Automated Skin Disease Detection Using an Optimized Xception-Based Framework with Comparative Deep Learning Analysis

V. Akhila
PG Scholar, Department of Computer Science and Engineering,Jawaharlal Nehru Technological University Hyderabad University College of Engineering Jagtial(Autonomous),Nachupally(Kondagattu), Telangana,India -505501.
Email: vallampatlaakhila123@gmail.com

Dr. B. Sateesh Kumar
Professor, Department of Computer Science and Engineering,Jawaharlal Nehru Technological University Hyderabad University College of Engineering Jagtial(Autonomous), Nachupally(kondagattu),Telangana, India - 505501.
Email: sateeshbkumar@jntuh.ac.in

***Abstract*— Skin diseases and hematological disorders are among the most prevalent health conditions worldwide, where timely and accurate diagnosis is essential for effective treatment and improved clinical outcomes. Conventional diagnosis relies heavily on manual visual assessment by medical experts, making the process time-consuming, subjective, and susceptible to inconsistent predictions, particularly when diseases exhibit similar visual characteristics. Four publicly available medical image datasets, namely HAM10000, ISIC 2019–2020 Melanoma, Skin Cancer (Benign vs. Malignant), and Blood Cell Images, were utilized for comparative training and evaluation. The framework incorporates comprehensive preprocessing, including hair removal, image resizing, data augmentation, and image normalization, to enhance image quality and improve feature representation. Multiple deep learning architectures, including EfficientNetV2, InceptionResNetV2, InceptionV3, MobileNet, VGG19, ResNet50, Xception, ConvNeXt-Tiny, and a hybrid EDA-ResNet50 integrating Convolutional Block Attention Module and Coordinate Attention, were comparatively evaluated. Performance was assessed using accuracy, precision, recall, F1-score, and confusion matrix analysis. Experimental results demonstrated that Xception achieved the highest classification accuracy of 98.70% on the Blood Cell dataset, 82.30% on HAM10000, and 89.70% on the Skin Cancer dataset, while InceptionResNetV2, InceptionV3, and MobileNetV1 each achieved 94.30% on the ISIC dataset. The framework further integrates Grad-CAM-based explainability and prediction confidence estimation, providing an accurate, interpretable, and reliable computer-aided medical image classification solution for automated healthcare support.**

***"Keywords*— Skin Disease Detection, Convolutional Neural Networks, Transfer Learning, Deep Learning, Explainable Artificial Intelligence, Dermoscopic Image Classification."**

I. Introduction

Skin diseases and hematological disorders are among the most prevalent health conditions worldwide, affecting millions of individuals across different age groups and geographical regions. Dermatological disorders such as melanoma, eczema, psoriasis, fungal infections, and other skin abnormalities often exhibit similar visual characteristics, making accurate diagnosis a challenging clinical task. Likewise, microscopic blood cell analysis plays a vital role in diagnosing numerous hematological conditions, yet conventional examination remains dependent on expert interpretation. Early and reliable diagnosis is essential for initiating timely treatment, preventing disease progression, and improving clinical outcomes. The growing availability of medical imaging technologies, together with recent advances in artificial intelligence (AI) and deep learning (DL), has created new opportunities for developing intelligent computer-aided medical image classification systems that assist healthcare professionals with efficient, consistent, and accurate disease analysis [1]. These systems have the potential to improve diagnostic performance while reducing dependence on manual interpretation and extensive clinical expertise [2], [3].

Despite remarkable progress in automated medical image analysis, several challenges continue to limit the practical adoption of existing diagnostic systems. Conventional diagnosis remains highly dependent on visual inspection performed by experienced specialists, making the process time-consuming and susceptible to subjective interpretation. In addition, many existing approaches primarily emphasize classification accuracy while providing limited attention to image preprocessing, comparative evaluation of multiple deep learning architectures, prediction transparency, and practical deployment. Medical images frequently contain artifacts, illumination variations, and quality inconsistencies that can adversely affect feature extraction and classification performance if not appropriately addressed. Furthermore, the limited interpretability of many deep learning models reduces user confidence, creating barriers to their adoption in real-world clinical environments. These limitations highlight the need for comprehensive and interpretable medical image classification frameworks capable of delivering accurate, reliable, and explainable decision support across diverse healthcare applications [4], [5].

This work presents an intelligent computer-aided medical image classification framework for automated analysis of skin lesion and blood cell images using transfer learning. The proposed framework integrates comprehensive image preprocessing, comparative evaluation of multiple state-of-the-art deep learning architectures, Grad-CAM-based explainability, and prediction confidence estimation within a unified workflow. Following extensive experimental evaluation on four publicly available datasets, Xception demonstrated the most consistent classification performance and was selected as the final deployment model. The developed framework is deployed as a secure web-based application that enables automated image classification, visual explanation of predictions, disease information, and remedy recommendations through an interactive user interface, thereby improving accessibility and practical usability for intelligent healthcare support [6], [7].

A. Main Contributions:

The major contributions of this work are summarized as follows:

- Development of an intelligent medical image classification framework integrating transfer learning for automated analysis of skin lesion and blood cell images.
- Comparative evaluation of multiple state-of-the-art deep learning architectures to identify the most suitable model for deployment under consistent experimental settings.
- Integration of comprehensive preprocessing, Grad-CAM-based explainability, and prediction confidence estimation to improve prediction transparency and user trust.
- Deployment of the selected Xception model as a secure web-based application supporting automated image classification, disease information, remedy recommendations, and interactive visualization for practical healthcare assistance.

