

SKINLENS: SKIN DISEASE DETECTOR

Zaid Akhtar Mohammad, Zhuowen Yan, Tanzil Mohammed,

Naganjali Pujitha, Bangling Yin, Abhikumar Patel

Seidenberg School of Computer Science and Information Systems

Pace University, New York, NY, USA

Abstract— Skin diseases are common worldwide, but diagnosing them accurately and quickly is still a significant challenge. This paper introduces SkinLens, a web application that helps users diagnose skin conditions on their own, easily and accurately. The platform uses a deep learning image classification model, trained on diverse datasets, to identify 20 types of skin conditions. Users upload a photo of their skin condition, and the app analyzes it using a fine-tuned ResNet-50 model. After analysis, the app provides a detailed diagnosis, a brief description of the condition, and recommended treatments. SkinLens addresses problems like limited access to dermatologists and inefficiencies in traditional diagnosis methods, showing how AI can improve dermatological care.

Keywords— Machine Learning, TensorFlow, Firebase, Convolutional Neural Network (CNN), Image Classification, Dermatological Diagnosis, Skin Disease Detection.

I. INTRODUCTION

Skin conditions are one of the most common health concerns globally, affecting nearly 900 million people each year, according to the World Health Organization (WHO) [1]. These conditions range from minor irritations to serious diseases like melanoma and have significant effects on physical health, mental well-being, and quality of life. Despite their prevalence, diagnosing skin conditions accurately and quickly remains a challenge for many patients, making skin health a critical area for technological innovation.

Diagnosing skin conditions is complex due to the wide variety of symptoms and their similarities across different diseases. Studies show that even experienced dermatologists can struggle to differentiate between visually similar conditions, with diagnostic accuracy averaging 60%-70% for many diseases [2]. Access to dermatologists is also limited, especially in underserved areas, and the time and costs involved in traditional consultations create further barriers. As a result, patients often experience delays or receive inadequate treatment, which can worsen their conditions.

Artificial intelligence (AI) offers a promising way to address these challenges. Research shows that

deep learning models, such as convolutional neural networks (CNNs), can analyze medical images with an accuracy comparable to that of human experts [3]. By using diverse datasets and advanced algorithms, AI can overcome the limitations of traditional methods, delivering faster and more consistent diagnoses. These capabilities position AI as a transformative tool to improve both accessibility and efficiency in dermatological care.

This paper introduces SkinLens, a web-based application that uses AI to overcome the limitations of traditional diagnostic methods. SkinLens allows users to upload images of their skin conditions, which are analyzed by a fine-tuned ResNet-50 deep learning model. The platform provides a diagnosis, a brief description, and treatment recommendations, offering a self-help, fast, and free experience. The study explores key research questions, including: How can AI improve diagnostic accuracy and accessibility? What features are needed to encourage widespread adoption? How can AI tools integrate effectively with existing healthcare systems? By answering these questions, SkinLens aims to improve dermatological care and enhance skin health outcomes worldwide.

II. LITERATURE REVIEW

Artificial intelligence (AI) has significantly advanced dermatology, particularly in diagnosing skin conditions. Deep learning models, especially convolutional neural networks (CNNs), have shown accuracy levels comparable to, and in some cases exceeding, those of experienced dermatologists. For example, a study published in *Nature Medicine* reported that a CNN achieved a sensitivity of 95% and specificity of 91% for detecting melanoma, outperforming human experts in some cases [4].

The success of AI in dermatology relies heavily on access to diverse and comprehensive datasets. Key resources include the DermNet NZ image database,

which offers a wide range of dermatological images covering various conditions and skin types [5]. Another resource, the DermAtlas project, provides an extensive collection of clinical images aimed at improving medical education and research [6]. These datasets are essential for training AI models to recognize a broad spectrum of skin conditions, enhancing their diagnostic capabilities.

AI applications in dermatology have also proven effective in real-world healthcare settings. Mobile apps like SkinVision allow users to assess skin conditions using AI, making early detection of health risks more accessible [7]. These tools help reduce the burden on healthcare systems by providing fast and affordable diagnostic support. Similarly, professional platforms like DermAI integrate AI into clinical workflows, assisting dermatologists in evaluating patients. These applications demonstrate how AI can scale and adapt to address real-world challenges in dermatology.

