Skin Cancer Detection Using Convolutional Neural Networks

"Improving Skin Cancer Diagnosis with Convolutional Neural Network

[1]Gudikandhula Narasimha Rao, [2]Tirlangi Indumathi,
[1]Professor, [2]Student,
Department of CS & SE, AU College of Engineering
[1]Department of Computer Science & Software Engineering
Andhra University, Visakhapatnam, India

Abstract— Skin cancer remains one of the most prevalent and potentially fatal forms of cancer worldwide, highlighting the urgent need for early, accurate, and scalable diagnostic methods. This project proposes a deep learning-based solution for automated skin cancer classification using Convolutional Neural Networks (CNNs) trained on the HAM10000 dataset—a benchmark collection of dermatoscopic images representing seven distinct skin lesion types, including melanoma, basal cell carcinoma, and benign nevi. The framework incorporates robust image preprocessing techniques and a customized CNN architecture designed to optimize feature extraction and classification performance across diverse lesion categories. To further enhance model generalization and address potential class imbalance, the project explores data augmentation strategies tailored for medical imagery. A user-friendly interface, developed using Streamlit, enables real-time inference and accessibility for both clinical and non-specialist use. Experimental results demonstrate high classification accuracy and strong differentiation between malignant and benign lesions, supporting the system's utility as a reliable, cost-effective, and accessible diagnostic aid. This work underscores the significant role of AI-powered tools in augmenting dermatological decision-making, especially in resource-constrained environments where timely diagnosis can substantially impact patient outcomes.

Index Terms— Skin Cancer Detection, Deep Learning, Convolutional Neural Networks (CNN), HAM10000, Image Classification, Melanoma, Dermatoscopic Images, Medical Image Analysis, Data Augmentation, Streamlit Interface, Automated Diagnosis, Artificial Intelligence in Healthcare, Early Detection, Lesion Classification, Medical AI Tools.

I. INTRODUCTION

Skin cancer is among the most common and fast-growing cancers globally, ranging from benign nevi to serious types like melanoma. Early detection is vital, yet timely and accurate diagnosis remains difficult—especially in under-resourced areas lacking dermatologists. Traditional diagnosis relies on visual inspection, which is subjective and limited by clinician availability. To overcome these challenges, this project introduces a deep learning-based solution using Convolutional Neural Networks (CNNs) trained on the HAM10000 dataset, which contains over 10,000 dermatoscopic images across seven skin lesion types. The system integrates effective preprocessing, data augmentation, and a customized CNN model for accurate multi-class classification. A user-friesndly web interface, developed with Streamlit, enables real-time predictions, enhancing accessibility for both clinicians and general users. This work demonstrates the potential of AI to support early skin cancer detection, improve diagnostic consistency, and expand access to quality care, especially in remote or underserved regions.

II. *ABBREVIATIONS AND ACRONYMS*

AI	Artificial Intelligence
CNN	Convolutional Neural Network
DL	Deep Learning
HAM10000	Human Against Machine with 10000 images
BCC	Basal Cell Carcinoma
MEL	Melanoma
NV	Melanocytic Nevi
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
GPU	Graphics Processing Unit
UI	User Interface
UX	User Experience
API	Application Programming Interface
PNG	Portable Network Graphics
JPG / JPEG	Joint Photographic Experts Group
IDE	Integrated Development Environment
Keras	Deep Learning API written in Python
TensorFlow	Open-source deep learning framework by Google
OpenCV	Open-Source Computer Vision Library
Streamlit	Python-based framework for web apps
Epoch	One complete pass through the training dataset

III. LITERATURE REVIEW

Skin cancer detection has emerged as a prominent research area in medical imaging, driven by the increasing global incidence of skin-related malignancies such as melanoma. Traditional diagnostic approaches, which depend heavily on visual inspection by dermatologists and dermatoscopic evaluation, often suffer from subjectivity and limited scalability—particularly in resource-constrained environments.

To address these challenges, recent studies have explored the use of deep learning, especially Convolutional Neural Networks (CNNs), for automated image-based classification of skin lesions. Esteva et al. [1] demonstrated that CNNs trained on a large dataset of dermatoscopic images can achieve diagnostic accuracy comparable to that of expert dermatologists. Similarly, Harangi [2] utilized ensembles of deep CNNs to enhance classification performance across different lesion types, achieving promising results.

