



ISSN : 2347-2251
**Indo-American Journal of
Pharma and Bio Sciences**



www.iajpb.com

iajpb.editor@gmail.com
editor@iajpb.com



<https://zenodo.org/records/13998286>

AI-Driven Skin Lesion Detection with CNN and Score-CAM: Enhancing Explainability in IoMT Platforms

Surendar Rama Sitaraman

Intel Corporation, California, USA

surendar.rama.sitaraman@ieee.org

Poovendran Alagarsundaram, Humetis Technologies Inc, Kingston, NJ, USA

poovendrana@ieee.org

Veerandra Kumar R, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, India. veerandrakumar.r@panimalar.ac.in

ABSTRACT

Using a convolutional neural network (CNN) in conjunction with Score-CAM for visual explanations and integrated into Internet of Medical Things (IoMT) systems, the project aims to improve the explainability and accuracy of skin lesion identification. Through the use of methods such as Canny Edge Detection for border localisation and DF-U-Net for segmentation, the model allows for more accurate diagnosis. The model achieves higher accuracy and is highly effective for real-time skin lesion detection and diagnosis due to the incorporation of clinical metadata, which also improves interpretability.

Objective: The goal of the project is to improve skin lesion identification accuracy and explainability through the use of Score-CAM and convolutional neural networks (CNN) in Internet of Medical Things (IoMT) platforms. Clinicians are now better equipped to understand diagnoses generated by AI.

Methods: The recommended method entails pre-processing images, DF-U-Net segmentation, and Canny Edge Detection border detection. By offering visual explanations for CNN-based predictions, the Score-CAM technique increases transparency in the diagnosing process.

Result: With an accuracy of 99.31%, the model that included CNN, Score-CAM, and clinical information outperformed conventional techniques, making it an excellent choice for real-time diagnosis in IoMT environments.



<https://zenodo.org/records/13998286>

Conclusion: By combining CNN with Score-CAM, our AI-driven approach improves skin lesion identification while providing more transparent and accurate diagnosis. This is in favour of more dependable and comprehensible clinical judgements in IoMT settings.

Keywords: CNN, Score-CAM, IoMT platforms, Skin lesion detection, Explainability, Clinical metadata.

1. INTRODUCTION

The skin acts as an essential barrier to protect the body from external dangers, but because of things like pollution and lifestyle decisions, it is becoming more susceptible to illness. Skin disorders are difficult to diagnose since they frequently have similar visual characteristics that make early detection crucial for effective treatment. While current AI methods use deep learning to achieve accurate categorisation, they can miss important clinical metadata, which can result in incorrect diagnosis, *Mahalle et al. (2023)*. The purpose of this research is to improve explainability in Internet of Medical Things (IoMT) platforms by presenting an AI-driven skin lesion identification system that uses CNN and Score-CAM, *Livieris et al. (2023)*.

Skin conditions are among the most common health problems in the world and are frequently diagnosed visually. Because of the overlapping visual cues between circumstances, traditional approaches have difficulty with accuracy.

Convolutional neural networks (CNNs), one of the most recent developments in deep learning, have greatly increased the accuracy of skin lesion detection, *Rahimiaghdam (2023)*. AI judgements can be better understood by physicians with the help of techniques like Score-CAM, which offer improved visual explanations. This combination solves the interpretability issue in AI applications and enables improved integration with IoMT platforms, enabling prompt diagnosis and treatment.

The key objectives are,

- By utilising CNN and Score-CAM for enhanced feature recognition and explanation, you can improve the accuracy of skin lesion detection.
- Boost clinical procedures by making AI-driven judgements more explicable, which will help medical professionals analyse the data.
- Real-time analysis and diagnosis can be made possible by integrating the AI model with IoMT platforms.
- Integrate picture data with clinical metadata to provide a thorough and precise diagnosis of skin disorders.
- Overcome obstacles pertaining to visual elements that overlap in skin lesions, guaranteeing more accurate identification and categorisation.