Despite these advancements, significant challenges remain. One major issue is dataset bias, as many datasets underrepresent darker skin tones and rare conditions, leading to reduced accuracy for these populations [8]. Ethical concerns, such as data privacy and the lack of transparency in AI decision-making, further hinder adoption. Moreover, many existing tools focus solely on diagnosis and do not provide comprehensive solutions, such as treatment recommendations or communication with healthcare professionals. This lack of integration limits the effectiveness of AI-driven tools in real-world scenarios.

This research aims to address these gaps with SkinLens, an AI-powered web application designed to improve dermatological care. By using diverse datasets, SkinLens enhances inclusivity and diagnostic accuracy across a wide range of skin tones and conditions. Unlike existing tools, it goes beyond diagnosis by offering treatment recommendations and enabling real-time communication between patients and dermatologists. This approach bridges the gap between diagnosis and care, ensuring patients receive both accurate information and timely professional guidance. SkinLens sets a new standard for accessibility, inclusivity, and quality in AI-driven dermatology.

III. DATASET

The development of the **SkinLens** model relies on three prominent dermatology datasets: **HAM10000** [9], **SCIN** [10], and **Fitzpatrick17k** [11]. These datasets bring unique strengths and challenges to the training and validation process. Below, we detail each dataset, their contributions, and the preprocessing steps required to address their limitations.

A. HAM10000 Dataset

The HAM10000 (Human Against Machine with 10000 Training Images) dataset contains 10,015 dermoscopic images of pigmented lesions. It includes seven lesion categories, such as melanoma, basal cell carcinoma, and benign keratosis. HAM10000 is one of the most widely used datasets in dermatology for training AI models, providing consistent and well-annotated data.

Strengths:

- **Rich Labeling:** The dataset includes clear annotations, ensuring high-quality training data.
- **Lesion Variety:** Covers both benign and malignant lesions, supporting robust classification.

Challenges:

- **Limited Scope:** Focused primarily on pigmented lesions, which may reduce generalizability for other skin conditions.
- **Skin Tone Bias:** Most images represent lighter skin tones, potentially affecting the model's inclusivity for darker skin tones.

B. SCIN Dataset

The SCIN dataset, developed by Google Health and Stanford Medicine, consists of over 10,000 images of skin, nail, and hair conditions. It emphasizes allergic, inflammatory, and infectious conditions, distinguishing it from other datasets that focus mainly on lesions. Images were collected via crowdsourcing, with contributions from individuals diagnosed with various conditions. It stands out for its use of weighted labels, where multiple conditions may be assigned to a single image with varying confidence levels.

Strengths:

- **Diverse Conditions:** Includes unique conditions like fungal infections, eczema, and psoriasis, complementing datasets that focus on lesions.

- **Dermatologist Validation:** Expert-reviewed labels ensure reliability.

Challenges:

- **Weighted Labels:** Multiple condition annotations create ambiguity. To resolve this, we assigned the highest-confidence label to each image for clarity in model training.
- **Imbalanced Classes:** Certain conditions have significantly fewer images, which could skew model performance.
- **Quality Variability:** The crowdsourced nature of the dataset leads to variations in image resolution and quality.

C. Fitzpatrick17k Dataset

The Fitzpatrick17k dataset includes 16,577 clinical images sourced from the DermaAmin and Atlas Dermatologico resources. It represents 114 skin conditions across Fitzpatrick skin types I–VI, addressing the need for diverse skin tone representation in dermatology datasets.

Strengths:

- **Skin Tone Inclusivity:** Explicitly designed to represent six Fitzpatrick skin types, improving the model's fairness and accuracy across populations.
- **Wide Condition Coverage:** Encompasses both common and rare skin conditions, supporting a comprehensive diagnostic system.