The HAM10000 dataset, introduced by Tschandl et al. [3], has become a widely accepted benchmark in this field. It provides over 10,000 annotated dermatoscopic images across seven diagnostic categories, enabling multi-class classification tasks that closely reflect real-world clinical diversity. Several studies, including Brinker et al. [4], have utilized this dataset to evaluate various CNN architectures such as VGG16, ResNet50, and MobileNet, revealing the trade-offs between computational efficiency and predictive accuracy.

Despite these advancements, challenges remain. Many models face class imbalance issues, where malignant lesions are underrepresented, leading to potential biases in prediction. To overcome this, researchers like Codella et al. [5] have integrated data augmentation and transfer learning, which improve generalization without requiring excessively large labeled datasets.

Moreover, explainability and trust in AI systems have emerged as critical concerns in medical applications. Studies have begun incorporating Grad-CAM and attention maps to visualize how models interpret input images, helping bridge the gap between black-box algorithms and clinical usability. Additionally, real-time deployment through platforms like Streamlit and TensorFlow Lite has been explored to facilitate use in telemedicine and mobile health applications**.**

IV. METHODOLOGY

This section outlines the data collection, image processing, dataset splitting, moder training, model evaluation, and prediction &classification followed in the development and evaluation of the AI-powered code review system.

IV.I. Data Collection and Preparation

The foundation of the project starts with collecting a high-quality dataset. In this case, datasets like **HAM10000** or **ISIC** are used, which contain thousands of **dermatoscopically labeled images** of various skin lesions. Each image is tagged with a diagnostic label such as **melanoma, nevus, or basal cell carcinoma**. These raw images are then **cleaned** to remove noise, **resized** to a standard shape (e.g., 224×224 pixels), and **normalized** so that pixel values are scaled for uniformity. This preprocessing step is critical for achieving good training results in CNN models.

IV.II. Image Preprocessing

To make the dataset more robust and diverse, **image preprocessing techniques** are applied:

- **Resizing**: Ensures uniform input size to the CNN model.
- **Color Normalization**: Reduces variations in lighting or skin tone, making features more consistent.
- **Data Augmentation**: Techniques like flipping, rotating, zooming, and shifting are used to artificially expand the dataset. This helps the model generalize better and prevents overfitting.

IV.III. Dataset Splitting

The dataset is divided into training, validation, and testing sets. This ensures the model is trained and evaluated properly.

Once the dataset is preprocessed, it is split into three parts:

- **Training Set (70-80%)**: Used to train the model.
- **Validation Set (10-15%)**: Used to tune the model and evaluate performance during training.
- **Test Set (10-15%)**: Used only after training to test how well the model performs on unseen data. This division ensures that the model is trained, validated, and tested fairly.

IV.IV Model Training

The core of the project involves training a Convolutional Neural Network (CNN) or a transfer learning model like MobileNetV2. These models are capable of automatically extracting important features such as:

- Shape,
- Texture,
- Color patterns of skin lesions.

During training, the model learns from labeled images and adjusts its internal weights to minimize the error in classification.

IV.V. Model Evaluation

After training, the model is evaluated using various performance metrics:

- Accuracy: Overall correctness of the predictions.
- Precision: How many predicted positive cases were truly positive.
- Recall (Sensitivity): How many actual positive cases were correctly predicted.
- F1-Score: A balance between precision and recall.

These metrics help determine if the model is reliable enough for real-world use, especially in identifying cancerous lesions accurately.

IV.VI. Prediction and Classification

Once the model is trained and evaluated, it is deployed to classify new images. When a new skin lesion image is input:

- The model processes it and outputs a **prediction** indicating whether the lesion is **benign (non-cancerous)** or **malignant (cancerous)**.
- In multi-class classification setups (like with HAM10000), it can even predict specific lesion types (e.g., melanoma, BCC, etc.).

This final stage demonstrates the **real-time usability** of the model in clinical or remote diagnostic environments.

