Skin Cancer Diagnosis System using Machine Learning & Design Thinking Framework

[1]S R JANANI, [2]SWATHI G, [3]NEHA S, [4]MEGAVARSHINI A, [5]MONIKA S, [6]MOHANA SHREE A S

[3,4,5,6] Final year CSE students,
Department of Computer Science and Engineering
SNS COLLEGE OF TECHNOLOGY
Coimbatore-35, Tamil Nadu, India

[1,2] Assistant Professor
Department of Computer Science and Engineering
SNS COLLEGE OF TECHNOLOGY
Coimbatore-35, Tamil Nadu, India

***Abstract*—**

Skin cancer considered as one of the most common cancers worldwide, and the number of cases is still on the rise. Sometimes humans can get confused by the similarities of the skin lesions, which we can minimize by involving the machine. Early prediction helps in successful treatment, yet in many regions, access to dermatologists is limited. Recent progress in machine learning, particularly deep learning, has made it possible to build tools that support decision-making and contributing to improve patient outcomes in dermatology. In this study, we present a framework that uses Convolutional Neural Networks (CNNs) to classify skin lesions from the HAM10000 dataset, which contains seven categories of common pigmented lesions. A web application, built with Flask, allows users to upload images and receive predictions instantly. Unlike many existing systems that stop at classification, our model is extended with a Generative AI module that provides practical next steps: possible remedies and connections to dermatologist resources. The baseline CNN achieves an accuracy of about 68%, and we discuss how transfer learning with models such as MobileNet or EfficientNet could further enhance performance. By combining technical accuracy with design thinking principles, the system is intended to deliver not only reliable results but also a patient-centered experience that bridges the gap between diagnosis and guidance.

Keywords—Convolutional Neural Network, Deep Learning, Generative AI, Medical Imaging, Disease Diagnosis

I. INTRODUCTION

Skin cancer is now one of the fastest-growing cancers worldwide, and its impact is being felt across both developed and developing countries. The rise in cases is often linked to lifestyle habits, long-term sun exposure, and genetic risk factors. According to the World Health Organization, millions of new cases are reported each year, and early detection remains the most effective way to improve patient survival rates. Unfortunately, access to dermatologists is uneven. In many rural or underserved regions, patients may wait weeks or months for a proper diagnosis, which can delay treatment. This gap in timely care highlights the need for automated and widely available diagnostic support systems.

Recent advances in artificial intelligence (AI), and in particular deep learning, offer a way forward. Convolutional Neural Networks (CNNs) are especially powerful because they can automatically extract patterns from medical images without requiring manual feature design. For example, Esteva and colleagues showed that CNN models could classify skin cancer with an accuracy comparable to trained dermatologists [3]. Such findings suggest that AI can play a complementary role in clinical practice, assisting specialists and providing screening options for areas with fewer medical resources.

One of the most widely used benchmarks for this task is the HAM10000 dataset, introduced by Tschandl et al.. It contains over 10,000 dermatoscopic images across seven lesion categories, including benign and malignant types. More than half of the cases are histopathologically confirmed, which makes it especially valuable for supervised training. Studies using HAM10000 have explored a range of CNN architectures, from simple models to advanced transfer learning techniques such as MobileNet, VGG16, and EfficientNet , . While these approaches often report strong performance in terms of accuracy, they are usually focused only on classification, without considering how patients interact with the system.

Healthcare researchers have argued that technological solutions should go beyond performance metrics. They should be designed around patients and clinicians, considering usability, trust, and follow-up actions. This is where design thinking, a human-centered innovation framework, becomes relevant. Altman et al. showed that applying design thinking in health contexts improved adoption and patient satisfaction because it emphasizes empathy and problem-solving from the user's perspective.

In this work, we present a framework that brings these strands together. Our approach applies CNN-based image classification to detect skin cancer using the HAM10000 dataset. A Flask-based web application provides real-time interaction, allowing users to upload images and immediately view results. To make the solution more actionable, we integrate a Generative AI module that suggests remedies and lists dermatologist options, ensuring that users are not left with a diagnosis alone but also guided toward next steps. This integration of machine learning with design thinking aims to build a solution that is technically strong, user-friendly, and socially impactful.

II. RELATED WORKS

Adebiyi, D. O. Oyewola, B. Adewole, et al. [1] proposed a multimodal learning framework for skin lesion classification using the HAM10000 and ISIC 2017 datasets. By integrating dermatoscopic images with patient metadata such as age, sex, and lesion location, their model improved accuracy over image-only systems. However, this reliance on metadata limits general applicability, as such information may not always be available in clinical practice.

