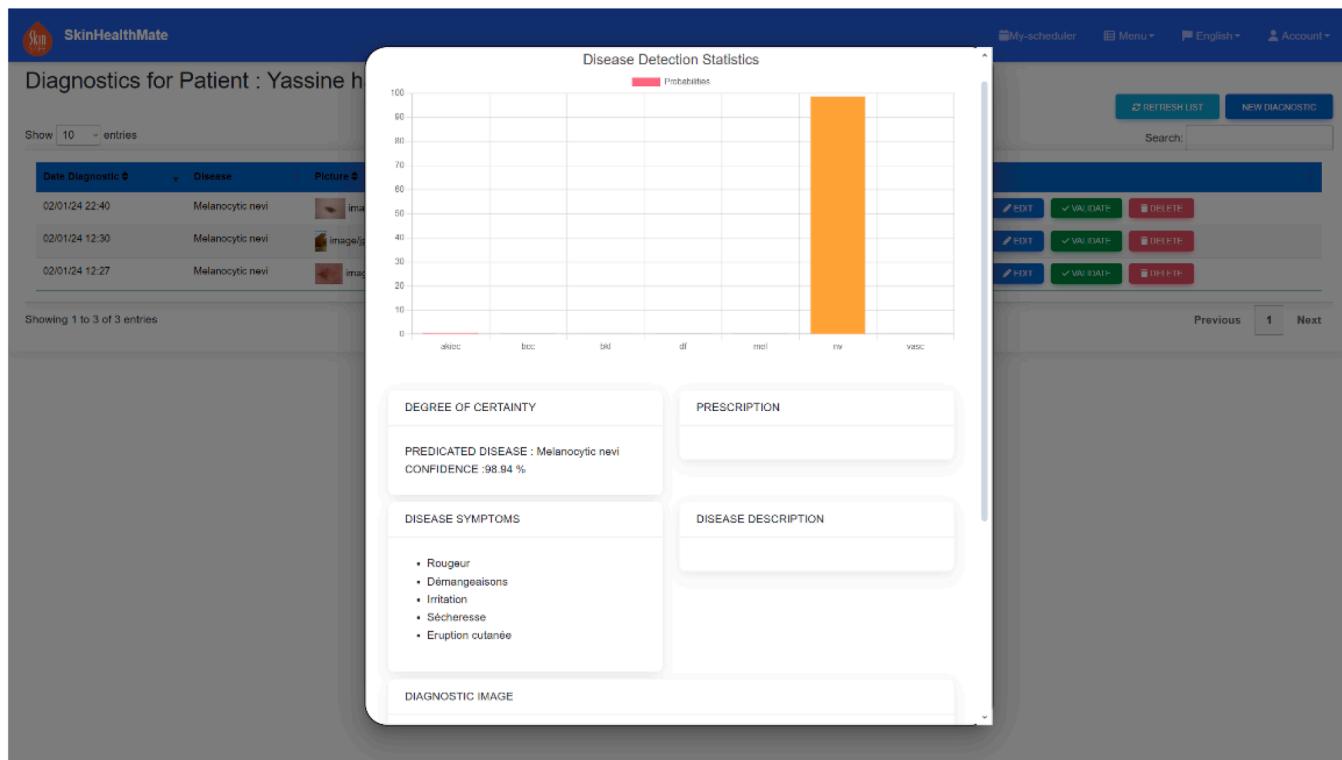


Fig. 12. Diagnostic details.

personalized treatment, highlighting the broad impact of such technologies in clinical settings.

The application's ability to continuously improve through the collection of correctly labeled data and the retraining of its CNN model ensures that it remains effective and relevant. This adaptability is crucial for handling a variety of medical specialties beyond dermatology, making it a versatile tool in healthcare.

