As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 03, pages 31-36.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue03.html
DOI: https://doi.org/10.5281/zenodo.18870866

Web-Based Explainable Skin Cancer Detection Using YOLOv8 and Grad-CAM

Ch Swathi | Sd Jabbar | Sk. Sajeed Ali | J Janaki Ram | Ch Yochan | T Vamsi

Department of ECE, NRI Institute of Technology, Vijayawada, AP, India

To Cite this Article
Ch Swathi, Sd Jabbar, Sk. Sajeed Ali, J Janaki Ram, Ch Yochan & T Vamsi (2026). Web-Based Explainable Skin Cancer Detection Using YOLOv8 and Grad-CAM, 12(03), 31-36. https://doi.org/10.5281/zenodo.18870866

Article Info
Received: 28 January 2026; Revised: 26 February 2026; Accepted: 02 March 2026.

KEYWORDS

Skin Cancer Detection, Skin Lesion Classification, Deep Learning, YOLOv8, Explainable Artificial Intelligence, Grad-CAM, Medical Image Analysis, Flask Web Application, Telemedicine

ABSTRACT

Skin cancer is among the most prevalent and rapidly increasing forms of cancer worldwide, where early and accurate diagnosis is essential for improving patient survival rates. Conventional diagnostic approaches rely heavily on visual inspection and expert dermatological evaluation, which can be subjective, time-consuming, and often inaccessible in remote or resource-constrained regions. To address these challenges, this work presents a web-based automated skin lesion detection and classification system using deep learning techniques. The proposed system employs the YOLOv8 model to detect and classify skin lesions from dermo-scopic and clinical images into categories such as melanoma, basal cell carcinoma, and benign lesions. A Flask-based web application enables both image upload and real-time webcam analysis, offering an intuitive and practical interface for clinical decision support. To enhance model transparency and interpretability, the system integrates Gradient-weighted Class Activation Mapping (Grad-CAM) to generate visual heatmaps that highlight diagnostically relevant regions influencing the model's predictions. The overall framework consists of image preprocessing, deep learning inference, explainability visualization, and result presentation. Experimental evaluations demonstrate that the proposed system achieves reliable and real-time performance with minimal computational overhead. Owing to its accuracy, explainability, and ease of deployment, the proposed solution is well suited for telemedicine platforms and preliminary skin cancer screening applications

INTRODUCTION

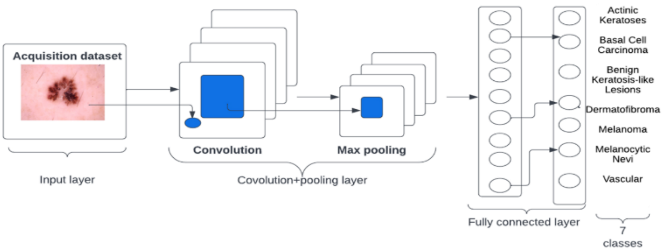
Skin cancer is one of the most common and rapidly increasing forms of cancer worldwide, with melanoma and non-melanoma skin cancers contributing significantly to global mortality rates. Early diagnosis plays a crucial role in improving patient survival;

however, traditional diagnostic methods rely heavily on visual inspection and expert dermatological assessment, making the process subjective, time-consuming, and often inaccessible in remote or resource-limited regions. These limitations have motivated the development of automated computer-aided diagnostic systems for skin lesion analysis.

Recent advances in deep learning have significantly improved performance in medical image analysis, particularly for skin lesion detection and classification. Convolutional Neural Networks (CNNs) have demonstrated strong capability in extracting discriminative visual features from dermoscopic images. However, despite their high accuracy, most deep learning models operate as black-box systems, offering limited interpretability, which hinders their acceptance in clinical decision-support systems [1], [4].

To address this challenge, explainable artificial intelligence (XAI) techniques have been increasingly incorporated into dermatological applications. Gradient-weighted Class Activation Mapping (Grad-CAM) has emerged as a widely adopted visualization technique that highlights image regions contributing most to model predictions [9]. Recent studies have demonstrated that Grad-CAM–based explanations improve transparency and align model outputs with expert dermatological reasoning [1], [5].

Several works have focused on improving classification accuracy through advanced feature learning and attention mechanisms. Adaptive spatial feature fusion methods have been proposed to capture multi-scale lesion characteristics more effectively [2]. Attention-based dual-network architectures integrating clinical metadata have further enhanced lesion detection performance by emphasizing diagnostically relevant regions [3]. While effective, such approaches often involve complex architectures and increased computational cost.