II. RELATED WORK

Recent advances in artificial intelligence have significantly improved automated medical image analysis through the application of deep learning techniques. Alotaibi and AlSaeed proposed a transfer learning framework integrated with deep attention mechanisms to improve skin cancer detection by emphasizing clinically significant lesion regions. Their approach demonstrated enhanced classification performance and better localization of abnormal skin patterns, highlighting the effectiveness of attention-guided feature learning. However, the framework primarily focused on skin cancer detection and offered limited evaluation across diverse dermatological conditions [11]. Similarly, Yasir and Kim introduced an attention-based deep feature aggregation network that enhanced feature representation by combining multiple discriminative feature maps. The proposed architecture achieved improved classification accuracy, although its computational complexity and dependence on large-scale datasets may limit deployment in resource-constrained environments [12]. Shu et al. presented a multi-stage and multi-attention framework for multimodal skin lesion classification that effectively fused complementary visual information from different image modalities. While the approach improved diagnostic accuracy, the increased architectural complexity may restrict practical implementation in lightweight clinical applications [13].

To further improve diagnostic performance, several researchers have incorporated advanced convolutional architectures with attention mechanisms and ensemble learning strategies. Younas et al. developed an attention-based inception-residual convolutional network capable of extracting discriminative lesion features while reducing information loss during feature propagation. Their findings demonstrated improved classification performance, although the model primarily emphasized predictive accuracy rather than prediction interpretability [14]. Akilandasowmya et al. proposed a diagnostic framework that combined deep hidden feature extraction with ensemble classifiers for early skin cancer detection. The integration of multiple classifiers improved predictive robustness; however, the framework remained focused on cancer-specific diagnosis without addressing broader multi-class medical image classification challenges [15]. Houssein et al. introduced a deep convolutional neural network for multiclass skin cancer classification that achieved competitive performance across multiple disease categories. Despite encouraging results, limited emphasis was placed on explainable decision support, reducing transparency for clinical interpretation [16].

Earlier studies also demonstrated the effectiveness of lightweight deep learning architectures for automated dermatological analysis. Toğaçar et al. combined autoencoders, MobileNetV2, and spiking neural networks to develop an intelligent skin cancer detection framework with reduced computational requirements while maintaining reliable classification performance. Although the architecture demonstrated computational efficiency, the interpretability of prediction outcomes remained limited, which is an important consideration for clinical acceptance [17]. More recently, Aruk et al. proposed a hybrid ConvNeXt-based framework that enhanced skin lesion classification through improved feature representation and architectural optimization. Their findings indicated superior classification performance compared with conventional convolutional networks; nevertheless, broader integration with user-oriented diagnostic support, explainability, and practical deployment remained relatively unexplored [18].

Comprehensive review studies have further highlighted the strengths and remaining challenges of existing approaches. Shakya et al. conducted an extensive comparative analysis of deep learning and transfer learning techniques for skin cancer classification and concluded that modern architectures substantially improve diagnostic accuracy but continue to face challenges related to dataset diversity, generalization capability, and explainability [19]. Likewise, Lyakhova and Lyakhov presented a systematic review of artificial intelligence methods for skin cancer detection and classification, emphasizing the growing need for reliable, interpretable, and clinically deployable diagnostic systems that extend beyond prediction accuracy alone [20]. Collectively, these studies demonstrate remarkable progress in automated medical image classification while indicating that further improvements are required to achieve robust, transparent, and practically deployable computer-aided diagnostic systems.

A. Research Gap

Although existing studies have significantly advanced automated skin disease analysis through deep learning, several important limitations remain. Most approaches primarily focus on improving classification accuracy while providing limited attention to comprehensive image preprocessing, comparative evaluation of multiple deep learning architectures, prediction interpretability, and practical deployment. In addition, many frameworks evaluate only a small number of models, making it difficult to identify the most suitable architecture under consistent experimental conditions. Explainability is frequently limited, reducing user confidence and restricting adoption in real-world healthcare environments. Furthermore, most reported systems remain research prototypes and lack integrated web-based applications that provide secure user interaction, automated prediction, visual explanation, disease information, and practical diagnostic assistance.

To address these limitations, the proposed framework presents an intelligent computer-aided medical image classification system that integrates comprehensive image preprocessing, comparative evaluation of multiple state-of-the-art transfer learning models, Grad-CAM-based explainability, prediction confidence estimation, and practical web-based deployment within a unified workflow. By selecting the most consistent model through comparative experimentation and deploying it in an interactive application, the framework provides an accurate, interpretable, and user-friendly solution for automated medical image classification and intelligent healthcare support.

III. Materials And Methods

A) Overall Framework

The proposed framework presents an intelligent computer-aided medical image classification system for automated analysis of skin lesion and blood cell images using transfer learning. Four publicly available datasets, namely HAM10000, ISIC 2019–2020 Melanoma, Skin Cancer (Benign vs. Malignant), and Blood Cell Images, are utilized for comparative model training and evaluation. The framework integrates comprehensive image preprocessing, multiple state-of-the-art deep learning architectures, Grad-CAM-based explainability, and prediction confidence estimation to improve classification accuracy, interpretability, and reliability. Following comparative performance evaluation, Xception is selected as the final deployment model and integrated into a secure web-based application for practical medical image classification and intelligent healthcare support.