Challenges:

- **Label Consolidation:** Conditions with overlapping labels, such as "Atopic Dermatitis" and "Contact Dermatitis," were grouped under the broader "Eczema" category to simplify classification. This approach reduces complexity but may obscure distinctions between closely related subtypes.
- **Class Imbalance:** Some conditions are underrepresented.

D. Analysis and Integration

The combination of these datasets was designed to leverage their individual strengths while addressing their weaknesses. **HAM10000** provides high-quality data for lesion classification, **SCIN** broadens the scope to include inflammatory and infectious conditions, and **Fitzpatrick17k** ensures diversity in both skin conditions and tones. Through data

augmentation and careful label consolidation, we balanced class representation and minimized biases.

Each condition included in the model has at least 300 images, ensuring sufficient representation for effective training. The integration of these datasets results in a robust foundation for the SkinLens model, capable of providing accurate and inclusive diagnostic results.

IV. METHODS

The SkinLens project is a comprehensive platform that integrates cutting-edge technologies in web development, machine learning, and cloud infrastructure to deliver a robust and user-friendly solution for skin condition analysis. By incorporating an extensive set of completed user stories from the product backlog, the methodology reflects a holistic and feature-complete system designed to cater to patients and dermatologists.

A. User Frontend

The frontend of the SkinLens application is built using React.js, a robust JavaScript library for creating dynamic and interactive user interfaces. Initial designs were prototyped using Figma, emphasizing intuitive navigation and accessibility. Tailwind CSS ensures responsive, visually appealing styling across devices. All code development is managed in VSCode, with collaborative version control via GitHub. The patient interface allows users to register, log in, and securely manage their accounts. Patients can start new cases, upload images, view diagnostic results, download reports, and communicate with dermatologists via real-time chat. They can also access past cases with status tracking and notifications for reviewed reports. The dermatologist interface enables users to log in, access patient reports, review diagnostic results, and provide professional feedback through comments and diagnoses. Additionally, dermatologists can interact with patients via real-time chat, manage active cases, review chat histories, and update case statuses. The modular architecture of React ensures scalability, allowing seamless implementation of additional features in the future.

B. API Gateways

The backend is implemented using Flask, which serves as the intermediary between the frontend and

the machine learning model. The RESTful API architecture abstracts complexities, enabling modularity and scalability. The backend processes uploaded images, which are stored in Google Cloud Storage (GCS) and linked to unique case IDs. It retrieves and preprocesses the images for model inference and facilitates real-time diagnosis by handling communication with the CNN model, ensuring that diagnostic results are promptly delivered to patients. The API is designed to handle errors securely and efficiently, providing user-friendly error messages and retry options for invalid file uploads or failed predictions.

C. Cloud Infrastructure

The application leverages Google Cloud Platform (GCP) to ensure scalability, reliability, and cost efficiency. Firebase services are integrated to manage authentication, storage, and database operations. Firebase Authentication provides secure role-based login for patients and dermatologists, while Firestore Database handles structured data such as user information, case details, chat history, and diagnosis data. Firebase Storage securely stores patient-uploaded images, with an automatic deletion mechanism for temporary files uploaded by guest users. This infrastructure enables real-time updates to notify patients when reports are reviewed and facilitates seamless communication between patients and dermatologists.

D. Machine Learning Model

At the core of SkinLens is a Convolutional Neural Network (CNN) developed using TensorFlow and Keras, designed to classify 20 skin conditions. The training dataset, stored in Google Cloud Storage, combines images from HAM10000, SCIN, and Fitzpatrick17k datasets, ensuring diverse coverage of skin conditions and skin tones. Each condition is represented by at least 300 images, resulting in a dataset of approximately 8,000 labeled images.

Before training, the images are preprocessed using dynamic data augmentation techniques, including resizing, flipping, cropping, and normalization, to enhance the model's robustness and generalization capabilities. The ResNet-50 architecture serves as the backbone of the model. This architecture is well-suited for image classification tasks, as its residual connections effectively address the vanishing

gradient problem, enabling deeper layers to learn complex patterns. [12] The model consists of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.

