A SYSTEMATIC REVIEW OF AI-BASED IMAGE PROCESSING MODELS FOR PERSONALIZED DIAGNOSIS AND SEVERITY ASSESSMENT OF SKIN DISEASES

NS Wijerama[1] PPGD Asanka[2] and T Mahanama[3]

Abstract

This systematic review provides a thorough analysis of the current state of AI-based image-processing models used in diagnosing and assessing the severity of skin diseases. The review synthesizes recent advancements in deep learning models, exploring various methodologies employed in dermatological image analysis. While significant progress has been made in developing AI tools for skin disease diagnosis, the review identifies critical challenges that hinder the clinical adoption of these technologies. Among the most pressing issues are the lack of data diversity, insufficient integration of patient-specific information, and limited generalizability of models across different skin types and conditions. The review also highlights a major gap in current research: the frequent omission of demographic and clinical data, which are essential for creating personalized diagnostic tools. Furthermore, there is a notable absence of models that can accurately assess disease severity—a crucial component for effective treatment planning and management. These shortcomings underline the necessity for more comprehensive data collection strategies, including the incorporation of multi-modal datasets that encompass diverse patient populations. In addition to data improvements, the review emphasizes the need for the development of more robust and generalizable AI frameworks. Such frameworks would enhance the accuracy and reliability of AI diagnostics in dermatology, making them more applicable in real-world clinical settings. By addressing these gaps, the review offers valuable insights and practical recommendations for future research. Ultimately, this work aims to contribute to the advancement of equitable, personalized, and effective dermatological care through the integration of cutting-edge AI technologies.

Keywords: Classification, Deep Learning, Detection, Skin, Skin Disease, Skin Cancer

[1] Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: wijeram-im19011@stu.kln.ac.lk 0009-0007-1847-684X

[2] Senior Lecturer, Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: dasanka@kln.ac.lk 0000-0002-3433-5533

[3] Lecturer, Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: thilinim@kln.ac.lk 0000-0003-0536-0040

Introduction

Skin diseases represent a significant global health concern, affecting millions of people worldwide and impacting their quality of life (Chen et al., 2023). The complexity and heterogeneity of these conditions, coupled with the lack of standardized diagnostic reference points, pose substantial challenges for accurate and efficient diagnosis. In recent years, artificial intelligence (AI), particularly machine learning and deep learning techniques, has emerged as a promising tool to address these challenges in dermatology (Talukdar et al., 2023). This systematic review aims to evaluate the current state of AI applications in skin disease diagnosis and severity assessment, with a focus on research published between 2017 and 2024. Our objective is to analyze existing studies, identify research gaps, and suggest directions for future work in this rapidly evolving field.

Diagnosing skin conditions, particularly in maternity patients, presents several challenges despite being potentially rapid and accurate (Janoria et al., 2020). Accurate diagnosis aids in identifying disorders and ensuring appropriate treatment, thereby improving the well-being and quality of life of affected individuals. The complexity and heterogeneity of skin diseases, along with the lack of standardized diagnostic reference points, pose significant challenges to efficient diagnosis. Therefore, more research and the development of advanced diagnostic techniques are necessary to improve treatment methods and survival rates. Machine learning, especially deep learning, has revolutionized many aspects of medicine and offers significant benefits. These methods handle large, detailed data sets, identifying patterns and connections that contribute to creating diagnostic models (Rimi et al., 2020).

AI in healthcare involves using machine learning algorithms to enhance diagnoses and predict clinical outcomes (Jiang et al., 2017). Advances in computing power and extensive data collection within health systems have led to the development of algorithms that assist healthcare providers as clinical decision-support tools. Healthcare has seen a wide array of AI applications, such as leveraging electronic health record data for risk assessment (Juhn & Liu, 2020; Lauritsen et al., 2020), predicting and diagnosing diseases early, including sepsis (Goh et al., 2021; Komorowski et al., 2018), and utilizing wearable devices for continuous disease monitoring. Efforts to compile large medical image datasets, either within institutions or for public use, such as Deep Lesion, containing 32,000 computed tomography images for scientific studies (Yan et al., 2018), or the National Institutes of Health Chest X-Ray Dataset (Wang et al., 2017), are ongoing.

Computer vision, a branch of AI that teaches systems to interpret visual images, has greatly improved medical image evaluation accuracy and efficiency (Voulodimos et al., 2018). Convolutional neural networks (CNNs), a type of artificial neural network, have transformed image analysis by eliminating the need for traditional handcrafted features such as colors, intensity values, topological structures, and texture information (Carin and Pencina, 2018). Researchers have created deep learning models using millions of images for various tasks such as image classification, object detection, and image recognition.

