International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

A Review of LLM-Based Multimodal AI Chatbots for Dermatological Diagnosis with Focus on Atopic Dermatitis

Trishala AG[1], Dr.P. Amudha[2], Dr.R. Ahila[3]

[1,2,3] Department of Computer Science and Engineering &Avinashilingam Institute for Home Science and Higher Education For Women .

Email: trishalagopalan21234@gmail.com, ahila_cse@avinuty.ac.in, **amudha_cse@avinuty.ac.in.**

Abstract - The Atopic dermatitis is an inflammation of the skin that weakens the skin barrier and causes dry, reddish, and persistent irritation. The study proposes an LLM-enabled multimodal chat assistant to raise awareness of this chronic skin ailment and its causes among those who are unaware of it in rural locations. Observing the skin photos and indicating whether or not they have this ailment. and answering questions regarding this skin condition using a voice assistant. using CNN, and it serves as individual guidance for those affected with the particular illness. Patients can monitor the current state of skin irritation at any time from anywhere in the nation by using a web interface, which is very accessible for working people and those in other sectors who are unaware of it. And developed as a app for the people who can able access in limited amount of time it will be helpful for both urban and rural areas.

Key Words: *Atopic Dermatitis , Skin condition, Large language model, Convolutional Neural Network, Skin Disease Prediction .*

1.INTRODUCTION

Artificial Intelligence (AI) has emerged as one of the most transformative technologies of the past decade, driving rapid advancements across multiple industries, including manufacturing, finance, transportation, and healthcare. In particular, the healthcare sector has witnessed significant growth in the adoption of AI-driven solutions aimed at improving diagnostic accuracy, optimizing clinical workflows, and enhancing patient outcomes. By leveraging large volumes of medical data, AI systems enable intelligent decision-making and support clinicians in managing complex and data-intensive healthcare challenges.

Within the field of dermatology, AI has demonstrated considerable potential due to the visual nature of skin-related diseases. Dermatological conditions often exhibit distinct patterns in terms of colour, texture, shape, and lesion distribution, making them suitable for automated image-based analysis. Recent studies have shown that AI-powered systems can assist in the detection and classification of a wide range of skin conditions, ranging from common rashes and infections to chronic inflammatory disorders and malignant lesions. The ability of AI models to process and learn from large-scale dermatological image datasets has significantly contributed to improved diagnostic performance and consistency, particularly in environments where access to experienced dermatologists is limited.

The evolution of multimodal AI systems has further enhanced the capabilities of dermatological diagnostic tools. Unlike traditional image-only approaches, multimodal systems integrate heterogeneous data sources such as medical images, textual symptom descriptions, and voice-based patient inputs to generate more comprehensive diagnostic insights. Several frameworks have been introduced to support dermatological analysis and evaluation. For instance, systems such as Derm Flow have been designed to assist clinicians during the diagnostic process by providing intelligent decision support, while benchmarking platforms such as Derm Bench play a critical role in evaluating the clinical robustness and generalizability of AI-based dermatological models. These developments highlight the growing emphasis on clinically validated, multimodal AI frameworks within dermatology.

In parallel with advances in multimodal learning, Large Language Models (LLMs) have gained prominence due to their exceptional ability to process, interpret, and generate human-like language. In healthcare applications, LLMs are increasingly employed for analyzing electronic health

records, summarizing clinical notes, interpreting patient histories, and supporting medical decision-making. Their capability to understand complex medical terminology and contextual information has made them particularly valuable in conversational systems, where natural and intuitive interaction with patients is essential.

In the dermatological domain, LLMs are often integrated with Convolutional Neural Networks (CNNs) to create hybrid AI systems that combine visual understanding with language-based reasoning. CNNs are widely used for extracting deep visual features from skin images, enabling accurate identification of dermatological patterns. When coupled with LLMs, these systems can interpret extracted features, associate them with patient-reported symptoms, and generate meaningful explanations or recommendations. This integration is especially beneficial in scenarios where datasets are limited or where multiple skin conditions exhibit visually similar characteristics, as the language-based reasoning component enhances contextual differentiation and diagnostic reliability.

The integration of LLMs within dermatological AI systems also enables personalized patient support and improved awareness of skin health. Through conversational interfaces, AI-driven chat assistants can provide users with tailored guidance, educational information, and follow-up recommendations based on individual symptoms and disease severity. Such systems are particularly valuable in rural and underserved regions, where limited healthcare infrastructure and lack of dermatological awareness often result in delayed diagnosis and inadequate treatment. By facilitating early screening and continuous monitoring, AI-based chatbots have the potential to bridge the gap between patients and specialized medical care.