7. Limitations and future work

The SkinHealthMate application has certain limitations that need to be acknowledged. Firstly, the application cannot ensure accurate diagnosis for darker-skinned people, as the model has been trained on the HAM10000 database, which consists mainly of dermatoscopic images of fair skin. Also, the accuracy of AI models can be affected by the quality of the input images. Variations in lighting, angle, and resolution can impact the performance of the model, leading to incorrect diagnoses. Ensuring consistent image quality across different devices and settings can be challenging [37]. In addition, while the AI algorithm can classify a range of skin diseases, it may struggle with rare or complex conditions that deviate from the training data, leading to misdiagnoses or reduced accuracy. Complex cases may require a more comprehensive approach involving various medical tests and expert opinions, which the application may not adequately take into account. Furthermore, the use of AI in healthcare raises regulatory and ethical issues, including concerns about patient privacy, data security, and the potential for algorithmic bias. Ensuring compliance with healthcare regulations and maintaining ethical standards are critical challenges that must be addressed [38]. Given these limitations, the app should be seen as a complementary tool to medical expertise, rather than as a standalone solution. Future work consists of refining the skin disease classification system to ensure its effectiveness across a broader spectrum of skin types and a wider range of dermatological conditions. These advances are intended to enhance the usefulness of the application, encouraging inclusion and improving its impact on various user demographics.

8. Quality assurance

Quality assurance (QA) is critical in ensuring that web and mobile applications function effectively and meet the desired standards of performance, security, reliability, and maintainability. Both types of applications require robust testing frameworks and methodologies to guarantee their quality, but there are key differences between the two in terms of challenges and approaches.

For web applications, QA efforts emphasize compatibility across different browsers, responsiveness for various screen sizes, and handling different internet speeds. Performance optimization, particularly in reducing loading times and maintaining a secure codebase, is a priority. Tools like SonarQube are essential for maintaining high code quality by identifying issues related to reliability, maintainability, security, and code duplication.

In contrast, mobile applications require additional considerations, such as device fragmentation, different operating system versions, and the need for optimized performance on devices with varying hardware capacities. Ensuring offline functionality, responsive gesture recognition, and low battery consumption are crucial factors in mobile QA. Moreover, meeting the respective standards of both the Apple App Store and Google Play is key to successful mobile app deployment.

Table 3 highlights the key quality assurance metrics for web and mobile applications based on a SonarQube analysis. The data illustrates how each platform presents its own distinct challenges and how QA strategies must be tailored accordingly.

The table shows that while both applications handle their respective challenges well, the web application prioritizes cross-browser functionality and overall web performance, whereas the mobile application focuses on delivering a highly optimized experience tailored to mobile devices. Each platform has been optimized to ensure seamless performance, whether users are interacting with the app via a web browser or on a mobile device.

In addition to this analysis, performance tests were conducted on both web and mobile applications. Table 4 presents the response times

Table 3

Key differences in quality assurance for web and mobile applications based on a recent analysis using SonarQube.

Quality Metric	Web Application	Mobile Application
Security Issues	32 Open Issues	1 Open Issue
Reliability Issues	14 Open Issues	48 Open Issues
Maintainability	238 Open Issues	201 Open Issues
Coverage	0.0% (No lines covered)	0.0% (451 lines to cover)
Duplications	6.4% (on 7.1k lines)	2.6% (on 2.5k lines)
Code Size	7.1k Lines of Code	2.5k Lines of Code
Platform Focus	Optimized for multiple browsers	Optimized for multiple devices (iOS, Android)
Unique Challenges	Handling different screen sizes, browser compatibility	Device fragmentation, OS version compatibility

for different image sizes processed by each platform, giving an overview of how both types of applications handle varying data loads.

Table 4 showcases the response times for processing images of varying sizes in both web and mobile applications. The data highlights how both platforms can efficiently handle smaller images, with processing times well within acceptable ranges.