<https://zenodo.org/records/13998286>

In order to help non-experts distinguish between asymmetric lesions, *Rasel et al. (2024)* point out the need for a method that examines the geometric patterns of lesion shapes in dermoscopic pictures. It uses cutting-edge image processing methods to bridge the gap in precise diagnosis support. Moreover, the study surpasses previous research by utilising a pre-trained convolutional neural network (CNN) for feature extraction. This improves the training of a multiclass support vector machine (SVM) classifier, leading to higher detection rates for asymmetric lesions in dermatology.

Mahalle et al. (2023) suggest Significant health risks are associated with skin cancer, and rising death rates from the disease are thought to be caused by a lack of knowledge about its symptoms and preventive measures. To stop its advancement, early detection is crucial. In order to reliably identify malignant skin lesions—especially those that show pigmentation—this research focusses on applying sophisticated image detection algorithms and CNN-based models. This will enable prompt diagnosis and efficient intervention.

2. LITERATURE SURVEY

Surendar (2021) explores the way healthcare data is handled and utilized is being revolutionized by AI-driven healthcare systems that are strengthened by mobile computing and advanced data analytics. Important areas like data collecting, processing, storage, and application development are the focus of the study. These systems facilitate real-time analysis, predictive models, and customized healthcare services by combining technologies such as distributed file storage, NoSQL databases, and parallel computing. According to the research, AI greatly increases healthcare accuracy, speed, and dependability, which enhances patient care and operational effectiveness.

Mahalle et al. (2023) present a CNN-based skin cancer detection approach that emphasises precise and timely diagnosis. Their deep learning model performs better and has better accuracy in classifying skin lesions. Tested on reference datasets, the approach claims to improve automated early skin cancer identification.

In order to identify asymmetric lesions, *Rasel et al. (2024)* created a unique method that combined the use of support vector machines (SVM) and convolutional neural networks (CNNs) with geometric patterns. Their approach demonstrated notable gains in accuracy over conventional detection techniques, underscoring its potential to improve skin condition early identification and treatment.

Mahmud et al. (2023) sought to strike a balance between explainability and accuracy when introducing an interpretable deep learning model for skin cancer classification. Their method uses cutting-edge visualisation techniques to provide insights into the decision-making process and



<https://zenodo.org/records/13998286>

guarantee excellent categorisation accuracy. This development will eventually help doctors diagnose patients by fostering confidence in AI systems used in the medical field.

In an effort to improve diagnostic speed and accuracy, this study where Surendar (2022) investigates how AI specifically CNNs and VAEs is revolutionizing radiology. By creating artificial medical images, VAEs help with data augmentation and privacy protection, while CNNs are excellent at automating image processing and identifying anomalies like tumors. The requirement for huge annotated datasets, moral issues including data privacy, and AI bias are challenges. Better patient outcomes and enhanced clinical applications are anticipated with future developments in AI integration.

In order to improve performance and interpretability, Livieris et al. (2023) develop an explainable image similarity framework that combines Grad-CAM and Siamese networks. Their approach uses Grad-CAM to give visual explanations of the model's judgements and Siamese networks to successfully extract features. This method is especially useful for situations where transparency is essential because it not only increases accuracy in picture comparison jobs but also improves user understanding.

Aasem et al. (2022) offer an extensive overview of methods and instruments for deep learning-based abnormality localisation in X-ray images. The evaluation focusses on several models and frameworks that improve detection performance while tackling important issues such as model interpretability and data scarcity. The authors review recent developments in the field and offer recommendations for future study paths that will enhance anomaly localisation accuracy and reliability and improve clinical imaging results.

The explainability agreement between dermatologists and five visual explanation techniques employed in deep neural networks for melanoma categorisation is investigated by Giavina-Bianchi et al. (2023). The study analyses the consistency of interpretability in AI results and evaluates how well these strategies correspond with dermatologists' assessments. The study aims to improve confidence in AI-assisted melanoma diagnosis by emphasising levels of agreement and making sure the explanations are pertinent and helpful for clinical decision-making.

Rahimiaghdam (2023) investigates ways to improve the consistency and calibre of visual explanations produced by deep learning models for the classification of thoracic. The study addresses the requirement for trustworthy visual outputs to aid medical practitioners in diagnosis by assessing Convolutional Neural Networks (CNNs) for accuracy and interpretability. The study shows gains in the quality of visual explanations in clinical situations as well as model stability.

