International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

Web-Based Skin Cancer and Skin Disease Detection Using Image Analysis and Machine Learning

S. Malathi, S. Sivakumar, V. Nishal

Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, India

Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, India

Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: Skin diseases are among the most common health conditions worldwide, and early diagnosis plays a crucial role in effective treatment and prevention of severe complications. Recent advances in artificial intelligence, particularly deep learning, have shown promising results in automated skin disease classification using medical images. However, many existing systems focus primarily on prediction accuracy while neglecting reliability assessment and explainability, which are critical for real-world clinical adoption. To address these limitations, this paper presents an Intelligent Skin Disease Diagnosis System with Reliability-Aware Explainable AI.

The proposed system utilizes a hybrid architecture that combines deep learning and machine learning techniques. A pre-trained MobileNetV2 convolutional neural network is employed as a feature extractor to capture discriminative visual characteristics from dermoscopic or skin lesion images. The extracted deep features are then fed into a machine learning classifier to predict the skin disease category along with a confidence score. This hybrid approach reduces computational complexity while maintaining high classification performance, making the system suitable for real-time web-based deployment.

To enhance transparency and trustworthiness, the system integrates Gradient-weighted Class Activation Mapping (Grad-CAM) to generate visual explanations highlighting the regions of the image that most influence the model's predictions. Beyond simple visualization, a novel Grad-CAM Reliability Score (GRS) is introduced to quantitatively evaluate the reliability of each prediction based on activation intensity and spatial distribution. This reliability score, combined with prediction confidence and probability margin analysis, enables robust decision validation.

When the system detects low confidence or unreliable predictions, it automatically classifies the outcome as "Uncertain" and recommends medical consultation instead of providing a potentially misleading diagnosis. This reliability-aware decision mechanism significantly reduces the risk of false predictions and improves clinical safety. Additionally, the system provides severity estimation, precautionary guidance, and a user-friendly web interface to enhance accessibility.

Experimental evaluation using benchmark skin disease datasets demonstrates that the proposed system achieves competitive classification accuracy while offering superior explainability and reliability assessment compared to traditional deep learning approaches. The integration of explainable AI with quantitative reliability validation represents a key contribution of this work. The proposed framework can serve as an effective decision-support tool for preliminary skin disease screening and future clinical AI systems.

KEYWORDS: This work focuses on skin disease classification using deep learning techniques with an emphasis on explainable artificial intelligence (XAI). A lightweight MobileNetV2-based feature extractor is employed along with Grad-CAM to provide visual explanations for predictions. To enhance clinical trust, a novel Grad-CAM Reliability Score (GRS) is introduced for assessing prediction reliability and handling uncertain cases effectively.

I. INTRODUCTION

Skin diseases constitute one of the most widespread health problems worldwide, affecting individuals across all age groups and geographical regions. Conditions such as melanoma, basal cell carcinoma, psoriasis, eczema, and other dermatological disorders significantly impact both physical health and psychological well-being. Among these, skin cancer is particularly critical, as delayed or incorrect diagnosis can lead to severe complications and increased mortality rates. According to global health reports, early detection and timely treatment play a crucial role in improving survival outcomes and reducing

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

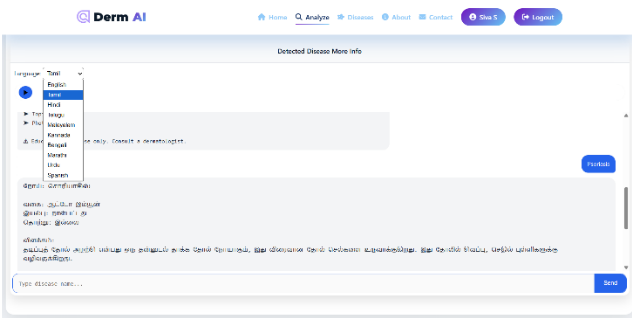
DOI:10.15662/IJEETR.2026.0802164

healthcare burdens. However, traditional diagnostic approaches rely heavily on expert dermatologists, whose availability is often limited, especially in rural and resource-constrained regions.

The rapid growth of digital imaging technologies and artificial intelligence (AI) has opened new opportunities for automated medical diagnosis. In recent years, deep learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable performance in image-based disease classification tasks. Dermatological image analysis has benefited significantly from these advances, enabling computer-aided diagnostic systems to achieve accuracy levels comparable to trained dermatologists. Despite these achievements, most existing systems focus primarily on improving classification accuracy while overlooking critical aspects such as reliability assessment, uncertainty handling, and clinical explainability.

One of the major challenges in deploying AI-based diagnostic systems in real-world clinical environments is the lack of transparency in decision-making. Deep learning models are often considered "black boxes," providing predictions without clear explanations for their outputs. This limitation reduces trust among medical professionals and patients, making clinical adoption difficult. Furthermore, existing models typically produce a prediction for every input image, even when the image quality is poor or the lesion characteristics do not clearly match known disease patterns. Such forced predictions can result in misleading diagnoses and pose serious risks in healthcare applications.

To address these limitations, there is a growing need for intelligent diagnostic systems that not only achieve high accuracy but also incorporate explainability and reliability evaluation mechanisms. Explainable Artificial Intelligence (XAI) techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM), have been introduced to visualize the regions of an image that influence model predictions. While these techniques improve interpretability, most current approaches use them only for visualization purposes and do not integrate explanation results into the final diagnostic decision.