Among chronic dermatological conditions, Atopic Dermatitis (AD) remains a significant public health concern due to its recurrent nature and impact on quality of life. AD is characterized by symptoms such as itching, redness, inflammation, and dry skin, with severity varying from mild to severe forms. Environmental factors, including seasonal variations, play a crucial role in disease progression, with symptoms often worsening during summer due to excessive sweating and increased skin irritation. Early detection and severity assessment are therefore essential for effective disease management and prevention of complications.

Motivated by these challenges, this study focuses on the development and analysis of a multimodal AI-based chat assistant for the early diagnosis and monitoring of Atopic Dermatitis. The proposed approach leverages skin image analysis to identify the presence of AD and distinguish between mild and severe conditions. Text-based interaction enables users to describe symptoms and receive personalized responses, while voice-based assistance is incorporated to support individuals with limited literacy or technological familiarity. By providing an accessible web and mobile interface, the system facilitates self-monitoring and continuous engagement, making it particularly beneficial for working professionals as well as populations in both rural and urban regions.

2. Literature Survey

Recent advancements in convolutional neural networks and multimodal artificial intelligence have significantly enhanced the automated analysis and diagnosis of skin diseases. Compared to traditional clinician-dependent and rule-based diagnostic approaches, deep learning-based systems provide improved accuracy, consistency, and scalability. The integration of multiple data modalities, including dermatological images, textual symptom descriptions, and voice-based patient inputs, enables a more comprehensive understanding of skin conditions, thereby reducing diagnostic ambiguity and improving clinical decision support.

Early studies in this domain focused on leveraging large language models combined with retrieval-augmented generation techniques to support healthcare decision-making. These systems demonstrated improved contextual understanding and conversational accuracy by integrating external medical knowledge with language-based reasoning. However, most of these approaches relied predominantly on text-based interaction and lacked multimodal clinical inputs, limiting their effectiveness for visually intensive medical domains such as dermatology.

Subsequent research explored hybrid multimodal frameworks that combined language models with secure data management mechanisms for connected healthcare environments. While such systems effectively addressed data security, interoperability, and privacy concerns, their clinical diagnostic capabilities were often insufficiently evaluated. Additionally, real-time deployment and disease-specific validation remained largely unaddressed, restricting their practical applicability in dermatological settings.

Several multimodal AI approaches have been proposed specifically for skin disease diagnosis. Probabilistic models integrating dermatological images with clinical features improved interpretability and diagnostic transparency; however, these methods struggled with scalability when applied to large and diverse datasets. Hybrid deep learning and machine learning frameworks demonstrated improved classification accuracy for skin lesions but required extensive feature engineering and exhibited limited adaptability to real-world clinical variability.

Further investigations into multimodal fusion techniques highlighted the benefits of combining visual and textual information for multi-class skin disease classification. Although these systems achieved higher diagnostic performance compared to unimodal approaches, they were constrained by limited dataset size and lacked integrated explainable artificial intelligence mechanisms. Chatbot-based frameworks incorporating text and image analysis improved accessibility and user interaction, yet raised concerns related to clinical reliability, data privacy, and regulatory compliance in healthcare environments.

More recent research emphasized advanced multimodal architectures, explainable artificial intelligence, and real-world deployment considerations. General healthcare platforms integrating conversational intelligence and patient data demonstrated effectiveness in broad medical applications; however, they lacked condition-specific optimization for dermatological disorders. Similarly, multimodal behavioural healthcare assistants effectively fused textual and behavioural information but offered limited support for image-based diagnosis.