- For smaller images (4.91 KB - 6.38 KB): Both platforms handle these sizes efficiently, with web applications showing slightly faster response times. This suggests that smaller media assets are processed efficiently in real-time scenarios, allowing both web and mobile applications to perform optimally under normal usage conditions.
- For medium-sized images (28.9 KB - 84.7 KB): The response times show a gradual increase as image size grows, with both platforms handling the additional load effectively. Web applications demonstrate strong performance in managing these image sizes, and mobile applications are similarly well-equipped to process such assets, making them suitable for media-rich content.
- For larger images (352 KB): Even when faced with larger image sizes, both web and mobile applications demonstrate solid performance. The web application maintains a high level of responsiveness, while the mobile application also handles these larger assets efficiently, demonstrating its ability to manage high-resource tasks, particularly in environments where performance is a priority.

This analysis emphasizes the importance of image size and optimization in application performance. Both platforms have shown the capacity to handle images of varying sizes effectively, making them suitable for a wide range of content-heavy use cases.

Table 4

Comparison of response times between web and mobile applications.

Image	Image Size (KB)	Response Time Web App (ms)	Response Time Mobile App (ms)
	4.91	399.7	753.8
	6.38	343.5	771.2
	28.9	478.2	863.8
	84.7	406.4	769.1
	352	497.6	822.1

9. Conclusions

Artificial intelligence (AI) is revolutionizing the field of dermatology, offering unprecedented improvements in diagnostic accuracy, efficiency, and patient care. SkinHealthMate Application signifies a pivotal advancement in dermatological care, by leveraging technology to overcome the limitations of traditional dermatological methods, this application aims to streamline diagnosis, improve communication among medical professionals, and ultimately enhance patient care and outcomes.

Its availability in two versions (web and mobile) facilitates access to and use of the services offered, enabling healthcare professionals to reduce diagnosis times and optimize their expertise for complex cases. With comprehensive patient records, including patient medical history, symptom description, clinical images, dermatologist notes and comments, it facilitates informed decision-making and leads to more effective interventions. Beyond its benefits for dermatologists, the application extends its usefulness to secretaries responsible for booking and managing appointments. Additionally, the platform's appointment management system ensures seamless coordination between dermatologists and patients, reducing scheduling errors, optimizing the use of healthcare resources and enabling patients to access their medical records and previous treatments. The centralized system fosters a sense of collective ownership. Also, access to dermatologists to validate or refine initial diagnoses, can help collect more correctly labeled data, which can be used to retrain and strengthen the model.

Future research and development should focus on overcoming existing challenges to maximize the application's effectiveness and ensure equitable access to its benefits.

CRediT authorship contribution statement

Amina Aboulmira: Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Mohamed Lachgar:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Hamid Hrimech:** Writing – review & editing, Validation, Supervision, Project administration, Formal analysis. **Aboudramane Camara:** Visualization, Software, Formal analysis, Conceptualization. **Charafeddine Elbahja:** Visualization, Software, Formal analysis, Conceptualization. **Amine Elmansouri:** Visualization, Software, Conceptualization. **Yassine Hassini:** Visualization, Software, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the github link to my code and data in the manuscript