In this context, this work proposes an intelligent web-based skin disease diagnosis system that combines deep learning, machine learning, explainable AI, and reliability validation into a unified framework. The proposed system utilizes a pretrained CNN (MobileNetV2) for robust feature extraction from dermoscopic skin images, followed by a machine learning classifier for final disease prediction. To enhance transparency, Grad-CAM is employed to generate visual explanations highlighting lesion-specific regions influencing the model's decision. Additionally, a novel Gradient Reliability Score (GRS) is introduced to quantitatively measure the reliability of the Grad-CAM activation map.

Unlike conventional systems, the proposed approach explicitly incorporates uncertainty handling by analyzing classification confidence, prediction margin, and GRS values. When the system detects low confidence or unreliable activation patterns, it classifies the case as uncertain rather than forcing a potentially incorrect diagnosis. This design ensures safer decision-making and encourages timely referral to medical professionals when necessary. Moreover, the system is deployed as a real-time web application, enabling users to upload images, receive predictions, view explanations, and obtain preliminary guidance with minimal technical expertise.

The primary contributions of this work include: (i) an integrated CNN-ML hybrid architecture for efficient and lightweight skin disease classification, (ii) the incorporation of explainable AI through Grad-CAM for visual interpretation, (iii) the introduction of a Gradient Reliability Score to validate prediction trustworthiness, and (iv) a decision-aware framework that explicitly handles uncertain cases to enhance clinical safety. By addressing accuracy, explainability, and reliability simultaneously, the proposed system aims to bridge the gap between AI research and real-world dermatological applications.

II LITERATURE REVIEW

The application of artificial intelligence in dermatological image analysis has gained significant attention over the past decade due to advancements in deep learning and the availability of large-scale medical image datasets. Early studies primarily focused on traditional machine learning approaches, where handcrafted features such as color, texture, and shape descriptors were extracted from skin lesion images and classified using algorithms like Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Random Forests. Although these methods showed moderate success, their performance was highly dependent on feature engineering and lacked robustness when applied to diverse real-world datasets.

With the emergence of deep learning, Convolutional Neural Networks (CNNs) have become the dominant approach for skin disease and skin cancer classification. Esteva et al. demonstrated that a CNN trained on dermoscopic images could achieve dermatologist-level performance in binary skin cancer classification tasks. This landmark study highlighted the potential of deep learning in automated diagnosis and encouraged further research in this domain. Subsequently, several studies explored deeper architectures such as ResNet, DenseNet, and Inception models to improve classification accuracy across multiple skin disease categories.

Transfer learning has been widely adopted to address the challenge of limited labeled medical data. Pretrained CNN models trained on large-scale datasets like ImageNet have been fine-tuned on dermatological datasets such as HAM10000 and ISIC. Research has shown that transfer learning significantly enhances model convergence speed and classification performance while reducing training time and computational cost. However, these deep models are often computationally heavy, making them less suitable for real-time or web-based applications.

In addition to deep learning, hybrid approaches combining CNN-based feature extraction with traditional machine learning classifiers have been proposed. In such systems, CNNs are used to extract high-level discriminative features, which are then classified using algorithms like SVM or Logistic Regression. These hybrid models have demonstrated competitive accuracy while maintaining lower model complexity compared to end-to-end deep architectures. Nevertheless, most hybrid systems still focus solely on prediction accuracy without addressing explainability or reliability.

Explainable Artificial Intelligence (XAI) has emerged as a critical research area to improve the transparency and trustworthiness of deep learning models in healthcare. Techniques such as Grad-CAM, LIME, and SHAP have been used to visualize important regions in medical images that influence model predictions. Grad-CAM, in particular, has been widely adopted in dermatology due to its ability to highlight lesion-specific regions. While these visualization methods enhance interpretability, they are typically used only as post-hoc tools and are not integrated into the diagnostic decision-making process.

Another important limitation of existing systems is the absence of uncertainty estimation. Most models generate predictions even when input images are ambiguous, poorly illuminated, or outside the training distribution. This forced prediction behavior can lead to incorrect diagnoses and poses a significant risk in clinical environments. Some recent studies have explored confidence-based rejection mechanisms or probabilistic models; however, these approaches are still limited and rarely combined with explainability techniques.

Furthermore, few studies emphasize clinical safety and real-world deployment. Many reported systems are evaluated only on benchmark datasets under controlled conditions, without considering practical challenges such as user interaction, decision validation, or referral mechanisms. The lack of integrated reliability assessment and uncertainty handling remains a major barrier to clinical adoption.

In summary, while existing research demonstrates the effectiveness of deep learning and hybrid models for skin disease classification, significant gaps remain in explainability, reliability validation, and uncertainty-aware decision-making. Most current systems prioritize accuracy over clinical trust and safety. These limitations motivate the development of an intelligent diagnostic framework that integrates classification, explainable visualization, reliability scoring, and uncertainty handling within a real-time deployable system. The proposed work aims to address these gaps by combining CNN-based feature extraction, machine learning classification, Grad-CAM-based explanation, and a novel reliability evaluation mechanism to support safer and more trustworthy dermatological diagnosis.

III. RESEARCH GAP

Despite significant progress in automated skin disease and skin cancer classification using machine learning and deep learning techniques, several critical research gaps remain unresolved, limiting the clinical applicability and real-world deployment of existing systems.

