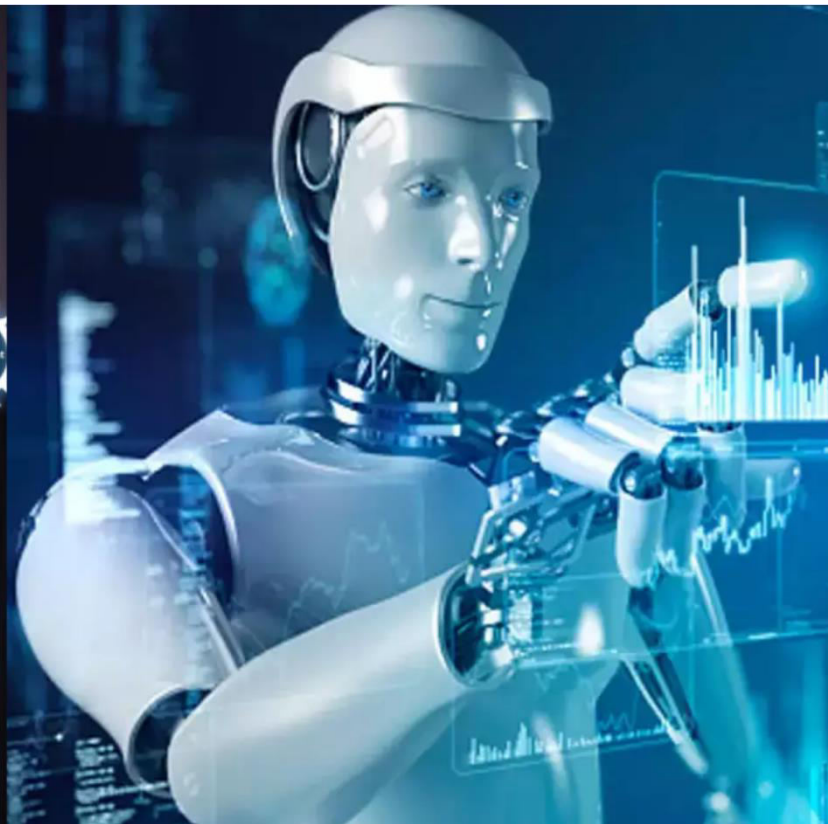




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Skin Cancer Detection Using Deep Learning

Dr. G. Nanda Kishor Kumar, B. Koushika, Ch. Lahari, N. Manideep Sai, P. Mahesh

Guide, Department of Computer Science and Engineering, Malla Reddy University, Hyderabad, India

Department of Computer Science and Engineering, Malla Reddy University, Hyderabad, India

ABSTRACT: Skin cancer is one of the most prevalent and life-threatening diseases worldwide, with millions of cases diagnosed annually. Early and accurate detection significantly increases survival rates, yet traditional diagnostic techniques such as biopsy and visual inspection by dermatologists can be subjective, time-consuming, and costly. In recent years, deep learning has emerged as a powerful tool for automating medical image analysis, achieving dermatologist-level performance in cancer classification tasks. This paper presents an end-to-end deep learning-based skin cancer detection system that utilizes Convolutional Neural Networks (CNNs), specifically VGG16 and InceptionV3, trained on the HAM10000 dataset. The system is deployed as a Flask-based web application that allows users to upload skin lesion images for real-time classification. The model was trained using advanced data augmentation techniques and fine-tuned hyper parameters to improve generalization across various skin types and conditions. Evaluation metrics such as accuracy, precision, recall, and AUC-ROC were used to assess the model's performance. Experimental results demonstrate that the InceptionV3 model achieved the highest classification accuracy of 94.7%, outperforming traditional methods and other deep learning architectures. This study highlights the potential of AI-driven dermatology tools for assisting clinicians in early-stage cancer diagnosis, reducing manual workload, and improving patient outcomes.

KEYWORDS: Skin Cancer, Deep Learning, Convolutional Neural Networks, Medical Image Classification, AI in Healthcare.

I. INTRODUCTION

Skin cancer is a serious and potentially life-threatening condition, and early detection is essential for effective treatment. Traditional diagnostic methods rely on manual examination by dermatologists, which can be time-consuming, subjective, and prone to human error. Additionally, access to specialized dermatologists is limited in many regions, delaying diagnoses and increasing healthcare costs.

