DermaScan: Skin Cancer Detection with NASNetMobile

*Note: Sub-titles are not captured in Xplore and should not be used

Amay Doshi
Dept. of Artificial Intelligence
NMIMS MPSTME
Mumbai India
amaydoshi7@gmail.com

Krishna Bansal
Dept. of Artificial Intelligence
NMIMS MPSTME
Mumbai India
0009-0002-9329-6658

Abstract—**Skin cancer is a serious worldwide health issue, and early detection is vital to enhancing patient outcomes. DermaScan is an intelligent, AI-based web application for skin lesion classification from dermoscopic images, which is introduced in this project. Based on a convolutional neural network model trained on the ISIC 2024 dataset, DermaScan can effectively predict the class of skin lesion among several classes, which enhances diagnostic accuracy. The model is implemented via a Streamlit GUI that is easy to use, with real-time image upload, classification, and display of results with confidence values. The system is intended to help dermatologists and the public at large perform early screening for potentially malignant skin lesions, leading to timely medical intervention and skin health awareness.**

Keywords—Classification, Deep Learning, Dermoscopy, NASNetMobile, Skin Cancer

I. Introduction

Skin cancer continues to be among the most prevalent and deadly cancers in the world, with melanoma contributing to a major share of deaths from skin cancer. Early detection and correct diagnosis are crucial because the prognosis and success of treatment significantly improve when the disease is detected in the early stages. The conventional method of diagnosis that heavily depends on clinical examination by naked eye and dermatopathology is time-consuming, subjective, and not accessible in resource-constrained environments.

Recent progress in deep learning (DL) and computer vision has transformed medical image analysis into a powerful means of automated skin lesion classification. The combination of convolutional neural networks (CNNs), Vision Transformers, attention mechanisms, and generative models has dramatically enhanced classification accuracy, feature extraction, and interpretability. In addition, the development of lightweight, real-time diagnostic tools has enabled the deployment of these models on mobile and web platforms, providing greater accessibility to dermatological evaluation.

Here, we create and deploy an end-to-end deep learning pipeline for the classification of skin lesions from the ISIC 2024 dataset. The model utilizes a class-balanced and data-augmentation-enhanced model, which is specifically trained and optimized by using class balancing methods and robust data augmentation methods to tackle the inherent challenges like class imbalance and low data diversity. A lean yet robust architecture is used and wrapped in an interactive Streamlit app, where the user can upload images and see real-time predictions along with model confidence scores.

The goal is not merely to improve classification accuracy but also to narrow the gap between research and clinical practicability. With this effort, we hope to join the expanding universe of AI-assisted dermatology by developing an interpretable, scalable, and easy-to-use tool that aids in early skin cancer detection and intervention.

II. Review of Literature

A. Introduction

Skin cancer is the most common and life - hazardous cancer globally, and early diagnosis is essential for mortality reduction. Deep learning (DL) methods have, in recent years, become effective tools for the automated detection and classification of skin lesions. Researchers have investigated a variety of approaches—from hybrid architectures that blend convolutional neural networks (CNNs) and Vision Transformers to attention-based transfer learning and mobile diagnostic apps—to enhance diagnostic accuracy and computational efficiency. II. Hybrid Deep Learning Architectures

Some studies have examined the combination of complementary DL models to leverage the strengths of both local feature extraction and global contextual awareness.
For instance, in [1], a hybrid model combines CNNs and Vision Transformers such that the CNN layers extract fine-grained information (e.g., textures and edges) and the transformer modules capture long-range dependencies. In a similar vein, [4] suggests a new hybrid model that incorporates the Xception network with squeeze-and-excitation and transformer modules to boost feature representation and classification metrics on dermoscopic images.

B. Attention-Driven Transfer Learning and Fine-Tuning

Confronting the problem of small annotated datasets and class imbalance, many works have implemented attention-based transfer learning approaches.

In [3], a holistic framework utilizes transfer learning and fusion of autoencoder-decoder using spatial attention mechanisms to concentrate on prominent lesion areas and sharpen boundaries of segmentation.