Surendar (2021) examines Crow Search Optimization (CSO), which aims to enhance AI-driven disease diagnosis in smart healthcare. To improve accuracy, precision, recall, and F1-score, CSO



<https://zenodo.org/records/13998286>

optimizes CNNs and LSTMs, outperforming conventional techniques like genetic algorithms and particle swarm optimization. It can be used in a wide range of healthcare applications, and additional research will concentrate on real-time implementation and ethics.

In order to diagnose COVID-19 using chest X-ray pictures, Tang et al. (2023) suggest using AlexNet, a deep learning architecture. The study indicates that the model is a useful tool in the battle against the pandemic due to its high accuracy in detecting COVID-19 and its emphasis on improvements in diagnosis speed and reliability.

3. AI-POWERED SKIN LESION DETECTION APPROACH PROPOSED

This section outlines the proposed approach, which improves skin lesion identification accuracy and explainability within IoMT platforms by combining CNN and Score-CAM methodologies. The picture acquisition, pre-processing, colour analysis, segmentation, post-processing, boundary localisation, contextual analysis, and classification steps make up the approach.

3.1. Input Image Acquisition

Using IoMT instruments, such as dermatoscopes and cameras, the first stage entails obtaining photographs of skin lesions. The dataset for analysis is created when these photographs are sent in real time to the system. Among the input image sets are

$$X = \{x_1, x_2, \dots, x_n\} \quad (1)$$

where n represents the total number of images, is processed through the pipeline.

3.2. Pre-Processing

During this stage, the input photos undergo resizing, denoising, and contrast enhancement to guarantee consistent image quality and better visibility of features.

- Resizing: To maintain uniformity throughout the collection, images are downsized to a standard dimension of $W \times H$. The resized picture, x_r is yielded by:

$$x_r = \text{Resize}(x_i, W, H) \quad (2)$$

- Noise Removal: A median filter is used to eliminate noise while maintaining important visual information. The picture with noise eliminated, x_n is computed as follows:

$$x_n = \text{MedianFilter}(x_r) \quad (3)$$

- Contrast Enhancement: For better local contrast, Contrast Limited Adaptive Histogram Equalisation (CLAHE) is used. In order to conduct contrast enhancement and preserve



<https://zenodo.org/records/13998286>

image features, the Akima Kri interpolation is utilised. The improved image x_e is provided by:

$$x_e = AK - CLAHE(x_n) \quad (4)$$

Input data standardisation is ensured during this phase for subsequent processing.