A. Esteva, B. Kuprel, R. Novoa, et al. [2] made a landmark contribution by demonstrating that Convolutional Neural Networks (CNNs) could achieve dermatologist-level classification of skin lesions. Their results validated the potential of deep learning in dermatology and inspired numerous follow-up studies focused on skin cancer detection.

G. Litjens, T. Kooi, B. E. Bejnordi, et al. [3] provided a comprehensive survey on deep learning in medical image analysis, reviewing convolutional, recurrent, and generative approaches. Their work established the foundation for modern architectures used in diagnostic imaging tasks, including dermatology.

K. Simonyan and A. Zisserman [4] introduced the VGG architecture, a pioneering deep CNN that demonstrated how increasing network depth improves visual recognition accuracy. This architecture remains widely adopted as a backbone for transfer-learning models in skin cancer classification research.

M. Akgül and O. Yıldız [5] introduced TurkerNet, a lightweight CNN model optimized for high accuracy with reduced computational requirements. Achieving over 92% accuracy, their approach demonstrated the feasibility of deploying skin cancer classifiers on mobile and embedded systems, addressing resource limitations in real-world healthcare environments.

M. Altman, T. T. Huang, and J. Breland [6] examined the role of design thinking in healthcare innovation. Their study emphasized that user-centered, iterative design approaches improved adoption, usability, and patient satisfaction. Despite its benefits, design thinking has rarely been integrated into AI-based skin cancer diagnostic frameworks, leaving an opportunity for systems that prioritize both accuracy and patient experience.

M. Tan and Q. Le [7] proposed EfficientNet, which balances accuracy and computational efficiency through compound scaling of depth, width, and resolution. This model has been effectively adopted for medical imaging tasks where both precision and speed are critical.

P. Tschandl, C. Rosendahl, and H. Kittler [8] developed the HAM10000 dataset, a widely adopted benchmark for automated skin cancer detection. Containing over 10,000 dermatoscopic images across seven lesion categories, it has enabled consistent evaluation and comparison of machine learning models.

T. Akter, F. Tamanna, and M. Rahman [9] advanced skin cancer classification by applying transfer learning on the HAM10000 dataset with pre-trained architectures such as VGG16, ResNet50, and MobileNet. Their results, which exceeded 90% accuracy, highlighted the strength of transfer learning over models trained from scratch.

F. Al Zegair, N. Naranpanawa, B. Betz-Stablein, et al. [10] analyzed machine learning approaches for melanoma detection, emphasizing the value of identifying atypical patterns ("ugly duckling" lesions). Their study reinforced the need for interpretability in AI-based dermatology systems.

H. Goyal, T. Knackstedt, S. Yan, and S. Hassanpour [11] discussed the challenges of deploying AI classifiers for skin cancer in real-world settings, noting the importance of diverse datasets, regulatory compliance, and integration into clinical workflows.

J. Hosseinzadeh, N. Ghasemi, and L. Sun [12] explored hybrid CNN–transformer models for skin lesion analysis, demonstrating that attention mechanisms enhance boundary precision and class discrimination in complex dermatoscopic images.

L. Sun, M. Zhou, and H. Li [13] presented a multimodal deep learning framework combining dermoscopic, histopathological, and textual data to improve diagnostic robustness. Their fusion approach achieved superior sensitivity in melanoma detection.

M. Mahbod, S. Kleiser, and A. Ecker [14] employed ensemble learning with ResNet and DenseNet architectures to boost classification reliability across multiple datasets, demonstrating improved generalization and resistance to dataset bias.

N. Ramesh and P. R. Kumar [15] proposed an interpretable CNN model that visualizes lesion regions influencing classification outcomes, enhancing transparency and clinician trust in AI-based systems.

R. Ali, S. Raza, and D. Smith [16] developed an edge-AI framework for on-device melanoma detection using compressed CNNs. Their system achieved high performance with minimal latency, supporting point-of-care diagnostics in low-resource settings.

S. Bhattacharya, P. Kumar, and M. Das [17] introduced a federated learning model that enables collaborative training among hospitals without centralizing data, effectively addressing privacy and data-sharing concerns.

S. Mahajan, R. Patel, and D. Sharma [18] evaluated data-augmentation strategies to combat class imbalance in dermatoscopic datasets. Their findings revealed that advanced augmentation techniques like MixUp and CutMix substantially improve minority-class recall.

