ENHANCING TRUST IN SKIN LESION DIAGNOSIS THROUGH GRAD-CAM INTEROPERABILITY AND MEDICAL KNOWLEDGE INTEGRATION

Sanjeev Sharma and Akhtar Husain

Department of CSIT, MJPRU University,
Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

Abstract

Machine learning models have shown remarkable precision in skin lesion classification, but their acceptance in clinical settings is still restricted due to insufficient interpretability and consistency with medical reasoning. This study presents a framework that combines Gradient-weighted Class Activation Mapping (Grad-CAM) with knowledge from dermatology to improve model transparency and confidence.Grad-CAM is employed to produce visual explanations, which are systematically evaluated against recognized clinical features to determine alignment.The method put forward is assessed on benchmark dermoscopic datasets, indicating that the integration of medical priors enhances both interpretability and diagnostic dependability. The accuracy of the BCC/non-BCC classification was 91%. Regarding clinically meaningful XAI results, 98.9% accuracy was attained in identifying clinician-relevant BCC patterns. The average Grad-CAM normalized value for the human-assisted seprated clinical characteristics in the Clinically-inspired Visual XAI findings is 0.57.

Keywords

Basal cell Carcinoma (BCC), Squamous cell Carcinoma (SCC), XAI , Artificial Intelligence, Machine Learning, Deep Learning

1. INTRODUCTION

Cancer is the life-threatning disease, and skin cancer is one of them. Early detectionis the only way which gives excellent results in recovery of the patient. ML techniques show excellent results when combined with some pre-trained learning models. Detection of melanoma skin cancer achieves an advanced cure rate by using surgical treatment. Manual separartion of doubtful lesions abettance early melanoma disease.

Diagnosing various forms of skin cancer is the primary aim of numerous research projects. The creation of public databases has also led to a remarkable expand in the quantity of issued papers in recent years. Despite the fact that these databases are comprehensive and available to all, for identifying pupose the methods which are specific defined are not easily defined.. Providing a thorough diagnosis that explains the identified clinical features in addition to classification metrics is essential for creating a tool that is beneficial from a medical standpoint. This will yield a more complete outcome. Here, Serrano used the identification of dermoscopic criteria for BCC to create a clinically inspired skin lesion classification tool [1, 2, 3,4].

In any type of medical diagnostics, XAI methods like Grad-CAM are essential for enhancing CNNs' comprehensibility and transpicuousness. These approaches help medical practitioners

 DOI:10.5121/ijci.2026.150134

understand the prognosis principles of AI models By illustrating the rationale behind the identification decisions[5, 6, 7, 8]. This promotes acceptance and confidence in AI-driven diagnostics. Research on XAI has accelerated in the last couple of years, and a huge number of review papers released[8].

Specifically, XAI in skin image analysis has been the subject of multiple recent works[9]. Attentional modules are used by Barata et al. to enhance interpretability in the hierarchical diagnosis of skin lesions. To assist dermatologists in diagnosing melanoma, Chanda et al. created a multimodal XAI system that offers both textual and geographical explanations alongside their diagnostic predictions. Using clinical photos, Rez et al. have created an explicable skin cancer diagnostic model that enables general practitioners to identify and refer patients sooner. Their method uses Grad-CAM++ to provide visual explanations of model decisions[10, 11, 12].

In previous study, Explainable Machine learning is utilized to show the locations where the patterns have been targeted to identify key attributes. Providing clinical labeling and patient diagnosis explanation by highlighting areas of clinical importance for diagnosis, such as the dermoscopic characteristics of BCC, is an additional step in the proposed method.