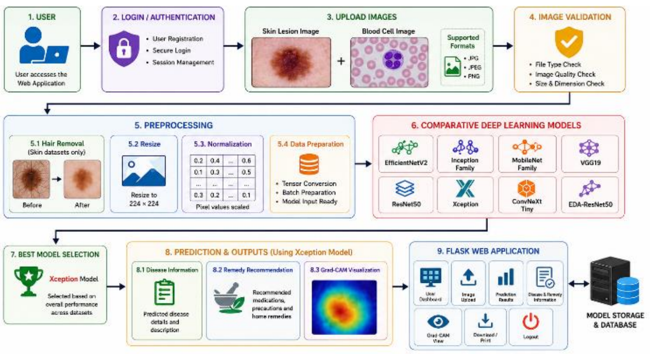

Fig. 1. System Architecture

As illustrated in Fig. 1, the workflow begins with dataset preparation and image preprocessing, followed by transfer learning-based model training and comparative evaluation. The best-performing Xception model is selected for deployment, where uploaded skin lesion and blood cell images are automatically classified. The system further generates Grad-CAM visualizations, prediction confidence, disease information, and remedy recommendations through a secure web-based interface, providing an interpretable and user-friendly computer-aided diagnostic solution.

B) Dataset Collection

The proposed framework utilizes four publicly available medical image datasets to perform comprehensive comparative evaluation of deep learning models for skin lesion and blood cell image classification. The datasets represent diverse disease categories, image characteristics, and classification complexities, enabling robust assessment of model generalization and performance across multiple medical imaging scenarios.

i) HAM10000: The HAM10000 (Human Against Machine with 10,000 Training Images) dataset contains 10,015 dermoscopic skin lesion images belonging to seven diagnostic categories, including melanoma, basal cell carcinoma, benign keratosis, and melanocytic nevi. Its multi-class nature and class imbalance make it suitable for evaluating automated skin lesion classification models.

ii) ISIC 2019–2020 Melanoma: The ISIC 2019–2020 Melanoma dataset comprises dermoscopic images collected from the International Skin Imaging Collaboration for melanoma recognition and skin lesion analysis. The dataset provides diverse lesion appearances acquired from multiple clinical sources, supporting comparative evaluation of deep learning architectures under realistic dermatological imaging conditions.

iii) Skin Cancer (Benign vs. Malignant): The Skin Cancer (Benign vs. Malignant) dataset contains dermoscopic images categorized into benign and malignant skin lesions. Its binary classification structure enables evaluation of deep learning models for accurate differentiation between cancerous and non-cancerous lesions, supporting early skin cancer identification and diagnostic decision-making.

iv) Blood Cell Images: The Blood Cell Images dataset consists of microscopic blood cell images representing multiple hematological cell categories captured under standardized laboratory conditions. The dataset enables automated blood cell classification and extends the proposed framework beyond dermatological analysis, providing comprehensive evaluation across diverse medical image classification tasks.

C) Image Preprocessing

The preprocessing stage prepares medical images for effective deep learning analysis by improving data consistency, reducing image artifacts, and enhancing feature representation before model training and inference. The implemented preprocessing pipeline consists of hair removal, image resizing, data augmentation, and image normalization, ensuring reliable feature extraction and robust classification performance across diverse medical image datasets.

i) Hair Removal: Hair removal is applied only to dermoscopic skin lesion images to eliminate hair artifacts that may partially obscure lesion boundaries and clinically significant visual patterns. By preserving lesion structures while removing unwanted interference, this preprocessing step enables the classification models to concentrate on diagnostically relevant regions, thereby improving feature extraction and classification reliability.

ii) Image Resizing: Image resizing converts all input images into a uniform spatial resolution compatible with the deep learning architectures. Standardizing image dimensions ensures consistent feature representation, facilitates efficient batch processing, reduces computational complexity, and enables fair comparative evaluation of multiple transfer learning models across different medical image datasets.

iii) Data Augmentation: Data augmentation is employed during model training to increase dataset diversity through realistic geometric transformations, including rotation, flipping, and scaling. These augmented samples expose the models to varied image appearances without additional data collection, reducing overfitting, improving generalization capability, and enhancing classification robustness for previously unseen medical images.

iv) Image Normalization: Image normalization standardizes pixel intensity values to reduce variations caused by different imaging conditions and acquisition devices. Maintaining a consistent intensity distribution improves training stability, accelerates model convergence, and enables the deep learning architectures to learn discriminative visual features more effectively, resulting in reliable and consistent classification performance.

D) Deep Learning Models

EfficientNetV2 (B0–B3 Variants): EfficientNetV2 is a family of transfer learning-based convolutional neural networks comprising the B0, B1, B2, and B3 variants, which progressively increase model capacity while maintaining computational efficiency through compound scaling. These variants enable effective feature extraction with varying architectural complexity and were comparatively evaluated to analyze their classification performance and generalization capability across diverse medical image datasets.

InceptionResNetV2: InceptionResNetV2 combines multi-scale feature extraction with residual learning to improve information propagation throughout the network. The architecture efficiently captures both local and global image characteristics while mitigating gradient degradation in deeper networks. Its hybrid design enhances feature representation, improves classification stability, and provides robust performance for complex medical image classification involving multiple disease categories.

InceptionV3: InceptionV3 utilizes factorized convolutional operations and multi-scale feature extraction to learn discriminative visual representations with reduced computational complexity. The architecture effectively captures fine-grained image features while maintaining efficient model training and inference. Its optimized design enables reliable classification performance and serves as an effective transfer learning model for comparative evaluation across diverse medical image datasets.

MobileNet (V1–V3 Variants): MobileNet is a family of lightweight convolutional neural networks comprising the V1, V2, and V3 variants, designed to achieve efficient feature extraction using depthwise separable convolutions. These progressively improved architectures provide an effective balance between computational efficiency and classification accuracy, making them suitable for resource-constrained environments. All three variants were comparatively evaluated to assess lightweight deep learning performance for automated medical image classification.