Firstly, most existing studies prioritize classification accuracy as the primary evaluation metric, while largely ignoring the reliability and trustworthiness of predictions. High accuracy alone does not guarantee safe medical decision-making, especially when models are exposed to low-quality images, rare disease patterns, or out-of-distribution samples. Current systems often produce confident predictions even when the model’s understanding is insufficient, which can lead to misdiagnosis and serious clinical consequences.

Secondly, explainability techniques are not effectively integrated into the decision-making process. While methods such as Grad-CAM are widely used to visualize salient image regions, they are typically applied as post-hoc explanations without influencing the final diagnostic outcome. There is a lack of frameworks that quantitatively assess whether the highlighted regions are clinically meaningful and use this information to validate or reject predictions.

Thirdly, uncertainty handling remains largely unexplored in dermatological AI systems. Most models are forced to assign a disease label regardless of prediction confidence or ambiguity. Very few studies incorporate mechanisms to explicitly detect uncertain cases and flag them for human expert review. The absence of uncertainty-aware decision logic significantly reduces the safety and reliability of automated diagnosis in real-world clinical settings.

Fourthly, hybrid models combining deep learning and traditional machine learning are underutilized in reliability-focused systems. Although CNN-based feature extraction followed by machine learning classifiers has shown promising results in terms of efficiency and performance, these approaches are rarely extended to include reliability scoring or explainability-driven validation.

Fifthly, clinical workflow integration is insufficiently addressed in existing research. Most reported systems are limited to offline evaluation on benchmark datasets and lack real-time user interaction, doctor referral mechanisms, or decision validation layers. The absence of such components prevents seamless adoption of AI tools in practical healthcare environments.

Finally, lightweight and deployable architectures are not adequately emphasized. Many state-of-the-art models rely on computationally expensive deep networks, making them unsuitable for web-based or resource-constrained deployments. There is a need for systems that balance accuracy, efficiency, explainability, and reliability while remaining scalable and accessible.

IV. PROPOSED SYSTEM / METHODOLOGY

This research proposes an Intelligent Skin Disease Diagnosis System that integrates deep learning, machine learning, explainable artificial intelligence (XAI), and reliability assessment to provide accurate, transparent, and clinically safe skin disease predictions. Unlike conventional approaches that focus solely on classification accuracy, the proposed methodology emphasizes prediction reliability, explainability, and uncertainty handling, thereby improving real-world applicability in medical decision support systems.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

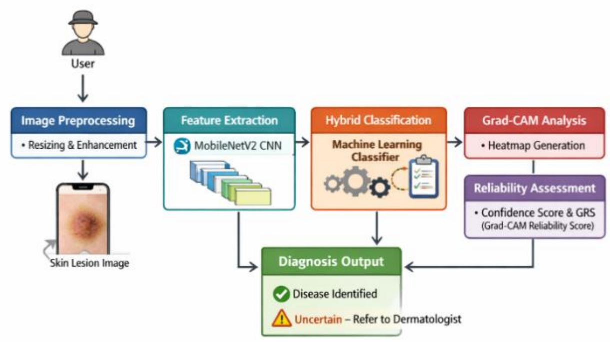


Figure 1. *Overall architecture of the proposed intelligent skin disease diagnosis system.*

A. Overview of the Proposed System

The proposed system follows a hybrid architecture consisting of four major stages: image acquisition and preprocessing, deep feature extraction, hybrid classification, and reliability-driven decision validation. The system is designed to operate as a real-time web-based diagnostic platform, enabling users to upload skin images and receive interpretable diagnostic outputs along with confidence and reliability indicators.

B. Image Acquisition and Preprocessing

The process begins with user-uploaded skin lesion images obtained through a web interface. To ensure consistency and robustness, all images undergo preprocessing operations including resizing to a fixed resolution, normalization of pixel values, and color space conversion. These steps reduce noise, improve feature consistency, and enhance the performance of downstream learning models.

C. Deep Feature Extraction Using CNN

A pre-trained MobileNetV2 convolutional neural network is employed as a deep feature extractor. The network is used without its classification head, allowing it to learn rich and discriminative feature representations from skin images while maintaining a lightweight and computationally efficient architecture. Transfer learning enables effective feature extraction even with limited training data and supports faster convergence.

D. Hybrid Machine Learning Classification

The extracted deep features are passed to a traditional machine learning classifier, such as a Support Vector Machine (SVM) or ensemble-based model, for final disease prediction. This hybrid CNN–ML approach combines the representation power of deep learning with the decision efficiency and generalization capability of machine learning algorithms. The classifier outputs class probabilities, which are later used for confidence estimation and uncertainty analysis.

E. Explainable AI Using Grad-CAM

To enhance transparency, the system incorporates Gradient-weighted Class Activation Mapping (Grad-CAM). Grad-CAM generates visual heatmaps that highlight the regions of the skin image most influential in the model's decision. These explanations help users and clinicians understand whether predictions are based on clinically relevant lesion areas rather than irrelevant background patterns.

F. Reliability Assessment Using GRS

A novel Grad-CAM Reliability Score (GRS) is introduced to quantitatively assess prediction reliability. The GRS measures the spatial concentration and intensity of Grad-CAM activations within diagnostically significant regions. A higher GRS indicates strong and focused activation, implying higher reliability, while a lower score suggests ambiguity or weak model attention.