Training was conducted over 10 epochs using the Adam optimizer [13], which adjusts learning rates dynamically for faster convergence. A categorical cross-entropy loss function was used, given the multi-class nature of the classification problem. The batch size and learning rate were fine-tuned during initial experiments to balance training speed and performance. Iterative retraining was performed to incorporate feedback and improve the model's diagnostic accuracy for diverse skin conditions.

The trained model is integrated with the Flask API, allowing real-time analysis of user-uploaded images. This integration ensures users receive diagnostic results within seconds, enabling a responsive and efficient user experience.

E. Workflow

The workflow of SkinLens involves a systematic process that begins with patients uploading skin condition images. These images are securely stored in Firebase Storage and linked to unique case IDs for tracking. The uploaded images are then processed and analyzed by the CNN model through the Flask API, and the results are returned to the frontend. Patients receive detailed PDF reports that include the diagnosis, a brief description, and treatment recommendations. Dermatologists can review these reports, add comments, and update case statuses in real time. A real-time chat feature further enhances communication, enabling direct interaction between patients and dermatologists.

F. Deployment and Services

The SkinLens application is fully deployed on Google Cloud Platform, ensuring high availability and scalability. The React application serves as the frontend, while the Flask APIs are containerized using Docker and orchestrated with Kubernetes on GCP. Firebase services manage authentication, database operations, and file storage in real time. Continuous Integration and Delivery (CI/CD) pipelines streamline updates and feature rollouts, enabling the efficient deployment of improvements.

This methodology highlights a feature-rich application that fulfills all the user stories in the product backlog. It ensures a seamless, efficient, and secure experience for patients and dermatologists while achieving the project goals of accurate skin condition analysis and improved health outcomes.

V. RESULTS

The SkinLens application demonstrates significant progress in leveraging artificial intelligence for the diagnosis of skin conditions. The core of the system, a Convolutional Neural Network (CNN), achieved a validation accuracy of 96.29% with a validation loss of 0.12 during training, indicating high reliability and precision in recognizing diverse skin conditions. The model was trained on a meticulously curated dataset stored in Google Cloud Storage (GCS), comprising over 8,000 images. The training process utilized a custom GCS data generator, enabling efficient data handling, real-time processing, and scalability for future dataset expansions. Key improvements in model accuracy were achieved through the inclusion of optimized layers, consistent batch normalization, and Adam optimization, ensuring a robust framework for predictions.

The user-centric interface provides seamless functionality, including the ability to upload skin images securely, receive a detailed diagnosis, and access a comprehensive report. Each report includes the diagnosed condition, a user-friendly description, recommended treatments, uploaded images, timestamps, and metadata. These reports are securely stored in Firebase, ensuring privacy and accessibility for authenticated users. For guest users, images are processed temporarily and deleted automatically within a set timeframe, adhering to strict privacy protocols. The application supports the diagnosis of over 20 skin conditions, with predictions displayed in under a minute, addressing the critical need for timely medical insights. To further enhance patient understanding, SkinLens provides concise, readable descriptions and actionable treatment recommendations based on AI results.

The platform also features advanced user management, including patient dashboards for tracking health progress and case statuses, and a dermatologist dashboard for reviewing reports and

providing professional feedback. Notifications keep users updated when reports are reviewed, fostering real-time communication between patients and dermatologists. Additionally, a real-time chat feature allows for immediate consultations, facilitating the exchange of text and images for comprehensive discussions. The model's performance is improved through retraining, ensuring accuracy exceeds industry thresholds for AI diagnosis systems. SkinLens stands out as an innovative, secure, and efficient tool that bridges the gap between technology and dermatological care, offering an enhanced experience for both patients and healthcare professionals. Below figures shows the functionality mentioned