VGG19: VGG19 is a deep convolutional neural network composed of sequential convolutional layers that learn hierarchical visual representations from input images. Its uniform network architecture facilitates stable feature extraction and effective transfer learning for image classification tasks. The model provides reliable baseline performance and enables comprehensive comparison with more advanced convolutional neural network architectures.

ResNet50: ResNet50 is a deep residual learning architecture that introduces shortcut connections to facilitate efficient gradient propagation during training. These residual connections enable the network to learn complex visual representations while avoiding performance degradation in deeper architectures. The model effectively extracts discriminative image features and provides robust classification performance across challenging medical imaging applications.

ConvNeXt-Tiny: ConvNeXt-Tiny is a modern convolutional neural network designed by incorporating architectural improvements inspired by Vision Transformers while preserving the computational efficiency of convolutional models. It enhances feature extraction through optimized convolutional blocks, large kernel operations, and improved normalization strategies. The architecture delivers strong classification performance and excellent generalization across diverse medical image datasets.

EDA-ResNet50: EDA-ResNet50 is an enhanced deep learning architecture developed by integrating ResNet50 with the Convolutional Block Attention Module (CBAM) and Coordinate Attention (CA) mechanisms. CBAM refines channel and spatial feature representations, while Coordinate Attention preserves positional information to emphasize diagnostically significant image regions. This integrated attention mechanism improves feature discrimination, robustness, and classification performance for automated medical image analysis.

$$F' = CA\left(CBAM\left(ResNet50(F)\right)\right) \tag{1}$$

where:

F = Input feature map

ResNet50(F) = Deep extracted feature representation

CBAM = Convolutional Block Attention Module

CA = Coordinate Attention

F′ = Refined output feature map

Xception: Xception is a deep convolutional neural network that utilizes depthwise separable convolutions to efficiently learn complex visual representations while significantly reducing computational complexity. The architecture improves feature extraction by independently performing spatial and channel-wise convolution operations, enabling accurate recognition of subtle disease characteristics. Owing to its superior comparative performance across multiple datasets, Xception is selected as the final deployment model for the proposed medical image classification framework.

$$Y = PW\left(DW(X)\right) \tag{2}$$

where:

X = Input feature map

DW = Depthwise convolution

PW = Pointwise (1×1) convolution

Y = Output feature map

E) Comparative Model Selection

All deep learning architectures were trained and evaluated under identical experimental settings to ensure a fair and unbiased performance comparison. Each model utilized the same preprocessed datasets, training strategy, optimization settings, and evaluation protocol. Performance was assessed using accuracy, precision, recall, F1-score, and confusion matrix analysis to comprehensively measure classification effectiveness across the four medical image datasets. Based on the comparative experimental results, Xception consistently demonstrated the most balanced and reliable performance by effectively learning discriminative image features while maintaining strong generalization capability. Consequently, Xception was selected as the final deployment model and integrated into the proposed web-based medical image classification framework to provide accurate, robust, and interpretable computer-aided diagnostic support.

F) Explainable Artificial Intelligence

To improve the transparency and interpretability of automated predictions, the proposed framework incorporates Explainable Artificial Intelligence (XAI) through Gradient-weighted Class Activation Mapping (Grad-CAM) and prediction confidence estimation. Grad-CAM generates class-specific activation heatmaps by utilizing the feature maps and gradients of the final convolutional layer, highlighting the image regions that contribute most significantly to the predicted class. These visual explanations enable users to better understand the model's decision-making process and support clinical interpretation. In addition, prediction confidence is obtained from the Softmax probability distribution, providing a quantitative measure of the model's certainty for each prediction. The integration of visual explanations and confidence estimation enhances user trust, improves prediction transparency, and supports reliable computer-aided medical image classification in practical healthcare applications.

G) Web-Based Deployment

The proposed framework is deployed as a secure web-based application using the Flask framework to provide an interactive and user-friendly interface for automated medical image classification. SQLite is employed for user registration, authentication, and secure management of user information. After successful login, users upload skin lesion and blood cell images, which undergo input validation to verify supported image formats before preprocessing and classification. The deployed Xception model performs automated prediction, while Grad-CAM generates visual explanations highlighting diagnostically significant image regions. The application further presents prediction confidence, disease information, and corresponding remedy recommendations, enabling transparent and practical computer-aided medical image classification for intelligent healthcare support.

IV. Experimental Results

A) Evaluation Metrics

The performance of the evaluated deep learning models is assessed using four widely adopted classification metrics: accuracy, precision, recall, and F1-score. These metrics provide a comprehensive evaluation of predictive performance by measuring overall correctness, positive prediction reliability, completeness of positive class identification, and the balance between precision and recall.