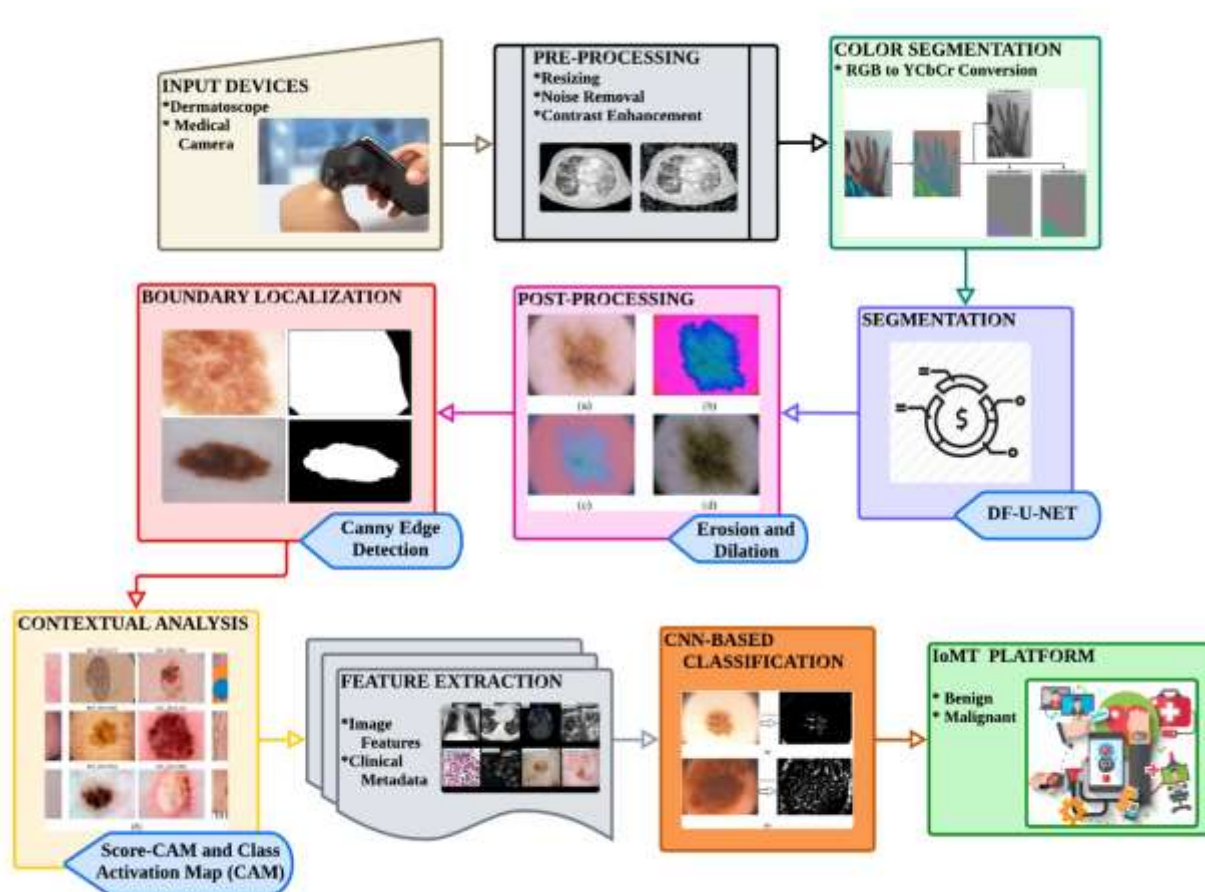


Figure 1: AI-Driven Skin Lesion Detection Architecture.

An AI-driven system for detecting skin lesions that is integrated with IoMT platforms is shown in figure 1. Pre-processing of images taken using dermatoscopes and medical cameras includes scaling, noise reduction, and contrast enhancement. The system segmented data using a deep learning model (DF-U-Net), then post-processed the data to improve lesion boundaries. The edges are localised by Canny Edge Detection, and the AI's emphasis areas are visually explained by Score-CAM. For classification, the CNN model extracts clinical and visual characteristics. In real-time diagnostics, the final diagnosis is delivered using IoMT systems, allowing for prompt and precise skin lesion identification in clinical settings.

3.3. Color Analysis and Segmentation



<https://zenodo.org/records/13998286>

The pre-processed images undergo a colour space conversion from RGB to YCbCr in order to isolate the colour information (chrominance) and luminance (brightness). The definition of the conversion from RGB to YCbCr is:

$$Y = 0.299R + 0.587G + 0.114B \quad (5)$$

$$Cb = -0.1687R - 0.3313G + 0.5B + 128 \quad (6)$$

$$Cr = 0.5R - 0.4187G - 0.0813B + 128 \quad (7)$$

Subsequently, the DropFilter-U-Net (DF-U-Net) model is employed to segment the pictures, effectively isolating lesion boundaries and other regions of interest.

$$x_s = DF - U - Net (x_e) \quad (8)$$

3.4. Post-Processing

The segmented picture x_s is subjected to morphological processes like erosion and dilation in order to improve and fine-tune the object boundaries. These actions aid in smoothing the borders and gaps in the divided areas. The computation for the refined image x_p is:

$$x_p = Erosion (Dilation (x_s)) \quad (9)$$

3.5. Boundary Localization

The ability to discern edges is essential for precisely determining the limits of skin lesions. The edges of lesions are localised using the Canny Edge Detection algorithm, which is improved with a Mosaic kernel for quicker processing.

