International Journal of Advanced Mathematical Sciences, 11 (1) (2025) 73-76

International Journal of Advanced Mathematical Sciences

Website: www.sciencepubco.com/index.php/IJAMS
https://doi.org/10.14419/w0p98q32
Research paper

ViT-enhanced ai-powered deep learning framework for skin disease diagnosis

Dhiraj Chavan [1] *, Prathamesh Sonawane [1], Neha Jadhav [1], Janhavi Naidu [1], Mrs. Mital Kadu [2], Mrs. Vibhavari Jawale [2]

[1] *Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune - 44, India*
[2] *Dr. D. Y. Patil International University, Akurdi, Pune - 44, India*
**Corresponding author E-mail: dhirajchavan9421@gmail.com*

Abstract

Skin concerns are a rising health issue globally, and accurate detection and diagnosis of these issues are key in preventing serious consequences. We provide a complete overview of deep learning approaches to dermatoscopic image classification, specifically focusing on the newly developed Vision Transformer (ViT) approaches. We discuss the advantages of using ViT approaches in skin disease classification versus prior deep learning approaches, specifically, convolutional neural networks (CNNs). We analyse four commonly reported skin conditions: Basal Cell Carcinoma, Benign Keratosis, Eczema, and Melanoma. In doing so, we explore the current literature and datasets available and summarize the advancements of artificial intelligence (AI) in dermatology, identify potentially the most effective designs, and consider their incorporation into clinical populations. This review is intended to provide insight into the current developments in the design of AI-AI-assisted automated skin disease diagnosis processes, including important trends, performance, and efficiencies of models, and the current trends in skin disease diagnosis. In an ideal world, this review will provide a foundation for the development of more accurate and ultimately less expensive diagnostics to enhance patient care in dermatology.

Keywords: *Vision Transformer (ViT); Skin Disease Classification; Basal Cell Carcinoma; Benign Keratosis; Eczema; Melanoma; Deep Learning; Medical Imaging; Dermatology; Transformer Models.*

1. Introduction

The skin is the largest organ in the body, and it plays an important role as a natural barrier against environmental risks and pathogens. Despite its protective function, the skin is still at risk for various skin diseases. Skin diseases can become serious health issues if they are not diagnosed and treated appropriately; therefore, accurate diagnosis is critical. A patient's susceptibility to skin disease is influenced by genetic risk factors, environmental exposures, aging, infections, and immune function. The rising global burden of skin diseases has created a need for timely, efficient, and accurate diagnostics.

Recent advances in deep learning have changed the landscape of medical imaging with innovative automated systems for the detection and classification of complex illnesses. The current review will compare AI models for skin disease classification, with an emphasis on Vision Transformers (ViT), a recent architecture that has been successful for image-based diagnostics. The overall aim is to assess the effectiveness of models to improve accuracy and efficiency in dermatological practice.

Historically, the diagnosis of skin diseases has been performed through clinical examination, dermoscopy, and histopathological examination by expert dermatologists. While this remains the gold standard in dermatology, these are time-consuming and somewhat subjective, meaning that there is variability in diagnosis between dermatology practitioners. Furthermore, many rural and underserved areas have diminished access to trained providers of skin disease care. In this situation, AI-capable diagnostic tools - especially those utilizing deep learning and large annotated image datasets - are a strong alternative to expert performance.

This study will explore the performance of current AI architectures, especially ViT with its self-attention mechanism that can enhance performance over CNNs, to explore the AI ability to present expert performance. Utilizing publicly available skin image datasets that ensured balanced disease class representation, the experience will evaluate the AI ability to distinguish skin conditions, including Basal Cell Carcinoma, Benign Keratosis, Eczema, and Melanoma.

The purpose of this investigation is to specifically assess the capabilities of Vision Transformer models in dermatological diagnostic applications. The study showcases the ability of these Vision Transformer models to differentiate between disease diagnosis categories. This highlights the increasing possibilities for the utilization of artificial intelligence (AI) to assist traditional diagnostic methods and improve outcomes for patients in dermatology.