2. METHODOLOGY

Figure 1 illustrates the Skin lesion XAI framework, which uses a multimodal approach to address the challenges of skin lesion diagnosis. Because of its methodical approach and extensive improvement, this innovative approach offers exceptional transparency and interpretability in the categorization of skin lesions, representing a significant advancement in the field. In order to improve model training and guarantee the framework's stable generalizability, the first step entails data curation. HAM10000, which has a sizable collection of dermoscopic images, is used in the current study. This dataset, which includes roughly 10,000 photos of various benign and malignant skin lesion types, is well-known in the domains of dermatology and machine learning. Following the acquisition of the data set, a thorough data preparation procedure is started, involving a number of methods meant to improve the precision and consistency of the input data. To guarantee that the input features have consistent scales and distributions, preprocessing involves normalizing and standardizing the data.

Maintaining homogeneity in the data is crucial for enabling the best possible model development, and this includes rescaling feature values using methods like Min-Max scaling and Z-score standardization. Using techniques like edge detection, texture analysis, and morphological operations, feature engineering in the context of dermatoscopic pictures entails extracting pertinent features from the images, such as color, texture, shape, and boundary characteristics. The qualities obtained from this procedure are crucial inputs for deep learning systems, which enable accurate identification and categorization of important characteristics linked to skin diseases.

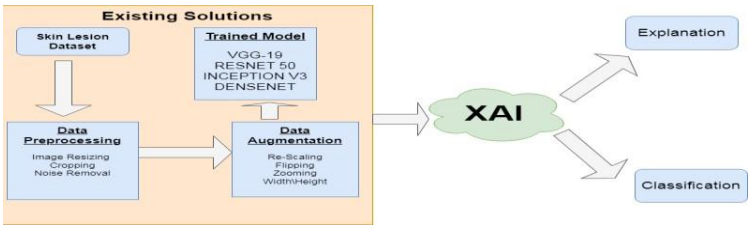

Figure 1. A proposed framework for Skin lesion classification XAI, Skin lesion Dataset; VGG-19, Resnet 50, Inception v3,densenet.

Rotation, translation, scaling, flipping, and color changes are all part of the extensive augmentation process that the data undergoes during preprocessing. Pixel values are standardized by scaling them between using image rescaling. Next, picture resizing makes sure that every image fits the input dimensions of the model. Various expectations techniques, like length, rotation, height, width, and zoom, are utilized to create different visual viewpoints that can be used to enhance learning. These include random rotation and height and width shifts, which enable translations in both the directions. The use horizontal flip, which twists images horizontally, and the zoom range, which randomly zooms in and out, can further diversifies the collected data.

3. DATASET COMPOSITION

The collection of dataset was a highly challenging and time taken task due to ensure reliability for research and development. However, this method presents several benefits in biomedical research, especially through the use of detected newly skin images that remain unexamined included in the strategies of other researchers. On the other hand, an important challenge arises because self-collected images do not undergo prior preprocessing, making the model training process more complexThere are so many world reconize datasets related to skin images. There are two big datasets of melonama related images are ISIC2020 and HAM10000. These two datasets contains lots of images conerned to melanoma skin cancer. These collections cover a wide variety of melanocytic nevus, SCC , different type of skin conditions, BCC, actinic keratosis, benign keratosis, dermatofibroma, melanoma, squamous cell carcinoma, and vascular lesions. In contrast, our dataset features images of several uncommon skin diseases. Dataset description is presented in Table 1."

Name	Description
Total Label	7470
Feature	224*224
Graphic format	JPG

Table1. Skin Disease Dataset Details

Each image can exhibit a range of dermoscopic patterns. To facilitate the labeling process during image annotation and to enable subsequent analysis of the dermatologists' annotations, a one-hot encoding scheme was employed. Each label associated with an image is represented as a binary vector, where each digit corresponds to a specific dermoscopic pattern associated with basal cell carcinoma (BCC). The presence of a pattern is denoted by a value of 1, while its absence is represented by 0. In a BCC lesion, there are several distinct patterns can be observed.