References

- [1] I. Zalaudek, et al., Dermoscopy in General Dermatology, *Dermatology* 212 (1) (2006) 7–18, <https://doi.org/10.1159/000089015>.
- [2] G. Argenziano, et al., Dermoscopy features of melanoma incognito: indications for biopsy, *J. Am. Acad. Dermatol.* 56 (3) (2007) 508–513, <https://doi.org/10.1016/j.jaad.2006.10.029>.
- [3] S.A. Nemeth, N. Lawrence, Site identification challenges in dermatologic surgery: a physician survey, *J. Am. Acad. Dermatol.* 67 (2) (2012) 262–268, <https://doi.org/10.1016/j.jaad.2012.03.016>.
- [4] D. Shoieb, S. Youssef, W. Aly, Computer-aided model for skin diagnosis using deep learning, *J. Image Graph.* 4 (2016) 116–121, <https://doi.org/10.18178/jig.4.2.122-129>.
- [5] R. Pratiwi, S. Nurmaini, M. dian rini, Naufal Rachmatullah, A. Darmawahyuni, Deep ensemble learning for skin lesions classification with convolutional neural network, *IAES Int. J. Artif. Intell.* 10 (2021) 563–570, <https://doi.org/10.11591/ijai.v10.i3.pp563-570>.
- [6] N. Kausar, et al., Multiclass skin cancer classification using ensemble of fine-tuned deep learning models, *Appl. Sci.* 11 (2021) 10593, <https://doi.org/10.3390/app112210593>.
- [7] X. Dong, Z. Yu, W. Cao, Y. Shi, Q. Ma, A survey on ensemble learning, *Front. Comput. Sci.* 14 (2) (2020) 241–258, <https://doi.org/10.1007/s11704-019-8208-z>.
- [8] E. Göceri, Vision transformer based classification of gliomas from histopathological images, *Expert. Syst. Appl.* 241 (2024) 122672, <https://doi.org/10.1016/j.eswa.2023.122672>.
- [9] F. Idlahcen, A. İdrı, E. Göceri, Exploring data mining and machine learning in gynecologic oncology, *Artif. Intell. Rev.* 57 (2) (2024) 20, <https://doi.org/10.1007/s10462-023-10666-2>.
- [10] E. Göceri, Convolutional neural network based desktop applications to classify dermatological diseases, in: 2020 IEEE 4th international conference on image processing, applications and systems (IPAS), IEEE, 2020, pp. 138–143, <https://doi.org/10.1109/IPASS0080.2020.9334956>.
- [11] E. Göceri, Polyp segmentation using a hybrid vision transformer and a hybrid loss function, *J. Imaging Inform. Med.* (2024) 1–13, <https://doi.org/10.1007/s10278-023-00954-2>.
- [12] E. Göceri, Automated skin cancer detection: where we are and the way to the future, in: 2021 44th International Conference on Telecommunications and Signal Processing (TSP), IEEE, 2021, pp. 48–51, <https://doi.org/10.1109/TSP52935.2021.9522605>.
- [13] A. Aboulmira, H. Hrimech, M. Lachgar, Comparative study of multiple CNN models for classification of 23 skin diseases, *Int. J. Online Biomed. Eng.* 18 (11) (2022), <https://doi.org/10.3991/ijoe.v18i11.32517>.
- [14] E. Göceri, Impact of deep learning and smartphone technologies in dermatology: automated diagnosis, in: 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, 2020, pp. 1–6, <https://doi.org/10.1109/IPTA50016.2020.9286706>.
- [15] Q.D. Buchlak, et al., Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review, *Neurosurg. Rev.* 43 (5) (2020) 1235–1253, <https://doi.org/10.1007/s10143-019-01163-8>.
- [16] S. Gerke, T. Minssen, G. Cohen, Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare, in: A. Bohr, K. Memarzadeh (Eds.), *Artificial Intelligence in Healthcare*, Academic Press, 2020, pp. 295–336, <https://doi.org/10.1016/B978-0-12-818438-7.00012-5>.
- [17] G. Litjens, et al., A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88, <https://doi.org/10.1016/j.media.2017.07.005>.
- [18] P. Pathak, Y. Punetha, Kratika, Identification of skin diseases using convolutional neural network, in: *Soft Computing: Theories and Applications: Proceedings of SoCTA 2020*, 2, Springer, 2021, pp. 171–180, https://doi.org/10.1007/978-981-16-1696-9_16.