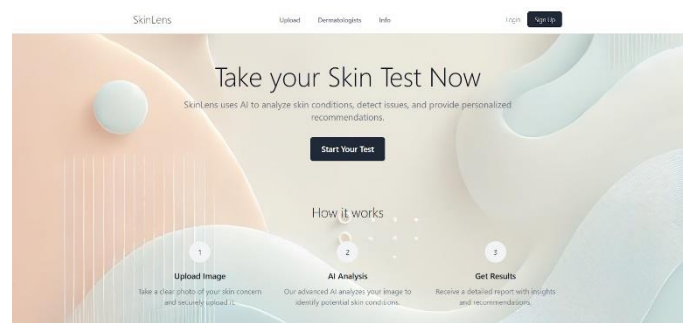


Fig. 1. Home page

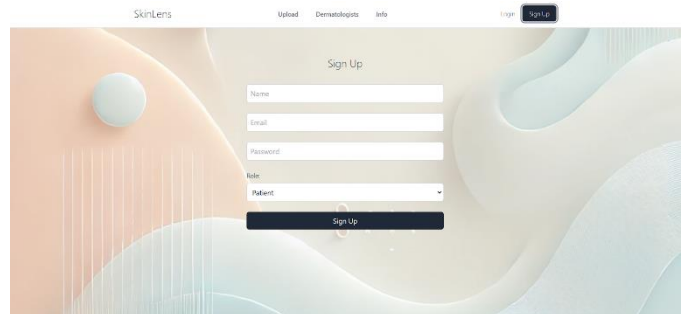


Fig. 2. SignUp Page

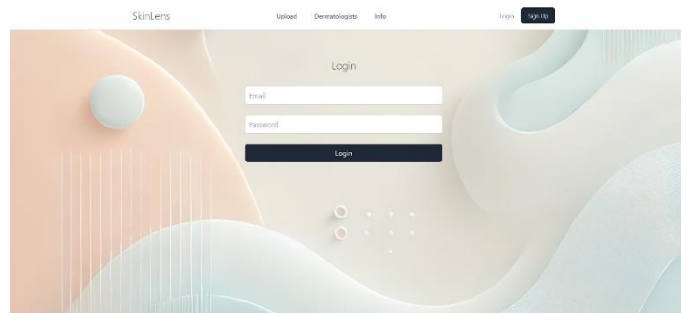


Fig 3. LogIn Page

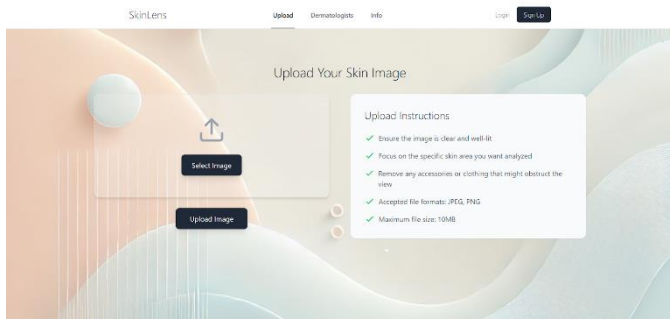


Fig. 4. Upload Image Page

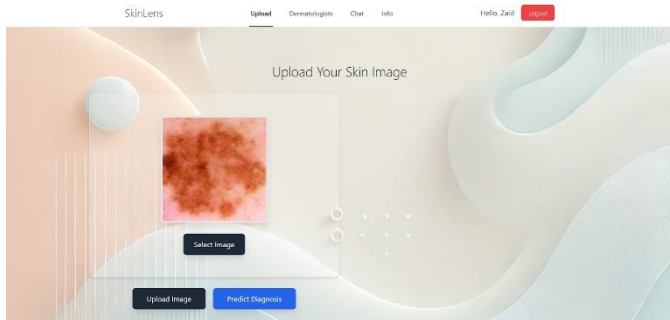


Fig. 5. Preview

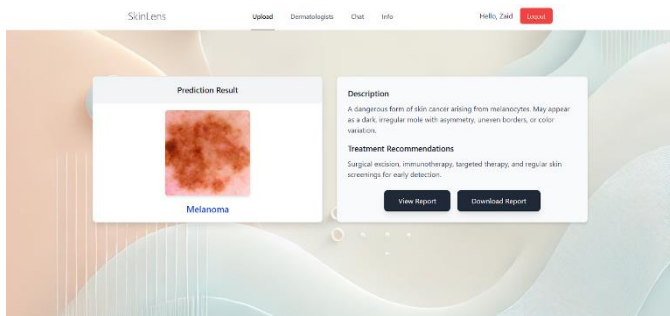


Fig. 6. Description and Recommended Treatments

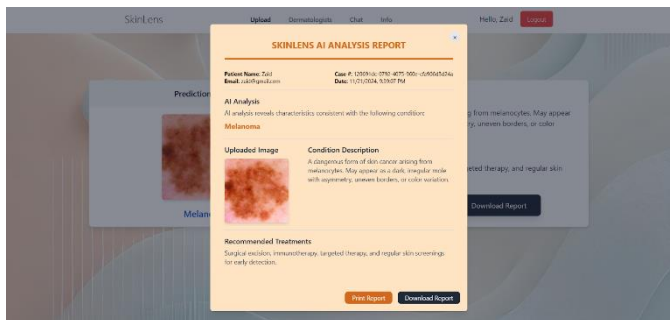


Fig. 7. Report

VI. CONCLUSIONS

SkinLens addresses the pressing challenge of accurate and timely diagnosis of skin conditions, an issue that affects millions globally. By leveraging artificial intelligence, particularly a fine-tuned

Convolutional Neural Network (CNN), SkinLens offers a modern, web-based solution to improve accessibility and efficiency in dermatological care. The model is trained on a carefully curated dataset of 8,000 images across 20 skin conditions, ensuring robust diagnostic capabilities.

The platform empowers users by providing a self-help, fast, and accurate diagnostic experience. It goes beyond traditional methods by delivering detailed condition descriptions, treatment recommendations, and enabling real-time communication with dermatologists. This integrated approach bridges gaps in access to dermatologists, reduces diagnostic delays, and offers a user-friendly alternative to conventional care.

SkinLens exemplifies the transformative potential of AI in healthcare. Future work could focus on expanding the scope of recognized conditions, incorporating additional features such as symptom tracking or continuous learning, and further enhancing model accuracy to address a wider range of skin tones and conditions. By prioritizing accessibility, accuracy, and user experience, SkinLens sets a new benchmark for AI-driven dermatological care and paves the way for improved skin health outcomes worldwide.