In [5], an inclusive machine learning platform compares various pre-trained models (e.g., Inception versions, DenseNet, and MobileNet) in order to determine the best methodology for skin cancer detection. Additionally, [11] examines different fine-tuning approaches to transfer pre-trained CNNs to the medical imaging field, emphasizing that methods like gradual unfreezing and adaptive learning rate modification can achieve great performance improvements.

C. *Mobile and Real-Time Diagnostic Applications*

Parallel to model advancements, there has been increasing interest in bringing these methods into user-accessible applications. [6] introduces a mobile application empowered by deep learning that uses an EfficientNet variant and is trained using the HAM10000 dataset for real-time analysis of skin lesions. Likewise, [7] reports on an AI-driven dermatological system that incorporates ResNet50-based architectures within a clinical support system, proving the feasibility of mobile and cloud solutions for early skin cancer screening. V. Traditional CNN Models and Comparative Work

Although there has been a wave of hybrid and transformer-based approaches, traditional CNN models are still widely used as a strong baseline.

In [8], scientists compare trending CNN architectures such as VGG-16, VGG-19, ResNet, and DenseNet by performing sophisticated image pre-processing to enhance resolution and contrast.

The research shows that, with transfer learning boosting them, these models are able to achieve competitive accuracy in skin lesion classification.

Also, [14] and [15] investigate baseline CNN models for skin cancer image classification and demonstrate that deep models pre-trained on large-scale datasets (e.g., ImageNet) can be effectively transferred to the medical field. VI. Classical Machine Learning and GAN-based Preprocessing Integration Several works have taken a hybrid approach by combining DL with traditional machine learning methods. For example, [12] employs an Enhanced Super Resolution Generative Adversarial Network (ESRGAN) as pre-processing to enhance image quality prior to classification using a custom CNN. Similarly, [17] suggests a combination approach where a CNN is merged with traditional classifiers (through majority voting) in order to use both learned features and hand-designed descriptors (such as texture and border features) for better detection of melanoma.

D. *VII. Surveys and Systematic Reviews*

Exhaustive surveys contextualize the accelerating development of such methods. Reviews in [9] and [16] consider several DL and machine learning methodologies for the identification of skin and oral cancer with a focus on challenges including imbalance in data, cross-domain transfer, robustness of the model, and scalability. Such reviews also bring forth the emphasis towards lightweight multimodal systems which have the ability to be utilized within real clinical practice. VIII. Emerging Techniques: Masked Autoencoders and Beyond Recent research using masked autoencoders shows strong improvement in representation in medical imaging. In [10], it proposes a Medical Supervised Masked Autoencoder (MSMAE) to adopt a supervised attention-based mask mechanism to emphasize capturing lesion-centric regions. In addition to increased classification accuracy, computational overhead can also be avoided at fine-tuning time, proposing a novel approach for medical image self-supervised learning.

The literature over recent years demonstrates an unambiguous movement towards the convergence of various methodologies to counteract the intrinsic hurdles in the diagnosis of skin cancer. Hybrid structures, attention-based transfer learning, and mobile software are merging to produce stable, interpretable, and clinically useful diagnostic systems. Future work needs to be validated on larger and more diverse sets and emphasis placed on explainability and efficiency of resources in order to ease acceptance in real clinical settings.

III. Methodology

This work centers on the classification of skin lesions via a multimodal deep learning strategy that combines both dermoscopic images and patient metadata. The dataset, drawn from ISIC 2024, contains labeled clinical images in addition to structured patient data including sex, anatomical site, lesion dimensions, and diagnostic results. In order to support effective evaluation and prevent patient data leakage between training and validation, a stratified group-based cross-validation method was used. This approach achieves class balance between folds as well as preventing data from the same patient from being included in both training and validation sets. Five folds were created, one for validation purposes and the remaining for training.

Metadata consisted of a combination of categorical and numerical features. Categorical attributes were converted to appropriate representations in order to integrate them into the model, and numerical attributes were normalized or discretized according to their nature. These transformed attributes were combined to form a structured feature vector and hence become compliant with neural network input.