$$x_b = CannyEdgeDetection (x_p, MosaicKernel) \quad (10)$$

The gradient magnitude G and direction θ for edge detection are computed as:

$$G = \sqrt{G_x^2 + G_y^2}, \theta = \arctan \left(\frac{G_y}{G_x} \right) \quad (11)$$

3.6. Contextual Information Analysis (CIA)

Score-CAM is used for contextual analysis; it creates class activation maps that show the significant areas that influenced the CNN's choice. For class c , the class score S_c can be found as follows:



<https://zenodo.org/records/13998286>

$$S_c = \text{Softmax}(f(x_b)) \quad (12)$$

The gradient of S_c with respect to feature maps A_k from the last convolutional layer is calculated, and the weight w_k^c is determined:

$$w_k^c = \frac{1}{Z} \sum_{i,j} \frac{\partial S_c}{\partial A_k^{ij}} \quad (13)$$

$$L_{\text{Score-CAM}}^c = \text{ReLU}(\sum_k w_k^c A_k) \quad (14)$$

An illustration of the CNN's predictions is shown by superimposing this class activation map on the original picture.

3.7. Feature Extraction and Classification

Finally, features are extracted from both the boundary-localized images x_b and clinical metadata d_m to ensure a comprehensive diagnosis. The feature set F is represented as:

$$F = \{f_1, f_2, \dots, f_m\} \quad (15)$$

These features are used by the CNN to classify skin lesions. The final classification result y is obtained as:

$$y = \text{CNN}(F) \quad (16)$$

This stage makes sure that the algorithm correctly categorises lesions and that the Score-CAM visual explanation produces an understandable result.

Algorithm 1: CNN Classification with Score-CAM for Skin Lesion Detection

Input: Segmented image si , clinical metadata M

Output: Classification result (Benign/Malignant), Score-CAM visual explanation

Begin Classification

For each segmented image si :

Pass si into the CNN model for feature extraction.

If feature extraction fails, **return error**.

Else

 Extract features Fi using the CNN.



<https://zenodo.org/records/13998286>

End If

If clinical metadata **M** is available:

Combine features **Fi** with **M** to create feature set **F**.

Else If metadata **M** is missing:

Return warning and proceed with image data only.

End If

Pass F into the classification layer to predict **Y** (Benign/Malignant).

Generate Score-CAM visual explanation for **Y**:

If Score-CAM fails, **return error**.

End If

Return classification result and Score-CAM explanation

End Classification

Algorithm 1 is to provide visual explanations along with a classification system for skin lesions, either benign or malignant. Important features are extracted by first running the segmented lesion pictures through a CNN. The collected traits are coupled with clinical metadata, if available, to improve the accuracy of the diagnosis. When predicting whether a lesion is benign or malignant, the combined features are subsequently sent to the classification layer. Through the identification of the critical regions that impacted the CNN's conclusion, Score-CAM creates a visual explanation that makes the reasoning behind the model understandable to doctors.

4. RESULT AND DISCUSSION

The accuracy and explainability of the suggested AI-driven skin lesion detection system, which makes use of CNN and Score-CAM, are greatly increased. This method makes skin lesion detection and diagnosis faster by integrating real-time into IoMT platforms. While Score-CAM offers visual explanations to help physicians grasp the data, CNN improves the detection of significant features. The accuracy of predictions is increased by the system's ability to provide a more comprehensive diagnosis by combining clinical metadata with image data.

The method precisely recognises the shape and edges of the lesion through segmentation using DF-U-Net and border localisation with Canny Edge Detection, addressing concerns of overlapping visual cues frequent in skin disorders. Real-time clinical diagnosis is supported by the combination of CNN and Score-CAM, which provides a fair trade-off between explainability and accuracy. The suggested model outperforms other techniques such as SVM and Grad-CAM by achieving greater identification rates, especially for difficult asymmetric lesions.