Accuracy: Accuracy measures the overall proportion of correctly classified samples among all evaluated instances and reflects the overall classification performance of the model.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{3}$$

Precision: Precision measures the proportion of correctly predicted positive samples among all samples classified as positive, indicating the reliability of positive predictions.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{4}$$

Recall: Recall measures the ability of the model to correctly identify all actual positive samples, indicating the completeness of positive class detection.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{5}$$

F1-Score: F1-score is the harmonic mean of precision and recall, providing a balanced evaluation of classification performance, particularly for imbalanced datasets.

$$F1 - Score = \frac{2 \times (Recall \times Precision)}{Recall + Precision} \tag{6}$$

B) Comparative Performance Evaluation

The proposed framework was comparatively evaluated using four publicly available medical image datasets under identical experimental settings. All deep learning models were trained and assessed using the same preprocessing pipeline and evaluation metrics to ensure a fair and unbiased comparison. The experimental results demonstrate the classification capability of each model across different medical image classification tasks, providing the basis for selecting the most suitable architecture for deployment.

i) ISIC 2019–2020 Dataset

Table 1. Performance Evaluation on ISIC 2019–2020 Dataset

Model	**Accuracy**	**Macro F1-Score**	**Recall**	**Precision**
Xception	0.925	0.926	0.925	0.926
EfficientNet V2B0	0.932	0.931	0.932	0.932
EfficientNet V2B1	0.932	0.932	0.932	0.932
EfficientNet V2B2	0.931	0.930	0.931	0.931
EfficientNet V2B3	0.932	0.932	0.932	0.932
InceptionResNet V2	0.943	0.942	0.943	0.943
InceptionV3	0.943	0.942	0.943	0.943
MobileNetV1	0.943	0.942	0.943	0.943
MobileNetV2	0.931	0.931	0.931	0.931
MobileNetV3	0.913	0.913	0.913	0.914
VGG19	0.939	0.939	0.939	0.945
ResNet50	0.941	0.941	0.941	0.943
EDA-ResNet50	0.941	0.940	0.941	0.944
ConvNeXt-Tiny	0.942	0.942	0.942	0.947

Table 1 summarizes the comparative performance of the evaluated deep learning models on the ISIC 2019–2020 dataset using accuracy, macro F1-score, recall, and precision. InceptionResNetV2, InceptionV3, and MobileNetV1 achieved the highest classification accuracy of 94.3%, demonstrating consistent feature learning and balanced performance across all evaluation metrics.

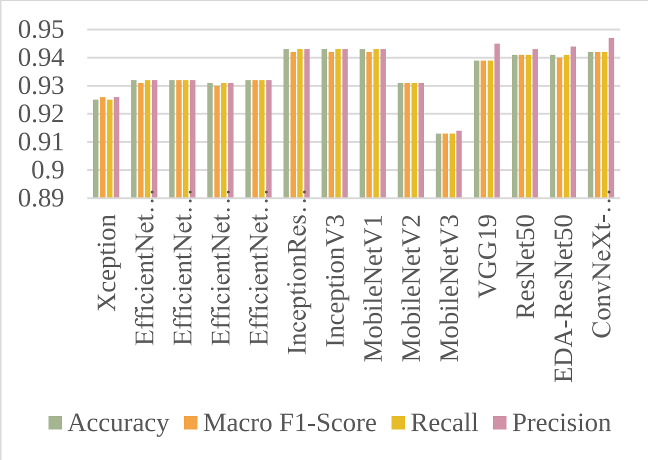

Figure 2. Performance Comparison on ISIC 2019–2020 Dataset

As illustrated in Fig. 2, InceptionResNetV2, InceptionV3, and MobileNetV1 consistently outperformed the remaining models on the ISIC dataset, indicating their effectiveness in learning discriminative representations for melanoma classification under identical experimental conditions.

ii) Skin Cancer (Benign vs. Malignant) Dataset

Table 2. Performance Comparison on Skin Cancer (Benign vs. Malignant) Dataset

Model	**Accuracy**	**Macro F1-Score**	**Recall**	**Precision**
Xception	0.897	0.897	0.897	0.897
EfficientNet V2B0	0.870	0.870	0.870	0.875
EfficientNet V2B1	0.880	0.880	0.880	0.882
EfficientNet V2B2	0.894	0.893	0.894	0.896
EfficientNet V2B3	0.862	0.862	0.862	0.862
InceptionResNet V2	0.882	0.882	0.882	0.882
InceptionV3	0.882	0.882	0.882	0.882
MobileNetV1	0.882	0.882	0.882	0.882
MobileNetV2	0.836	0.836	0.836	0.837
MobileNetV3	0.886	0.886	0.886	0.889
VGG19	0.895	0.896	0.895	0.899
ResNet50	0.873	0.872	0.873	0.874
EDA-ResNet50	0.877	0.877	0.877	0.878
ConvNeXt-Tiny	0.877	0.877	0.877	0.878

Table 2 presents the comparative classification performance of the evaluated models on the Skin Cancer dataset. Xception achieved the highest classification accuracy of 89.7% while maintaining superior macro F1-score, recall, and precision, demonstrating the most balanced performance for binary skin lesion classification.

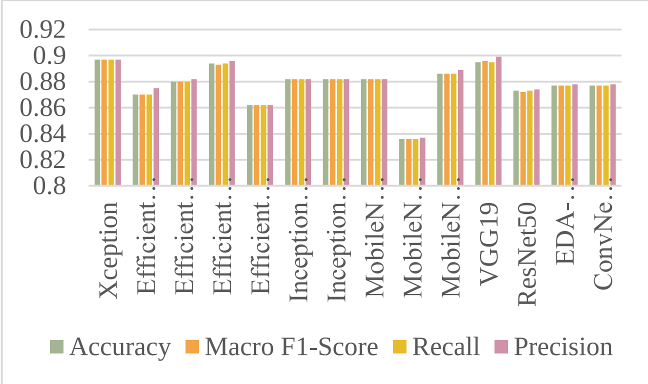

Figure 3. Performance Comparison on Skin Cancer Dataset

Figure 3 shows that Xception consistently achieved the best overall performance among the evaluated architectures, indicating its capability to effectively distinguish benign and malignant skin lesions with improved classification reliability.