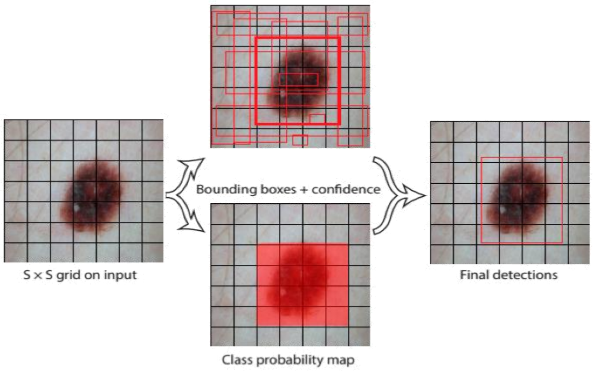


Figure 1 Conceptual Overview of Explainable Skin Lesion Detection Pipeline

Object detection–based frameworks have recently gained attention for skin lesion analysis due to their ability to perform simultaneous localization and classification. The YOLO (You Only Look Once) family of models introduced a unified real-time object detection framework with high inference speed [7]. Subsequent improvements, such as YOLOv4, enhanced accuracy while maintaining real-time performance [8]. These advancements make modern YOLO architectures suitable for medical imaging applications requiring fast and reliable predictions.

Publicly available datasets have played a vital role in advancing automated skin lesion research. The HAM10000 dataset provides a diverse and clinically validated collection of dermoscopic images representing common pigmented skin lesions, enabling robust training and evaluation of deep learning models [10].

Motivated by these developments, this work proposes a web-based, explainable skin cancer detection system using the YOLOv8 model integrated with Grad-CAM visualization. The system enables real-time lesion detection and interpretation through a Flask-based web application, aiming to provide an accurate, transparent, and deployable solution for telemedicine and preliminary skin cancer screening.

RELATED WORK

Explainable deep learning has become a major research focus in automated skin cancer diagnosis, addressing the lack of transparency in traditional CNN-based models. Matas et al. proposed an interpretable skin cancer diagnosis framework using Grad-CAM combined with expert dermatological annotations, demonstrating improved trust and clinical relevance of model predictions [1]. Similarly, Paccotacya-Yanque et al. performed a comparative analysis of explainability techniques and reported that Grad-CAM consistently produces clinically meaningful visual explanations for skin lesion classification tasks [4].

Recent studies have also explored advanced feature extraction and attention mechanisms to enhance diagnostic performance. Liu et al. introduced an adaptive spatial feature fusion technique that improves classification accuracy by effectively capturing lesion structures at multiple spatial scales [2]. Atiq and Fattah proposed a dual-network attention-based model that integrates clinical metadata with dermoscopic images, allowing the network to focus on diagnostically important lesion regions [3]. While these approaches yield improved accuracy, they often require increased computational complexity.

Explainable deep learning architectures specifically designed for skin cancer diagnosis have also been investigated. Hamim et al. developed SmartSkin-XAI, an explainable DenseNet-based framework that generates visual explanations alongside predictions, achieving high diagnostic accuracy [5]. Mukherjee et al. further enhanced model interpretability by incorporating Swish-activated convolutional layers, leading to improved feature representation and clearer activation maps [6]. However, these models are primarily evaluated in offline settings and are not optimized for real-time deployment.

Object detection models offer an alternative paradigm by enabling simultaneous lesion localization and classification. The original YOLO framework introduced a unified real-time object detection approach with significant speed advantages [7]. YOLOv4 further optimized detection accuracy and computational efficiency through architectural improvements [8]. Despite these advancements, the integration of modern YOLO variants with explainability techniques for skin lesion detection remains relatively limited.

Benchmark datasets have significantly contributed to the evaluation and comparison of automated skin lesion systems. The HAM10000 dataset, introduced by Tschandl et al., provides a large-scale, multi-source collection of dermoscopic images and has become a standard benchmark for skin lesion classification research [10].

In contrast to existing methods, the proposed work integrates a state-of-the-art YOLOv8 object detection model with Grad-CAM–based explainability in a web-based framework. This combination enables real-time lesion detection, transparent decision visualization, and practical deployment, addressing key limitations of prior approaches in explainable skin cancer diagnosis.

Proposed System

The proposed system presents a web-based, explainable skin cancer detection framework that integrates deep learning–based lesion detection with visual interpretability. The system is designed to perform accurate and real-time skin lesion localization and classification using the YOLOv8 object detection model, while enhancing transparency through Grad-CAM–based visual explanations. A Flask-based web application serves as the deployment platform, enabling easy access for clinical support and telemedicine applications.