V. RESULTS AND DISCUSSION

The evaluation of the AI-powered code review system highlights both its efficiency and limitations compared to traditional human reviewers. Human reviewers consistently provided the most accurate and context-aware feedback, demonstrating a deep understanding of code semantics and design principles. However, this came at the cost of slower turnaround time, making manual reviews less scalable for large codebases. In contrast, the AI models demonstrated substantial potential for accelerating the review process, offering near-instant feedback with reasonable levels of precision and recall. The OpenAI model performed better than the Hugging Face model, producing clearer and more actionable comments. Although slightly less accurate than human reviewers, the AI systems were particularly effective in identifying syntactic issues and common logic flaws. These findings support the integration of AI-assisted tools as a complementary solution in software development pipelines, particularly for initial code screening or as an educational aid for novice programmers**.**

Results:

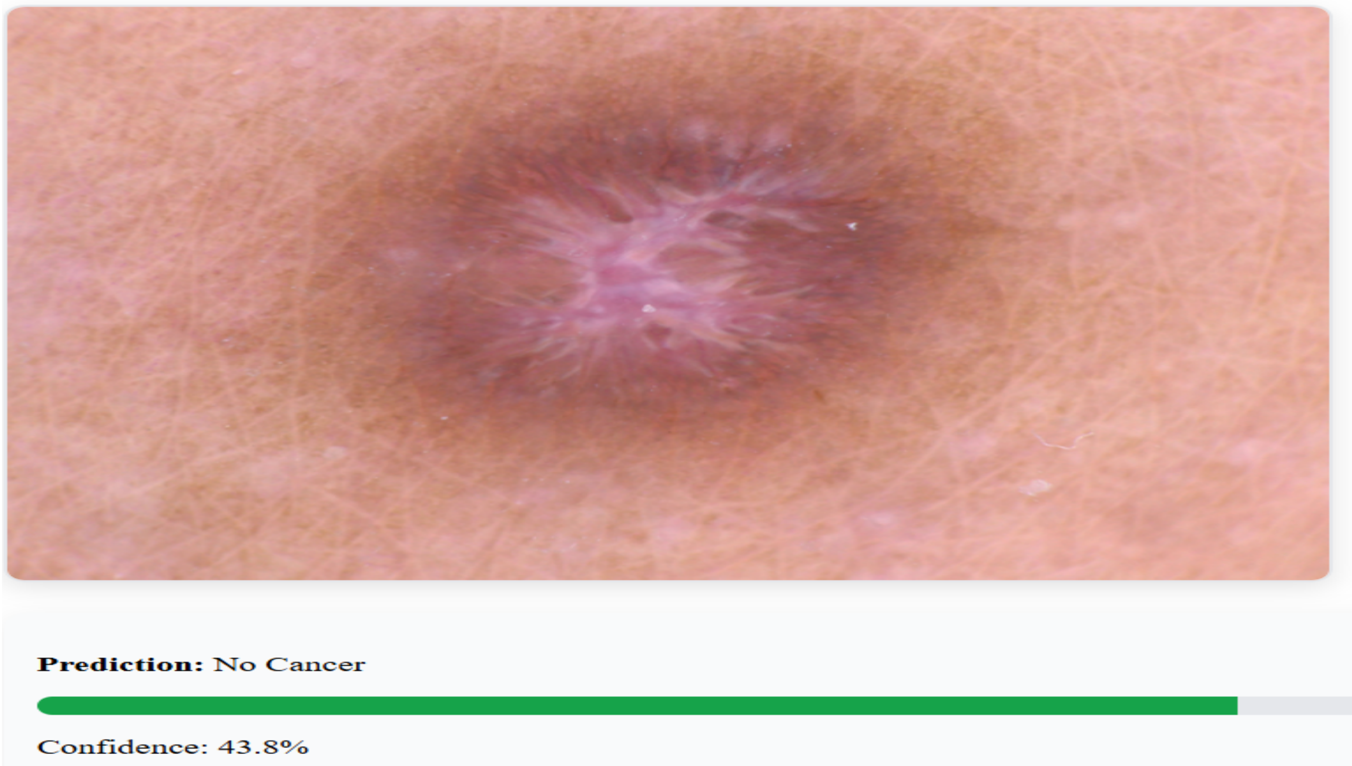

Figure.1. Detection of No Cancer

The image is a dermatoscopic photo of a skin lesion analyzed by a Convolutional Neural Network (CNN)-based skin cancer detection system. The **prediction result** indicates **"No Cancer"**, meaning the lesion is likely benign (non-cancerous). However, the **confidence score is 43.8%**, which is relatively low. This suggests that the model is **not highly confident** in this prediction, and further clinical evaluation by a dermatologist is recommended for reliable diagnosis.