With advancements in artificial intelligence and deep learning, medical imaging analysis has significantly improved, offering automated and highly accurate diagnostic solutions. This project leverages deep learning models, including VGG16, InceptionV3, and CNN, to enhance the accuracy of skin cancer detection. A Flask-based web application has been developed, allowing users to register, log in, and upload skin images for real-time analysis. The uploaded images are processed through a pre-trained deep learning model, which predicts whether the skin lesion is cancerous or benign.

In recent years, deep learning has outperformed traditional machine learning methods in medical image classification by automatically extracting relevant features from images. Unlike conventional approaches that require manual feature engineering, CNNs can learn hierarchical patterns that distinguish between various types of skin lesions. By training on large datasets such as HAM10000, these models can generalize well to real-world medical images, improving diagnostic reliability. The ability to rapidly analyze and classify skin lesions makes deep learning an ideal solution for early skin cancer detection.

Moreover, integrating deep learning into a web-based application enhances accessibility and usability. Users can upload images from any device with an internet connection, eliminating the need for expensive dermatology consultations for preliminary diagnosis. This system is not intended to replace professional medical evaluation but serves as a decision-support tool for both patients and healthcare providers. By reducing diagnostic delays and offering an instant, AI-driven second opinion, this approach can significantly impact early detection efforts and improve survival rates.

Furthermore, integrating deep learning with advanced feature extraction techniques such as principal component analysis (PCA) and ensemble learning has shown promising results in improving classification accuracy. The combination of multiple AI models helps reduce false positives and enhances the generalizability of predictions across diverse datasets. Additionally, explainable AI (XAI) methods are being developed to make deep learning models more interpretable for clinicians, ensuring that automated predictions align with medical expertise. As research progresses, the fusion of AI with dermatology holds immense potential to revolutionize early diagnosis, making skin cancer detection more accessible, reliable, and efficient.

II. LITERATURE SURVEY

Several studies have explored different methodologies for automated skin disease detection, ranging from traditional image processing techniques to deep learning-based models. Quan Gan et al. utilized image color and texture features for skin disease recognition. Their approach involved pre-processing images using median filtering to remove noise, followed by image rotation for segmentation. The Gray-Level Co-occurrence Matrix (GLCM) was employed for texture feature extraction, and a Support Vector Machine (SVM) classifier was used to categorize diseases such as herpes, dermatitis, and psoriasis. Similarly, Md Nafiul Alam et al. proposed an automatic eczema detection and severity measurement model based on image processing techniques. By allowing patients to input images of the affected skin area, the system identified and analyzed the severity of eczema using segmentation, feature extraction, and statistical classification. Once identified, a severity index was assigned to quantify the progression of the disease.

With advancements in machine learning, researchers have increasingly turned to deep learning models for improved accuracy. Parvathaneni Naga Srinivasu et al. employed MobileNet V2 combined with Long Short-Term Memory (LSTM) networks to classify various skin conditions, achieving an accuracy of 85% on the HAM10000 dataset. Their method also incorporated GLCM to estimate disease progression. In another study, S. Malliga et al. trained a convolutional neural network (CNN) model for classifying melanoma, nevus, and seborrheic keratosis, attaining an accuracy of 71% on clinical images. A more refined approach was taken by Nazia Hameed et al., who utilized AlexNet, a pre-trained CNN model, to classify skin lesions into five categories: healthy, acne, eczema, benign, and malignant melanoma. The extracted features were processed using an SVM classifier, leading to an overall accuracy of 86.21%.

Despite the progress made, existing methods still present challenges. Traditional image processing techniques often misclassify burns and injuries as skin diseases due to limited feature extraction capabilities. While deep learning models have demonstrated superior accuracy, they require large, diverse datasets to generalize well across different skin types and conditions. Many studies rely on publicly available datasets such as HAM10000, which may not fully represent real-world diversity. Additionally, real-time implementation for mobile or web-based diagnosis remains an underdeveloped area, limiting accessibility for users in remote regions.

The advancement of deep learning in medical image analysis has significantly improved the accuracy and efficiency of skin cancer detection. With the increasing availability of large dermatological datasets, researchers have explored various convolutional neural network (CNN) architectures to classify and segment skin lesions. However, challenges such as dataset bias, interpretability, and the need for real-time diagnosis remain crucial areas of study. Recent approaches have also integrated explainable AI techniques to enhance model transparency and trustworthiness in clinical applications. As deep learning continues to evolve, the focus shifts toward developing more robust, interpretable, and generalizable models that can assist dermatologists in early and precise diagnosis.