2. Application

This study's overarching implications have great potential for practical use in dermatological diagnostics. By incorporating findings from the assessment of transformer-driven deep learning architectures, particularly Vision Transformers, this review paves the way for the development of sophisticated automated diagnostic systems for deployment in dermatological departments. If implemented, these developments could substantially enhance the speed, reliability, and accuracy of skin disease diagnosis, which would lead to improved patient outcomes and more efficient processes.
The architectural models developed in this study could, in theory, be the basis for the development of integrated diagnostic systems that enable healthcare professionals to improve their diagnostic processes across health disciplines. Such systems would enhance the diagnostic capabilities of dermatologists and seamlessly extend to disciplines related to dermatology, such as oncology, general medicine, and emergency medicine, where skin concerns indicate the status of internal processes.
The incorporation of advanced models within existing medical imaging platforms provides a promising approach toward improving the precision and efficiency of diagnostic processes. The combination of artificial intelligence technologies with established diagnostic workflows can provide better decision-support systems that leverage the interpretation of medical professionals with the computational benefits of machine learning algorithms.
These technological innovations are consistent with the wider movement toward AI-driven solutions in medical diagnostics, in which deep learning tools are playing an increasing role in screening, diagnosing, and monitoring several ailments. For example, the recent advancement of AI systems in dermatology shows that models like ViT’s can achieve diagnostic accuracies comparable to, and sometimes better than, trained providers in detecting various diseases, especially in early-stage disease presentation.
AI’s use in dermatology presents plenty of potential benefits related to AI’s consistency in diagnosing, reduced human error, increased access to care, and ultimately improving health outcomes. However, it can also raise important challenges concerning data privacy, interpretability, bias reduction, and clinical accountability, all of which will need to be resolved for AI to be responsibly and equitably utilized in clinical practice.
Ultimately, the successful integration of AI into dermatological practice and the forefront of clinical medicine will require purposeful, multi-disciplinary collaboration which encompasses technology, ethics, regulation, and clinical education. As AI tools become increasingly sophisticated, lucid designs will be warranted—tools should be designed to supplement, rather than to supplant, clinical judgement and reasoning to enable clinicians to provide higher accuracy and timely patient-centred care promptly.

3. Literature survey

In recent years, there is an increasing interest in the application of artificial intelligence (AI) in dermatology, especially with deep learning models for the classification of skin diseases. Convolutional Neural Networks (CNNs) are often referenced as the longstanding core of medical image analysis, with Vision Transformers (ViTs) emerging as a strong alternative for modeling complex global patterns in images. CNNs are commonly used for dermatological image classification as they are competent in extracting local features and spatial hierarchies. Esteva et al. (2017) demonstrated that CNNs could classify skin lesions with performance comparable to dermatologists. Nevertheless, CNNs perform poorly in their modeling of long-range dependencies in high-resolution dermoscopic images, as they are limited by their inductive biases — locality and translation invariance.To resolve these issues, Dosovitskiy et al. (2020) proposed a Vision Transformer (ViT), which implements self-attention to sequences of image patches, with no inductive bias to enable the model to learn global representations. ViT has been shown to perform competitively, if not better than CNNs in image classification benchmarks, including in medical images. Multiple recent studies have now suggested ViT's ability to better extract long-range dependencies and contextual relationships will improve the accuracy of dermatologist diagnoses.
The goal of this review is to analyze and compare the classification performance of ViT and CNN architectures for the categorization of skin diseases, namely Basal Cell Carcinoma, Benign Keratosis, Eczema, and Melanoma. Throughout the review, we will try to assess the capability of ViT as a trustworthy and effective diagnostic instrument in the emerging field of AI-assisted dermatology.

3.1. Literature survey papers

1) Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review (2018)

Pomponio and colleagues adopted a pre-trained AlexNet model to extract features from 399 images obtained by a standard camera, and then classified the features with a k-nearest-neighbour classifier, obtaining an overall accuracy of 93.64% (sensitivity = 92.1%, specificity = 95.18%). As a downside, this study had a small dataset and required the input of a human to label the lesion.

