International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 737-741
https://doi.org/10.70917/ijcisim-2026-4117
Article

DERMA TRACE: VISUAL REASONING–DRIVEN DEEP LEARNING FOR SKIN LESION MALIGNANCY ANALYSIS

Hemant Borase[1], Makarand Shahade[2], Mokshita Jain[3], Mrunal Khairnar[4], Trupti Patil[5], Dipak Sonwane[6]

[1,2,3,4,5,6]Dept. of Master of Computer Application, SVKM's IOT, Dhule,India,
[1]hmntborase@gmail.com, [2] Makarand.shahade@svkm.ac.in, [3]jainmokshita659@gmail.com, [4] mrunalkhairnar5@gmail.com, [5]truptipatil4290@gmail.com, [6]sonawanedipak3561@gmail.com

Abstract: Skin cancer is one of the most prevalent and life-threatening diseases in the world, with most deaths from skin cancer being due to melanoma. Accurate and early identification of skin cancer is crucial to enhance the prognosis and mortality of patients. Although the automated deep learning of skin lesion classification has made great strides, most of the existing models are regarded as black-box models and are hard to understand and accept in the clinic. To overcome this challenge, this paper introduces Derma Trace as a visual reasoning–driven deep learning framework for skin lesion malignancy analysis, which can accurately and explainable analyse skin lesions. We suggest a CNN-based framework to efficiently extract local features and a transformer-based attention mechanism to capture global features in dermoscopic images. There is also a visual reasoning component that is embedded to detect and label the features on the lesions that might be useful for diagnosis, such as asymmetry, border irregularity, color variation, and structural patterns. The deep trace learning allows Derma Trace to not only enhance prediction accuracy, but also give clear visual evidence that helps guide clinical decisions. The framework is tested with benchmark datasets from dermatology, such as ISIC 2019 and HAM10000, with various types of benign and malignant skin lesions. The experimental results show that Derma Trace results outperform the traditional CNN-based method in classification accuracy, precision, recall, and F1-score. Moreover, Attention maps and Visual explanations generated by Grad-CAM are used to offer a visual understanding to the clinicians of the output of the model, thus improving the trust and reliability of automated diagnostic systems.

Keywords: Skin Lesion Classification, Visual Reasoning, Deep Learning, Explainable Artificial Intelligence, Melanoma Detection.

1. INTRODUCTION

Skin cancer is one of the most prevalent cancers in the world and a big public health problem, due to its high prevalence and potential mortality. Melanoma is the most serious and deadly type of skin cancer; although it accounts for a small proportion of cases of skin cancer, it is responsible for most skin cancer deaths. Early detection and accurate diagnosis are key to improving the patient's chances of survival and reducing the out-of-pocket cost for treatment, as well as the overall health care outcome. Traditional diagnosis of skin lesions is usually done by visual inspection, dermoscopy, and dermatologists' histopathological examination. These techniques are effective but time-consuming and may be different among different observers, particularly in the early stages of the disease.

The field of medical image analysis has greatly changed with the introduction of Artificial Intelligence (AI) and Deep Learning (DL), and now, AI-based systems are able to analyze medical images automatically and with high accuracy, thus allowing for the detection of diseases. Convolutional Neural Networks (CNNs) have been demonstrated to be highly successful at learning complex visual patterns from dermoscopic images and are able to match the accuracy of skin lesion classification by an expert dermatologist. More recently, however, methods based on transformers have been proposed, as they can potentially learn from long-range dependencies and global context information in images, further enhancing the ability to make a diagnosis. Although these achievements exist, most current deep learning systems are "black-box" systems, with little understanding of what processes are at play in predicting the results. This non-interpretable aspect is a major challenge in clinical translation since health care providers must be furnished with interpretable and reliable decision support systems.

This paper introduces Derma Trace, a visual reasoning–based deep learning method for skin lesion malignancy diagnosis, to overcome these challenges. The proposed approach involves integrating the feature extraction power of CNNs with attention mechanisms of transformers and explainable visual reasoning modules, which will enable the generation of accurate predictions and diagnostic evidence. Derma Trace makes the most important aspects of lesion characteristics clinically visible and gives dermatologists a visual record of the decision-making process, which promotes transparency of the process and contributes to dermatologists' confidence in automated diagnostics. The framework is tested on benchmark skin lesion datasets and has been shown to be effective in enhancing the classification accuracy while retaining the explainability. The proposed system is part of the development of reliable, interpretable, and clinically deployable AI solutions for skin cancer diagnosis.