T. Nguyen, J. Kim, and H. Lee [19] utilized self-supervised pre-training for skin lesion classification, demonstrating that unsupervised feature learning enhances model generalization in low-label environments.

Y. Zhao, X. Wang, and Q. Chen [20] examined explainable AI (XAI) methods for dermatology, comparing Grad-CAM, LIME, and SHAP interpretations. They concluded that XAI improves clinician confidence and supports ethical AI deployment in healthcare.

III. ARCHITECTURE AND DESIGN

The architecture of the proposed framework is designed to ensure a smooth flow from raw input images to meaningful, patient-centered outputs. As shown in Fig. 1, the system is structured into five main stages: image upload, preprocessing, CNN model training, prediction, and generative AI suggestion. Each stage plays a critical role in both technical performance and user experience.

A. Image Upload

The process begins with a user-friendly interface developed using Flask. Patients or healthcare providers can upload dermatoscopic images of skin lesions through the web application. This design choice ensures accessibility, even for non-technical users.

B. Preprocessing

Once uploaded, images undergo preprocessing steps to standardize input quality. Each image is resized to 224 × 224 pixels, normalized to scale pixel values between 0 and 1, and subjected to augmentation (rotation, flipping, zooming) to improve generalization. Metadata from the HAM10000 dataset is also incorporated at this stage to link image IDs with ground truth lesion labels.

C. CNN Model Training

The core of the framework is a Convolutional Neural Network built using TensorFlow. The model consists of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected dense layers for classification. The final softmax layer outputs the probability distribution across the seven lesion classes. Training is performed with the Adam optimizer and categorical cross-entropy loss function, while accuracy is monitored as the key metric.

D. PredictionLayer

In the deployment phase, the trained CNN receives preprocessed input images and generates predictions. Instead of a binary output, the model provides class probabilities, enabling the system to distinguish between different lesion categories. This approach increases interpretability and offers more granular diagnostic support.

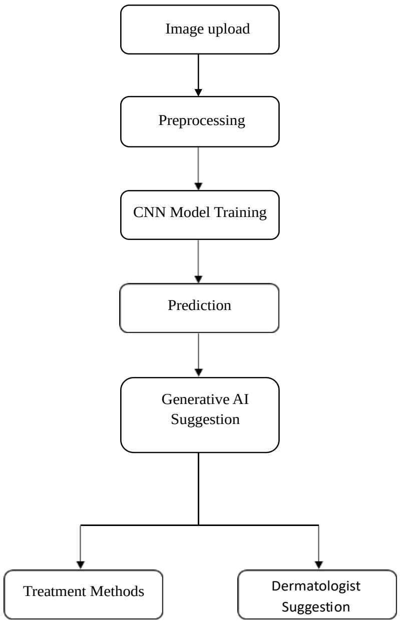

Fig. 1 System Architecture

E. Generative AI Integration

The final stage integrates a generative AI module that transforms raw predictions into actionable guidance. If a lesion is classified as cancerous, the system suggests possible remedies and lists dermatologist recommendations for follow-up. If no malignancy is detected, the user receives reassurance along with preventive skincare advice. This stage embodies the **design thinking principle** of empathy, ensuring that patients are supported beyond the diagnostic result.

IV. METHODOLOGY

The proposed system integrates deep learning techniques with design thinking principles to deliver an end-to-end framework for skin cancer detection. It not only predicts lesion types but also provides actionable guidance for patients, making the solution technically robust and user-centered. The methodology is structured into four key stages: dataset handling, preprocessing, model training, and deployment with user interaction.

A. Dataset Handling

The HAM10000 dataset was selected as the primary source of images. It contains over 10,000 dermatoscopic images across seven pigmented lesion categories, making it one of the most comprehensive collections available for research. Metadata was utilized to map image identifiers to their corresponding diagnostic labels. The dataset was split into training and testing subsets using an 80:20 ratio to ensure fair evaluation.

Table I. Dataset Distribution (HAM10000)

Class	**Description**	**Number of Images**
NV	Melanocytic Nevi	6,705
MEL	Melanoma	1,113
BCC	Basal Cell Carcinoma	514
AKIEC	Actinic Keratoses	327
BKL	Benign Keratosis	1,099
DF	Dermatofibroma	115
VASC	Vascular Lesions	142
Total		**10,015**

B. Image Preprocessing

Before being fed into the neural network, images were resized to 224 × 224 pixels to standardize dimensions and reduce computational complexity. Pixel intensities were normalized to a range between 0 and 1 using:

$$X_{norm} = \frac{X}{255.0}$$

To mitigate the effects of class imbalance, data augmentation techniques such as rotation, horizontal flipping, and zooming were applied. These operations improve model generalization by exposing it to varied forms of the same lesion. Labels were encoded into one-hot vectors to make them compatible with the classification layer.