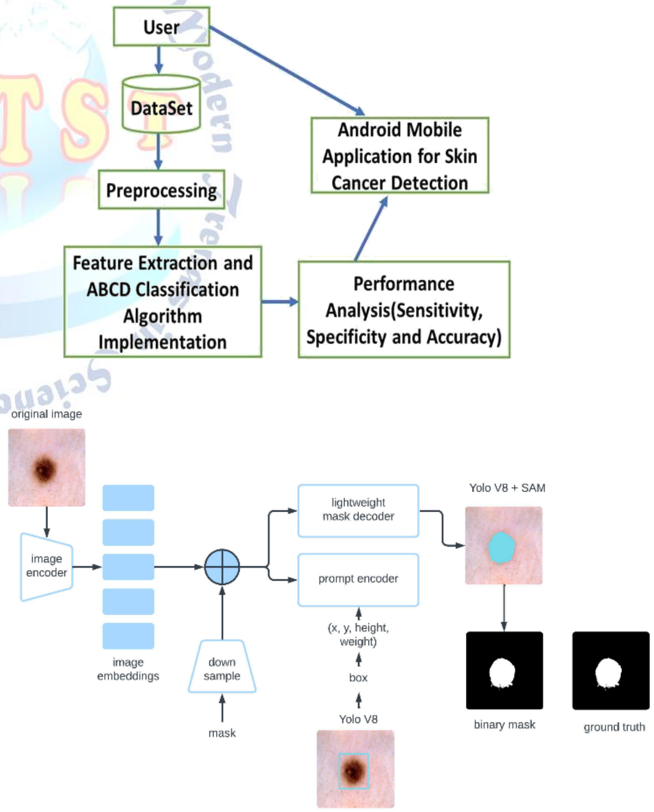


Figure 2 Block Diagram of the Proposed Explainable Skin Cancer Detection System

- Overall Architecture

The architecture of the proposed system consists of five main modules: image acquisition, preprocessing, lesion detection and classification, explainability generation, and result visualization. The modular design ensures scalability, real-time performance, and ease of integration with existing healthcare workflows.

- Image Acquisition

Skin images are acquired either through user-uploaded dermoscopic or clinical images or via a real-time webcam feed. This dual-mode input allows the system to support both offline image analysis and live screening scenarios, making it suitable for diverse medical environments.

- Image Preprocessing

The acquired images undergo preprocessing operations such as resizing, normalization, and format conversion to match the input specifications of the YOLOv8 model. These steps help improve detection robustness by reducing noise and handling variations in image resolution and illumination.

- Skin Lesion Detection and Classification

The preprocessed images are passed to the YOLOv8 deep learning model, which performs simultaneous lesion localization and classification in a single forward pass. The model outputs bounding boxes, class labels (e.g., melanoma, basal cell carcinoma, benign lesions), and confidence scores. The single-stage architecture of YOLOv8 ensures high inference speed with reliable detection accuracy.

- Explainability Module

To enhance interpretability and clinical trust, Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to the detected lesion regions. Grad-CAM generates heatmaps that highlight the most influential regions contributing to the model's predictions. These visual explanations assist clinicians in understanding and validating the automated diagnosis.

- Web-Based Result Visualization

The final outputs, including predicted lesion class, confidence score, bounding box visualization, and Grad-CAM heatmap overlay, are displayed through a Flask-based web interface. The intuitive user interface enables rapid interpretation of results and supports real-time interaction with minimal computational overhead.

METHODOLOGY

The proposed methodology describes the systematic process followed to achieve accurate, real-time, and explainable skin cancer detection. The system combines deep learning–based object detection with explainable AI techniques and web-based deployment. The complete workflow consists of dataset preparation, model training, inference, explainability generation, and result visualization.

A. Dataset Preparation

The system is trained and evaluated using dermoscopic skin lesion images obtained from publicly available datasets such as HAM10000. The dataset contains multiple lesion categories, including melanoma, basal cell carcinoma, and benign lesions. Images are annotated with bounding boxes and class labels to support object detection–based learning.

Prior to training, the dataset is split into training, validation, and testing sets. Data augmentation techniques such as horizontal flipping, rotation, scaling, and brightness adjustment are applied to improve model generalization and reduce overfitting.

B. Image Preprocessing

Each input image undergoes preprocessing to ensure compatibility with the YOLOv8 model. The preprocessing steps include resizing images to a fixed input resolution, pixel normalization, and color space conversion where necessary. These steps enhance detection robustness under varying image acquisition conditions.

