iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79572

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

An Explainable AI-Driven Clinical Decision Support System for Skin Cancer Detection Using Deep and Texture Feature Fusion

Mrs. G. Archana[1], Pasham Manikanta Reddy[2], Pamulaiah gari Uday Kumar[3], Pichika Venkatesh[4]
[1]*M.C.A, M.Phill, Assistant Professor, AIDS, DSU, Tamil Nadu*
[2, 3, 4]*Department of Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan University, Tamil Nadu, India*

Abstract: ***Skin cancer is among the most common forms of cancer worldwide, and its early detection plays a crucial role in improving patient survival rates. However, existing automated systems based purely on deep learning often lack interpretability and do not incorporate clinical reasoning. In this work, we propose an explainable Clinical Decision Support System (CDSS) for binary skin cancer classification using a hybrid feature fusion approach. The model combines deep features extracted from a pretrained ResNet18 network with handcrafted texture features obtained from Gray-Level Co-occurrence Matrix (GLCM) and wavelet transforms. This combination enables better differentiation between benign and malignant lesions by capturing both high-level and fine-grained image characteristics. Additionally, the system integrates ABCDE-based clinical rules for risk assessment and utilizes Grad-CAM to provide visual explanations of model predictions. Experimental evaluation on the HAM10000 dataset shows that the proposed approach achieves an accuracy of 84.23% and a malignant recall of 80.89%, outperforming baseline CNN models. The results demonstrate that the system not only improves prediction performance but also enhances transparency, making it more suitable for real-world clinical decision support.***

I. INTRODUCTION

Skin cancer is one of the most frequently diagnosed cancers worldwide, with melanoma being the most aggressive and life-threatening form. Early detection plays a critical role in improving survival rates; however, traditional diagnosis through dermoscopic examination is often subjective and heavily dependent on the clinician's experience. Differences in visual interpretation, along with the growing number of patients, highlight the need for more reliable and automated diagnostic support systems.

In recent years, deep learning techniques—especially Convolutional Neural Networks (CNNs)—have shown strong performance in medical image analysis. Models such as ResNet, DenseNet, and EfficientNet have been widely applied for skin lesion classification and have achieved promising results. Despite their effectiveness, most of these models operate as black-box systems, providing limited insight into how decisions are made. Additionally, they primarily rely on deep image features and often overlook clinically relevant reasoning that is commonly used in dermatology.

In real-world clinical practice, dermatologists typically assess skin lesions using the ABCDE rule, which considers Asymmetry, Border irregularity, Color variation, Diameter, and Evolution. These criteria play an important role in estimating the risk of malignancy. However, many existing deep learning-based approaches do not incorporate such structured clinical knowledge into their prediction process. Furthermore, the lack of visual explanation techniques reduces transparency, making it difficult for healthcare professionals to fully trust and adopt these systems.

To overcome these challenges, this study presents an explainable Clinical Decision Support System (CDSS) for binary skin cancer detection using a hybrid feature fusion approach. The proposed framework combines deep features extracted from a pretrained ResNet18 model with handcrafted texture features derived from GLCM and wavelet transforms. In addition, ABCDE-based clinical analysis is integrated to provide structured risk assessment, while Grad-CAM is used to generate visual explanations of model predictions. The system is evaluated on the HAM10000 dataset using a stratified 70:15:15 split and achieves an accuracy of 84.23% along with a malignant recall of 80.89%.

These results indicate that the proposed approach not only improves classification performance but also enhances interpretability, making it more suitable for practical clinical applications.

II. LITERATURE SURVEY

Deep learning has brought significant progress in automated skin lesion analysis over the past few years. A notable study by Esteva et al. [1] demonstrated that deep convolutional neural networks can achieve performance comparable to dermatologists when trained on large-scale datasets. While this work highlighted the potential of deep learning in medical image analysis, it did not address the need for interpretability, which is essential for clinical adoption.

Subsequent research has focused on improving model performance using advanced architectures. ResNet-based models, due to their residual connections, enable better gradient flow and deeper feature learning. Tschandl et al. [2] utilized architectures such as ResNet and EfficientNet on the HAM10000 dataset and reported strong results in multi-class classification. However, these methods largely function as black-box models and do not incorporate clinical reasoning into their predictions.