2) Deep Learning–Based Methods for Automatic Diagnosis of Skin Lesions (2020)

The authors examined the performance of neural network architectures on the ISIC 2019 database, achieving a validation accuracy of 88.33%. The model had a sensitivity of 88.46% and specificity of 88.24%, indicating there is still room to improve the ability to identify different types of skin lesions.

3) Computer-Aided Clinical Skin Disease Diagnosis Using CNN and Object Detection Models (2019)

The authors compiled two datasets (Skin-10 and Skin-100), which had 10 and 100 classes of skin disease, respectively, based on images. The authors compared several state-of-the-art CNN models, resulting in the maximum accuracy of only 79.01% for Skin-10 and 53.54% for Skin-100, thus demonstrating the difficulties in classifying more classes of skin diseases.

4) Skin Diseases Classification Using Deep Learning Methods (2020)

A 4-convolutional neural network (CNN) model was used for the classification of skin diseases while achieving an average accuracy of 93.6%. The model had an accuracy, specificity, and sensitivity of 91%, 98.3%, and 95.9%, respectively. Despite this good result, the study faced limitations due to the imbalanced dataset and the use of non-standardized dermatoscopic images.

5) Multi-Class Skin Cancer Classification Using Vision Transformer Networks and Convolutional Neural Network-Based Pre-Trained Models (2023)

This study applied fine-tuned ResNet models for the classification of skin cancer, achieving a maximum validation accuracy of 77% and a training accuracy of 86%. The latter indicates that distinguishing visually alike skin cancer classes proves to be hard.

6) Deep Learning-Based Transfer Learning for Classification of Skin Cancer (2021)

Various transfer learning models, such as VGG19, InceptionV3, and ResNet50, were evaluated for skin cancer classification. The highest validation accuracy was 89.66%, while examining the Xception model. This indicates the usefulness of transfer learning in this area.

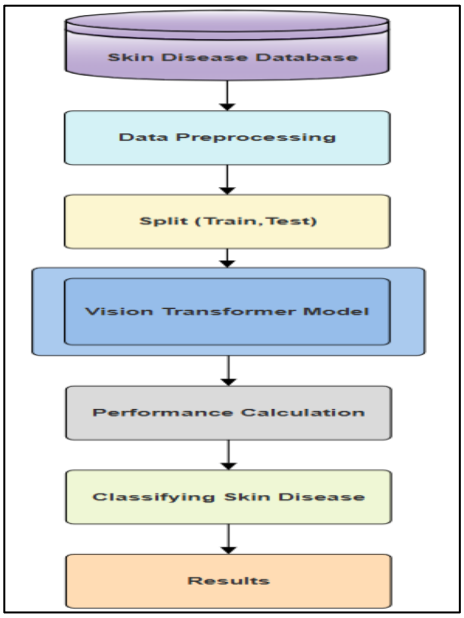
7) Deep Learning Techniques for Accurate Skin Cancer Classification and Mobile Application (2022)
The model utilized in this work was DenseNet169 to classify skin lesions as binary. The results of their work showed an accuracy of 91.10% with 95.67% specificity and 82.49% sensitivity. In addition to the performance evaluation, the feasibility of using the model in the mobile application for clinical utilization was discussed.

8) Skincare Disease Detection of Dermoscopic Images via Deep Learning Methods at High Accuracies (2023)
A basic CNN model with two convolutional blocks was examined in this work, as well as the use of data augmentations to enhance performance. The authors achieved an accuracy value of 86.69% with the use of data augmentations. They highlight, however, that further performance evaluation and additional data augmentations are needed to assess accuracy.

9) Skin cancer detection using deep learning - A review (2023)
This review paper assessed deep learning frameworks and algorithms used for skin cancer classification. They noticed several models did not report accuracies greater than 94%. The authors of this paper stress the need for larger, easily accessible data sets and accurate, robust models to improve diagnostic accuracy.

10) Assist-Dermo: A Lightweight Separable Vision Transformer Model for Multiclass Skin Lesion Classification (2023)
The authors of this study proposed a lightweight vision transformer model for dermoscopic segmentation called Squeeze-Light. Some of their variations for this model achieved higher accuracies depending on the metric. Other models like VGG16 and AlexNet reported accuracy values of 79% and 81.1%, respectively. The authors demonstrated the variability of models in their performance across different architectures.