C. YOLOv8-Based Skin Lesion Detection

The YOLOv8 model is employed for simultaneous skin lesion localization and classification. Unlike traditional two-stage detectors, YOLOv8 performs detection in a single forward pass, making it suitable for real-time applications. The model predicts bounding box coordinates, lesion class labels, and corresponding confidence scores.

During inference, Non-Maximum Suppression (NMS) is applied to eliminate redundant overlapping bounding boxes, ensuring precise lesion localization.

D. Explainability Using Grad-CAM

To enhance interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to the trained YOLOv8 model. Grad-CAM computes the gradients of the predicted class score with respect to feature maps in the final convolutional layers. These gradients are used to generate a heatmap that highlights the most influential regions contributing to the prediction.

The resulting heatmaps are overlaid on the original images, allowing clinicians and users to visually assess whether the model focuses on clinically relevant lesion regions.

E. Web-Based Deployment

The complete detection and explainability pipeline is deployed using a Flask-based web application. Users can upload images or perform real-time webcam analysis. The application displays the predicted lesion class, confidence score, bounding box visualization, and Grad-CAM heatmap, enabling intuitive and transparent interaction with the system.

Algorithm 1 Skin Lesion Detection Using YOLOv8

Require: Input skin image (uploaded image or webcam frame)
Ensure: Bounding boxes, lesion class labels, confidence scores

Acquire the input skin image
Resize the image to the YOLOv8 input resolution
Normalize pixel values
Feed the preprocessed image into the trained YOLOv8 model
Predict bounding boxes, class labels, and confidence scores
Apply Non-Maximum Suppression to remove overlapping detections
Output the final detected lesion regions with labels and scores

Algorithm 2 Grad-CAM Based Explainability Generation

Require: Detected lesion image and predicted class
Ensure: Grad-CAM heatmap

Select the final convolutional layer of the YOLOv8 model
Compute gradients of the predicted class score with respect to feature maps
Perform global average pooling on the gradients
Compute weighted sum of the feature maps
Apply ReLU to generate the Grad-CAM heatmap
Normalize the heatmap values
Overlay the heatmap on the original image for visualization

Algorithm 3 Web-Based Skin Cancer Detection and Visualization

Require: User image or live webcam frame
Ensure: Predicted lesion type with visual explanation

Receive image input through the Flask web interface
Perform preprocessing and lesion detection using Algorithm 1
Generate explainability heatmap using Algorithm 2
Overlay bounding boxes and heatmaps on the image
Display predicted class, confidence score, and explanation to the user

Transparent decision-making through Grad-CAM. Web-based deployment for accessibility and scalability.

Results and Discussions

This section presents the experimental results and performance analysis of the proposed YOLOv8-based explainable skin cancer detection system. The system was evaluated using dermoscopic skin lesion images, focusing on detection accuracy, classification performance, real-time inference capability, and explainability effectiveness.

- **Experimental Setup**

The proposed model was trained on annotated skin lesion images obtained from publicly available datasets such as HAM10000. The dataset was divided into training, validation, and testing subsets. Model training was carried out using transfer learning on the YOLOv8 architecture. The system was deployed on a Flask-based web application and tested using both uploaded images and real-time webcam input.

- **Performance Evaluation Metrics**

The system performance was evaluated using standard metrics commonly adopted in medical image analysis:

- Accuracy
- Precision
- Recall
- F1-score
- Mean Average Precision (mAP)

These metrics provide a comprehensive assessment of both detection reliability and classification effectiveness.

- **Quantitative Results**

The YOLOv8 model achieved high detection accuracy across multiple skin lesion categories, including melanoma, basal cell carcinoma, and benign lesions. The system demonstrated strong precision and recall values, indicating reliable lesion localization and reduced false positives. The high mAP score confirms the robustness of the object detection framework in identifying lesion boundaries accurately.

In addition, the inference time per image remained low, enabling real-time performance suitable for clinical screening and telemedicine environments.

- **Qualitative Results and Visual Analysis**

The qualitative analysis highlights the effectiveness of the proposed system in both lesion detection and explainability. Bounding boxes generated by YOLOv8 accurately localized lesion regions, while Grad-CAM heatmaps successfully highlighted diagnostically relevant areas such as irregular borders and color variations. These visual explanations align well with clinical diagnostic patterns, thereby improving trust in automated predictions.