REFERENCES

- [1] World Health Organization (WHO), "Global burden of skin diseases and evidence gaps," 2019. [Online]. Available: <https://www.who.int>
- [2] "Accuracy of Skin Cancer Diagnosis by Dermatologists: A Meta-analysis," *JAMA Dermatology*, 2019. DOI: 10.1001/jamadermatol.2019.1379
- [3] Esteva A., Kuprel B., Novoa R.A., et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, 2017. DOI: 10.1038/nature21056
- [4] A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature Medicine*, vol. 23, no. 8, pp. 926–930, 2017, DOI: 10.1038/nature21056.
- [5] DermNet NZ, "DermNet NZ Image Database," [Online]. Available: <https://dermnetnz.org>. [Accessed: 20-Nov-2024].
- [6] DermAtlas, "DermAtlas: Online Medical Image Repository," [Online]. Available: <https://en.wikipedia.org/wiki/DermAtlas>. [Accessed: 20-Nov-2024].
- [7] SkinVision, "SkinVision: AI-powered mobile app for skin condition detection," [Online]. Available: <https://www.skinvision.com>. [Accessed: 20-Nov-2024].
- [8] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019, DOI: 10.1126/science.aax2342.
- [9] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, pp. 1–9, 2018, DOI: 10.1038/sdata.2018.161.

- [10] Google Health Research, "Assessing the accuracy of AI in dermatology with the SCIN dataset," [Online]. Available: <https://health.google>. [Accessed: 20-Nov-2024].
- [11] M. Groh, C. Harris, L. R. Soenksen, et al., "Evaluating deep neural networks trained on clinical images in dermatology with the Fitzpatrick17k dataset," *NPJ Digital Medicine*, vol. 4, no. 1, pp. 1–9, 2021, DOI: 10.1038/s41746-021-00485-8.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, DOI: 10.1109/CVPR.2016.90.
- [13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations (ICLR)*, San Diego, CA, USA, 2015. [Online]. Available: <https://arxiv.org/abs/1412.6980>