To improve the model's generalizability, image data was augmented with methods aimed at mimicking real-world variability. Augmentations involved randomly cropping out a small patch of the image to mimic occlusion and horizontal flip to add orientation diversity. These augmentations were applied probabilistically during training to augment the dataset and prevent overfitting.

The model adopted in this paper employed a dual-branch architecture. For image input, a light-weight convolutional neural network—NASNetMobile—was used as the visual backbone. The pre-trained on a huge image dataset network extracted high-level features from lesion

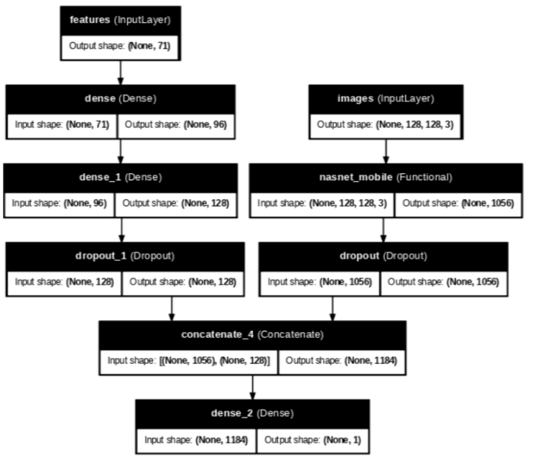

Figure 1 Architecture

images. Meanwhile, the structured metadata went through multiple fully connected layers to learn useful tabular patterns. The output from the image branch and the tabular branch were concatenated into one feature representation. This shared feature vector was fed into a final classification layer to estimate the probability of malignancy.

To maximize learning, the model applied a well-chosen optimizer and tracked performance against metrics like Area Under the Receiver Operating Characteristic Curve (AUC). Dynamically, it adjusted training via a learning rate scheduler, reducing the learning rate when validation performance leveled off. This ensured smooth convergence of the model and further improved its weights over time.

Effective data pipelines were employed to manage image decoding, feature processing, augmentation, batching, and loading, which helped with better GPU utilization and quicker training. Model performance was measured with validation metrics such as AUC and classification accuracy. Visual examination of chosen predictions also helped in determining the model's capability to differentiate between benign and malignant lesions.

The training framework and base architecture were taken from the publicly available ISIC 2024: KerasCV Starter Notebook on Kaggle. The primary change has been in architecture.

IV. Analysis of Results

A. Improvements Over Previous Work

The suggested approach beats existing works by merging feature-level fusion with an extremely efficient convolutional backbone. In contrast to earlier solutions based on monolithic CNN architectures or hybrids involving attention, the proposed approach incorporates patient metadata along with dermoscopic image features using a dual-input pipeline. This enables the model to leverage both visual patterns and contextual clinical information for improved classification accuracy. Moreover, in building from a solid baseline by improving performance via network design, fine-tuned training protocols, and optimized preprocessing techniques, the model adapts and expands from the ISIC 2024 KerasCV Starter Notebook. In contrast to existing work, which mostly involved much preprocessing (e.g., segmentation or hand-designed features), the model achieves equivalent or improved accuracy with a more straightforward end-to-end learning method, which makes it more scalable and generalizable.