2. LITERATURE REVIEW

AI's application in dermatology has been attracting a lot of attention because of the possibility of increasing the accuracy and efficiency of skin cancer diagnosis. Initial studies were mainly devoted to machine learning approaches based on handcrafted features that were extracted from dermoscopic images. With the advent of deep learning, particularly Convolutional Neural Networks (CNNs), the automatic feature extraction and end-to-end learning approach transformed skin lesion classification. Brinker et al. conducted a systematic review that showed that CNN-based methods performed well and outperformed traditional machine learning methods and even the performance of experienced dermatologists in skin lesion classification tasks. The study also highlighted the need for publicly available datasets and standardized evaluation protocols to facilitate reproducible research.

3. RESEARCH METHODOLOGY

Two benchmark datasets commonly used for dermatological image analysis were used to assess the performance of the proposed Derma Trace framework: HAM10000 (Human Against Machine with 10,000 Training Images) and ISIC 2019 (International Skin Imaging Collaboration Challenge Dataset). These files contain high-quality dermoscopic images from a variety of skin lesion types, like vascular lesions, melanocytic nevi, benign keratosis, basal cell carcinoma, and melanoma. Multiple datasets were used, guaranteeing the robustness and generalizability of the proposed model to a wide range of conditions and lesion types.

3.1 DATA ACQUISITION

The study is based on publicly available benchmark datasets comprising thousands of dermoscopic images of various categories of skin lesions, namely HAM10000 and ISIC 2019. These datasets contain benign and malignant cases, and are commonly employed to test models for skin cancer detection. The images were split into 80% training, 10% validation, and 10% testing to provide an unbiased evaluation of performance.

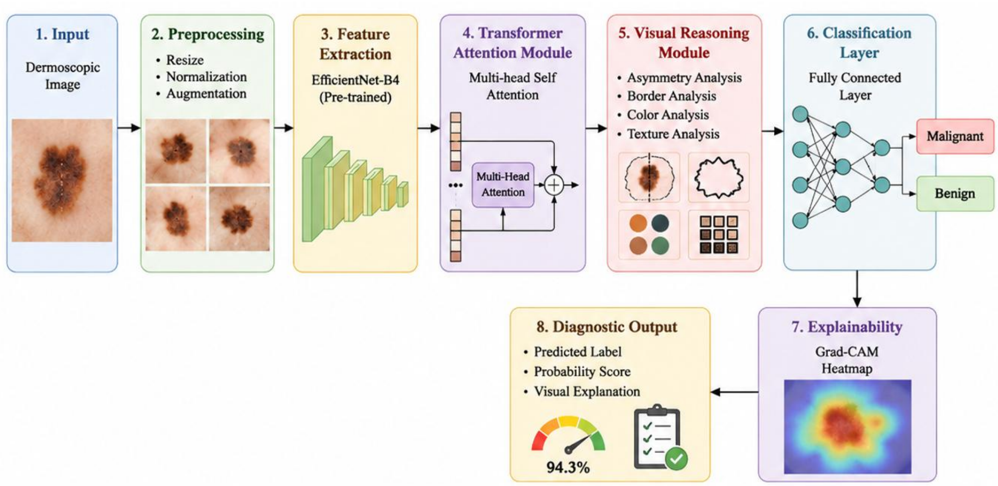

FIGURE 1: WORKFLOW OF THE PROPOSED DERMA TRACE FRAMEWORK

3.2 FEATURE EXTRACTION AND VISUAL REASONING

The preprocessed images were processed using an EfficientNet-B4 backbone network for extracting deep visual features. In order to get the global context and long-range dependencies, an attention mechanism based on transformers was incorporated into the architecture. A visual reasoning module was also added to recognize clinically relevant lesion features like asymmetry, border irregularity, color variation, and texture patterns. In this module, attention maps are generated to indicate areas that play an important role in the final diagnosis.