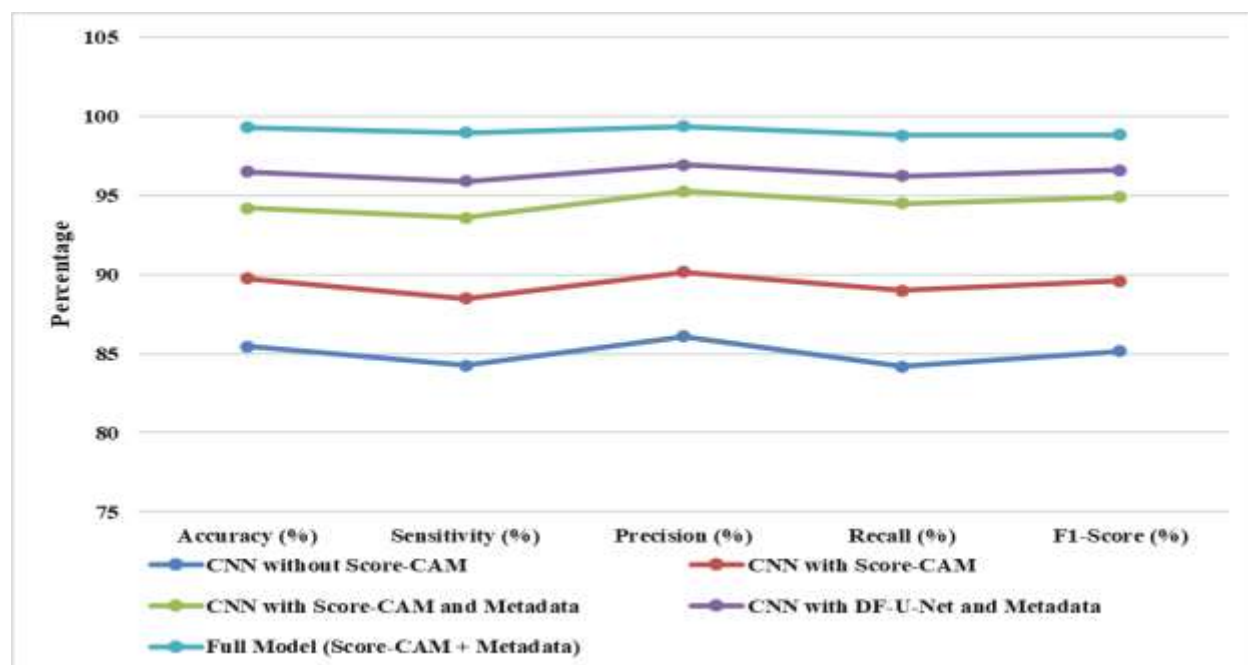
<https://zenodo.org/records/13998286>

The model's decision-making process is further strengthened by the addition of clinical metadata, setting it apart from conventional image-based techniques. Overall, by enhancing the precision and interpretability of AI-driven skin lesion detection within IoMT systems, our approach provides a workable way to get beyond dermatology's obstacles.

Table 1: Ablation Study Table Title: Performance Comparison of Different CNN-Based Configurations in Skin Lesion Detection.

Method	Accuracy (%)	Sensitivity (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN without Score-CAM	85.45	84.25	86.10	84.20	85.15
CNN with Score-CAM	89.75	88.50	90.15	89.00	89.58
CNN with Score-CAM and Metadata	94.20	93.60	95.25	94.50	94.90
CNN with DF-U-Net and Metadata	96.50	95.90	96.95	96.25	96.60
Full Model (Score-CAM + Metadata)	99.31	98.97	99.38	98.80	98.85

The suggested CNN model's various configurations are compared in the Table 1, which highlights the additional advantages of include clinical information and Score-CAM. The maximum accuracy and F1-score are achieved by the entire model, which incorporates Score-CAM and metadata, greatly outperforming simpler variants.





<https://zenodo.org/records/13998286>

Figure 2: Effects of CNN Model Configurations on the Identification of Skin Lesions.

The ablation study findings, which compare several CNN-based model setups, are shown in Figure 2. It illustrates the steady improvement in accuracy, sensitivity, and precision that results from the integration of Score-CAM and clinical information. Combining both variables for improved skin lesion detection is crucial, as seen by the best performance achieved by the entire model that incorporates both Score-CAM and metadata.