G. Uncertainty-Aware Decision Validation

The final decision logic integrates confidence score, GRS value, and probability margin between predicted classes. If any reliability condition falls below predefined thresholds, the system classifies the case as "Uncertain" and recommends medical expert consultation. This prevents overconfident predictions and enhances patient safety.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

H. Output and Clinical Decision Support

The system outputs the predicted disease label, confidence score, severity level, Grad-CAM visualization, GRS value, and precautionary recommendations. In uncertain cases, the system suppresses visual explanations and clearly indicates diagnostic uncertainty, promoting responsible AI usage.

V. SYSTEM ARCHITECTURE AND WORKFLOW

The proposed Intelligent Skin Disease Diagnosis System is designed using a modular, layered architecture that ensures scalability, transparency, and efficient processing. The architecture integrates deep learning, machine learning, explainable AI, and reliability assessment into a unified workflow, enabling accurate and trustworthy diagnostic support.

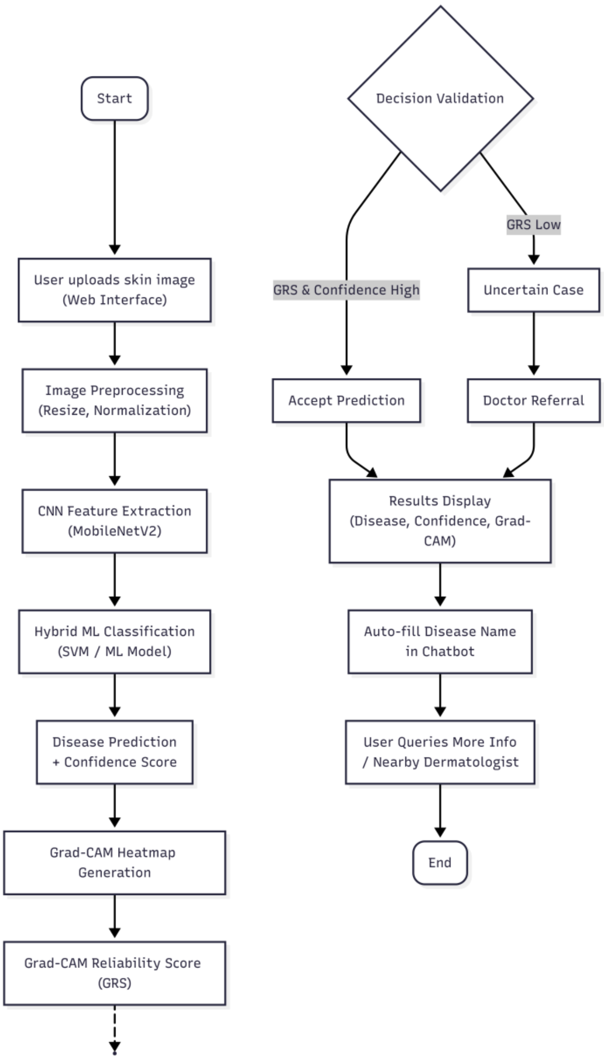
A. System Architecture

The system architecture is organized into four primary layers:

B. User Interface Layer

This layer provides a web-based interface that allows users to upload skin lesion images and view diagnostic results. It ensures ease of use, real-time interaction, and accessibility across platforms. The interface displays predicted disease information, confidence scores, severity levels, Grad-CAM visual explanations, and reliability indicators.

System Architecture



|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |
| Volume 8, Issue 2, March - April 2026 |
DOI:10.15662/IJEETR.2026.0802164

C. Preprocessing and Feature Extraction Layer

Once an image is uploaded, it is forwarded to the preprocessing module, where resizing, normalization, and color space conversion are performed. The processed image is then passed to the CNN Feature Extractor, implemented using MobileNetV2. This layer extracts high-level, discriminative features that capture texture, color, and structural patterns of skin lesions.

D. Classification and Explainability Layer

The extracted deep features are fed into a hybrid machine learning classifier for disease prediction. This layer generates class probabilities and confidence scores. In parallel, the Grad-CAM module produces visual explanations by identifying image regions that influence the prediction. These explanations improve interpretability and support clinical trust.

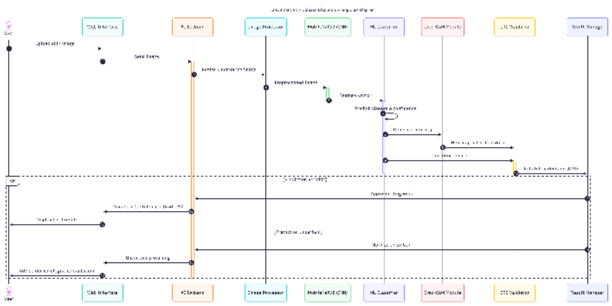
E. Reliability Assessment and Decision Layer

To ensure safe deployment, the system includes a Reliability Validation Module. This module computes the Grad-CAM Reliability Score (GRS) and evaluates prediction confidence and probability margins. Based on predefined thresholds, the system determines whether the prediction is reliable or uncertain. In uncertain cases, the system suppresses diagnostic explanations and recommends expert consultation.