In addition to deep learning approaches, traditional texture-based methods have also been explored for skin lesion analysis. Techniques such as Gray-Level Co-occurrence Matrix (GLCM) and wavelet transforms capture important statistical and frequency-domain characteristics of images [3]. These handcrafted features can complement deep features by providing fine-grained texture information. Harangi [4] proposed ensemble CNN models to improve classification performance, while Kassem et al. [5] combined MobileNet with SVM classifiers to achieve efficient and accurate results on the HAM10000 dataset. Similarly, Adegun and Viriri [6] focused on lesion segmentation using fully convolutional networks. Despite these contributions, most of these methods lack explainability and do not integrate clinically meaningful insights.

More recently, there has been growing interest in explainable AI (XAI) for medical applications. Techniques such as Grad-CAM [7] provide visual explanations by highlighting image regions that influence model predictions, thereby improving transparency and user trust. However, many existing studies still prioritize classification accuracy over interpretability and rarely incorporate structured clinical decision-making frameworks.

Overall, while prior research has made significant advancements in skin lesion classification, there remains a gap in combining high-performance models with explainability and clinically relevant reasoning. Table I presents a summary of existing approaches along with their key limitations.

Table I. Comparison of related works

Author/Year	Method	Dataset	Acc.(%)	Limitation
Esteva et al. (2017)	Deep CNN	ISIC/Clinic	91.0	No XAI
Harangi (2018)	Ensemble CNN	ISIC 2017	85.6	High complexity
Tschandl et al. (2019)	ResNet/EfficientNet	HAM10000	83.1	No clinical rules
Kassem et al. (2020)	MobileNet+SVM	HAM10000	82.4	No explainability
Xie et al. (2020)	DenseNet+GLCM	ISIC 2018	80.7	No Grad-CAM
Adegun & Viriri (2021)	FCN Segment.	ISIC 2017	79.2	Seg. only
Proposed System	ResNet18+GLCM+ Wavelet+Grad-CAM	HAM10000	84.23	Binary only

Most existing approaches either rely on deep learning models or handcrafted features independently, which limits their overall effectiveness. In contrast, this work presents a unified framework that combines ResNet18-based deep features with GLCM and wavelet texture descriptors, along with ABCDE-based clinical analysis and Grad-CAM explainability within a Clinical Decision Support System.

III. PROPOSED METHODOLOGY

The system is designed as a multi-stage hybrid framework that integrates deep learning, handcrafted feature extraction, clinical rule-based analysis, and explainability techniques. The overall workflow begins with image preprocessing, followed by deep feature extraction using ResNet18. In parallel, texture features are extracted using GLCM and wavelet transforms. These features are then combined through a fusion process and used for binary classification. In addition, ABCDE-based clinical analysis is incorporated for structured risk assessment, and Grad-CAM is applied to generate visual explanations. Finally, the system produces a comprehensive clinical report based on the model's predictions.

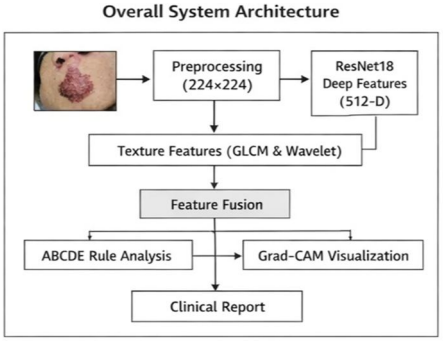


Fig. 1. Overall System Architecture of the Proposed CDSS.

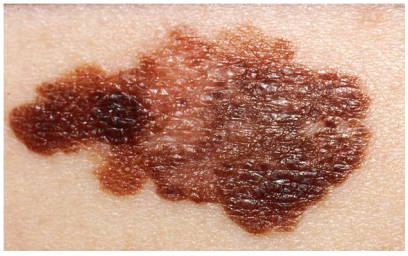


Fig. 2. Sample Input Dermoscopic Lesion Image (HAM10000 Dataset).

A. Image Preprocessing

The input dermoscopic images are first resized to 224×224 pixels to match the input requirements of the ResNet18 model. They are then normalized using the ImageNet mean ($\mu = [0.485, 0.456, 0.406]$) and standard deviation ($\sigma = [0.229, 0.224, 0.225]$) to ensure consistency with the pretrained network. During training, the data is processed in mini-batches of size 16 to improve computational efficiency and stability.