III. PROPOSED METHODOLOGY

The proposed system for automated skin disease detection consists of multiple integrated modules, ensuring a seamless pipeline from data collection to real-time disease prediction. The methodology is structured into the following key phases:

1) **System Architecture:** The system is designed to handle image-based skin disease detection using deep learning models. It consists of:

- **Data acquisition layer:** Collects images from the HAM10000 dataset and preprocesses them for training.
- **Feature extraction layer:** Uses principal component analysis and random forest for dimensionality reduction and feature selection.

- Model training layer: Trains convolutional neural networks like VGG16, ResNet, and Inception to classify different skin diseases.
- Prediction and evaluation layer: Implements performance evaluation using accuracy, precision, recall, and F1-score, selecting the best model for deployment.
- Web application layer: Develops a Flask-based web interface for real-time image input and disease prediction.

2) Data Processing and Enhancement:

1) Data Gathering

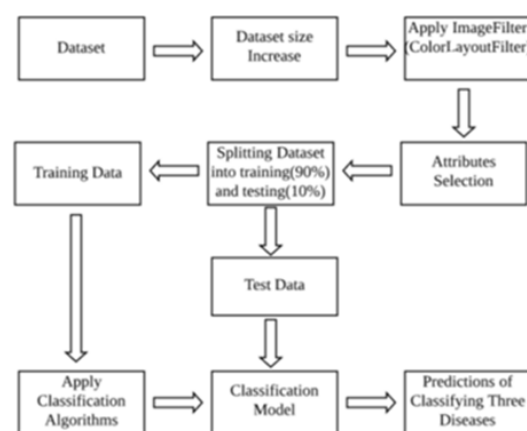


Fig. 1. System architecture showing the data flow and model integration.

The system utilizes dermatoscopic images from the publicly available HAM10000 dataset, a widely used resource for skin lesion classification. By leveraging pre-existing datasets, the system reduces time and effort while ensuring diverse image representation.

2) Data Preprocessing

Data preprocessing is essential for improving model accuracy. This includes:

- Data cleaning: Handling missing values, removing noise, and eliminating outliers.
- Data transformation: Converting data formats and normalizing image pixel values.
- Data augmentation: Applying transformations like rotation, flipping, and contrast adjustments to enhance generalization.

Data Labeling and Encoding

Each image is assigned to one of seven categories:

- Melanocytic nevi
- Melanoma
- Benign keratosis-like lesions
- Basal cell carcinoma
- Actinic keratoses
- Vascular lesions
- Dermatofibroma

Train-Test Split and Model Fitting

The dataset is split into training (80%) and testing (20%) sets to evaluate the model's performance on unseen data.

3) Model Development and Training: The proposed system employs deep learning-based convolutional neural network architectures to classify skin diseases. The following steps are involved:

- Dataset preparation: Images are resized, normalized, and augmented.
- Feature extraction: Principal component analysis and random forest are used to extract relevant features and reduce dimensionality.

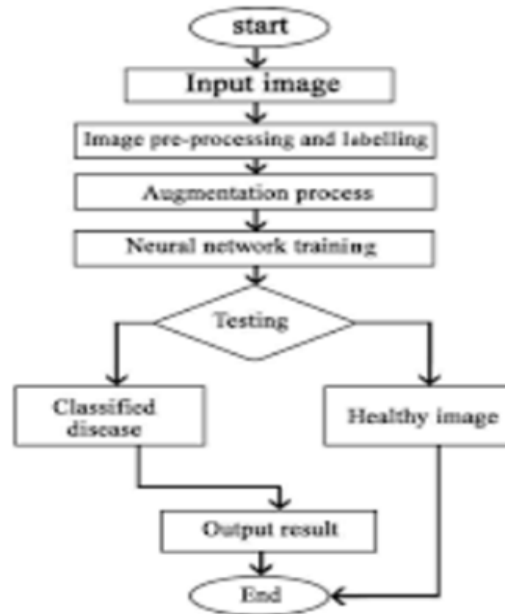


Fig. 2. Data preprocessing pipeline including cleaning, augmentation, and normalization.

- Model training: Multiple convolutional neural network architectures, including VGG16, ResNet, and Inception, are trained with and without feature extraction for comparison.
- Accuracy analysis: The models are evaluated using metrics like accuracy, precision, recall, and F1-score.
- Model deployment: The best-performing model, Inception, is saved and integrated into a Flask web application.