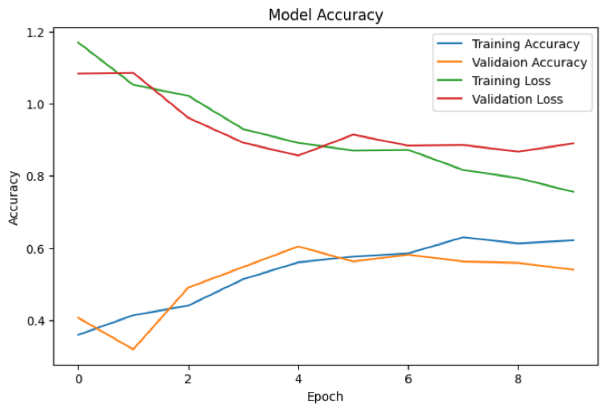


Figure 3 Training and Performance Graph

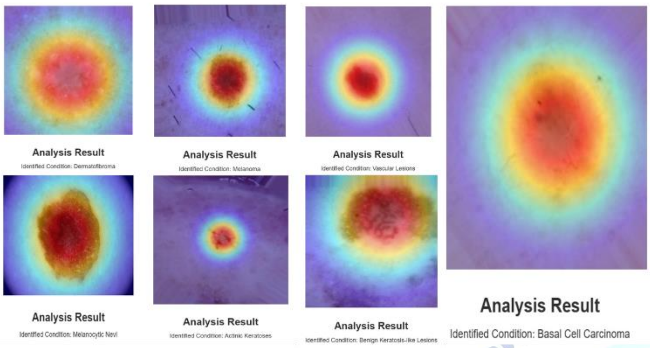


Figure 4 Sample Detection and Explainability Outputs

- **Graphical Performance Analysis**

The performance metrics were further analyzed using graphical representations. Accuracy and loss curves illustrate stable convergence during training, while class-wise precision and recall graphs demonstrate consistent performance across different lesion types. These graphs confirm that the proposed model generalizes well and avoids overfitting.

CONCLUSION

This work presented a web-based, explainable skin cancer detection system that integrates the YOLOv8 deep learning model with Grad-CAM–based visual interpretability. The proposed framework enables accurate and real-time skin lesion localization and classification while providing transparent visual explanations to support clinical decision-making. By combining object detection and explainable artificial intelligence, the system addresses key limitations of traditional black-box deep learning approaches in medical image analysis.

Experimental results demonstrate that the proposed system achieves reliable detection performance with minimal computational overhead, making it suitable for real-time screening and telemedicine applications. The integration of Grad-CAM enhances model transparency by highlighting diagnostically relevant regions, thereby improving trust and usability in clinical environments. Furthermore, the Flask-based web deployment ensures accessibility and ease of use, allowing the system to be effectively utilized in remote and resource-constrained settings.

Future work will focus on validating the system across multiple large-scale datasets, incorporating additional clinical metadata to improve diagnostic accuracy, and optimizing the framework for mobile and edge-device deployment. The proposed solution contributes toward accessible, efficient, and explainable AI-assisted skin cancer screening and holds significant potential for supporting early diagnosis and improved patient outcomes.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Matas, J., et al., "Interpretable Skin Cancer Diagnosis using Grad-CAM and Expert Annotations," arXiv preprint arXiv:2407.00104, 2024.

[2] Liu, H., et al., "Adaptive Spatial Feature Fusion for Improved Skin Lesion Classification," arXiv preprint arXiv:2510.03876, 2025.

[3] Atiq, S., & Fattah, S., "Dual-Network Attention Model for Skin Lesion Detection with Clinical Metadata," arXiv preprint arXiv:2510.17773, 2025.

[4] Paccotacya-Yanque, M., et al., "Comparative Analysis of Explainability Techniques for Skin Lesion Classification," arXiv preprint arXiv:2412.03166, 2024.

[5] Hamim, M., et al., "SmartSkin-XAI: Explainable DenseNet for Skin Cancer Diagnosis," Diagnostics, vol. 15, no. 1, pp. 64, 2025.

[6] Mukherjee, A., et al., "Explainable Deep Learning with Swish-Activated Convolutions for Skin Cancer Detection," Springer Neural Computing and Applications, 2026.

[7] Redmon, J., et al., "You Only Look Once: Unified, Real-Time Object Detection," CVPR, 2016.

[8] Bochkovskiy, A., Wang, C.Y., & Liao, H.Y.M., "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.

[9] Selvaraju, R.R., et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," International Journal of Computer Vision, vol. 128, pp. 336–359, 2020.

[10] Tschandl, P., et al., "The HAM10000 Dataset, a Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions," Scientific Data, vol. 5, no. 180161, 2018.