B. Deep Feature Extraction Using ResNet18

ResNet18 [8] is used as the primary deep feature extractor due to its residual learning structure, which helps address the vanishing gradient problem in deep neural networks. The model is initialized with pretrained ImageNet weights, allowing effective transfer learning for dermoscopic image analysis. To adapt the network for binary skin lesion classification, the original fully connected layer is replaced with a task-specific classifier. Each input image of size 224×224×3 is processed through a sequence of convolutional layers, batch normalization, ReLU activations, and residual blocks. After feature extraction, a global average pooling layer is applied to obtain a 512-dimensional feature vector that captures high-level semantic information of the lesion.

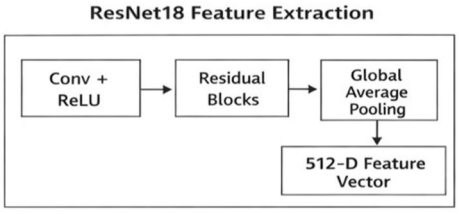


Fig. 3. ResNet18 Feature Extraction Architecture.

C. Texture Feature Extraction Using GLCM and Wavelet

In addition to deep features, handcrafted texture descriptors are used to capture fine-grained patterns in skin lesions that may not be fully represented by CNNs alone. The Gray-Level Co-occurrence Matrix (GLCM) is applied to the grayscale lesion image to extract second-order statistical features, including contrast, homogeneity, energy, and correlation. These features describe variations in intensity, structural consistency, and relationships between neighbouring pixels.

To further incorporate frequency-domain information, a two-dimensional discrete wavelet transform (DWT) using the Haar wavelet is performed. From the resulting sub-bands, statistical measures such as mean and standard deviation are computed for both approximation and detail components, resulting in eight wavelet-based features. Together, the GLCM and wavelet descriptors are combined to form a 12-dimensional texture feature vector used for subsequent processing.

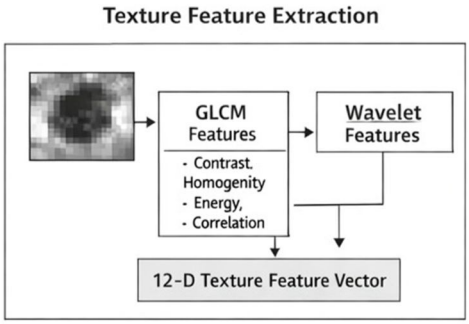


Fig. 4. Texture Feature Extraction Using GLCM and Wavelet Transform.

D. Feature Fusion Strategy

To combine complementary information from different feature sources, the deep features obtained from ResNet18 (512 dimensions) are concatenated with the 12-dimensional texture feature vector, resulting in a unified 524-dimensional representation.

$F\_f = [F\_d \,\|\, F\_t]$, $F\_d \in \mathbb{R}^{512}$, $F\_t \in \mathbb{R}^{12}$

Here, || denotes the concatenation operation between the deep feature vector (F_d) and the texture feature vector (F_t). The resulting fused representation is then passed through fully connected layers with ReLU activation for binary classification. This fusion approach allows the model to leverage both high-level semantic information and low-level statistical patterns, improving its ability to distinguish subtle differences between benign and malignant lesions.

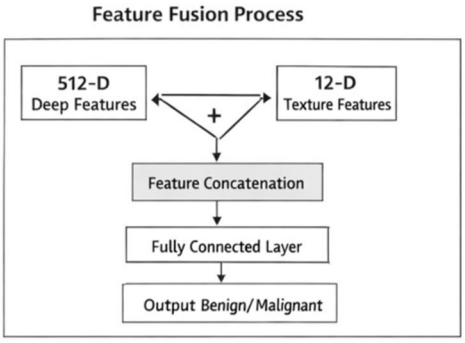


Fig. 5. Hybrid Feature Fusion Process.

E. ABCDE-Based Clinical Rule Integration

To enhance the reliability of predictions, clinically established diagnostic criteria are incorporated into the system. In dermatology, the ABCDE rule—Asymmetry, Border irregularity, Color variation, Diameter, and Evolution—is widely used to assess the risk of malignancy. In this work, each of these criteria is approximated using computational measures.

Asymmetry is estimated by comparing intensity differences between horizontally flipped halves of the lesion. Border irregularity is quantified using contour compactness derived from the lesion segmentation mask. Color variation is analyzed through K-means clustering to identify dominant color distributions within the lesion. Diameter is approximated using the minimum enclosing circle of the detected lesion region. Finally, an evolution score is derived by combining variations observed in asymmetry, border, and color characteristics.