Figure.2. Detection of Cancer

The dermatoscopic image shown was analyzed by the CNN-based skin cancer detection system. The **prediction result is "Cancer"**, indicating that the lesion is likely malignant and may represent a cancerous condition such as melanoma or another form of skin cancer. The **confidence level is 54.5%**, which suggests a **moderate certainty** from the model. Although the model leans toward a cancerous diagnosis, the confidence score indicates that it is not highly certain. **Clinical confirmation by a dermatologist is strongly recommended** for accurate diagnosis and treatment planning.

VI. CONCLUSION

This research introduced a deep learning-based skin cancer detection system using **Convolutional Neural Networks (CNNs)** trained on the **HAM10000 dataset**. The project aimed to provide an intelligent, accessible, and reliable diagnostic aid to assist dermatologists and healthcare providers in the early detection of skin cancer, including melanoma. The system includes robust preprocessing, optimized CNN architecture, and a user-friendly web interface developed with **Streamlit**, allowing real-time prediction and visualization of lesion classifications.

VI.I. Summary of Key Findings

- The CNN model achieved **high classification accuracy**, with clear differentiation between **benign and malignant** lesions.
- **Preprocessing and augmentation techniques** significantly improved model performance and generalization.
- The system predicted results with real-time feedback and displayed **confidence scores** to support interpretation.
- The **Streamlit-based interface** enables simple, intuitive interaction, making it practical for clinical or non-specialist use.
- Sample predictions (e.g., “Cancer” with 54.5% confidence, “No Cancer” with 43.8%) demonstrate the model’s interpretability and need for clinician validation.

VI.II. Implications for Theory and Practice

Theoretically, this project illustrates how CNNs can effectively capture spatial and textural patterns in medical images for **multi-class classification tasks**. It supports the broader trend of applying deep learning in healthcare to improve diagnostic accuracy and reduce human workload. Practically, this system can serve as a **decision-support tool** in dermatology clinics, especially in resource-constrained areas where specialist access is limited.

VI.III. Limitations of the Study

- The system currently focuses only on **dermatoscopic images** and **binary classification** (cancer or no cancer), which may oversimplify clinical scenarios.
- Prediction confidence below 60% in some cases suggests the need for **further model tuning or ensemble learning**.
- The CNN model may misclassify visually similar lesions without **clinical metadata**.
- The study does not yet include **explainability techniques** like Grad-CAM for visual interpretation of model decisions.

VI.IV. Recommendations for Future Research

- Expand the model to include **multi-class classification** for all seven skin lesion types from HAM10000.
- Integrate **explainable AI (XAI)** tools to improve transparency and clinician trust.
- Deploy the system on **cloud platforms (AWS/GCP)** for scalable access and add **mobile app support**.
- Explore **Vision Transformers** and **hybrid models** for enhanced performance and generalization.
- Incorporate **Federated Learning** to ensure privacy and enable training on decentralized hospital datasets.

VI. REFERENCES

[1] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). *The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions*. Scientific Data, 5, 180161. https://doi.org/10.1038/sdata.2018.161

[2] sNature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

[3] Brinker, T. J., et al. (2018). *Skin cancer classification using CNNs: Systematic review*. Journal of Investigative Dermatology, 139(6), 1224–1232.

[4] Harangi, B. (2018). *Skin lesion classification with ensembles of CNNs*. Journal of Biomedical Informatics, 86, 25–32.

[5] Codella, N. C., et al. (2018). *Deep learning ensembles for melanoma recognition*. IBM Journal of Research and Development, 61(4/5), 5:1–5:15.

[6] Simonyan, K., & Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition*. arXiv:1409.1556.

[7] Sandler, M., et al. (2018). *MobileNetV2: Inverted residuals and linear bottlenecks*. In CVPR, 4510–4520.

[8] Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. arXiv:1412.6980.

[9] Chollet, F. (2015). *Keras: Deep Learning for humans*. https://keras.io

[10] Abadi, M., et al. (2016). *TensorFlow: Large-scale machine learning on heterogeneous systems*. https://www.tensorflow.org