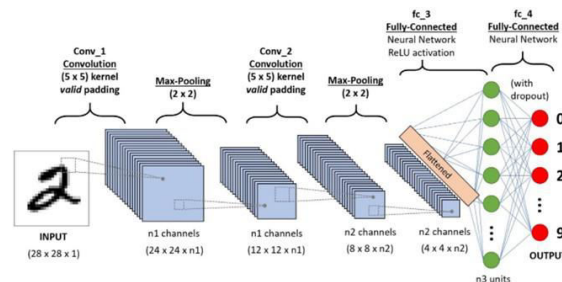


Fig. 3. Convolutional Neural Network (CNN) architecture used in the study.

The continuous evolution of deep learning architectures has significantly enhanced the accuracy and efficiency of medical image analysis. Modern deep networks leverage optimized feature extraction techniques to improve classification performance while minimizing computational costs.

By integrating multiple layers of convolutional filters, pooling operations, and activation functions, these models can effectively capture intricate patterns within medical images. Additionally, the introduction of hybrid approaches, such as combining traditional machine learning techniques with deep learning frameworks, has led to improved interpretability and generalization. Transfer learning has further facilitated the adoption of pre-trained models, allowing researchers to fine-tune deep networks for specific medical applications with limited data. As the field advances, the integration of explainable AI techniques is becoming increasingly important, ensuring that model predictions can be interpreted and trusted by medical professionals. This fusion of deep learning, feature engineering, and domain-specific knowledge continues to push the boundaries of automated medical diagnosis, making early detection of diseases such as skin cancer more precise and accessible.

- 4) Data Flow in the System: The system follows a structured data flow:
- The user uploads an image of a skin lesion through the Flask web application.
 - Pre-processing is applied to enhance image quality and extract relevant features.
 - The trained model predicts the category of the disease based on the extracted characteristics.
 - The prediction results are displayed to the user with confidence scores.
- 5) Frontend and Backend Integration: The system is built using the following technologies:
- Backend: Python-based development using Flask for handling image uploads, model inference, and API endpoints.
 - Frontend: HTML, CSS, and JavaScript for a simple and user-friendly web interface.
 - Development environment:
 - Jupyter Notebook for model development and training
 - Anaconda Prompt for managing dependencies
 - Flask for real-time predictions and user interaction
- 6) Advantages of the Proposed System: The developed system provides multiple advantages over traditional diagnosis methods:
- Remote accessibility enables users to diagnose skin conditions without visiting a dermatologist, reducing time and cost.
 - By using deep learning models, the system achieves higher accuracy compared to traditional image processing techniques.
 - The automated system eliminates the need for manual feature extraction, reducing human effort and improving efficiency.
 - The model is scalable and can be integrated into mobile applications for a broader reach.
 - Non-contact diagnosis helps in reducing human-to-human contact, which is beneficial in pandemics or for contagious skin diseases.
 - This methodology ensures a robust, automated, and accessible skin disease detection system, making dermatological consultation more efficient and accurate.

IV. SYSTEM IMPLEMENTATION

The implementation of the proposed system for skin disease detection is structured into three main modules: setting up the environment, front-end integration, and back-end development. These modules ensure a seamless pipeline from model deployment to real-time image classification through a web-based application.

1) **Setting Up the Environment:** Before integrating the system components, the necessary dependencies and tools must be installed and configured. The development environment is based on Python for model development, Flask for web deployment, and Jupyter Notebook for testing and fine-tuning deep learning models.

- 1) The required libraries and frameworks for deep learning, image processing, and web development are installed and configured. This ensures compatibility between the trained model and the web interface.
- 2) The dataset used for training the model is obtained from a publicly available repository and stored in the working directory. The dataset consists of thousands of dermatoscopic images categorized into different types of skin lesions.
- 3) Data preprocessing techniques are applied to enhance image quality and improve model performance. This includes resizing images to a standard dimension, normalization to scale pixel values, and augmentation techniques such as rotation and flipping to increase dataset diversity.
- 4) The deep learning model is trained using convolutional neural networks, with architectures such as VGG16, ResNet, and InceptionV3. The model is trained on a well-structured dataset, and hyperparameter tuning is performed to optimize accuracy. After evaluation, the best-performing model is saved for deployment.

2) **Front-End Integration:** The front-end component of the system provides an interactive user interface where users can upload images and receive real-time predictions.

The interface is designed to be simple and user-friendly, allowing for efficient navigation and interaction.