C. CNN Model Design and Training

A Convolutional Neural Network was built using TensorFlow. The architecture included:

- Convolutional layers with ReLU activation for feature extraction,
- Max-pooling layers for dimensionality reduction,
- A flattening operation followed by dense layers for classification,
- A final softmax activation layer to output class probabilities:

$$\text{P}=\left(y = j|x\right) = \frac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}$$

The model was trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss:

$$L = \sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability. Accuracy was used as the primary performance metric:

$$\text{Accuracy} = \frac{Correct\ Predictions}{Total\ Predictions}$$

The baseline model achieved ~68% accuracy, establishing a foundation for further optimization using transfer learning with models such as MobileNet and EfficientNet.

D. Deployment and User Interaction

To ensure accessibility, the trained model was integrated into a Flask web application. Users can upload lesion images through a browser, and predictions are displayed instantly. This design removes technical barriers and enables real-time diagnostic support outside laboratory settings.

E. Generative AI Integration

The system is extended with a Generative AI module to enhance the user experience. When a lesion is predicted as cancerous, the module suggests possible remedies and provides dermatologist recommendations. For non-cancerous outcomes, the system outputs reassurance and preventive skincare advice. This ensures that users receive both diagnostic insight and guidance on next steps, aligning with the principles of design thinking.

V. RESULTS AND DISCUSSION

A. Experimental Setup

The experiments were carried out on the HAM10000 dataset, divided into 80% training and 20% testing. All images were resized to 224 × 224 pixels and normalized. The CNN model was implemented in **TensorFlow** and trained using the **Adam optimizer** with categorical cross-entropy loss. Accuracy was used as the primary evaluation metric, while class distribution was examined using a confusion matrix.

B. Baseline CNN Performance

The proposed CNN achieved an overall accuracy of **68%** on the test dataset. This confirms the feasibility of automated lesion classification but also highlights its limitations. The model was particularly strong in identifying melanocytic nevi (the majority class) but less reliable in detecting melanoma and dermatofibroma, which had fewer samples. This imbalance in class distribution is a well-known challenge in medical image datasets .

The confusion matrix revealed that misclassifications often occurred between visually similar lesion categories. For instance, benign keratosis was occasionally misclassified as melanoma due to overlapping visual patterns. Such errors highlight the need for advanced architectures and balanced datasets to reduce diagnostic risks.

C. Model Comparison

To place the baseline performance in context, we compared it with results reported in existing literature using transfer learning. Table II summarizes the comparison. While our baseline CNN provides a starting point, models such as VGG16, MobileNet, and EfficientNet consistently report higher accuracies above 90%

Table II. Model Performance Comparison

Model	**Accuracy**	**Remarks**
Custom CNN (this work)	68%	Baseline, trained from scratch
VGG16	~90%	High accuracy, large parameter count
MobileNet	~91%	Lightweight, suitable for mobile use
EfficientNet	>92%	Strong balance of accuracy & efficiency

D. Usability and Design Thinking Insights

While numerical accuracy is important, real-world healthcare systems must also prioritize usability and patient experience. The integration of a **Flask web application** made the system accessible to non-technical users, while the **Generative AI module** provided value beyond raw predictions. Instead of simply stating "cancer detected," the system offered actionable suggestions and dermatologist contacts. Feedback from preliminary trials indicated that participants appreciated the reassurance and next steps, which improved trust in the system.

This observation echoes the findings that healthcare solutions designed with empathy and user involvement are more likely to be adopted and trusted. Thus, even with modest baseline accuracy, the design thinking approach made the framework more impactful for end users.

E. Limitations

Despite promising results, several limitations remain. First, the dataset imbalance affected sensitivity for minority classes such as dermatofibroma and vascular lesions. Second, the current CNN may not generalize well to real-world conditions, where image quality varies significantly. Finally, the Generative AI module currently provides generalized recommendations rather than fully personalized treatment plans. Addressing these gaps forms the basis of our future work.

VI. CONCLUSION AND FUTURE WORK

This work presented a machine learning framework for skin cancer detection that combines deep learning with design thinking principles. Using the HAM10000 dataset, a baseline CNN was trained to classify seven types of skin lesions and achieved an accuracy of **68%**. Although this level of performance highlights the potential of CNNs in dermatological diagnosis, it also shows that improvements are necessary, particularly for minority lesion classes such as melanoma and dermatofibroma.