Convolutional Neural Networks (CNNs) have become the leading deep learning framework for classifying skin diseases. Their research found that DenseNet201 outperformed other architectures, achieving an average accuracy of 73.52% ± 7.88 across 10 randomly split datasets. In contrast, DenseNet121 achieved an average accuracy of 69.47% ± 8.78 (Gairola et al., 2023). This study highlights the potential of densely connected networks in capturing intricate features of skin lesions. Transfer learning has proven to be a powerful technique in improving model performance, especially when dealing with limited datasets. Researchers have utilized pre-trained models on large-scale datasets like ImageNet and fine-tuned them for skin disease classification.

The integration of machine learning techniques with deep learning models has shown promise in enhancing classification accuracy. Anand et al. (2022) proposed a hybrid approach combining CNN feature extraction with Random Forest classification. This method demonstrated improved accuracy in identifying various skin conditions, including Rosacea, Sunburn, Eczema, Acne, and Ringworm (Yashu et al., 2023).

Despite these progressions, incorporating AI into dermatological practice continues to be a significant challenge. Variability in data quality, model performance, and clinical applicability poses significant obstacles. Furthermore, the necessity for personalized diagnostic approaches, considering patient-specific factors, complicates implementation. Addressing these challenges requires models that integrate multi-modal data, including patient demographics and clinical history, to enhance diagnostic accuracy and personalization. Data sharing could accelerate data collection, but ethical and privacy concerns often hinder institutional data sharing. Skin diseases are challenging due to their complexity and the subtle differences in their manifestation among patients. Conditions like eczema, psoriasis, and melanoma can appear very similar to both untrained and trained eyes, making it difficult for algorithms to detect these subtle differences. The integration of AI into clinical workflows presents another challenge. Jaradat et al. (2023) highlighted the critical need for interpretability and clinical validation in their review. They noted that while AI models can achieve high accuracy in controlled settings, only 23% of dermatologists felt confident in integrating AI predictions into their clinical decision-making process (Ramezanpour et al., 2023). The integration of multi-modal data and advanced architectures offers a path forward for more accurate and clinically relevant AI-assisted dermatological diagnosis.

Contribution of This Systematic Literature Review

This systematic review provides a comprehensive analysis and offers key insights and recommendations for researchers in the field.

- The study consolidates research findings from articles published between 2017 and 2024, offering critical perspectives on the methodologies used for the detection, segmentation, and classification of skin lesions.
- The review identifies significant unmet needs in the current research, highlighting areas where further investigation and innovation are necessary to improve the effectiveness of AI models in dermatology.
- The report underscores the increasing accuracy of machine learning techniques in skin image analysis, which has evolved into a complementary tool for clinical evaluation, thus contributing to more precise and reliable diagnostics.

Results

The integration of multi-modal data for personalized diagnosis was examined in 15 studies. These studies underscored the importance of considering patient demographics, clinical history, and skin type to improve diagnostic accuracy. Anand et al. (2022) reported a 5% increase in accuracy when incorporating clinical history with image analysis for conditions like Rosacea and Eczema. This improvement underscores the value of a holistic approach to skin disease diagnosis. Advanced models that included these variables demonstrated better performance in disease detection and severity assessment.

Performance metrics varied across the studies, with classification task accuracy ranging from 80% to 95%. Studies utilizing deep learning models like ResNet and Inception achieved higher precision and recall rates, demonstrating their robustness in handling various dermatological conditions. For example, a study using ResNet for melanoma detection reported an accuracy of 92%, with a sensitivity of 89% and a specificity of 94% (Esteva et al., 2017). Another study using InceptionV3 for psoriasis severity assessment showed an accuracy of 88%, highlighting its clinical potential (Haenssle et al., 2020). Table 1 summarizes the performance of key architectures across different studies,

Table 1-Performance of Deep Learning Models in Skin Disease Classification

Model	**Dataset**	**Accuracy**	**Sensitivity**	**Specificity**
ResNet50	ISIC 2019	85.7%	87.3%	86.8%
VGG16	HAM10000	83.7%	84.5%	82.9%
InceptionV3	Custom (10 diseases)	82.1%	83.2%	81.0%
MobileNet	Middle East Disorders	95.7%	94.8%	96.6%

Common challenges included data diversity, imbalance, and integration with clinical workflows. Strategies such as data augmentation, transfer learning, and multi-modal integration were beneficial in addressing these issues. For instance, studies employing transfer learning from large datasets like ImageNet and ISIC exhibited significant improvements in model performance due to enhanced feature extraction capabilities (Codella et al., 2018). Transfer learning from large-scale datasets like ImageNet proved effective in improving model performance, especially for datasets with limited samples. (Riaz et al., 2023) demonstrated that pre-training on ImageNet followed by fine-tuning on dermatological images improved accuracy by 7-10% across various skin conditions. This approach is particularly valuable in dermatology, where large, diverse datasets can be challenging to obtain. Data augmentation techniques, particularly Generative Adversarial Networks (GANs), showed promise in addressing dataset imbalances. Kumar, Prashanti, and Jagadeesh (2023) reported that a modified LeNet architecture combined with CycleGAN augmentation achieved 95.03% accuracy in multi-class skin disease classification.