- [19] T.-Y. Yang, T.-W. Chien, F.-J. Lai, et al., Web-based skin cancer assessment and classification using machine learning and mobile computerized adaptive testing in a rasch model: development study, *JMIR. Med. Inform.* 10 (3) (2022) e33006, <https://doi.org/10.2196/33006>.
- [20] M.A. Hashmani, S.M. Jameel, S.S.H. Rizvi, S. Shukla, An adaptive federated machine learning-based intelligent system for skin disease detection: a step toward an intelligent dermoscopy device, *Appl. Sci.* 11 (5) (2021) 2145, <https://doi.org/10.3390/app11052145>.
- [21] Y. Liu, et al., A deep learning system for differential diagnosis of skin diseases, *Nat. Med.* 26 (6) (2020) 900–908, <https://doi.org/10.1038/s41591-020-0842-3>.
- [22] E. Göceri, An application for automated diagnosis of facial dermatological diseases, *İzmir Katip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi* 6 (3) (2021) 91–99.
- [23] I. Oztel, G.Yolcu Oztel, V.H. Sahin, Deep learning-based skin diseases classification using smartphones, *Adv. Intell. Syst.* 5 (12) (2023) 2300211, <https://doi.org/10.1002/aisy.202300211>.
- [24] V. R. Pai, S. G. Pai, P. Suhasi, and P. Rekha, “Identification and classification of skin diseases using deep learning techniques,” 2023, [10.21203/rs.3.rs-2628782/v1](https://doi.org/10.21203/rs.3.rs-2628782/v1).
- [25] A. Singh, S. Srinath, V. Arasu, N.K. Thomas, et al., Machine learning on web: skin lesion classification using CNN, in: *2022 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, 2022, pp. 260–265, <https://doi.org/10.1109/ICICT54344.2022.9850506>.
- [26] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, *Sci. Data* 5 (1) (2018) 180161, <https://doi.org/10.1038/sdata.2018.161>.
- [27] E. Göceri, Medical image data augmentation: techniques, comparisons and interpretations, *Artif. Intell. Rev.* 56 (11) (2023) 12561–12605, <https://doi.org/10.1007/s10462-023-10453-z>.
- [28] F. Perez, C. Vasconcelos, S. Avila, E. Valle, Data Augmentation for Skin Lesion Analysis, 2018, https://doi.org/10.1007/978-3-030-01201-4_33.
- [29] D.A. Shoieb, S.M. Youssef, W.M. Aly, et al., Computer-aided model for skin diagnosis using deep learning, *J. Image Graph.* 4 (2) (2016) 122–129, <https://doi.org/10.18178/jig.4.2.122-129>.
- [30] H.K. Latif, M. Aljanabi, Analysing and evaluation of the effectiveness of different filters on segmentation skin tumors images, in: *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 2021 012068.
- [31] K. Shunmuga Priya, V. Selvi, Enhanced skin disease image analysis using hybrid CLAHE-median filter and salient k-means cluster, in: *International Conference on Multi-Strategy Learning Environment*, Springer, 2024, pp. 459–471.
- [32] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” arXiv, arXiv:1905.11946, 2020. Accessed: Jun. 04, 2022. [Online]. Available: <http://arxiv.org/abs/1905.11946>.
- [33] T.J. Brinker, et al., Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task, *Eur. J. Cancer* 113 (2019) 47–54.
- [34] A. Esteva, et al., Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542 (7639) (2017) 115–118.
- [35] X. Du-Harpur, F. Watt, N. Luscombe, M. Lynch, What is AI? Applications of artificial intelligence to dermatology, *Br. J. Dermatol.* 183 (3) (2020) 423–430.
- [36] S. Chan, V. Reddy, B. Myers, Q. Thibodeaux, N. Brownstone, W. Liao, Machine learning in dermatology: current applications, opportunities, and limitations, *Dermatol. Ther.* 10 (2020) 365–386.
- [37] M. Sikorska, A. Skalski, M. Wodzinski, A. Witkowski, G. Pellacani, J. Ludzik, Learning-based local quality assessment of reflectance confocal microscopy images for dermatology applications, *Biocybern. Biomed. Eng.* 41 (3) (2021) 880–890.
- [38] K.H. Keskinbora, E. Kumbasar, Ethical concerns in dermatology and cosmetic applications, *Med. Res. Arch.* 11 (2) (2023).