4. DERIVING STANDARD REFERENCE

Biopsies are used to establish the standard ground truth (GT) for Basal Cell Carcinoma analysis. Nevertheless, the dermoscopic patterns of BCC, which are assessed subjectively by dermatologists, do not yet have a recognized GT. When evaluating inter-expert consistency regarding multiple dermoscopic findings seen in lesions, a number of studies have shown a low kappa coefficient. As a result, a trustworthy single reference (SR) that is based on the agreement of several dermatologists is crucial. Multiple specialized labels were synthesized into a single SR for model training using an Expectation-Maximization (EM) based technique. By combining

multiple category labels provided by many dermatologists, our system derives a single SR estimation that represents the evaluators' combined competence. Biopsies are used to establish the standard ground truth (GT) for Basal Cell Carcinoma analysis. Nevertheless, the dermoscopic patterns of BCC, which are assessed subjectively by dermatologists, do not yet have a recognized GT. when we see the different patterns given by different dematologists and try to evaluate these patterns the maximum studies shows low cofficients. As a result, a trustworthy single reference (SR) that is based on the agreement of several dermatologists is crucial. Multiple specialized labels were synthesized into a single SR for model training using an Expectation-Maximization (EM) based technique. Combine the outcomes of many dermatologists and the system produce a single evaluated scheme.This algorithm was used by Silva et al. to extract an SR from the BCC pattern.

5. XAI INSPIRED BY CLINICAL PRACTICES

From a clinical standpoint, achieving complete precision in recognizing dermoscopic patterns of basal cell carcinoma (BCC) is not crucial for making an accurate diagnosis. A dermatologist only needs to identify one valid dermoscopic pattern of BCC to confirm the diagnosis. Therefore,Clinically focused XAI is producing any one outcome: if we are not found particular kind it means it is non-BCC diagnosis and if Pigmented Nevus(PN) pattern find then it support the prediction of non-BCC and produce negative indicators.

Symbolization	Legion	Unicoding	Distinctive
Image	[0101100]	1	BCC Menifestation
Image	[1000000]	0	Negative for BCC
Image	[0000000]	1	Negative for BCC

Table 2 : Illustration of Legion and Unicoding for Detection of BCC

The MobileNet-V2 architecture, which consists of three classifier layers, is used in the suggested artificial intelligence tool for skin lesion diagnosis. Three separate stages made up the optimization and training process for this model. The classifier weights were first trained using transfer learning from ImageNet. The final three layers of the feature encoder and classification module were then fine-tuned using a lower learning rate and more training epochs. In the last stage, retraining was limited to the classifier with a very low learning rate, and ML was employed from the Unicoding model to address the detection of basal cell carcinoma (BCC) patterns.

6. CLINICAL VISUAL EXPLAINABLE ARTIFICIAL INTELLIGENCE

To further explore to know more about the facts use segmentations.One expert create more segmentations for a particular part to know more specially one dermatologist create more segmentaion of given image to find BCC pattern.

To understand the working of CNN to take decisions, where metrics importance are more required. There are so many visualization techniques explainable AI and gradient weighted class mapping are the best for mapping. Feature maps capture patterns from various network layers, attention of the model is ; however, these outputs don't always lead to predictable results. described by Zhou et al. [20].

In this research, the decomposition of distinct BCC patterns were merged into a unified decomposed image, as illustrated in Fig. 2. Manually executed partitions were then

superimposed with actuation maps to corroborate the medical understanding facilitated by Grad-cam. This approach elucidates not only the areas of focus for the AI tool but also the dermoscopic characteristics of those regions.