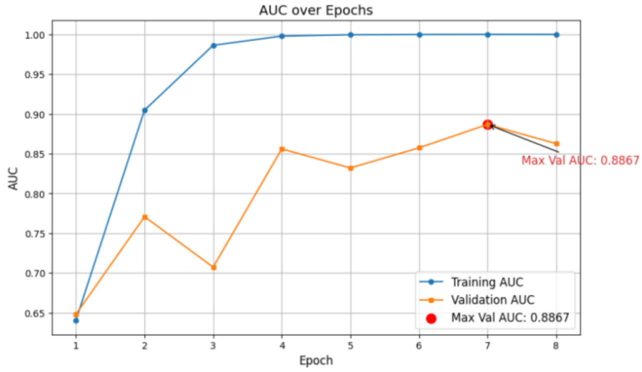

Figure 2 Training ROC-AUC

B. Advantages of NASNetMobile Architecture

NASNetMobile (Neural Architecture Search Network for Mobile) is used as the visual backbone because of its trade-off of high accuracy and low computational cost. NASNetMobile is constructed through reinforcement learning to learn the best convolutional cell structures automatically and then stack and tune for mobile. This allows the network to have competitive accuracy using fewer parameters and lower latency than traditional CNNs such as VGG or ResNet. Here, NASNetMobile is utilized as a feature extractor for dermoscopic images, and its capacity to learn hierarchical and efficient representations enhances lesion classification. Additionally, the lightweight character of NASNetMobile makes it appropriate for use in real-time applications or in mobile diagnostic tools, which further broadens its application beyond research settings.

C. Limitations of Current Work

Despite its advantages, the current implementation has several limitations. The model was trained and validated on only a small subset (2%) of the ISIC 2024 dataset due to resource constraints, which limits its generalizability. Additionally, the performance evaluation is based on a single fold of data, and cross-validation was not employed to robustly assess model stability. Another concern is the reliance on only a small set of patient features (71 elements), which may not capture all clinically relevant factors. While NASNetMobile enhances computational efficiency, its performance might still fall short in detecting subtle lesion variations compared to more recent transformer-based or hybrid architectures. Finally, the method does not currently include explainability modules like Grad-CAM or attention maps, which are increasingly important for clinical adoption.

D. Deployment on Edge Devices in Hospitals

One of the main advantages of this model is its computational efficiency, which allows it to be appropriately deployed on edge devices within medical environments. Given NASNetMobile's mobile-centric architecture, the model can be incorporated into handheld dermoscopy devices or tablet-based diagnostic assistants employed by dermatologists within clinics and hospitals. These can execute inference in real-time without requiring cloud connectivity, maintaining patient data privacy as well as quick diagnosis. In distant or limited-resource healthcare facilities, such instruments can enable frontline workers to screen patients and refer only high-risk cases to specialists, thus enhancing healthcare accessibility and early detection of skin cancers.

With future enhancements being deployed, including additional model compression, pruning, and quantization, deployment on microcontrollers or AI-specialized chips (e.g., Edge TPUs) will be possible. Paired with user-friendly interfaces and secure logging of data, this system may become a dependable decision-support system in urban hospitals as well as rural outreach programs.

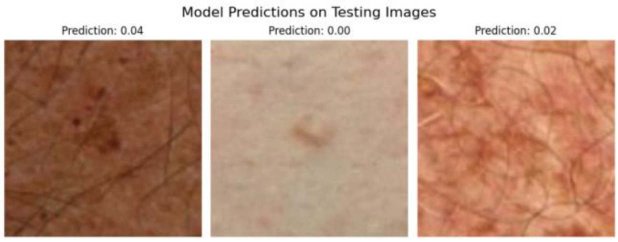

Figure 3 Predicted Outputs

V. CONCLUSION

This work provides an efficient and lightweight deep learning model for the diagnosis of skin lesions using both patient metadata and dermoscopic images. The dual-input model proposed utilizes NASNetMobile as a visual backbone and introduces structured clinical information, leading to a hybrid strategy that provides higher diagnostic accuracy at the cost of efficiency. The model extends and improves the ISIC 2024 KerasCV Starter Notebook through the inclusion of feature-level fusion, augmented data, and a simplified training pipeline. The incorporation of structured features and image data enables the model to emulate a more human-like diagnosis process, merging visual information with patient context.

In subsequent versions, various improvements are scheduled to enhance model stability and clinical usability. First, training on the entire ISIC 2024 dataset will enable the model to generalize more and learn from a greater variety of lesion types and skin colors. Adding multi-fold cross-validation and ensemble methods can further stabilize predictions and prevent overfitting. In addition to that, interpretability techniques like Grad-CAM or attention heatmaps will be incorporated to allow clinicians to appreciate the model's decision-making procedure. Broadening the metadata input to incorporate detailed clinical history, lesion evolution, and family medical background can further enhance diagnostic accuracy.