Based on these computed measures, an overall risk level is assigned as Low, Moderate, or High using predefined thresholds. This clinical assessment complements the neural network output and aligns the system's predictions with established dermatological practices.

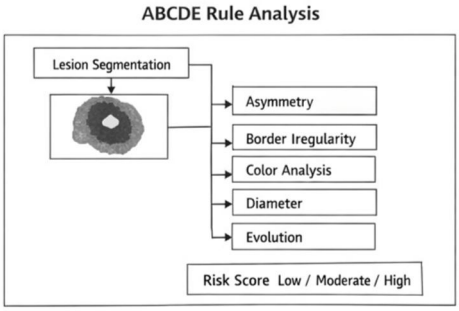


Fig. 6. ABCDE-Based Clinical Rule Analysis with Risk Scoring.

F. Explainability Using Grad-CAM

To improve interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is integrated into the system to highlight the regions of the image that contribute most to the model's predictions. It works by analyzing the gradients of the target class with respect to the feature maps of the final convolutional layer.

The resulting heatmap emphasizes the most important areas within the lesion, allowing better understanding of the model's decision-making process. This ensures that the network focuses on clinically relevant regions rather than background noise, thereby improving transparency and trust in the system.

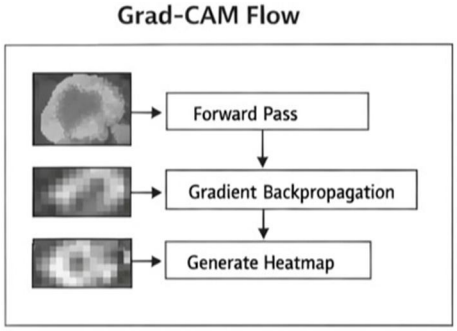


Fig. 7. Grad-CAM Flow: Forward Pass → Backpropagation → Heatmap.

IV. EXPERIMENTAL SETUP

The framework is evaluated using the HAM10000 (Human Against Machine with 10,000 training images) dataset, which consists of dermoscopic images representing various types of pigmented skin lesions. For this study, the dataset is simplified into a binary classification task, where melanoma, basal cell carcinoma, and actinic keratosis are grouped under malignant cases, while the remaining categories are treated as benign. To ensure a balanced evaluation, a stratified data splitting strategy is applied, maintaining the class distribution across training, validation, and test sets, as presented in Table II.

All input images are resized to 224×224 pixels and normalized using ImageNet statistics to align with the pretrained ResNet18 model requirements. The ResNet18 backbone is initialized with pretrained ImageNet weights and fine-tuned for the classification task using the Adam optimizer with a learning rate of 0.001. To address class imbalance, a weighted cross-entropy loss function is employed, giving higher importance to minority class samples. The model is trained for 15 epochs, with validation performance monitored at each epoch to prevent overfitting and ensure stable convergence.

All experiments are implemented using the PyTorch framework and executed in the Google Colab environment with GPU support. Performance evaluation is carried out using standard metrics, including Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. Particular emphasis is placed on malignant recall, as minimizing false negatives is critical in clinical screening scenarios. Additionally, the evaluation process considers both overall performance and class-wise behavior to ensure the robustness of the proposed system.

TABLE II. HAM10000 DATASET PARTITION

Split	Total Samples	Benign	Malignant	Ratio
Training	7,010	5,640	1,370	70%
Validation	1,502	1,208	294	15%
Test	1,503	1,210	293	15%
Total	10,015	8,058	1,957	100%

V. RESULTS AND EVALUATION

The hybrid ResNet18 and texture feature fusion approach demonstrates improved performance compared to conventional CNN-based methods. A detailed comparison of the proposed model with baseline and ablation variants is presented in Table III, highlighting the effectiveness of the feature fusion strategy.

Table III. Performance Comparison Of Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Baseline CNN	~75.00	~52.00	~68.00	~59.00
ResNet18 (Deep only)	80.10	53.20	76.50	62.80
ResNet18 + GLCM	82.30	54.80	78.20	64.50
ResNet18 + Wavelet	81.70	54.10	77.40	63.90
Proposed (Full Fusion)	84.23	56.70	80.89	66.67

Table IV. Confusion Matrix On Test Set (1,503 Samples)

Metric	Predicted Benign	Predicted Malignant	Total
Actual Benign	TN = 1029	FP = 181	1210
Actual Malignant	FN = 56	TP = 237	293

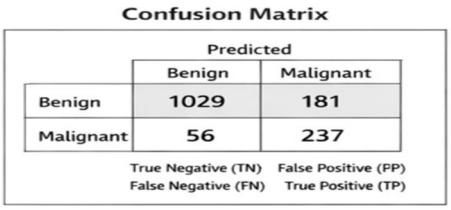


Fig. 8. Confusion Matrix Visualization on Test Dataset.