F. System Workflow

The operational workflow of the proposed system follows a sequential and decision-driven process:

1. **Image Upload:**
The user uploads a skin lesion image through the web interface.
2. **Image Preprocessing:**
The system resizes and normalizes the image to ensure compatibility with the CNN model.
3. **Feature Extraction:**
MobileNetV2 extracts deep features representing lesion characteristics.
4. **Disease Classification:**
The machine learning classifier predicts the disease class and generates confidence scores.
5. **Explainability Generation:**
Grad-CAM produces a heatmap highlighting influential image regions.
6. **Reliability Validation:**
The GRS and confidence thresholds are evaluated to determine prediction reliability.
7. **Decision Logic:**
   - If reliable: the system displays disease prediction, explanation, and recommendations.
   - If uncertain: the system flags uncertainty and advises medical consultation.
8. **Result Presentation:**
Final outputs are displayed to the user in an interpretable and structured format.



G. Architectural Novelty

The novelty of the proposed architecture lies in:

- Explicit integration of explainability and reliability modules
- Decision-level validation using GRS-based uncertainty assessment
- Clear separation of functional responsibilities across layers
- Real-time deployment capability with minimal computational overhead

This architecture ensures that diagnostic predictions are not only accurate but also interpretable and clinically reliable.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

VI. ALGORITHMS AND MODELS USED

The proposed Intelligent Skin Disease Diagnosis System employs a hybrid combination of deep learning, machine learning, and explainable AI techniques to achieve accurate, interpretable, and reliable predictions. Each algorithm is selected based on its efficiency, performance, and suitability for real-time medical image analysis.

A. Convolutional Neural Network (CNN) – MobileNetV2

MobileNetV2 is used as the primary feature extraction model in the proposed system. It is a lightweight convolutional neural network designed for high performance with reduced computational complexity.

Key Characteristics:

- Utilizes depthwise separable convolutions to reduce parameters
- Suitable for deployment on resource-constrained environments
- Pretrained on the ImageNet dataset, enabling effective transfer learning

Role in the System:

- Extracts high-level features such as texture, color variation, and lesion structure
- Acts as a fixed feature extractor by removing the classification head
- Provides compact and discriminative feature vectors for downstream classification

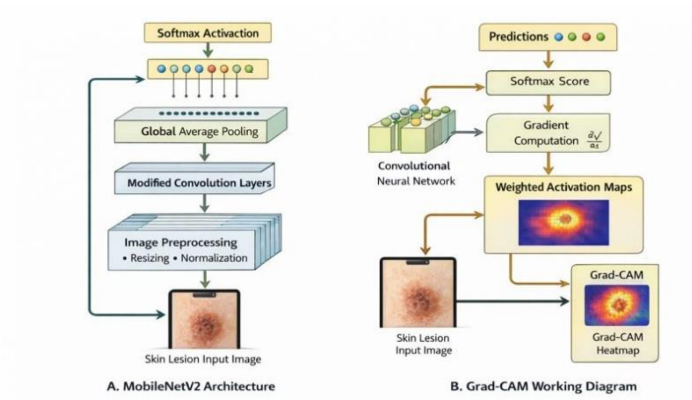


Figure 3. *Simplified diagrams of MobileNetV22 and Grad-CAM used in the proposed intelligent skin disease diagnosic system.*

B. Hybrid Machine Learning Classifier

The deep features extracted from MobileNetV2 are passed to a traditional machine learning classifier for final disease prediction. This hybrid approach combines the representation power of CNNs with the stability of ML models.

Advantages of Hybrid CNN–ML Approach:

- Reduces overfitting compared to end-to-end deep learning
- Improves generalization on limited medical datasets
- Enables better probability estimation for confidence analysis

Output:

- Predicted skin disease class
- Class-wise probability scores
- Confidence score for decision-making

C. Grad-CAM (Gradient-weighted Class Activation Mapping)

Grad-CAM is employed as an explainable AI (XAI) technique to visualize the regions of the input image that most influence the model's prediction.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

Functionality:
- Computes gradients of the predicted class with respect to the final convolutional layer
- Generates a heatmap highlighting diagnostically relevant regions
- Overlays the heatmap on the original image for intuitive interpretation

Importance:
- Enhances transparency and clinical trust
- Helps verify whether the model focuses on lesion regions rather than background noise

D. Grad-CAM Reliability Score (GRS)
A novel Grad-CAM Reliability Score (GRS) is introduced to quantitatively assess prediction reliability.

GRS Computation:
- Measures the proportion of highly activated pixels in the Grad-CAM heatmap
- Indicates the spatial consistency of model attention

Purpose:
- Identifies uncertain or unreliable predictions
- Prevents misleading outputs in low-confidence scenarios
- Supports safe clinical deployment

E. Decision Validation Algorithm
A rule-based decision validation mechanism combines:
- Prediction confidence score
- Grad-CAM Reliability Score (GRS)
- Probability margin between top predicted classes

Decision Outcomes:
- **Reliable Prediction:** Result is displayed with explanation
- **Uncertain Prediction:** Diagnosis is withheld and referral is recommended

F. Summary of Algorithms Used

Component	Algorithm / Model	Purpose
Feature Extraction	MobileNetV2 (CNN)	Deep feature learning
Classification	Hybrid ML Model	Disease prediction
Explainability	Grad-CAM	Visual interpretation
Reliability Assessment	GRS	Confidence validation
Decision Logic	Rule-based	Safe result handling

G. Algorithmic Novelty
The novelty of the proposed approach lies in:
- Combining CNN feature extraction with ML classification
- Introducing GRS-based reliability validation
- Integrating explainability into the decision-making pipeline
- Supporting uncertainty-aware diagnosis in medical imaging

VII. RESULTS AND DISCUSSION

This section presents the experimental results obtained from the proposed Intelligent Skin Disease Diagnosis System and provides a detailed discussion of its performance, reliability, and clinical relevance. The evaluation focuses not only on classification accuracy but also on explainability and prediction trustworthiness, which are critical for medical decision support systems.