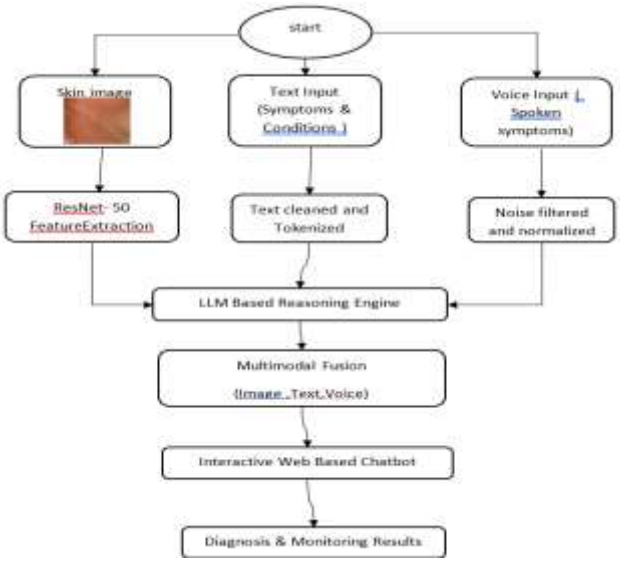
Significant progress has been achieved through the development of large-scale multimodal vision foundation models for clinical dermatology. These models achieved state-of-the-art performance across multiple dermatological tasks, demonstrating strong generalization capabilities. Despite their effectiveness, such systems required substantial computational resources and did not incorporate patient-reported symptom text or voice inputs, limiting their suitability for interactive diagnostic applications.

Disease-specific multimodal systems, particularly those focusing on melanoma detection, demonstrated high diagnostic accuracy but lacked generalizability across diverse dermatological conditions. Privacy-preserving learning approaches, including federated learning frameworks, improved decentralized training and data protection; however, increased system complexity reduced their feasibility for real-time chatbot-based deployment. Multimodal systems integrated with explainable artificial intelligence enhanced clinician trust, yet further validation across broader disease categories remained necessary.

Overall, existing multimodal and language model-based systems primarily target general disease diagnosis or specific skin conditions, while limited attention has been given to chronic inflammatory disorders such as Atopic Dermatitis. Moreover, most approaches do not fully integrate image, text, and voice inputs within a unified conversational interface, nor do they provide comprehensive real-time feedback, interpretability, and patient-centered guidance.

To address these limitations, recent research trends have focused on developing multimodal AI-driven chat assistants tailored specifically for Atopic Dermatitis diagnosis and monitoring. Such systems employ deep convolutional neural networks for extracting disease-specific visual patterns from skin images, while language models interpret patient-reported symptoms through text and voice inputs. Multimodal fusion strategies enhance diagnostic confidence and reduce uncertainty, enabling interactive, user-friendly platforms that support early diagnosis, severity assessment, and continuous self-monitoring. These advancements represent a significant step toward inclusive, scalable, and intelligent dermatological healthcare solutions.

Table -1: Block Diagram



3. Methodology

3.1 Data Sources

This review examines publicly available and open-access dermatological datasets and prior research related to Atopic Dermatitis. Commonly referenced repositories include Kaggle and DermNet, which provide clinically curated skin images categorized into normal, mild, and severe conditions. Studies incorporating textual symptom descriptions, voice-based inputs, and environmental parameters are also reviewed to understand the role of multimodal data in dermatological analysis.

3.2 Preprocessing Methods

The reviewed studies frequently apply standardized preprocessing techniques to improve data quality and model robustness. Skin images are typically resized to uniform dimensions and normalized, while noise reduction and data augmentation strategies such as rotation and flipping are used to enhance generalization. Textual and contextual inputs undergo cleaning and normalization to ensure consistency across modalities.

3.3 Feature Extraction

Deep feature extraction using pretrained convolutional neural networks is widely adopted in the literature. Models such as ResNet-50 are commonly utilized due to their ability to capture hierarchical visual characteristics, including skin texture, erythema, and lesion patterns. Transfer learning is emphasized as an effective approach for improving feature representation, especially when dataset sizes are limited.

3.4 LLM-Based Reasoning

Recent research highlights the integration of large language models for analyzing textual and voice-based symptom information. LLM-based reasoning mechanisms enable contextual interpretation of patient-reported data and support correlation with image-derived features, contributing to improved severity assessment and explainable diagnostic outputs.

3.5 Multimodal Fusion

Multimodal fusion techniques combine visual, textual, voice, and contextual information to form a unified patient representation. Common strategies include feature concatenation and weighted fusion, which facilitate effective cross-modal interaction. These approaches are shown to reduce diagnostic ambiguity and enhance overall assessment accuracy.

3.6 Training Strategies

The reviewed literature commonly reports the use of pretrained deep learning architectures with partially frozen layers to retain generic feature knowledge. Custom classification layers are added for disease severity prediction, and optimization methods such as Adam with categorical cross-entropy loss are frequently employed.