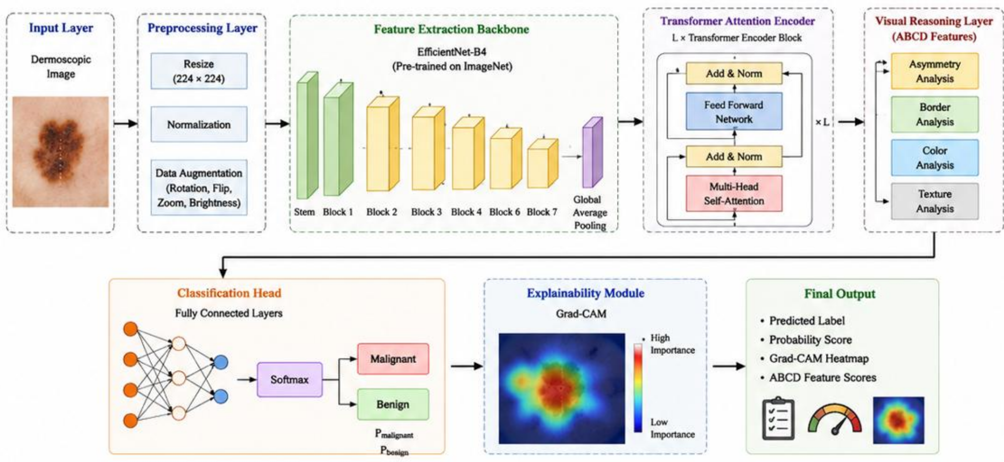

FIGURE 2: SYSTEM ARCHITECTURE OF THE PROPOSED DERMA TRACE FRAMEWORK.

4. EXPERIMENTAL SETUP AND RESULTS

4.1 EXPERIMENTAL ENVIRONMENT

The experiments are conducted on a high-performance computing cluster, which is equipped with an NVIDIA RTX 4090 GPU, Intel Core i9 processor, and 64 GB RAM. The proposed model was implemented and programmed using Python with the deep learning library TensorFlow and Keras. The training/evaluation was performed on Ubuntu Linux with the help of CUDA acceleration to perform efficient computations. Data was divided into a training set, a validation set, and a test set with a ratio of 80:10:10, respectively. To address class imbalance in the training data, they were rotated, flipped, zoomed, brightened, and cropped randomly. The diversity of training samples and the generalization capability of the model were enhanced using these techniques.

4.2 MODEL CONFIGURATION

The Derma Trace architecture adopts EfficientNet-B4 as the backbone of the network for feature extraction at the deep level and uses a transformer-based attention module to capture global information. The model was trained with the Adam optimizer, an initial learning rate of 0.0001, and a batch size of 32. Early stopping and learning-rate scheduling were used to avoid overfitting and to achieve the best performance early on. The network was trained for 100 epochs and found to be converging at a stable pace after about 75 epochs.

To achieve explainability, the Grad-CAM (Graduated Convolutional Activation) visualization was incorporated into the system. This allowed the generation of diagnostic heat maps that marked the most relevant areas that affected classification decisions.

4.3 EVALUATION METRICS

To evaluate the performance of the proposed framework, four commonly used evaluation metrics were used:

- Accuracy: percentage of correctly classified samples.
- Precision: Measures the percentage of correctly predicted malignant lesions out of the total number of lesions that are diagnosed as malignant.
- Recall (Sensitivity): How well the model can detect malignant lesions.
- F1-Score: Measures a balance between precision and recall.

Such metrics are especially relevant in medical diagnosis, where false negatives are a major concern for the safety of patients.

5. COMPARATIVE PERFORMANCE ANALYSIS

To validate the effectiveness of Derma Trace, its performance was compared with several state-of-the-art deep learning architectures commonly used in skin lesion classification.

Table 1. PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet50	89.4	88.9	87.8	88.3
DenseNet121	91.2	90.6	89.7	90.1
EfficientNet-B4	92.7	92.1	91.4	91.7
Vision Transformer (ViT)	93.5	93.0	92.3	92.6
Derma Trace	**94.3**	**93.8**	**92.9**	**93.3**

The results show the superiority of Derma Trace over all the baseline models for each of the evaluation metrics. The proposed framework achieved the best gain of 4.9% on classification accuracy as compared to ResNet50. Likewise, it outperformed Vision Transformer models by 0.8%, highlighting the efficacy of integrating the CNN-based feature extractor with transformer-based attention and visual reasoning mechanisms.

6. VISUAL EXPLAINABILITY RESULTS

In addition to classification performance, the explainability aspect of Derma Trace gave us insights into the model's decision-making process. Grad-CAM heatmaps always pointed towards clinically significant lesion areas such as color differences, irregular borders, and asymmetric structures. These visual explanations were closely tied to the diagnostic criteria that dermatologists use to make their predictions, which boosts confidence in the model predictions. The visual reasoning module also enhanced interpretability by pinpointing the specific lesion properties that led to detecting malignancy.