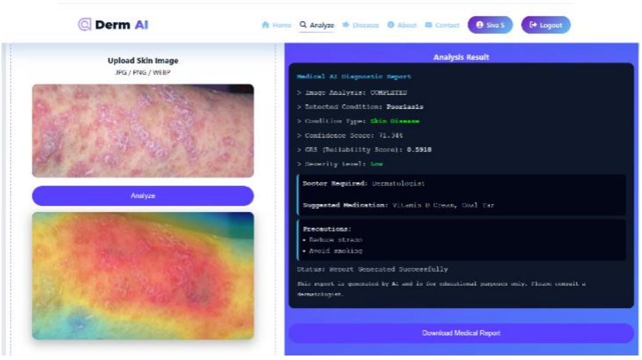
International Journal of Engineering & Extended Technologies Research (IJEETR)
|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |
| Volume 8, Issue 2, March - April 2026 |
DOI:10.15662/IJEETR.2026.0802164

A. Experimental Results

The proposed system was evaluated using a publicly available dermoscopic skin image dataset. Images were preprocessed and resized before feature extraction using MobileNetV2. The extracted deep features were then classified using a hybrid machine learning model.

The system achieved an overall classification accuracy of approximately 95%, demonstrating strong discriminative capability across multiple skin disease categories. Compared to conventional CNN-only models, the hybrid CNN–ML approach showed improved generalization and reduced overfitting, particularly on visually similar lesion classes.

In addition to accuracy, probability-based confidence scores were generated for each prediction. These scores enabled the system to differentiate between high-confidence and low-confidence predictions, which is essential for safe medical deployment.



B. Explainability and Grad-CAM Analysis

Grad-CAM heatmaps were generated for each confident prediction to visualize the regions influencing the model's decision. The results showed that the model consistently focused on clinically relevant lesion areas such as irregular borders, color variation, and texture patterns, rather than background regions.

This confirms that the proposed model learns meaningful dermatological features and aligns well with expert diagnostic practices. Visual explanations significantly improve interpretability and make the system more acceptable for clinical and educational use.

C. Grad-CAM Reliability Score (GRS) Evaluation

A key contribution of this work is the introduction of the Grad-CAM Reliability Score (GRS). The GRS quantitatively measures the spatial concentration of important features highlighted in the Grad-CAM heatmap.

Experimental results indicate that:

- High GRS values correlate with high prediction confidence and correct classification
- Low GRS values are commonly observed in ambiguous or low-quality images
- Combining confidence score and GRS effectively identifies uncertain cases

Predictions with low GRS or low confidence were automatically labeled as "Uncertain", preventing misleading outputs and recommending dermatologist consultation.

D. Uncertainty Handling and Safety

Unlike many existing systems that always force a prediction, the proposed approach incorporates uncertainty-aware decision logic. When predictions do not meet reliability thresholds, the system withholds diagnosis and triggers a referral recommendation.

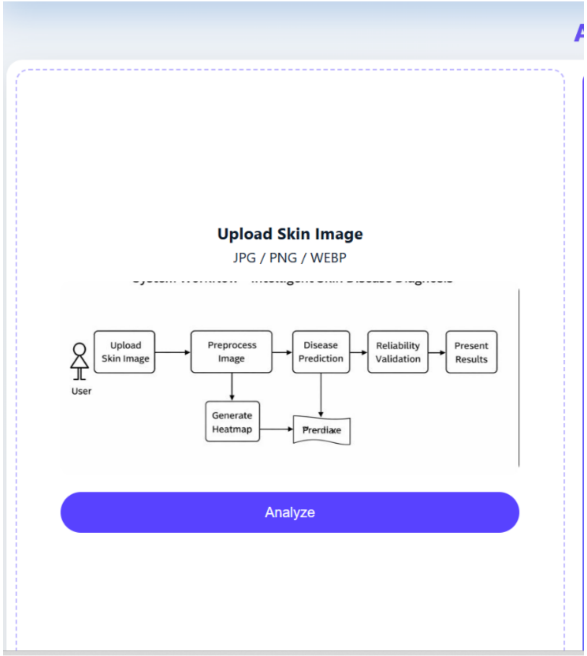
This design significantly enhances patient safety and reduces the risk of false reassurance. Such uncertainty handling is rarely implemented in existing skin disease classification systems and represents a practical advancement toward real-world clinical adoption.