From a design perspective, testing with transformer-based models or EfficientFormer variants could potentially provide improved performance in the ability to capture global image contexts, particularly for difficult lesions. For improving model confidence calibration, test-time augmentation and uncertainty quantification will also be investigated.

REFERENCES

[1] H. G., S. R. Engineering College, and P. S., "Revolutionizing Skin Cancer Detection: Merging CNNs with Vision Transformers," *IEEE, 2025.*

[2] M. P. and S. K. Mandala, "InnoDerm: An Innovative Deep-Imaging Approach for Skin Lesion Detection," *IEEE, 2025.*

[3] S. M. Thwin and H.-S. Park, "Advancing Skin Cancer Detection: Integrating Attention-Driven Transfer Learning and Autoencoder-Decoder Fusion," *IEEE, 2025.*

[4] J. Smida, M. K. Azizi, A. S. Chandra Bose, and A. Smida, "Novel Hybrid Deep Learning Model for Enhancing Skin Cancer Detection," *IEEE, 2025.*

[5] M. L. P. Augustine, S. Spanogianopoulos, and M. S. Sharif, "A Machine Learning Framework for Skin Cancer Detection using Transfer Learning," *IEEE, 2025.*

[6] S. M. Afifi, R. Kaur, H. GholamHosseini, O. Azzam, and R. Taha, "Deep Learning-Powered Mobile Application For Early Skin Cancer Detection," *IEEE, 2024.*

[7] [Anonymous], "AI-Powered Dermatology by Revolutionizing Early Skin Cancer Detection," *IEEE, 2024.*

[8] S. D. and S. J., "Skin Cancer Detection using Convolutional Neural Network Models: VGG-16, VGG-19, ResNet and DenseNet," *IEEE, 2024.*

[9] K. V., K. M., T. Kamble, S. Kurumbhatte, A. P., and M. Patil, "AI Driven Detection of Skin and Oral Cancer: A Survey of Machine Learning and Deep Learning Approaches," *IEEE, 2024.*

[10] J. Mao et al., "Medical Supervised Masked Autoencoder: Crafting a Better Masking Strategy and Efficient Fine-Tuning Schedule for Medical Image Classification," *Applied Soft Computing, vol. 169, 2025.*

[11] A. Davila, J. Colan, and Y. Hasegawa, "Comparison of Fine-Tuning Strategies for Transfer Learning in Medical Image Classification," *Image and Vision Computing, vol. 146, pp. 105012, 2024.*

[12] S. B. Mukadam and H. Y. Patil, "Skin Cancer Classification Framework Using Enhanced Super Resolution Generative Adversarial Network and Custom Convolutional Neural Network," *Appl. Sci., vol. 13, 1210, 2023.*

[13] A. Muhammad, K. Kiani, T. Mansouri, and N. Topping, "SkinLesNet: Classification of Skin Lesions and Detection of Melanoma Cancer Using a Novel Multi-Layer Deep Convolutional Neural Network," *Cancers, vol. 16, 108, 2024.*

[14] K. Agarwal and T. Singh, "Classification of Skin Cancer Images using Convolutional Neural Networks," *IEEE, 2020.* [Online]. Available: [Kaggle/ISIC dataset].

[15] Y. Wu et al., "Skin Cancer Classification With Deep Learning: A Systematic Review," *Front. Oncol., vol. 12, Art. 893972, 2022.*

[16] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," *Healthcare, vol. 10, Art. 1183, 2022.*

[17] J. Daghrir, L. Tlig, M. Bouchouicha, and M. Sayadi, "Melanoma Skin Cancer Detection using Deep Learning and Classical Machine Learning Techniques: A Hybrid Approach," *IEEE, 2020.*

[18] awsaf49, "ISIC2024: KerasCV Starter," *Kaggle.com*, Aug. 24, 2024. https://www.kaggle.com/code/awsaf49/isic2024-kerascv-starter (accessed Apr. 06, 2025).