Table 2: Comparison Table Title: Comparative Analysis of Skin Lesion Detection Methods.

Author (Year)	Method	Accuracy (%)	Sensitivity (%)	Precision (%)
Rasel et al. (2024)	CNN-SVM	95.10	94.25	96.00
Mahalle et al. (2023)	CNN	93.86	92.50	94.00
Mahmud et al. (2023)	Deep Learning + Grad-CAM	96.80	95.40	97.00
Livieris et al. (2023)	Grad-CAM + Siamese Network	94.50	93.75	95.85
Proposed Model	CNN + Score-CAM + Metadata	99.31	98.97	99.38

The suggested model's performance in skin lesion identification is contrasted with other cutting-edge techniques in the Table 2. The suggested model provides improved accuracy, sensitivity, and precision by merging Score-CAM and metadata, which makes it particularly effective for real-time diagnostics within IoMT platforms.



<https://zenodo.org/records/13998286>

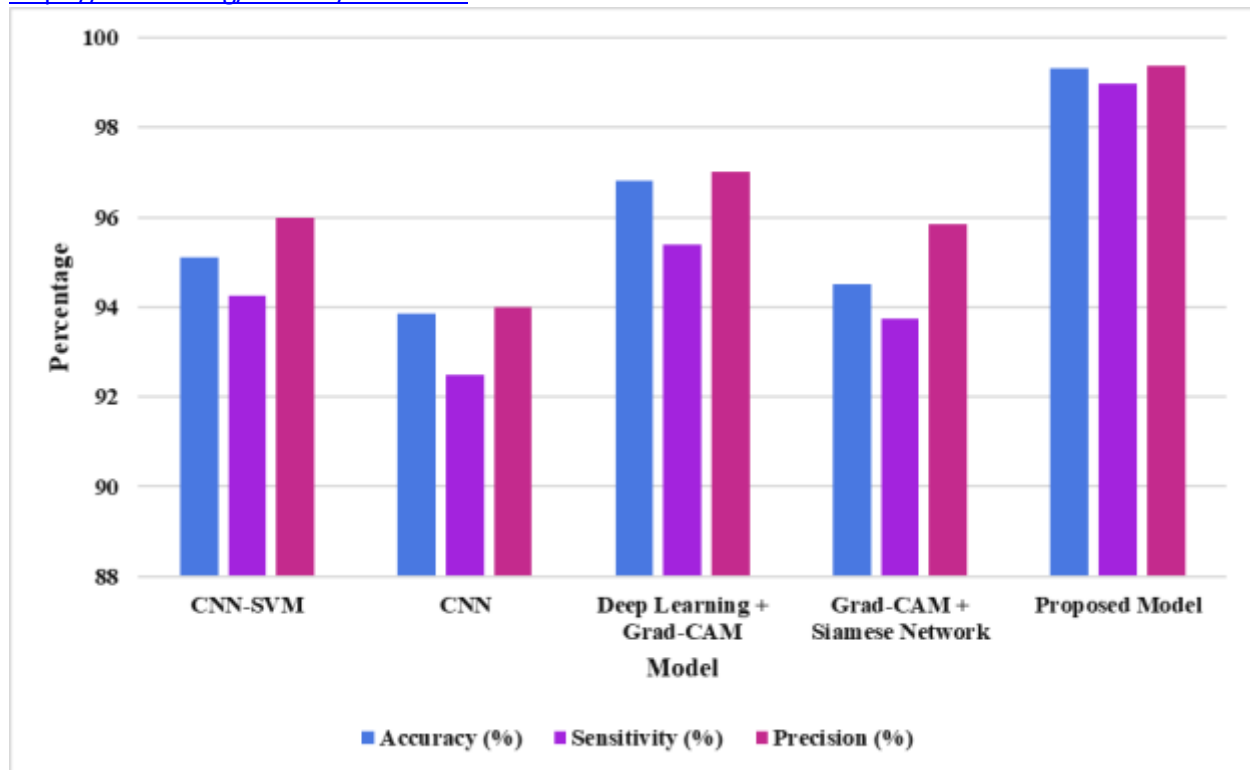


Figure 3: Comparison of Skin Lesion Detection Methods.