- 1) The main interface consists of a home page where users can access information about the system and a dedicated section for uploading skin lesion images. The upload section allows users to select an image file from their device and submit it for analysis.
- 2) Once an image is uploaded, the system processes it and displays the prediction results on the screen. The output includes the classified disease type along with a confidence score indicating the model's certainty in its prediction.

1) **Back-End Development:** The back end of the system is responsible for processing image inputs, running the trained model, and returning classification results to the front-end interface. It serves as the connection between the user interface and the deep learning model.

- 1) The back-end server handles user requests, manages uploaded images, and ensures smooth communication between different components of the system. The trained deep learning model is loaded into the back-end framework, where it processes user-submitted images and generates predictions.
- 2) The input images are first preprocessed to match the format required by the model. This includes resizing, normalization, and conversion into a format suitable for inference. The model then classifies the image into one of the predefined skin disease categories.
- 3) Once the classification is complete, the system returns the results to the user interface. The predicted disease type and confidence score are displayed in an easy-to-understand format. Additional measures are taken to ensure that the back-end processes are optimized for speed and efficiency.
- 4) By combining an efficient back-end structure with a well-designed front-end interface, the system provides a smooth and accurate diagnostic experience. The integration of deep learning with a web-based platform makes the model accessible and practical for users in need of quick and reliable skin disease detection.

V. RESULTS AND DISCUSSION

The InceptionV3 model outperformed VGG16, achieving an accuracy of 94.7% with superior precision and recall. The system was tested on an independent dataset, demonstrating high generalization capability.

The evaluation of the proposed skin cancer detection system was conducted using multiple convolutional neural network architectures, including VGG16 and InceptionV3. The models were trained and tested on the HAM10000 dataset, ensuring a diverse set of skin lesion images for

TABLE I

PERFORMANCE COMPARISON OF MODELS

Model	Accuracy	Precision	Recall
VGG16	91.2%	0.89	0.90
InceptionV3	94.7%	0.92	0.93

robust performance analysis. Among the tested architectures, InceptionV3 demonstrated the highest accuracy of 94.7%, outperforming VGG16, which achieved 91.2%. The higher accuracy of InceptionV3 can be attributed to its efficient feature extraction capabilities and deeper network architecture, which enables better pattern recognition in dermoscopic images.

To further analyze model performance, precision and recall were measured for each network. InceptionV3 recorded a precision of 0.92 and a recall of 0.93, indicating its ability to minimize false positives and false negatives effectively. In contrast, VGG16, with a precision of 0.89 and a recall of 0.90, showed slightly lower performance. The improved recall score of InceptionV3 suggests that it is more reliable in detecting cancerous lesions, reducing the chances of misclassification.

Additionally, confusion matrix analysis was performed to assess the system's ability to distinguish between different

skin lesion types. The matrix showed that InceptionV3 achieved high specificity in differentiating malignant and benign lesions, reducing misclassification rates. This is crucial in real-world applications where accurate diagnosis can significantly impact treatment decisions.

The overall efficiency of the proposed system was further validated through time-based performance analysis. The model inference time was optimized to provide real-time predictions within a few seconds, ensuring practical usability in clinical settings. The Flask-based web application successfully integrated the trained model, allowing users to upload images and receive immediate diagnostic results. Despite achieving high accuracy, the system faces some challenges, particularly in handling rare skin lesion categories with limited training data. Future improvements will focus on expanding the dataset, incorporating additional augmentation techniques, and integrating explainable AI methods to enhance interpretability. Further research will also explore the deployment of lightweight models for mobile-based skin cancer detection, making the system more accessible to users worldwide.

VI. CONCLUSION AND FUTURE WORK

This research successfully demonstrates the effectiveness of deep learning, particularly convolutional neural networks (CNNs), in the detection and classification of skin cancer.

The use of pre-trained models such as VGG16 and InceptionV3 has significantly improved diagnostic accuracy by extracting essential features from dermatoscopic images. Among the tested architectures, InceptionV3 outperformed VGG16, achieving an accuracy of 94.7%, making it the preferred choice for deployment in real-time applications. The system was successfully integrated into a web-based application, allowing users to upload skin lesion images and receive immediate classification results. This implementation not only improves accessibility but also reduces the dependency on manual diagnosis, which can be subjective and time-consuming.