4. System architecture diagram



5. Problem definition

Skin diseases rank among the most common health conditions globally, affecting millions and placing a large strain on health systems, particularly in dermatology. Maintaining timely and accurate diagnosis is critical for halting the progression of diseases and for providing treatment; however, the traditional approach to diagnosing skin diseases is often limited to trained, subjective dermatologists and is often not available in remote settings or regions with a lack of resources.
The introduction of artificial intelligence (AI) and deep learning methods represents a potential solution to these challenges. AI-based applications, particularly those involving images and analysis, can potentially facilitate greater diagnostic reliability, reduce human error, and increase access to adequate dermatologic care. However, there are key questions that remain for investigation: Which deep learning architectures are the most effective for classifying skin diseases? How do these models need to be validated for successful utilization in diverse clinical settings, and how do these models generalize across diverse skin tones, lesion types, and different image qualities?
To investigate these challenges, recent research examined the appropriateness of state-of-the-art neural network models, specifically Vision Transformers (ViTs), compared to conventional Convolutional Neural Networks (CNNs). The publicly available skin image data set (e.g., HAM10000, ISIC Archive) is a common benchmark. The common classes of skin disease examined from these data sets include Basal Cell Carcinoma, Benign Keratosis, Eczema, and Melanoma, which are some of the most representative and diagnostically complex diseases.
This review will consider the performance of transformer-based models in the context of skin disease classification to highlight the changing field of AI applications for dermatology. By reviewing the current efforts, strengths, weaknesses, and future directions, this review plays a role in realizing skin disease diagnosis that is more accurate, efficient, and broadly accessible to the population.

6. Conclusion

This review examined a new role for more advanced deep learning models in the automatic diagnosis of skin diseases. Skin disease diagnosis benefits from early, accurate diagnosis and safe decision-making. We highlighted the engagement with publicly available datasets, such as HAM10000 and the ISIC Archive, and centred our attention on those 3 common skin diseases (Basal Cell Carcinoma, Benign Keratosis, Eczema, and Melanoma). The findings indicate a shift to a more scalable AI-based diagnosis process.

As the exploratory section demonstrated, Vision Transformer (ViT) architectures are becoming a key component due to their unique approach to image analysis. Unlike the CNN-based approaches that we have examined, ViTs use self-attention to incorporate complex spatial relationships among pixels in the skin lesion images. Given the design and preliminary evidence from related fields, we anticipate that these architectures will result in comparable, and potentially better, classification performance, with more robustness to differing skin tones and disease stages.

ViTs have promise and will need to be examined for related costs, including computational cost and model interpretability, if the goal is to apply ViTs to real-world clinical practice. However, this review suggests that, in our opinion, ViTs dispositions toward generalizability and adaptability across various dermatological datasets demonstrated encouraging results for future research and practice integration.

Looking toward the future, hybrid approaches that effectively combine aspects of CNN and transformer-based architectures may strike a favourable balance between efficiency and diagnostic precision. In addition, developing these systems to operate effectively in resource-limited contexts would enhance inclusivity and applicability in AI-assisted dermatology.

In conclusion, the review highlights the transformative capabilities of ViTs and models like it for the classification of skin diseases/infections. As research and evidence develop, these technologies will likely change standard diagnostic practices toward a more equitable, accessible, and accurate dermatological healthcare approach across the world.