iii) HAM10000 Dataset

Table 3. Performance Comparison on HAM10000 Dataset

Model	**Accuracy**	**Macro F1-Score**	**Recall**	**Precision**
Xception	0.823	0.817	0.823	0.812
EfficientNet V2B0	0.792	0.792	0.792	0.795
EfficientNet V2B1	0.803	0.793	0.803	0.790
EfficientNet V2B2	0.807	0.800	0.807	0.798
EfficientNet V2B3	0.818	0.807	0.818	0.809
InceptionResNet V2	0.793	0.790	0.793	0.796
InceptionV3	0.793	0.790	0.793	0.796
MobileNetV1	0.793	0.790	0.793	0.796
MobileNetV2	0.788	0.782	0.788	0.780
MobileNetV3	0.806	0.804	0.806	0.802
VGG19	0.785	0.758	0.785	0.765
ResNet50	0.786	0.784	0.786	0.788
EDA-ResNet50	0.784	0.779	0.784	0.777
ConvNeXt-Tiny	0.762	0.730	0.762	0.740

Table 3 compares the performance of the evaluated deep learning models on the HAM10000 dataset. Xception achieved the highest classification accuracy of 82.3%, together with the best overall balance between macro F1-score, recall, and precision, demonstrating robust performance for multi-class dermoscopic image classification.

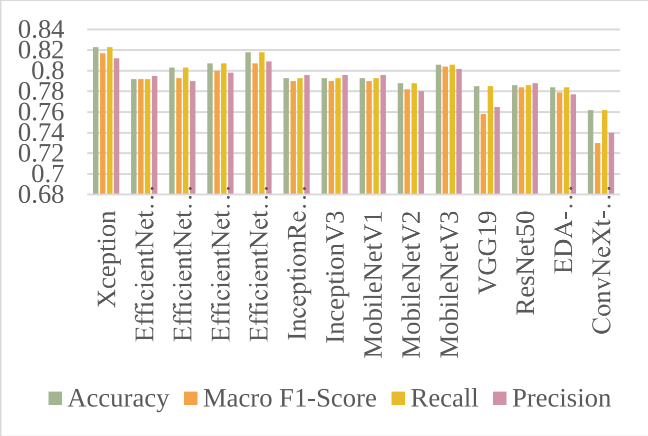

Figure 4. Performance Comparison on HAM10000 Dataset

As presented in Fig. 4, Xception consistently outperformed the remaining architectures on the HAM10000 dataset, highlighting its ability to learn discriminative lesion features across multiple skin disease categories while maintaining stable classification performance.

iv) Blood Cell Images Dataset

Table 4. Performance Comparison on Blood Cell Images Dataset

Model	**Accuracy**	**Macro F1-Score**	**Recall**	**Precision**
Xception	0.987	0.987	0.987	0.987
EfficientNet V2B0	0.979	0.979	0.979	0.980
EfficientNet V2B1	0.986	0.986	0.986	0.986
EfficientNet V2B2	0.985	0.985	0.985	0.985
EfficientNet V2B3	0.979	0.979	0.979	0.979
InceptionResNet V2	0.984	0.984	0.984	0.984
InceptionV3	0.984	0.984	0.984	0.984
MobileNetV1	0.984	0.984	0.984	0.984
MobileNetV2	0.973	0.973	0.973	0.974
MobileNetV3	0.983	0.983	0.983	0.984
VGG19	0.980	0.980	0.980	0.981
ResNet50	0.986	0.986	0.986	0.986
EDA-ResNet50	0.979	0.979	0.979	0.979
ConvNeXt-Tiny	0.979	0.979	0.979	0.979

Table 4 summarizes the comparative performance of the evaluated deep learning models on the Blood Cell Images dataset. Xception achieved the highest classification accuracy of 98.7%, together with superior macro F1-score, recall, and precision, demonstrating highly accurate and consistent blood cell image classification.

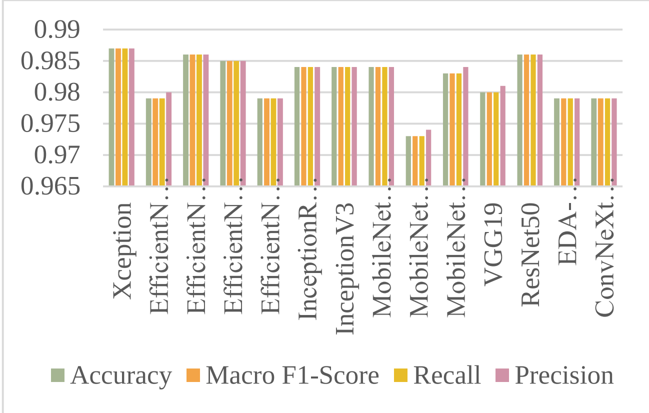

Figure 5. Performance Comparison on Blood Cell Images Dataset

Figure 5 demonstrates that Xception consistently achieved the best overall classification performance across all evaluation metrics, confirming its strong feature extraction capability and excellent generalization for automated blood cell image classification.

C) Comparative Performance Summary

The comparative experimental evaluation demonstrates that different deep learning architectures exhibit varying levels of performance across the four medical image datasets. InceptionResNetV2, InceptionV3, and MobileNetV1 achieved the highest classification accuracy of 94.3% on the ISIC 2019–2020 dataset, whereas Xception consistently outperformed the remaining models on the Skin Cancer (89.7%), HAM10000 (82.3%), and Blood Cell Images (98.7%) datasets. These results indicate that Xception provides the most balanced combination of classification accuracy, robustness, and generalization across diverse medical imaging tasks. Based on its superior overall performance and consistent predictive capability, Xception was selected as the final deployment model for the developed web-based medical image classification framework, supporting accurate, interpretable, and reliable computer-aided healthcare applications.