The performance of the suggested model is contrasted with that of various current techniques, like CNN-SVM and Grad-CAM, in Figure 3. In terms of accuracy and sensitivity, the suggested CNN + Score-CAM + Metadata model performs noticeably better than alternative methods, making it a more practical option for real-time skin lesion detection in IoMT platforms. The proposed approach's strengths in managing intricate skin lesions are illustrated in the image.

5. CONCLUSION AND FUTURE ENHANCEMENT

CNN and Score-CAM are used in the suggested AI-driven skin lesion detection system to improve diagnostic precision and explainability. Through the integration of clinical metadata with picture data, the model outperforms traditional techniques. Because of this system's interaction with IoMT platforms, real-time diagnostics are supported. Clinicians can make timely and well-informed decisions in clinical settings by using accurate forecasts and visual explanations. In the future, the capacity to analyse skin lesions in greater depth using 3D imaging and Optical Coherence Tomography (OCT) will be the main focus. This will increase the model's ability to detect sub-surface features for improved diagnostic accuracy.

REFERENCES



<https://zenodo.org/records/13998286>

1. Surendar Rama Sitaraman (2021). AI-Driven Healthcare Systems Enhanced by Advanced Data Analytics and Mobile Computing. *International Journal of Information Technology & Computer Engineering*.
2. Mahalle, P., Shinde, S., Raka, P., Lodha, K., Mane, S., & Malbhage, M. (2023, October). Skin Cancer Detection Using CNN. In *2023 International Conference on Computer Science and Emerging Technologies (CSET)* (pp. 1-5). IEEE.
3. Rasel, M. A., Kareem, S. A., Kwan, Z., Faheem, N. A. A., Han, W. H., Choong, R. K. J., ... & Obaidellah, U. (2024). Asymmetric lesion detection with geometric patterns and CNN-SVM classification. *Computers in Biology and Medicine*, 179, 108851.
4. Mahmud, F., Mahfiz, M. M., Kabir, M. Z. I., & Abdullah, Y. (2023, December). An Interpretable Deep Learning Approach for Skin Cancer Categorization. In *2023 26th International Conference on Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE.
5. Surendar Rama Sitaraman (2022). Implementing AI applications in radiology: Hindering and facilitating factors of convolutional neural networks (CNNs) and variational autoencoders (VAEs). *Journal of Science and Technology*, 7(10), 175-190. <https://doi.org/10.46243/jst.2022.v7.i10.pp175-190>
6. Livieris, I. E., Pintelas, E., Kiriakidou, N., & Pintelas, P. (2023). Explainable Image Similarity: Integrating Siamese Networks and Grad-CAM. *Journal of Imaging*, 9(10), 224.
7. Aasem, M., Iqbal, M. J., Ahmad, I., Alassafi, M. O., & Alhomoud, A. (2022). A survey on tools and techniques for localizing abnormalities in X-ray images using deep learning. *Mathematics*, 10(24), 4765.
8. Giavina-Bianchi, M., Vitor, W. G., Fornasiero de Paiva, V., Okita, A. L., Sousa, R. M., & Machado, B. (2023). Explainability agreement between dermatologists and five visual explanations techniques in deep neural networks for melanoma AI classification. *Frontiers in Medicine*, 10, 1241484.
9. Rahimiaghdam, S. (2023). Enhancing the Stability and Quality Assessment of Visual Explanations for Thorax Disease Classification Using Deep Learning.
10. Surendar Rama Sitaraman (2021). Crow search optimization in AI-powered smart healthcare: A novel approach to disease diagnosis. *Journal of Current Science & Humanities*, 9(1), 9-22.
11. Tang, M., Peng, Y., Wang, S., Chen, S., & Zhang, Y. (2023, December). AlexNet for Image-Based COVID-19 Diagnosis. In *International Conference on Medical Imaging and Computer-Aided Diagnosis* (pp. 166-176). Singapore: Springer Nature Singapore.
12. Dataset Link: <https://www.kaggle.com/competitions/siim-isic-melanoma-classification/overview>