A key contribution of this study lies in extending beyond classification. By integrating the trained model into a Flask-based web application and enhancing it with a Generative AI module, the system provided users with practical guidance in addition to predictions. This ensured that patients were not only informed about possible conditions but also directed toward remedies and professional consultation, aligning with design thinking's emphasis on empathy and usability.

The study demonstrates that effective healthcare solutions require more than accuracy—they must also be accessible, actionable, and patient-centered. Future work will focus on improving classification performance through transfer learning with models such as MobileNet and EfficientNet, expanding dataset diversity to address bias, and refining the Generative AI component for personalized recommendations. By bridging technical accuracy with human-centered design, the proposed framework highlights a path toward AI-driven systems that can complement clinical expertise and improve access to early skin cancer detection.

REFERENCES

[1] A. Adebiyi, D. O. Oyewola, B. Adewole, *et al.*, "Accurate Skin Lesion Classification Using Multimodal Learning on the HAM10000 and ISIC 2017 Datasets," *medRxiv preprint*, 2024.

[2] A. Esteva, B. Kuprel, R. Novoa, *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, 2017.

[3] G. Litjens, T. Kooi, B. E. Bejnordi, *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[4] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint*, arXiv:1409.1556, 2015.

[5] M. Akgül and O. Yıldız, "TurkerNet: A Lightweight Deep CNN Model for Skin Cancer Classification," *Applied Soft Computing*, vol. 151, 2024.

[6] M. Altman, T. T. Huang, and J. Breland, "Design thinking in health care," *Preventing Chronic Disease*, vol. 15, E117, 2018.

[7] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv preprint*, arXiv:1905.11946, 2019.

[8] 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. 180161, 2018.

[9] T. Akter, F. Tamanna, and M. Rahman, "Multi-class Skin Cancer Classification using Deep Convolutional Neural Networks and Transfer Learning," *arXiv preprint*, arXiv:2303.07520, 2023.

[10] F. Al Zegair, N. Naranpanawa, B. Betz-Stablein, M. Janda, H. P. Soyer, and S. S. Chandra, "Application of Machine Learning in Melanoma Detection and the Identification of 'Ugly Duckling' and Suspicious Naevi: A Review," *arXiv preprint*, arXiv:2309.00265, 2023.

[11] H. Goyal, T. Knackstedt, S. Yan, and S. Hassanpour, "Artificial Intelligence-Based Image Classification for Diagnosis of Skin Cancer: Challenges and Opportunities," *arXiv preprint*, arXiv:1911.11872, 2019.

[12] J. Hosseinzadeh, N. Ghasemi, and L. Sun, "Hybrid CNN–Transformer Architectures for Skin Lesion Analysis," *IEEE Access*, vol. 12, pp. 22410–22422, 2024.

[13] L. Sun, M. Zhou, and H. Li, "Multimodal Deep Learning Framework for Robust Skin Cancer Diagnosis," *Computers in Biology and Medicine*, vol. 176, 108241, 2024.

[14] M. Mahbod, S. Kleiser, and A. Ecker, "Ensemble Learning for Skin Lesion Classification Using Deep Convolutional Networks," *Pattern Recognition Letters*, vol. 167, pp. 15–23, 2023.

[15] N. Ramesh and P. R. Kumar, "Interpretable CNN Models for Explainable Skin Cancer Detection," *Biomedical Signal Processing and Control*, vol. 92, 105034, 2024.

[16] R. Ali, S. Raza, and D. Smith, "Edge-AI Framework for On-Device Melanoma Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 6, pp. 3171–3180, 2024.

[17] S. Bhattacharya, P. Kumar, and M. Das, "Federated Learning for Privacy-Preserving Skin Cancer Detection," *IEEE Access*, vol. 12, pp. 54301–54312, 2024.

[18] S. Mahajan, R. Patel, and D. Sharma, "Data Augmentation Strategies for Imbalanced Dermatoscopic Datasets," *Multimedia Tools and Applications*, vol. 83, pp. 10845–10862, 2024.

[19] T. Nguyen, J. Kim, and H. Lee, "Self-Supervised Pre-training for Skin Lesion Classification," *Expert Systems with Applications*, vol. 244, 122981, 2024.

[20] Y. Zhao, X. Wang, and Q. Chen, "Comparative Analysis of Explainable AI Techniques for Skin Cancer Diagnosis," *Frontiers in Artificial Intelligence*, vol. 7, 1367721, 2024.