D) Deployment Results

The developed framework is deployed as a secure web-based medical image classification application to demonstrate its practical usability. After successful user authentication, the application allows users to upload skin lesion and blood cell images for automated analysis. The uploaded images are validated, preprocessed, and classified using the deployed Xception model. The system presents prediction confidence, Grad-CAM visualizations, disease information, and remedy recommendations through an interactive interface, enabling accurate, interpretable, and user-friendly computer-aided medical image classification.

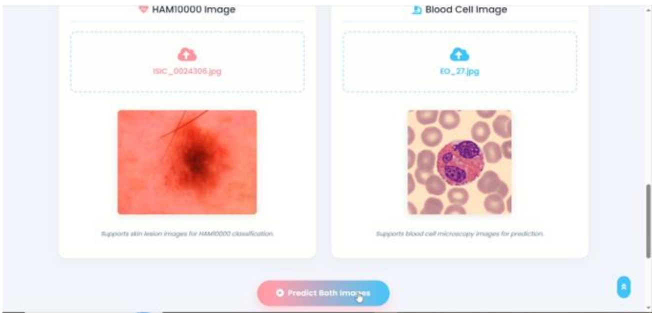

Fig.6 Upload the input Image – Case 1

Fig. 6 illustrates the input interface of the developed web application, where users upload HAM10000 skin lesion and blood cell images for automated analysis. The interface provides a simple, organized, and user-friendly environment for initiating disease prediction through image submission.

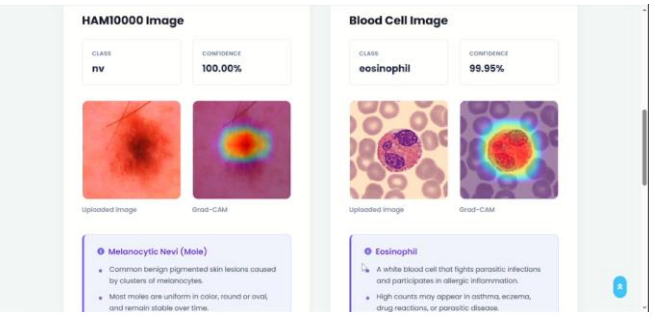

Fig.7 Predicted Result – Case 1

Fig. 7 presents the prediction results for both the HAM10000 skin lesion image and the blood cell image. The developed system correctly classifies the images with 100.00% and 99.95% confidence, respectively, while Grad-CAM visualizations highlight the most influential regions, providing accurate, reliable, and interpretable diagnostic predictions.

Fig.8 Upload the input Image – Case 2

Fig. 8 illustrates the input interface of the developed application, where users upload a HAM10000 skin lesion image and a blood cell image before prediction. The interface provides a simple, organized, and user-friendly environment for initiating automated medical image classification and analysis.

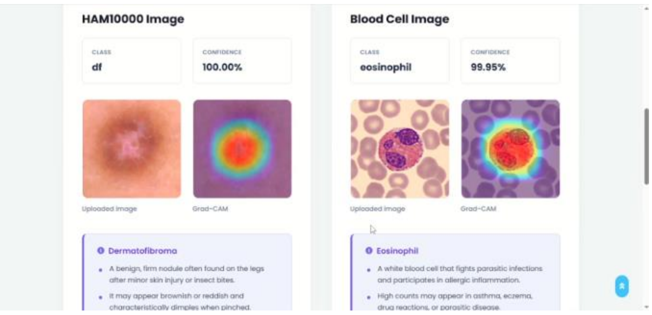

Fig.9 Predicted Result – Case 2

Fig. 9 presents the prediction results for the HAM10000 skin lesion image and the blood cell image. The developed system correctly identifies the classes with 100.00% and 99.95% confidence, respectively, while Grad-CAM visualizations highlight the most influential regions, providing accurate, interpretable, and reliable diagnostic predictions for both medical images.

V. Conclusion

In conclusion, the developed framework provides an accurate, reliable, and interpretable solution for automated medical image classification, supporting timely clinical diagnosis and decision-making for skin lesion and blood cell image analysis. The framework was developed using four publicly available medical image datasets, namely HAM10000, ISIC 2019–2020 Melanoma, Skin Cancer (Benign vs. Malignant), and Blood Cell Images, together with comprehensive image preprocessing, including hair removal, image resizing, data augmentation, and image normalization, to improve feature representation and classification performance. Multiple state-of-the-art deep learning architectures, including EfficientNetV2, InceptionResNetV2, InceptionV3, MobileNet, VGG19, ResNet50, Xception, ConvNeXt-Tiny, and the attention-enhanced EDA-ResNet50 integrating Convolutional Block Attention Module and Coordinate Attention, were comparatively evaluated under identical experimental settings. Based on the experimental results, Xception demonstrated the most consistent overall performance and was selected as the final deployment model, achieving a maximum classification accuracy of 98.70% on the Blood Cell Images dataset. Furthermore, the integration of Grad-CAM-based explainability, prediction confidence estimation, and deployment through a Flask-based web application with SQLite authentication enhances system transparency, usability, and accessibility. Overall, the developed framework provides an accurate, dependable, and interpretable computer-aided medical image classification solution for intelligent healthcare support.