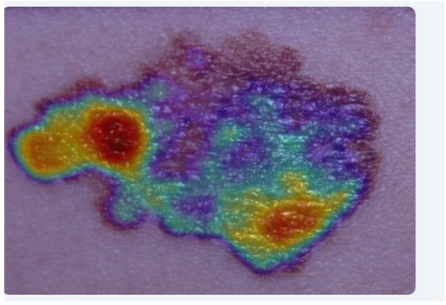


Fig. 9. Grad-CAM Attention Heatmap — Explainable AI Output.

On the test dataset, the model achieves an overall accuracy of 84.23%, with a precision of 56.70%, recall of 80.89%, and an F1-score of 66.67%. The relatively high recall for malignant cases (80.89%) indicates the model's ability to correctly identify cancerous lesions, which is particularly important in clinical screening where missing a positive case can have serious consequences.

In comparison to the baseline CNN model (approximately 75% accuracy), the fusion-based approach provides an improvement of around 9% in overall accuracy while maintaining better sensitivity toward malignant samples. This highlights the effectiveness of combining deep and texture features for improved classification performance.

Furthermore, Grad-CAM visualizations (Fig. 9) show that the model focuses primarily on lesion regions rather than surrounding skin, supporting the interpretability of the predictions. The confusion matrix further reflects this performance, with 1029 true negatives, 237 true positives, 181 false positives, and 56 false negatives.

VI. DISCUSSION

The experimental findings indicate that combining deep features with handcrafted texture descriptors leads to more robust classification compared to using CNN models alone. The hybrid fusion approach allows the model to capture both high-level semantic representations and fine-grained statistical patterns present in lesion images, resulting in improved discriminative capability.

The high recall achieved for malignant cases highlights the model's effectiveness in identifying cancerous lesions, making it suitable for screening applications where reducing false negatives is essential. Although the precision remains moderate (56.70%), this trade-off is acceptable in clinical contexts, as early detection is generally prioritized over strict specificity.

In addition, incorporating ABCDE-based rule analysis improves interpretability by aligning the system's predictions with established dermatological criteria. The Grad-CAM visualizations further support this by showing that the model focuses on clinically relevant regions of the lesion. Overall, the integration of feature fusion, clinical rule-based analysis, and explainable AI provides a balanced approach that connects data-driven predictions with practical clinical reasoning, enhancing its usability as a Clinical Decision Support System.

VII. CONCLUSION

This study presents an explainable Clinical Decision Support System for binary skin cancer detection using a hybrid feature fusion approach. By combining deep features from ResNet18 with handcrafted texture descriptors such as GLCM and wavelet transforms, the system effectively captures both high-level and fine-grained characteristics of skin lesions.

The experimental results on the HAM10000 dataset demonstrate an accuracy of 84.23% and a malignant recall of 80.89%, indicating strong performance in identifying cancerous cases. The inclusion of ABCDE-based clinical analysis and Grad-CAM visualization further enhances interpretability, making the system more aligned with real-world clinical practices.

Overall, the integration of feature fusion, clinical reasoning, and explainable AI provides a reliable and practical framework for AI-assisted dermatological screening, with potential to support early diagnosis and improve decision-making in healthcare settings.

VIII. FUTURE WORK

Future work can explore extending this framework to multi-class skin lesion classification to handle a wider range of dermatological conditions. Incorporating larger and more diverse datasets may further improve the model's generalization and robustness across different skin types and imaging conditions.