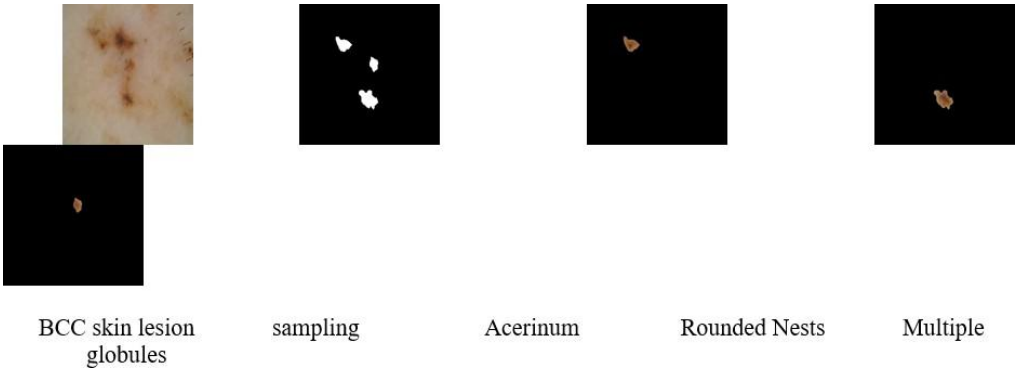

Figure2. An illustration of the segmentation of dermoscopic patterns in BCC

7. RESULTS

7.1. Deployment Section

The selected streamliner for the study given by AdamW streamliner [21], utilizing a small samples of 38 and a withdraw outlay of 0.3 overfitting risk. To tackle issue class imbalance, focal loss [22] was employed.

Among the data augmentation methods used were Gaussian blur, perspective transformation, and rotation [23, 24].

Given the restricted dimension of our database, we applied stratum k-fold subsampling to facilitate a thorough assessment of the framework effectiveness. Such method helps reduce the likelihood of unbalanced distributions in optimization and assesment, which is particularly important in datasets that exhibit unbalanced class ratios.

8. CLINICALLY GUIDED XAI: ENHANCED BCC DETECTION WITH RICH LABEL

8.1. Insights

The effectiveness of the AI tool in BCC detection identification is examined this part along With labeling information that are supplied to clarify this clustering. The considering yardsticks are outlined in Table 3.

The yardsticks are calculated as an average over all folds. This table is organized into three distinct sections. The first section presents the effectiveness of the AI tool in two-class categorization. Second section demonstrates its capability in recognizing Characteristic dermoscopic traits of BCC. The concluding segment evaluates the precision of the labels that offer clinical interpretations.

Table 3. Assessment by physician-guided analysis, a refined binary classifier, and binary and multilabel classification metrics.

	Recall	Specifity	Precision	Accuracy
BCC/NON - BCC	0.89	0.89	0.90	0.90
Pattern Detection				
Pigment Network	0.94	0.96	0.97	0.95
Ulceration	0.81	0.76	0.52	0.77
Ovid Nests	0.67	0.84	0.53	0.84
Multiglobules	0.61	0.81	0.32	0.80
Maple Leaf - Like	0.60	0.87	0.37	0.84
Spoke Wheel	0.89	0.76	0.61	0.80
Clinical Inspired XAI				
All 0's	- -	--	--	0.73
Pigment Network	0.94	0.96	0.97	0.95
BCC Pattern Detection	0.84	0.88	0.71	0.97

The diagnostic performance for BCC and non-BCC classifications is notably high, approximately 0.9 across all metrics. Nevertheless, a more thorough examination of the BCC pattern detection performance is necessary. Minority classes often exhibit low recall rates due to the AI tool's training on unbalanced datasets, which prefers the larger portion. Classes such as SW, MG , and machine learning are not protected. The implementation of strategies such as increased data, advanced sampling methods, and startum K- bell cross-checks, with the 1 VS – MALL approach, contributed to achieving more fair classification according to various prototypes. The overall performance of the prototype. Nevertheless, the obtained metrics cannot not be interpreted the performance same as BCC/ Non BCC. It exclusively concluded in terms of ability to explain binary classification correctly.

The significance lies not in the AI tool's failure to identify a specific BCC pattern, but rather in its ability to recognize any BCC pattern, as clinicians approach the diagnosis of skin lesions uniformly. This supplementary analysis is presented in the defined section of Table 3, demonstrating that 71% of non-Basal Cell Carcinoma (BCC) samples, 94% of non-BCC lesions exhibiting pigmentation, and 98% of BCC lesions displaying a recognizable BCC pattern are accurately segmented.