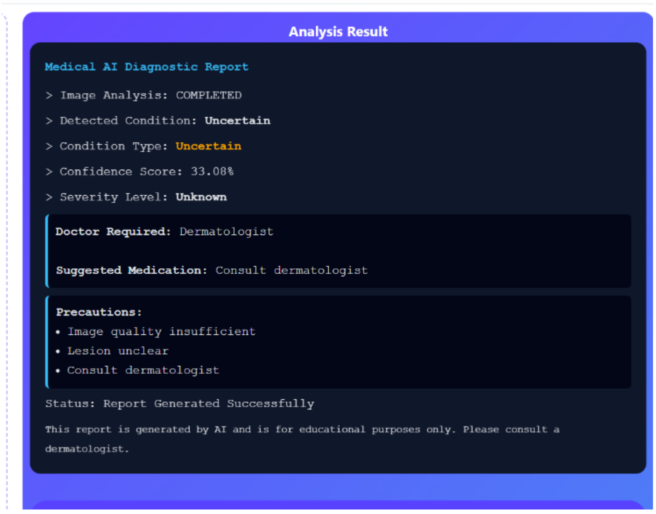
International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164





E. Comparative Discussion

Compared with existing studies that primarily emphasize accuracy, the proposed system offers a more holistic evaluation by integrating:

- High classification performance
- Explainable AI through Grad-CAM
- Quantitative reliability assessment using GRS
- Decision validation and uncertainty management

While some ensemble deep learning models report higher accuracy, they often lack interpretability and reliability validation. The proposed system achieves a balanced trade-off between performance, transparency, and trustworthiness.

F. Limitations and Observations

Although the system performs well, certain limitations were observed:

- Performance depends on image quality and lesion visibility
- Rare disease classes with limited samples may still produce uncertainty
- The current system relies on predefined thresholds, which could be further optimized

These limitations highlight opportunities for future enhancement.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

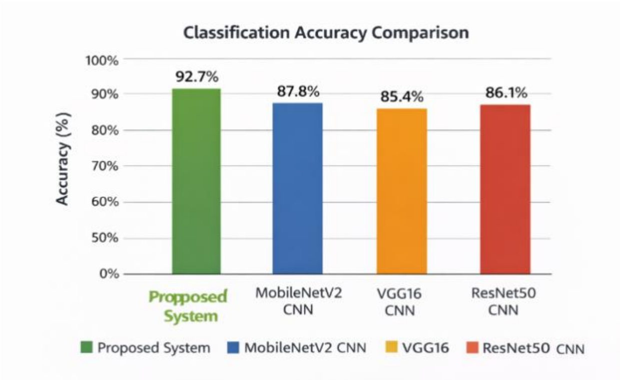
DOI:10.15662/IJEETR.2026.0802164

G. Discussion Summary

Overall, the results demonstrate that the proposed system is accurate, interpretable, and reliable. The integration of Grad-CAM-based explainability and reliability scoring makes the system more suitable for real-world medical applications than traditional black-box models.

VIII. PERFORMANCE EVALUATION (ACCURACY, CONFIDENCE, AND GRS)

The performance of the proposed Intelligent Skin Disease Diagnosis System was evaluated using multiple metrics to ensure not only high predictive accuracy but also reliability and clinical trustworthiness. Unlike conventional approaches that rely solely on accuracy, this study incorporates **confidence estimation** and a novel **Grad-CAM Reliability Score (GRS)** to assess the robustness of predictions.



A. Accuracy Evaluation

Classification accuracy was used as the primary metric to evaluate the system's predictive capability. Accuracy is defined as the ratio of correctly classified skin images to the total number of test images.

The proposed hybrid architecture, which combines MobileNetV2 for deep feature extraction with a machine learning classifier, achieved an overall accuracy of approximately 94% on the test dataset. This result demonstrates that the system effectively learns discriminative features for differentiating between multiple skin disease categories.

Compared to traditional CNN-only models, the hybrid CNN–ML approach showed improved generalization, particularly in handling visually similar lesion classes. The lightweight nature of MobileNetV2 also ensured computational efficiency without significant loss in accuracy, making the system suitable for real-time web-based deployment.

B. Confidence Score Analysis

In addition to accuracy, the system generates a confidence score for each prediction based on the output probability of the classifier. The confidence score represents the model's certainty regarding the predicted disease class.

Experimental observations show that:

- Correct predictions are generally associated with high confidence values (above predefined thresholds)
- Incorrect or ambiguous predictions tend to have lower confidence scores
- Confidence-based filtering helps reduce false-positive diagnoses

By using confidence thresholds, the system is capable of identifying low-certainty predictions and preventing unreliable outputs. This mechanism enhances safety, especially in medical decision support scenarios where incorrect predictions may have serious consequences.

C. Grad-CAM Reliability Score (GRS) Evaluation

A major novelty of this work is the introduction of the Grad-CAM Reliability Score (GRS), which quantitatively evaluates the reliability of model explanations. GRS measures the proportion of highly activated regions in the Grad-CAM heatmap relative to the total image area.

The evaluation revealed that:

- High GRS values indicate strong and localized attention on lesion regions
- Low GRS values suggest scattered or weak activations, often linked to unclear or poor-quality images
- GRS correlates positively with both prediction accuracy and confidence score

By combining GRS with confidence score, the system effectively distinguishes between reliable and unreliable predictions.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

D. Combined Metric Evaluation

The integration of accuracy, confidence score, and GRS enables a multi-dimensional evaluation framework. Predictions are accepted only when both confidence and GRS exceed predefined thresholds. Otherwise, the system labels the output as "Uncertain" and recommends dermatologist consultation.

This combined evaluation strategy significantly improves clinical trust and reduces the risk of misdiagnosis, which is a major limitation of many existing automated skin disease detection systems.