Despite achieving high accuracy, certain limitations need to be addressed to enhance the system's performance and reliability in real-world clinical settings. One major challenge is the interpretability of deep learning models. CNN-based architectures, while highly effective, operate as black-box systems, making it difficult for dermatologists and medical practitioners to understand how decisions are made. Future work will focus on integrating explainable AI (XAI) techniques to provide better visualization of model predictions. This will improve transparency and build trust among healthcare professionals using AI-assisted diagnosis tools.

Another important area of improvement is dataset expansion. The current system relies on the HAM10000 dataset, which, while diverse, may not fully represent all skin types and lesion variations across different demographic groups. To improve the generalization of the model, additional datasets from multiple sources will be incorporated, ensuring better classification performance on underrepresented skin lesion types. Additionally, synthetic data augmentation techniques can be explored to generate more training samples, further enhancing the robustness of the model.

Furthermore, deploying the model on mobile and cloud-based platforms will enhance accessibility, making skin cancer detection available to a larger audience. Implementing a lightweight version of CNN models, optimized for mobile devices, can enable real-time diagnosis through smartphone applications. This will be particularly beneficial for users in remote areas where access to dermatologists is limited. Cloud integration will also allow for centralized data collection and continuous model updates, ensuring that the system remains up-to-date with the latest advancements in medical AI.

Finally, future work will explore the integration of hybrid models combining VGG16, InceptionV3, and attention-based architectures such as Vision Transformers (ViTs) to further improve classification accuracy. By leveraging ensemble learning techniques, the system can enhance decision-making by aggregating predictions from multiple models. This will lead to more reliable and clinically relevant diagnoses, ultimately contributing to early detection and improved treatment outcomes for skin cancer patients.

REFERENCES

- [1] Esteva, et al., "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks," *Nature*, 2017.
- [2] H. A. Haenssle, et al., "Man Against Machine: Diagnostic Performance of a CNN for Melanoma Recognition," *Annals of Oncology*, 2018.
- [3] J. Li, et al., "Skin Lesion Classification Using Vision Transformers," *AI in Medical Imaging*, 2022.
- [4] J. Kawahara, et al., "Deep Learning-Based Skin Cancer Classification with Clinical Images," *IEEE Transactions on Biomedical Engineering*, 2016.
- [5] T. J. Brinker, et al., "Comparing Artificial Intelligence-Based Algorithms for Skin Cancer Classification," *JAMA Dermatology*, 2019.
- [6] P. Tschandl, et al., "Human-Computer Collaboration for Skin Cancer Recognition," *Nature Medicine*, 2020.
- [7] M. Nasr-Esfahani, et al., "Deep Learning-Based Lesion Segmentation and Classification in Skin Cancer Detection," *Artificial Intelligence in Medicine*, 2021.
- [8] G. Pacheco, et al., "The Impact of Dataset Bias on Deep Learning-Based Skin Cancer Classification Models," *Computerized Medical Imaging and Graphics*, 2021.
- [9] J. Yap, et al., "Automated Skin Lesion Analysis with Deep Learning," *IEEE Transactions on Medical Imaging*, 2018.
- [10] Q. Abbas, et al., "Deep CNN-Based Skin Lesion Segmentation and Classification for Medical Diagnosis," *IEEE Access*, 2019.
- [11] Matthews, et al., "Explainable AI for Skin Cancer Detection: A Step Toward Model Interpretability," *Journal of Biomedical Informatics*, 2020.
- [12] N. Codella, et al., "Deep Learning Ensembles for Skin Cancer Classification," *IEEE Journal of Biomedical and Health Informatics*, 2018.
- [13] S. Menzies, et al., "Dermoscopy-Based AI Algorithms for Skin Cancer Classification," *International Journal of Dermatology*, 2021.
- [14] H. Lui, et al., "Deep Learning in Skin Cancer Diagnosis: A Systematic Review," *The Lancet Digital Health*, 2020.
- [15] O. Carranza-Rojas, et al., "Convolutional Neural Networks for Skin Cancer Classification: An Empirical Analysis," *Pattern Recognition Letters*, 2021.
- [16] R. Kasmi, et al., "Transfer Learning for Skin Cancer Classification Using Pretrained Deep Networks," *Neural Computing and Applications*, 2019.
- [17] S. Daniel, et al., "A Mobile-Based AI Application for Skin Cancer Detection," *Journal of Medical Systems*, 2020.
- [18] M. Attia, et al., "Multi-Modal Deep Learning for Skin Lesion Classification," *Medical Image Analysis*, 2021.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details