The developed framework can be further extended by incorporating larger and more diverse medical image datasets to improve model generalization across different disease categories and imaging conditions. Integration of lightweight deep learning architectures can facilitate deployment on mobile and edge devices for real-time healthcare applications. Additional Explainable Artificial Intelligence techniques and multimodal clinical information may further enhance prediction reliability and clinical interpretability. Furthermore, integration with cloud-based healthcare platforms and telemedicine services can support remote consultation, continuous monitoring, and broader accessibility, extending the applicability of the framework for intelligent medical image classification and computer-aided clinical decision support.

References

[1] M. C. F. Simões, J. J. S. Sousa, and A. A. C. C. Pais, “Skin cancer and new treatment perspectives: A review,” Cancer Letters, vol. 357, no. 1, pp. 8–42, 2015.

[2] X. He, E.-L. Tan, H. Bi, X. Zhang, S. Zhao, and B. Lei, “Fully transformer network for skin lesion analysis,” Medical Image Analysis, vol. 77, Art. no. 102357, 2022.

[3] Dr.B Sateesh Kumar, Md Sirajuddin, Energy-Efficient and Secure Framework for IoMT-Based E-Healthcare Systems Using Intelligent Routing, IJCSE, ISBN No.2347-2693, Vol No.11, pp.10-16, March, 2023

[4] I. Pacal, B. Özdemir, J. Zeynalov, H. Gasimov, and N. Pacal, “A novel CNN-ViT-based deep learning model for early skin cancer diagnosis,” Biomedical Signal Processing and Control, vol. 104, Art. no. 107627, 2025.

[5] B. Özdemir and I. Pacal, “A robust deep learning framework for multiclass skin cancer classification,” Scientific Reports, vol. 15, no. 1, p. 4938, 2025.

[6] Dr.B.Sateesh Kumar, A Benchmark Survey on High Resolution Remote Sensing Images Learning Techniques, International Journal of Techno Engineering, IJTE, ISBN No.2057-5688, Vol No.XV, Issue No.III, pp.23-32, IJTE, JULY, 2023

[7] B. C. R. S. Furriel et al., “Artificial intelligence for skin cancer detection and classification for clinical environment: A systematic review,” Frontiers in Medicine, vol. 10, Art. no. 1305954, 2024.

[8] H. C. Reis and V. Turk, "Fusion of transformer attention and CNN features for skin cancer detection," Applied Soft Computing, vol. 164, Art. no. 112013, 2024.

[9] Dr.B Sateesh Kumar, Md Sirajuddin, Collaborative security schemes for Wireless Sensor Networks, ICCCE 2021, ISBN No.978-981-16-7984-1, pp.978-981, Springer, May, 2022

[10] T. Tuncer et al., "A lightweight deep convolutional neural network model for skin cancer image classification," Applied Soft Computing, vol. 162, Art. no. 111794, 2024.

[11] A. Alotaibi and D. AlSaeed, "Skin cancer detection using transfer learning and deep attention mechanisms," Diagnostics, vol. 15, no. 1, p. 99, 2025.

[12] S. M. Yasir and H. Kim, "Attention-based deep feature aggregation network for skin lesion classification," Electronics, vol. 14, no. 12, p. 2364, 2025.

[13] C. Shu, L. Yu, S. Tian, and X. Shi, "MSMA: A multi-stage and multi-attention algorithm for the classification of multimodal skin lesions," Biomedical Signal Processing and Control, vol. 93, Art. no. 106180, 2024.

[14] S. Younas et al., "Attention-based inception-residual CNN model for skin cancer diagnosis," Information, vol. 16, no. 2, p. 120, 2025.

[15] G. Akilandasowmya et al., "Skin cancer diagnosis: Leveraging deep hidden features and ensemble classifiers for early detection and classification," Biomedical Signal Processing and Control, vol. 88, Art. no. 105306, 2024.

[16] E. H. Houssein et al., "An effective multiclass skin cancer classification approach based on deep convolutional neural network," Cluster Computing, vol. 27, 2024.

[17] M. Toğaçar, Z. Cömert, and B. Ergen, "Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks," Chaos, Solitons & Fractals, vol. 144, Art. no. 110714, 2021.

[18] I. Aruk, I. Pacal, and A. N. Toprak, "A novel hybrid ConvNeXt-based approach for enhanced skin lesion classification," Expert Systems with Applications, vol. 283, Art. no. 127721, 2025.

[19] M. Shakya, R. Patel, and S. Joshi, "A comprehensive analysis of deep learning and transfer learning techniques for skin cancer classification," Scientific Reports, vol. 15, no. 1, p. 4633, 2025.

[20] U. A. Lyakhova and P. A. Lyakhov, "Systematic review of approaches to detection and classification of skin cancer using artificial intelligence," Computer Biology and Medicine, vol. 178, Art. no. 108742, 2024.

[21] M. Abdar et al., "Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning," Computer Biology and Medicine, vol. 135, Art. no. 104418, 2021.

[22] V. Anand et al., "An enhanced transfer learning based classification for diagnosis of skin cancer," Diagnostics, vol. 12, no. 7, p. 1628, 2022.

[23] S. Bechelli and J. Delhommelle, "Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images," Bioengineering, vol. 9, no. 3, p. 97, 2022.

[24] T. C.-E.-A. Tai, E. Janes, C. Czarnecki, and A. Wong, "Double-condensing attention condenser: Leveraging attention in deep learning to detect skin cancer from skin lesion images," Sensors, vol. 24, no. 22, p. 7231, 2024.

[25] 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. 1, pp. 1–9, 2018.