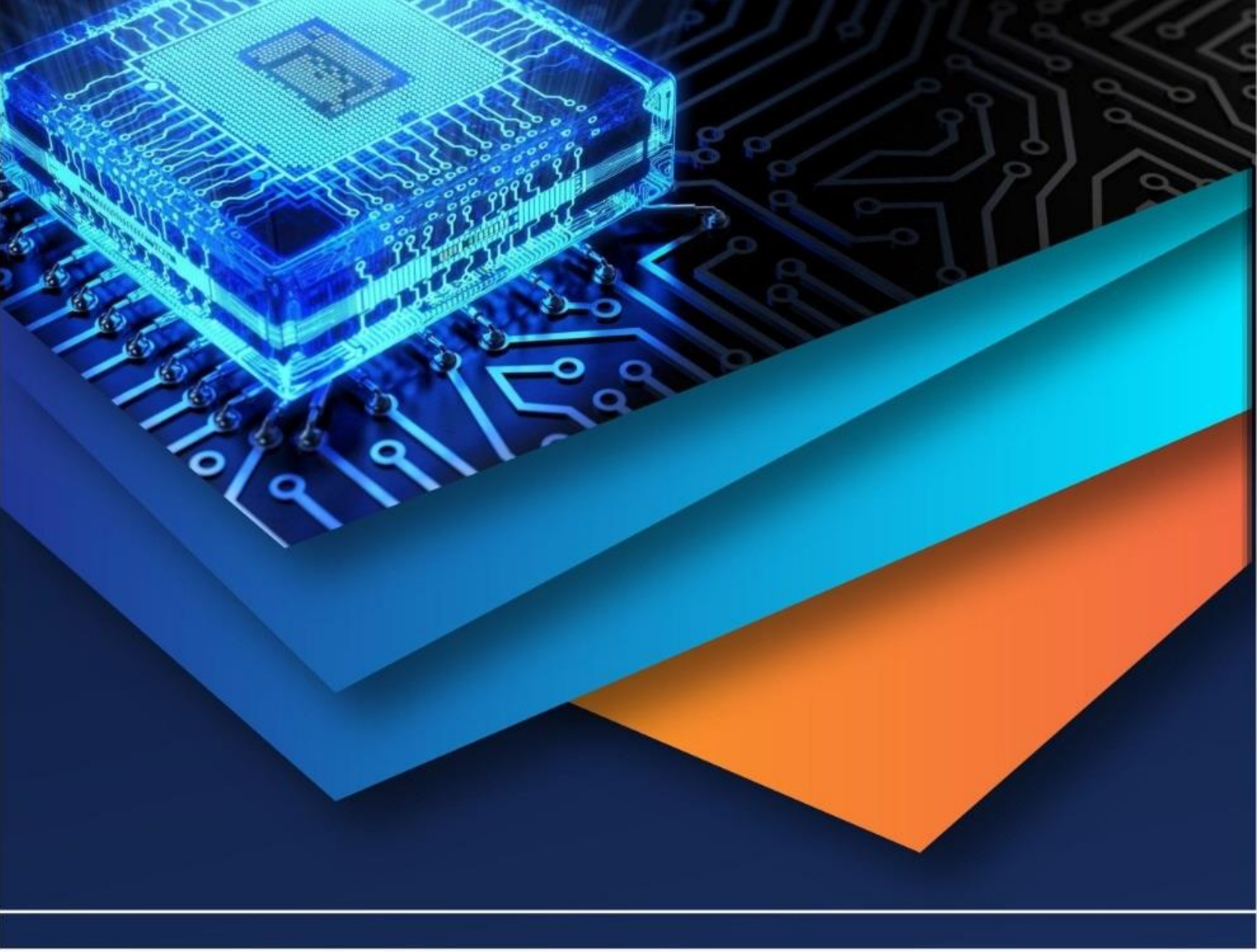
Additionally, the use of advanced architectures such as transformer-based models and attention mechanisms could enhance feature representation and improve classification performance. Clinical validation through collaboration with dermatologists and real-world testing will be essential to assess the system's reliability and practical usability.

Furthermore, integrating the system with mobile-based dermoscopic devices could enable scalable and accessible skin cancer screening, particularly in resource-limited settings.

REFERENCES

[1] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, Feb. 2017.
[2] P. Tschandl et al., "The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions," Scientific Data, vol. 5, no. 1, 2018.
[3] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," IEEE Transactions on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610–621, Nov. 1973.

[4] B. Harangi, "Skin lesion classification with ensembles of deep convolutional neural networks," Journal of Biomedical Informatics, vol. 86, pp. 25–32, Oct. 2018.
[5] M. A. Kassem et al., "Skin lesions classification into eight classes for HAM10000 dataset using deep CNN and transfer learning," IEEE Access, vol. 8, pp. 102851–102863, 2020.
[6] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," IEEE Access, vol. 8, pp. 7160–7172, 2020.
[7] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626.
[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY