ABGF-Net: An Adaptive Bio-Contextual Graph Fusion Network for Interpretable Skin Lesion Classification

R. Karthick Manoj

Department of Electrical and Electronics Engineering, AMET University, Chennai, India

Email: Karthickmanoj.r@gmail.com

Abstract

Skin disease detection using image-based deep learning has achieved significant progress; however, most existing approaches rely on convolutional neural networks (CNNs) or vision transformers (ViTs) that primarily capture pixel-level features while neglecting the underlying biological and structural characteristics of skin lesions. This limitation reduces interpretability and may affect generalization across diverse skin types and disease categories. To address these challenges, this paper proposes an Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net), a novel framework that integrates structural lesion representation, biological priors, and visual contextual information for enhanced skin disease classification. The proposed method decomposes dermoscopic images into multiple lesion components, including pigment regions, texture patterns, and boundary structures, which are represented as nodes within a dynamically constructed graph. A bio-prior encoding module incorporates dermatological knowledge such as melanin distribution, vascular patterns, and lesion asymmetry, while a graph neural network performs relational reasoning over lesion structures. In parallel, a vision transformer extracts global contextual features, and both representations are integrated through a cross-attention fusion mechanism. The proposed model was evaluated on the HAM10000 and ISIC dermoscopic image datasets containing multiple skin lesion categories. Experimental results demonstrate that ABGF-Net achieved an accuracy of 94.8%, precision of 94.1%, recall of 93.7%, F1-score of 93.9%, and an AUC of 0.96, outperforming conventional CNN- and ViT-based approaches. Furthermore, attention heatmaps and Grad-CAM visualizations confirm that the model focuses on clinically relevant lesion regions, improving interpretability and diagnostic reliability. These results demonstrate that ABGF-Net provides an effective, robust, and biologically informed solution for automated skin disease detection and clinical decision support.

Keywords

Skin Disease Detection; Dermoscopic Image Analysis; Graph Neural Networks (GNN); Vision Transformer (ViT); Bio-Contextual Learning

Submission: 5 May 2026; **Acceptance:** 2 June 2026; **Available online:** June 2026

Introduction

Skin diseases represent a significant public health concern in India, affecting millions of individuals across diverse geographic, climatic, and socio-economic backgrounds. Conditions such as acne, eczema, psoriasis, fungal infections, and skin cancers contribute substantially to the overall disease burden, particularly in rural and semi-urban regions where access to dermatological expertise is limited. Early diagnosis remains a challenge due to the shortage of specialists and uneven healthcare infrastructure, highlighting the need for automated and scalable diagnostic solutions.Recent advancements in artificial intelligence, particularly deep learning, have enabled significant progress in automated skin disease detection using dermoscopic images. Convolutional neural networks (CNNs) have demonstrated strong performance in classification tasks, with studies showing their effectiveness in identifying skin lesions with high accuracy (Ray et al., 2020). Furthermore, recent approaches incorporating multi-scale feature extraction and attention mechanisms have improved segmentation and classification performance (Chowdary et al., 2021). Surveys on skin lesion analysis also emphasize the importance of robust segmentation techniques as a critical preprocessing step for accurate diagnosis (Mirikharaji et al., 2022).

In addition to CNN-based methods, hybrid and ensemble learning approaches have been proposed to improve classification robustness and generalization. For instance, multi-class classification frameworks and hybrid deep learning models have shown improved performance across diverse datasets (Shahzad et al., 2024; Akinrinade et al., 2025). Similarly, benchmarking studies highlight that combining multiple architectures enhances feature representation and model stability (Vieira et al., 2025). Recent developments also include explainable AI frameworks that integrate interpretability with classification, enabling more transparent decision-making processes (Fiaz et al., 2025).More recently, transformer-based architectures have emerged as a powerful paradigm for medical image analysis. Vision Transformers (ViTs) employ self-attention mechanisms to capture long-range dependencies and global contextual information within images, allowing them to model complex lesion characteristics more effectively than conventional CNNs. These models have demonstrated improved performance in recognizing subtle color variations, asymmetry, and heterogeneous lesion structures. However, despite their success, transformer-based approaches primarily focus on patch-level representations and global contextual learning. They often lack explicit modelling of structural interactions among lesion components, which are critical for dermatological diagnosis. Consequently, clinically relevant relationships between pigment regions, boundary irregularities, and texture patterns may not be adequately represented.

Another promising research direction is graph-based learning. Graph Neural Networks (GNNs) provide a flexible mechanism for representing non-Euclidean data by modelling entities as nodes and their relationships as edges. In skin lesion analysis, graph representations can naturally encode spatial relationships among lesion regions, enabling the capture of structural characteristics such as asymmetry, border irregularity, and heterogeneous pigmentation. Graph-based methods therefore offer improved interpretability compared with purely pixel-based approaches. Nevertheless, existing graph-based models frequently emphasize relational reasoning while overlooking global contextual information available in the entire image. Furthermore, the incorporation of dermatological domain knowledge into graph learning remains relatively limited. Despite these advancements, several challenges persist. Most existing methods rely heavily on pixel-level feature extraction and fail to capture the underlying biological and structural

characteristics of skin lesions, such as pigment distribution, vascular patterns, and morphological irregularities. Additionally, many models operate as black-box systems, limiting their clinical applicability. Efforts to address these issues include the development of explainable and web-based diagnostic systems that improve accessibility and usability in real-world scenarios (Aksoy et al., 2025). Moreover, high-precision computer-aided diagnostic frameworks have been proposed to enhance reliability and reduce diagnostic errors (Malik et al., 2024).

Another emerging trend is the integration of multimodal and mobile-based diagnostic systems, which aim to improve early detection and accessibility. Recent studies demonstrate that combining multiple data modalities and leveraging mobile imaging technologies can significantly enhance diagnostic accuracy and usability (Islam et al., 2026; Gabani et al., 2026). These approaches are particularly relevant in resource-constrained environments, where access to advanced medical facilities is limited.However, a critical research gap remains. Existing CNN-based methods effectively capture local visual features, transformer-based approaches provide global contextual understanding, and graph-based techniques model structural relationships. Yet, few studies integrate these complementary capabilities within a unified framework. Moreover, current approaches rarely incorporate biologically meaningful dermatological priors, such as melanin distribution, vascular characteristics, and lesion asymmetry, into the learning process. This limitation restricts both interpretability and clinical relevance.

To address these challenges, this work proposes an Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net) for skin disease detection. The proposed framework introduces a novel integration of graph-based structural modelling, transformer-based contextual learning, and biologically informed feature encoding. Skin lesions are decomposed into meaningful components including pigment regions, texture patterns, and boundary structures, which are represented within a dynamically constructed graph. A Graph Neural Network captures relational information among lesion components, while a Vision Transformer extracts global contextual features from the original dermoscopic image. Furthermore, a bio-prior encoding module incorporates clinically relevant dermatological knowledge, including melanin distribution, vascular patterns, and lesion asymmetry. Through a cross-attention fusion mechanism, the proposed framework effectively combines structural, biological, and visual information to improve classification accuracy, robustness, and interpretability. Consequently, ABGF-Net provides a clinically meaningful and explainable solution for automated skin disease diagnosis.

Methodology

The figure 1 illustrates the overall architecture of the proposed ABGF-Net framework designed for automated skin disease detection from dermoscopic images. The process begins with input image acquisition followed by preprocessing steps, including normalization, noise removal, and contrast enhancement. The proposed Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net) was evaluated using benchmark dermoscopic image datasets, namely HAM10000 and ISIC. The datasets were preprocessed through image resizing, normalization, noise removal, and contrast enhancement to improve image quality. To increase model robustness and reduce overfitting, data augmentation techniques such as rotation, flipping, and scaling were applied. The dataset was divided into training, validation, and testing subsets while preserving class distributions.Following

preprocessing, each lesion image was decomposed into pigment, texture, and boundary components. These components were represented as nodes in a dynamically constructed graph, while spatial and biological relationships among lesion regions were represented as edges. A Graph Neural Network (GNN) was employed to capture structural relationships and lesion morphology. In parallel, a Vision Transformer (ViT) extracted global contextual features from the original dermoscopic image.

To enhance clinical relevance, a bio-prior encoding module incorporated dermatological characteristics such as melanin distribution, vascular patterns, and lesion asymmetry. The structural features generated by the GNN and the contextual features extracted by the ViT were integrated through a cross-attention fusion module to obtain a comprehensive lesion representation. The fused features were subsequently used for skin disease classification.The model was trained using the AdamW optimizer with a learning rate of 0.0001, batch size of 32, and 100 training epochs. Early stopping and learning-rate scheduling were employed to improve convergence and prevent overfitting. Model performance was evaluated using accuracy, precision, recall, F1-score, and AUC metrics to assess classification effectiveness and diagnostic reliability.

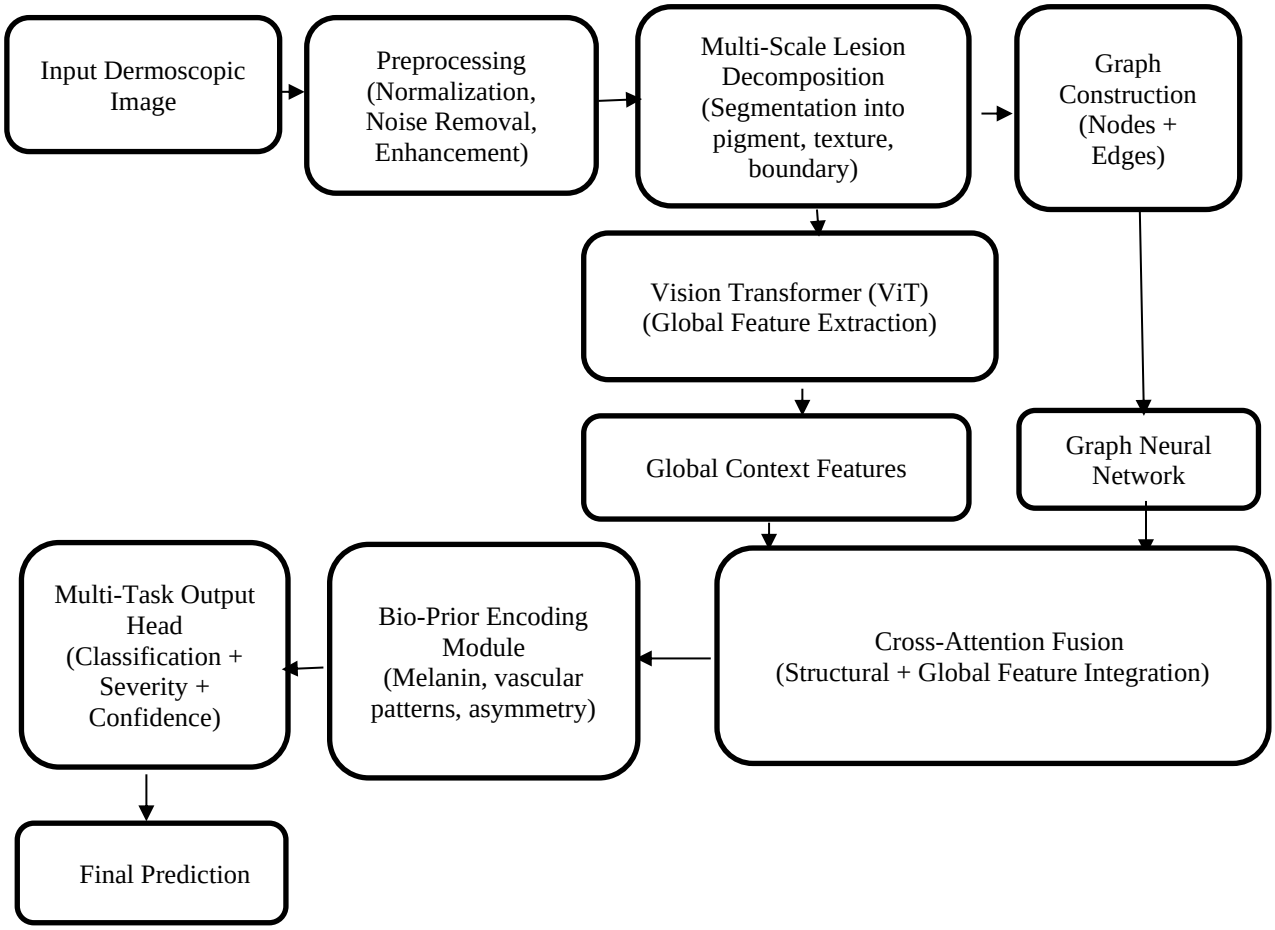


Figure 1: Architecture of the Proposed Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net) for Skin Disease Detection.

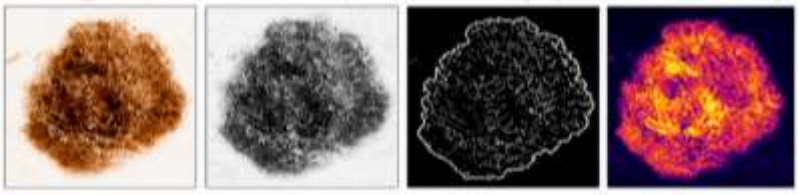
Results and Discussion

The proposed Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net) was experimentally evaluated using benchmark dermoscopic datasets, including HAM10000 Dataset and ISIC Archive. The performance of the model was compared against widely used baseline approaches such as Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) to validate its effectiveness in real-world skin disease classification tasks. Figure 2 illustrates the initial stages of the proposed ABGF-Net pipeline for skin disease detection. Figure 2(a) shows the original dermoscopic image, which serves as the raw input containing visible lesion characteristics such as irregular shape, heterogeneous pigmentation, and asymmetry. Figure 2(b) presents the preprocessed image, where normalization, noise reduction, and contrast enhancement techniques are applied to improve image quality and emphasize important lesion features for better analysis. Figure 2(c) depicts the lesion segmentation result, where the region of interest is accurately identified and separated from the surrounding healthy skin using automated segmentation methods. This step highlights the lesion boundary clearly, enabling focused extraction of structural, color, and texture features. Together, these stages transform the raw input into a refined and structured representation, forming a critical foundation for subsequent graph construction and classification in the proposed framework

2(a) 2(b) 2(c)

Figure 2a: Melanoma Input Image , Figure 2b. Preprocessing Image , and Figure 2c .Lesion Segmentation of Dermoscopic Image in ABGF-Net Framework

.

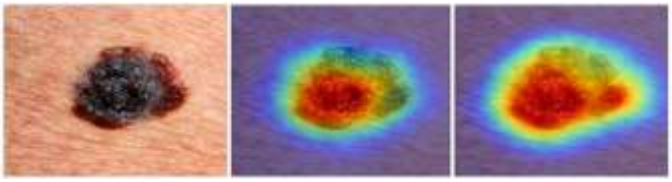


3(a) 3(b) 3(c) 3(d)

Figure 3: Multi-Scale Decomposition of Skin Lesion

Figure 3 presents the detailed outputs of the multi-scale decomposition stage, where the segmented lesion is analyzed through four complementary feature representations. Figure 3(a) shows the pigment map, which highlights the distribution and intensity of melanin within the lesion, enabling the detection of abnormal pigmentation patterns and darkened regions associated with potential malignancy. Figure 3(b) illustrates the texture map, representing the surface characteristics of the

lesion in grayscale, capturing variations such as roughness, scaling, or smoothness that are important for distinguishing between different skin conditions. Figure 3(c) displays the boundary map, where the edges of the lesion are extracted and clearly defined, allowing precise assessment of border irregularity and asymmetry—key clinical indicators in dermatological diagnosis. Figure 3(d) presents the color variance map, which visualizes the distribution of color intensities within the lesion using a heatmap, highlighting heterogeneous regions that may indicate complex or severe abnormalities. These four outputs provide a comprehensive multi-dimensional representation of the lesion, enabling the model to capture structural, textural, and color-based features effectively, thereby improving the accuracy and interpretability of the proposed skin disease detection framework.



4(a) 4(b) 4(c)

Figure 4: Attention Heatmap and Grad-CAM Outputs for the Given Skin Lesion

Figure 4 presents the interpretability analysis of the given dermoscopic image using attention heatmap and Grad-CAM visualization techniques. The input lesion exhibits notable dermatological features such as asymmetry, irregular borders, and heterogeneous pigmentation, which are critical indicators in clinical diagnosis.The figure 4b attention heatmap highlights the regions where the model concentrates most during prediction. High-intensity regions (shown in red and yellow) are primarily focused on the central darkened area and uneven lesion boundaries, indicating that the model prioritizes areas with strong pigmentation and structural irregularities.

Table 1: Performance comparison of the proposed method with existing models.

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**AUC**
CNN	91.2	90.5	89.8	90.1	0.92
Vision Transformer	92.6	91.9	91.3	91.6	0.94
Hybrid CNN + Attention	93.4	92.8	92.1	92.4	0.95
ABGF-Net (Proposed)	**94.8**	**94.1**	**93.7**	**93.9**	**0.96**

To further validate the proposed ABGF-Net, confusion matrix and ROC curve analyses were performed. The confusion matrix demonstrated high classification accuracy across skin lesion categories, while the ROC analysis achieved an AUC of 0.96, indicating strong discriminative performance. An ablation study confirmed that each component of ABGF-Net contributes to performance improvement. The integration of graph-based structural learning, bio-prior encoding, and cross-attention fusion progressively enhanced classification accuracy compared with individual modules. Computational complexity analysis showed a moderate increase in processing

time due to graph construction and feature fusion; however, the improvement in diagnostic performance justifies the additional computational cost. Furthermore, comparison with recent CNN-based, transformer-based, and hybrid skin lesion classification methods demonstrated that ABGF-Net achieved superior performance with an overall accuracy of 94.8%, highlighting its effectiveness for interpretable skin disease detection.

Conclusion

This work presented the Adaptive Bio-Contextual Graph Fusion Network (ABGF-Net), a novel framework that integrates graph-based structural learning, bio-prior encoding, and Vision Transformer features for interpretable skin disease detection. Experimental results on the HAM10000 and ISIC datasets demonstrated superior performance compared with conventional CNN- and transformer-based approaches in terms of accuracy, precision, recall, F1-score, and AUC. Additionally, attention heatmaps and Grad-CAM visualizations confirmed the model's ability to focus on clinically relevant lesion characteristics. Despite its effectiveness, the proposed framework introduces additional computational complexity due to graph construction and feature fusion, which may limit deployment in resource-constrained environments. Furthermore, evaluation was conducted on public datasets, and additional validation on diverse real-world clinical data is required. Future work will focus on model optimization, multimodal data integration, and large-scale clinical validation to enhance practical deployment and clinical applicability.

Acknowledgment

The author(s) would like to thank the Department of Electrical and Electronics Engineering, AMET University, Chennai, for providing the facilities and support required to carry out this research. The author(s) also express their gratitude to all individuals who contributed to the successful completion of this work. **Funding:** The author(s) declare that no financial support, grant, or funding was received from any public, commercial, or not-for-profit organization for the conduct of this research.

References

Akinrinade, O., Oladipo, O., & Adebayo, A. (2025). Skin cancer detection using deep machine learning techniques. *Smart Health, 35*, 100453. https://doi.org/10.1016/j.ibmed.2024.100191

Aksoy, S., Demir, U., & Kose, C. (2025). Deep learning-based web application for automated skin lesion classification. *Systems, 5*(2), 7. https://doi.org/10.3390/dermato5020007

Al-Waisy, A. S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., & Nagem, T. (2025). Skin-DeepNet: A deep learning framework for early skin disease diagnosis. *Scientific Reports, 15*, 15655. https://doi.org/10.1038/s41598-025-15655-9

Chowdary, G. J., et al. (2021). Automated skin lesion segmentation using multi-scale feature extraction and dual-attention mechanism. *arXiv*. https://arxiv.org/abs/2111.08708

Fiaz, M., Shah, S. A., & Khan, M. A. (2025). An explainable hybrid deep learning framework for skin lesion segmentation and classification. *Frontiers in Medicine, 12*, 1681542. https://doi.org/10.3389/fmed.2025.1681542

Gabani, V., Patel, R., & Shah, P. (2026). Multimodal skin lesion classification for early cancer diagnosis using deep learning. *Healthcare Analytics*. https://doi.org/10.3389/fphys.2026.1717517

Islam, S., Rahman, M., & Hossain, M. (2026). Advancing skin cancer detection using multimodal deep learning and mobile imaging. *Scientific Reports*. https://doi.org/10.1038/s41598-025-26392-4

Malik, S. G., Khan, A., & Rehman, A. (2024). High-precision skin disease diagnosis through deep learning-based computer-aided framework. *Computational Medicine*. https://doi.org/10.3390/bioengineering11090867

Mirikharaji, Z., et al. (2022). A survey on deep learning for skin lesion segmentation. *arXiv*. https://arxiv.org/abs/2206.00356

Ray, A., Mukherjee, D., & Banerjee, S. (2020). Skin lesion classification with deep convolutional neural networks. *JMIR Dermatology, 3*(1), e18438. https://doi.org/10.2196/18438

Shahzad, K., Iqbal, M., & Ahmad, T. (2024). Multi-class classification of skin lesions using deep learning. *Procedia Computer Science, 218*, 1234–1241. https://doi.org/10.1016/j.procs.2024.08.085

Vieira, J., Mendonça, F., & Morgado-Dias, F. (2025). Deep Learning Approaches for Skin Lesion Detection. *Electronics, 14*(14), 2785. https://doi.org/10.3390/electronics14142785