9. VISUAL XAI INSPIRED BY CLINICAL PRACTICES

This particular part seeks to analyse the precision of the learning model in targeting appropriate regions of the lesion, particularly the dermoscopic features of BCC adorned recognized by specialist. For achieving this the dermatologists already described BCC pattern compare with the area outined by the model.

We evaluated the model's performance by examining the conditional probability distribution functions of formalized GradCAM values for both manually labeled areas and non-significant areas. This analysis aimed to assess how well the model's activation zones corresponded with clinically relevant areas. In our notation, z(x, y) denotes the GradCAM value at the position (x, y). Define SA as the region marked by the dermatologist and DA as the overall area. The probability density function of GradCAM values for pixels within SA is denoted as P(z(x, y) | SA), while P(z(x, y) | Da) represents the same for pixels in Da.

Table 4 :Statistical analysis derived from the approximated dependent dissemination basis of Grad-CAM, clinically relevant regions.

Outcome	Overlap Score	Mean SA	Mean DA	Standard SA	Standard DA
Accurate	0.22	0.56	0.15	0.13	0.22
Inaccurate	0.31	0.32	0.13	0.01	0.21

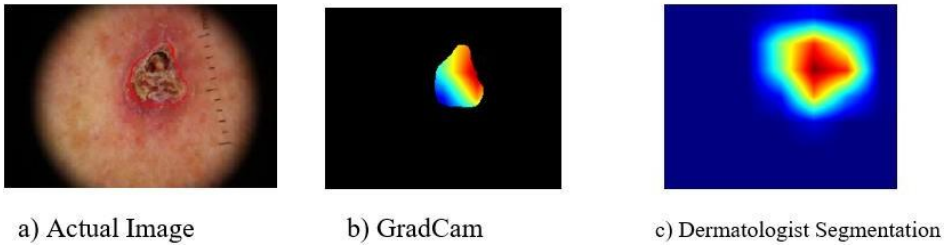

Figure 4 :A visually-driven interpretation of an accurate clinical prediction.

Table 4 presents a summary of the data derived from the probability density functions. It specifically details the mean and data variability of z (x, y) for the sets (x, y) ∈ SA and (x, y) ∈ DA, along with the converging area between P (z (x, y) | SA) and P (z (x, y) | DA). The findings indicate that samples predicted correctly exhibit a greater mean standard deviation compared to those predicted incorrectly. Furthermore, the intersection area is also more extensive in these instances. These observations highlight the model's inability to focus on clinically relevant areas when making incorrect predictions.

10. CONCLUSION

This paper presents the development of an AI tool designed for the diagnosis of BCC, which offers a clinical explanation. It accomplishes two key objectives that enhance its clinical relevance. Firstly, this tool is set to transform teledermatology protocols, significantly decreasing the waiting time for diagnosis. Secondly, the tool's dual clinical explanations enhance its practical applicability.

REFERENCES

[1] American Academy of Dermatology. Types of Common Skin Cancer. 2024.
[2] Skin Cancer Foundation. Skin Cancer Information. https://www.skincancer.org/ skin-cancer-information/#, 2024.
[3] Magdalena Cia¸z˙yn´ska, Joanna Narbutt, Anna Woz´niacka, and Aleksandra Lesiak. Trends in basal cell carcinoma incidence rates: a 16-year retrospective study of a population in central poland. Advances in Dermatology and Allergology/Post¸epyDermatologiiiAlergologii, 35(1):47–52, 2018.
[4] Ketty Peris, Maria Concetta Fargnoli, Claus Garbe, Roland Kaufmann, Lars Bastholt, Nicole Basset Seguin, Veronique Bataille, Veronique Del Marmol, Reinhard Dummer, Catherine A Harwood, et al. Diagnosis and treatment of basal cell carcinoma: European consensus–based interdisciplinary guidelines. European Journal of cancer, 118:10–34, 2019.
[5] Noel C. F. Codella, David A. Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin K. Mishra, Harald Kittler, and Allan Halpern. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (ISIC). CoRR, abs/1710.05006, 2017.
[6] Marc Combalia, Noel C. F. Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Cristina Carrera, Alicia Barreiro, Allan C. Halpern, Susana Puig, and Josep Malvehy. Bcn20000: Dermoscopic lesions in the wild. 2019.
[7] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-sourcedermatoscopic images of common pigmented skin lesions. Scientific Data, 5(1), August 2018.
[8] International skin imaging collaboration (isic) archive. 2024.
[9] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. International Journal of Computer Vision, 128(2):336–359, October 2019.