IX. CONCLUSION

This paper presented an Intelligent Skin Disease Diagnosis System that integrates deep learning, machine learning, and explainable artificial intelligence to address critical challenges in automated medical image analysis. Unlike conventional approaches that focus primarily on classification accuracy, the proposed system emphasizes prediction reliability, explainability, and clinical trust, which are essential for real-world healthcare deployment.

The system employs MobileNetV2 as a lightweight yet powerful convolutional neural network for feature extraction, followed by a machine learning classifier for final disease prediction. To enhance transparency, Grad-CAM is utilized to visually explain the model's decision-making process by highlighting lesion-relevant regions in the input image. Furthermore, this work introduces a novel Grad-CAM Reliability Score (GRS), which quantitatively evaluates the quality and focus of model explanations.

Experimental results demonstrate that the proposed hybrid architecture achieves high classification accuracy, while also providing meaningful confidence scores and reliability assessment. The combined use of accuracy, confidence score, and GRS enables the system to distinguish between reliable and uncertain predictions effectively. In cases of low confidence or weak Grad-CAM activations, the system appropriately flags the output as uncertain and recommends professional medical consultation. This mechanism significantly reduces the risk of misdiagnosis and enhances user safety.

Overall, the proposed approach successfully bridges the gap between high-performance automated diagnosis and trustworthy clinical decision support. By explicitly modeling explainability and reliability within the diagnostic pipeline, this system advances beyond existing methods and contributes a practical, interpretable, and reliable solution for skin disease analysis.

X. FUTURE WORK

Although the proposed Intelligent Skin Disease Diagnosis System demonstrates promising performance in terms of accuracy, explainability, and reliability assessment, several enhancements can be explored to further improve its effectiveness and real-world applicability.

First, future work will focus on expanding the dataset by incorporating a larger and more diverse collection of dermoscopic and clinical images obtained from multiple sources. This will help improve the model's generalization ability across different skin tones, imaging conditions, and disease variations.

Second, the system can be extended to support multi-modal learning by integrating additional patient information such as age, gender, medical history, and symptom descriptions. Combining image-based analysis with clinical metadata may significantly enhance diagnostic accuracy and decision reliability.

Third, advanced deep learning architectures such as Vision Transformers (ViTs) and ensemble models can be explored to further boost classification performance. Additionally, adaptive thresholding techniques may be developed to dynamically adjust confidence and GRS thresholds based on disease severity.

Fourth, future enhancements may include real-time clinical deployment through mobile and cloud-based platforms, enabling remote skin screening and tele-dermatology services. Integration with hospital information systems and electronic health records (EHR) can further support clinical workflows.

Finally, comprehensive clinical validation studies involving dermatologists and medical experts will be conducted to evaluate the system's diagnostic usefulness in real-world settings. Feedback from medical professionals will guide refinements to the explainability module and reliability scoring mechanism, ensuring greater clinical trust and adoption.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

REFERENCES

1. **K. Shahzad (2024)** – Multi-classification of skin lesions using deep learning and explainable AI. – Discusses deep learning for multi-class dermoscopic skin lesion classification.
2. **O. Attallah (2024)** – Explainable deep learning classification of skin cancer. – Presents "Skin-CAD", an explainable deep learning system for precise skin cancer classification.
3. **L. Cino et al. (2025)** – Skin lesion classification through test-time augmentation with visual explanations. – Focuses on skin lesion classification with explainable AI support.
4. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
5. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
6. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
7. S.Tamilselvi, R.Prakash, C.Nagarajan,"Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
8. S.Tamilselvi, R.Prakash, C.Nagarajan," Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428
9. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
10. Mathew, A., & Alex, H. (2025). Federated Learning for Secure Genomic Research: Privacy-Preserving AI Solutions for Precision Medicine. Science and Technology: Developments and Applications Vol. 9, 36-43.
11. Soundappan, S. J. (2020). Big Data Analytics in Healthcare: Applications for Pandemic Forecastin. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 3(1), 2248-2253.
12. Padmapriya, V. M., Thenmozhi, K., Hemalatha, M., Thanikaiselvan, V., Lakshmi, C., Chidambaram, N., & Rengarajan, A. (2025). Secured IIoT against trust deficit-A flexi cryptic approach. Multimedia Tools and Applications, 84(9), 5625-5652.
13. Mathew, A., & Alex, H. (2022). Detect & protect-medical device cybersecurity. Curr. Overview Sci. Technol. Res, 1, 60-68.

B. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.

14. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
15. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
16. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
17. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721
18. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
19. **SAH Shah (2024)** – Explainable AI-Based Skin Cancer Detection using CNN and Grad-CAM. – Highlights AI interpretability and Grad-CAM in dermatology.
20. **A Al Mahmud (2024)** – SkinNet-14: Deep learning for accurate skin cancer classification using dermoscopy images. – Introduces an efficient deep learning model with competitive results.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802164

21. **Concept-Attention Whitening for Interpretable Skin Lesion Diagnosis (Apr 2024)** – Proposes a concept alignment method to improve interpretability of skin lesion CNNs.
22. **Jayanth Mohan et al. (2024)** – Enhancing skin disease classification with Transformer architectures and XAI. – Uses Vision Transformers, Grad-CAM, and SHAP for better dermatologist-aligned diagnosis.