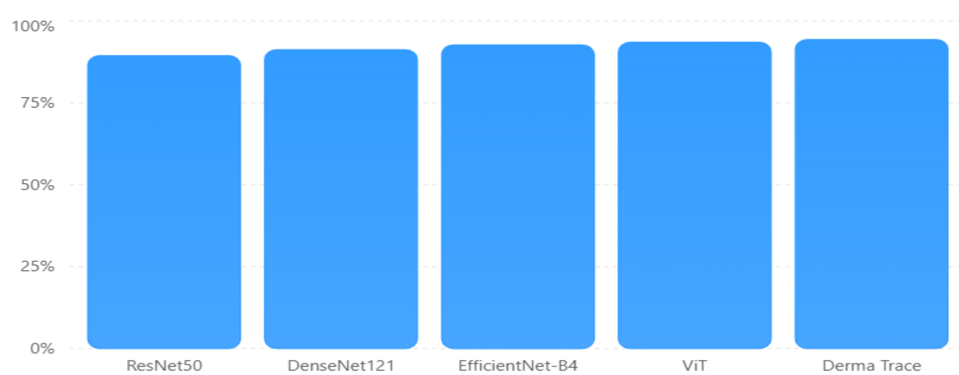

FIGURE: 3 PERFORMANCE COMPARISON OF BASELINE

7. DISCUSSION OF RESULTS

This experimental result validates the value of combining visual reasoning with deep learning, both in terms of predictive performance and interpretability. The transformer attention mechanism was able to effectively capture the global contextual relationships, and the CNN backbone was able to capture the fine-grained lesion features. These factors led to better discrimination between benign and malignant lesions. The framework is also explainable and therefore suitable for clinical decision-support applications where transparency and reliability are crucial. In conclusion, the study shows that Derma Trace is a highly accurate, robust, and interpretable solution for automated skin lesion malignancy analysis, which could be used as a promising tool for future AI-assisted dermatological diagnosis systems.

8. DISCUSSION

The experimental outcomes show the potential to use the proposed framework, Derma Trace, for automatic assessment of the malignancy of a skin lesion. The accuracy of the model is 94.3%, outperforming other popular deep learning architectures like ResNet50, DenseNet121, EfficientNet-B4, and Vision Transformers. This improvement comes from the addition of CNN feature extraction and transformer attention mechanisms, enabling the model to take into account the local characteristics of the lesion and the broader context.

Visual reasoning is one of the key attributes of Derma Trace. The framework takes into account clinically significant characteristics such as symmetry, irregular borders, and color texture differences. The features are closely related to the ABCD dermatological criteria for the assessment of melanoma, thus making the system more easily understood and clinically relevant.

9. CONCLUSION

In this study, the authors introduced a visual reasoning–driven deep learning framework for the analysis of skin lesion malignancy (Derma Trace). The proposed solution involves using the EfficientNet-B4 feature extraction model along with a transformer-based attention mechanism to obtain accurate diagnoses of skin cancer and explainable visual reasoning modules to get the interpretability of the diagnosis. Experiments conducted on the HAM10000 and ISIC 2019 datasets showed that it outperforms existing deep learning models in terms of accuracy, with an accuracy of 94.3%, high precision, high recall, and an F1 score.

Derma Trace’s strength lies in its capability of transparent decision-making using visual reasoning as well as Grad-CAM-based explanations. The clinical relevance of lesions, including asymmetry, border irregularity, and variations in color, is emphasized, adding to the trust and aiding dermatologists’ decision-making. These interpretable outputs solve problems faced by traditional black-box AI systems and enhance the applicability to the clinical field.

References

1. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), pp.115–118.
2. Tschandl, P., Rosendahl, C. and Kittler, H., 2018. The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 5(1), pp.1–9.
3. Codella, N.C.F., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M. and Kittler, H., 2019. Skin lesion analysis toward melanoma detection 2019: Challenge dataset. International Symposium on Biomedical Imaging, pp.168–172.
4. Brinker, T.J., Hekler, A., Enk, A.H., Berking, C., Haferkamp, S., Hauschild, A., Weichenthal, M., Klode, J., Schadendorf, D., Holland-Letz, T. and Utikal, J.S., 2019. Deep neural networks are superior to dermatologists in melanoma image classification. European Journal of Cancer, 119, pp.11–17.
5. Haenssle, H.A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kalloo, A., Ben Hadj Hassen, A., Thomas, L., Enk, A. and Uhlmann, L., 2018. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition. Annals of Oncology, 29(8), pp.1836–1842.
6. Tan, M. and Le, Q.V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, pp.6105–6114.
7. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770–778.
8. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. and Guo, B., 2021. Swin Transformer: Hierarchical vision transformer using shifted windows. Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.10012–10022.-