7. Future scope

1) Increasing Dataset Diversity: Future research could investigate larger and more diverse datasets with a broader representation of skin tone, age, and presentation of disease. This would support more generalizable models and ensure equitable diagnoses across the world's diverse populations.
2) Incorporating Multimodal Data: Weekly confounding clinical images with other data, such as patient symptoms, medical history, or histopathology reports, will highly improve diagnostic accuracy. Multimodal fusion represents a growing research area able to generate AI systems that are more holistic and contextually aware.
3) Increasing Model Understanding: Interpability will continue to limit adoption in clinical settings. Future efforts should focus on the development of visualization techniques, such as attention maps or saliency methods, to display all methods of predictions in a digestible format that increases dermatologists' trust and other healthcare providers.
4) Tracking Disease Stages: The development of models designed to analyse sequential images of skin over time will help track disease progression or response to treatment. Longitudinal tracking would create more individualized care and would allow for earlier intervention.
5) Deployment in Low-Resource Areas: ViT-based architectures should continue to be investigated in terms of their optimal performance toward real-time deployment on mobile or low-power devices. In addition to patient care, it would also allow for community-based researchers to perform AI-driven care in places with no previous access to xyz or qualified providers.
6) Integrating CNNs and transformers: There is an opportunity for future research on hybrid models that combine the strong local feature extraction of CNNs and the global attention of transformers to create more accurate and efficient diagnostic tools.

References

[1] Ali, L., Ahmad, M., Paul, A., et al. (2020). An Optimized Skin Cancer Classification Approach using Convolutional Neural Network. Diagnostics, 10(9), 718.

[2] Mahbod, A., Schaefer, G., et al. (2020). Transfer Learning Using a Multi-Scale and Multi-Network Ensemble for Skin Lesion Classification. Computer Methods and Programs in Biomedicine, 193, 105475. https://doi.org/10.1016/j.cmpb.2020.105475.

[3] Brinker, T. J., Hekler, A., Enk, A. H., et al. (2019). Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. European Journal of Cancer, 118, 91–96. https://doi.org/10.1016/j.ejca.2019.06.012.

[4] Talo, M., Baloglu, U. B., Yildirim, O., & Acharya, U. R. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images. Cognitive Systems Research, 54, 176-188. (For model benchmarking context). https://doi.org/10.1016/j.cogsys.2018.12.007.

[5] Salehahmadi, Z., Hajialiasghari, F., & Hajialiasghari, M. (2022). Skin Disease Detection Using Deep Learning: A Review. International Journal of Health Sciences, 6(S2), 3846–3857.

[6] Khan, M. A., Akram, T., Sharif, M., et al. (2021). Integrated deep learning model for skin lesion classification. Computers, Materials & Continua, 66(3), 3303–3319.

[7] Yu, L., Chen, H., Dou, Q., Qin, J., & Heng, P. A. (2017). Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Transactions on Medical Imaging, 36(4), 994–1004. https://doi.org/10.1109/TMI.2016.2642839.

[8] Harangi, B. (2018). Skin lesion classification with ensembles of deep convolutional neural networks. Journal of Biomedical Informatics, 86, 25–32. https://doi.org/10.1016/j.jbi.2018.08.006.

[9] Almaraz-Damian, J. A., Galván-Tejada, C. E., Gamboa-Rosales, H., et al. (2020). Melanoma detection using deep learning techniques on dermatoscopic images. Healthcare, 8(1), 29.

[10] Gessert, N., Nielsen, M., Shaikh, M., Werner, R., & Schlaefer, A. (2020). Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. MethodsX, 7, 100864. https://doi.org/10.1016/j.mex.2020.100864.

[11] Pacheco, A. G., & Krohling, R. A. (2019). The impact of patient clinical information on automated skin cancer detection. Computers in Biology and Medicine, 116, 103545. https://doi.org/10.1016/j.compbiomed.2019.103545.

[12] Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2020). Skin cancer classification using deep learning and transfer learning. Advances in Intelligent Systems and Computing, 1058, 499–509.

[13] Codella, N. C. F., Nguyen, Q. B., Pankanti, S., et al. (2017). Deep learning ensembles for melanoma recognition in dermoscopy images. IBM Journal of Research and Development, 61(4/5), 5:1–5:15. https://doi.org/10.1147/JRD.2017.2708299.

[14] Jinnai, S., Yamazaki, N., Hirano, Y., et al. (2020). Automatic detection of skin cancer using a smartphone: A systematic review. Journal of Dermatology, 47(9), 979–986.

[15] Tschandl, P., Codella, N., Akay, B. N., et al. (2019). Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. The Lancet Oncology, 20(7), 938–947. https://doi.org/10.1016/S1470-2045(19)30333-X.