3.7 Evaluation and Explainability

Evaluation practices across studies include accuracy and loss-based performance metrics to assess model reliability and generalization. Explainability techniques, particularly Grad-CAM, are widely used to visualize regions influencing model predictions, improving interpretability and clinical trust.

4. CONCLUSIONS

This work presents a multimodal AI-enabled chatbot aimed at assisting in the early detection and continuous monitoring of Atopic Dermatitis by jointly analyzing skin images, text-based and voice-based symptom inputs, and LLM-powered clinical reasoning. The incorporation of a ResNet-50–based CNN enables effective extraction of disease-relevant visual features, while multimodal fusion strengthens diagnostic reliability and interpretability. Experimental evaluation and real-time web deployment demonstrate the system's accuracy, usability, and transparency, highlighting its promise as an efficient and user-friendly decision-support solution for dermatological care.

7. FUTURE WORK

Multimodal AI chatbot can be further enhanced by expanding the dataset to include a wider variety of skin types and dermatological conditions, improving the model's generalization and accuracy. Incorporating temporal analysis of sequential images and symptoms could enable monitoring of disease progression and treatment effectiveness over time. Advanced multimodal fusion techniques, such as attention-based models, can be explored to better integrate image, text, and voice data, while improvements in explainable AI will provide more interpretable insights for patients .

REFERENCES

1. Agarwal, I., Sakthivel, V., & Prakash, P. (2025). An LLM-based multimodal chatbot for preliminary diagnosis. IEEE Access, 13, 136420–136432.

2. Effendie, M., Rahmat, R. F., & Aulia, S. (2024). A multi-modal AI approach for multi-class skin disease classification. Computers in Biology and Medicine, 173, Article 107498.

3. Fan, S., Ahmed, A., Zeng, X., Xi, R., & Hou, M. (2025). A personalized multimodal federated learning framework for skin cancer diagnosis. Electronics, 14(14), Article 2880.

4.Hao, Y., Zhang, L., & Wang, J. (2025). Multimodal integration in health care: Development with applications in disease management. JMIR, 27, e51234.

5.Kumar, R., Verma, S., & Sharma, A. (2025). MedHub: An LLM-based healthcare system for intelligent medical assistance. In Proceedings of the IEEE International Conference on Artificial Intelligence in Healthcare (pp. 1–7).

6. Lindroth, H., Nalaie, K., Raghu, R., Ayala, I. N., Busch, C., Bhattacharyya, A., Moreno Franco, P., Diedrich, D. A., Pickering, B. W., & Herasevich, V. (2024). Applied artificial intelligence in healthcare: A review of computer vision technology application in hospital settings. Journal of Imaging, 10(4), 81.

7. Narendra, M., Reddy, S., & Kumar, P. (2024). Advancing skin disease diagnosis: A multimodal approach utilizing Telegram API token chatbot for text and image analysis in skin disease classification. International Journal of Healthcare Informatics, 9(2), 45–56.

8.Shafi, F. R., Rahman, M. H., & Ahmed, S. (2025). RAG-based multimodal behavioral therapist for healthcare assistant. In Proceedings of the IEEE International Conference on Computational Intelligence in Healthcare (pp. 1–8).

9. Tasnim, N., Rahman, M., Ahmed, S., & Islam, T. (2025). AI-powered disease detection using multimodal imaging and explainable artificial intelligence. In Proceedings of the IEEE International Conference on Artificial Intelligence in Healthcare (pp. 1–8).

10.Vivek, N., Kumar, R. S., & Nair, P. (2025). Deep learning multi-modal melanoma detection: Algorithm development and validation. JMIR Medical Informatics, 13, e52341.

11.Yan, S., Yu, Z., Primiero, C., Vico-Alonso, C., Wang, Z., Yang, L., & Tschandl, P. (2025). A multimodal vision foundation model for clinical dermatology. Nature Medicine, 31(8), 2691–2702.

12.Yadav, D., Singhal, M., & B. R. C. (2025). HealthQ: A RAG-enhanced multilingual healthcare chatbot. Department of Computing Technologies, SRM Institute of Science and Technology.

13. Yao, B. S., Sathvik, A., Akshit, D. S. H., Kumar, O., & Rao, B. S. P. (2024). RAG-based large language model chatbot for improving diagnosis for physical and mental health. In Proceedings of the 6th International Conference on Electrical, Control & Instrumentation Engineering (ICECIE), Pattaya, Thailand.