[10] Bas H.M. van der Velden, Hugo J. Kuijf, Kenneth G.A. Gilhuijs, and Max A. Viergever. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Medical Image Analysis, 79:102470, July 2022.

[11] Ruey-Kai Sheu and Mayuresh Sunil Pardeshi. A Survey on Medical Explainable AI (XAI): Recent Progress, Explainability Approach, Human Interaction and Scoring System. Sensors, 22(20), 2022.

[12] [12] Shahab S Band, Atefeh Yarahmadi, Chung-Chian Hsu, Meghdad Biyari, Mehdi Sookhak, Rasoul Ameri, Iman Dehzangi, Anthony Theodore Chronopoulos, and Huey-Wen Liang. Application of explainable artificial intelligence in medical health: A systematic review of interpretability methods. Informatics in Medicine Unlocked, 40:101286,2023.

[13] Catarina Barata, M. Emre Celebi, and Jorge S. Marques. Explainable skin lesion diagnosis using taxonomies.Pattern Recognition, 110:107413, 2021.

[14] Tirtha Chanda, Katja Hauser, Sarah Hobelsberger, Tabea-Clara Bucher, Carina Nogueira Garcia, and et al. Wies. Dermatologist-like explainable AI enhances trust and confidence in diagnosing melanoma. Nature Communications, 2024. All Open Access, Gold Open Access, Green Open Access.

[15] Eman Rezk, Mohamed Eltorki, and Wael El-Dakhakhni. Explicable skin cancer classification based on incremental domain knowledge learning. Journal of Healthcare Informatics Research, 7(1):59 – 83, 2023. Cited by: 4.

[16] Ketty Peris, Emma Altobelli, Angela Ferrari, Maria Concetta Fargnoli, Domenico Piccolo, Maria Esposito, and Sergio Chimenti. Interobserver agreement on dermoscopic features of pigmented basal cell carcinoma. Dermatologic surgery, 28(7):643–645, 2002.

[17] Scott W Menzies, Karin Westerhoff, Harold Rabinovitz, Alfred W Kopf, William H McCarthy, and Brian Katz. Surface microscopy of pigmented basal cell carcinoma. Archives of dermatology, 136(8):1012–1016, 2000.

[18] Sam Polesie, Lisa Sundback, Martin Gillstedt, Hannah Ceder, Johan Dahlen Gyllencreutz, Julia Fougelberg Eva Johansson Backman, Jenna Pakka, ZAAR Oscar, and John Paoli. Interobserver agreement on dermoscopic features and their associations with in situ and invasive cutaneous melanomas. Acta Dermato-Venereologica, 101(10), 2021.

[19] Alexander Philip Dawid and Allan M Skene. Maximum likelihood estimation of observer error-rates using the em algorithm. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(1):20–28, 1979.

[20] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning deep features for discriminative localization. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2921– 2929, Los Alamitos, CA, USA, jun 2016. IEEE Computer Society.

[21] Aaron Defazio, Harsh Mehta, Konstantin Mishchenko, Ahmed Khaled, Ashok Cutkosky, et al. The road less scheduled. arXiv preprint arXiv:2405.15682, 2024.

[22] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. Focal loss for dense object detection. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 2999–3007, Los Alamitos, CA, USA, oct 2017. IEEE Computer Society.

[23] [23] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, Advances in Neural Information Processing Systems, volume 25. Curran Associates, Inc., 2012.

[24] P.Y. Simard, D. Steinkraus, and J.C. Platt. Best practices for convolutional neural networks applied to visual document analysis. In Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings., pages 958–963, 2003.