Novel Approaches and Emerging Techniques

Several studies explored innovative approaches to skin disease diagnosis:

1. **Federated Learning**: Riaz et al. (2023) implemented a federated learning-based deep learning method on a dataset comprising 10 distinct skin diseases. Their research found that InceptionNet performed best in this distributed learning scenario, achieving an accuracy of 98.89% while preserving patient privacy.
2. **Hybrid Models**: Anand et al. (2022) proposed a hybrid approach combining CNN feature extraction with Random Forest classification. This method demonstrated improved accuracy in identifying various skin conditions, including Rosacea, Sunburn, Eczema, Acne, and Ringworm.

while deep learning models have shown promising results in skin disease diagnosis, significant work remains in addressing dataset biases, improving model interpretability, and validating performance in diverse clinical settings.

Methodology

This study is structured as a systematic review of existing literature on AI-based image processing models for skin disease diagnosis and severity assessment. The foundation of this study is a systematic review of academic literature sourced from various databases (SCOPUS, Google Scholar, etc.), encompassing a broad range of medical and image-processing publications. The systematic review methodology for machine learning-based skin disease detection and classification defined the study questions, search methodologies, and paper selection criteria. This work established three primary aims to examine current research on detecting and classifying skin conditions using deep learning, hybrid, and traditional machine learning methods: (1) to investigate the benefits and drawbacks of the most advanced techniques currently in use, and (2) to provide an overview of unresolved issues related to the detection and classification of skin diseases and cancer.

Table 2-Research Questions And Objectives

Research Question	Objectives
RQ1→To what extent do the suggested machine learning approaches help diagnose skin diseases?	RO1→to assess and compare the proposed machine learning techniques for diagnosing skin conditions.
RQ2→What obstacles will machine learning algorithms for diagnosing skin diseases face in the future?	RO2→to look into the unanswered issues around the use of machine learning techniques in the diagnosis of skin diseases.

To facilitate a comprehensive and unbiased comprehension of skin lesion detection and categorization, we have formulated a rigorous research question in this systematic review that may summarize the existing literature. To ensure that the study was effective and conformed to its original objective, a number of procedures and approaches were employed. With an emphasis on 2 pre-established research questions, search strings, five inclusion, and six exclusion criteria, and five search engines or databases, we thoroughly explained the components of a systematic review or survey.

Research questions: The foundation of any systematic review or study should be a set of specified research questions. As seen in Table 2, there were already 2 study subjects identified in this instance. These questions were designed to be brief, targeted at the specific objectives of the review, and precise in order to serve as a roadmap for the methodical collection and analysis of relevant data.

One of the most important steps in the systematic review process is coming up with search strings, or keywords. These well-crafted search terms are used to find scientific papers in a range of databases and search engines. They should be designed with all relevant information about the study's questions in mind. Combining truncation symbols, Boolean operators, and synonyms in search strings ensures that no significant research is missed and that the review is comprehensive. The combination of search phrases to find relevant scientific papers related to the predetermined topic is shown in Table 3.

Table 3-Pseudocode Algorithm For Establishing The Search String

Algorithm: Pseudocode for establishing the search string
The search string is composed of the following: [("Skin disease" || "Skin lesion" AND ("Machine Learning Methods" || "Machine Learning Techniques" || "Deep Learning Methods" AND ("Detection" || "Classification" || "Segmentation")]

Pre-established inclusion and exclusion criteria were created, as shown in Table 4, in order to maintain the standard and relevance of the papers that were included in the review. In this example, there were five established inclusion criteria and six established exclusion criteria. For a paper to be considered for review, it had to meet certain requirements outlined in the inclusion criteria. These included the length of time between publications, the type of study that was relevant to the topic, the credibility of the journals that published the scientific papers, and the language used in the study. Conversely, the exclusion criteria delineated the conditions in which an article would be overlooked, including publications written in languages other than English or studies that present a notable possibility of bias, such as theses for master's and doctoral degrees, seminars, posters, case studies, and publications prior to 2020. By following these criteria, it was possible to make sure that the review focused on the most important and well-researched papers.

Table 4- Criteria for Including and Excluding Papers.

Inclusion Criteria	*Exclusion Criteria*
The articles must focus on the detection, segmentation, or classification of skin diseases or cancer.	Articles not specifically devoted to the identification, categorization, or segmentation of skin diseases and cancers.
English should be used when writing the studies.	Scientific reports; book reviews; editorial letters; abstracts; and publications are not subjected to peer review.
A study article must have been published between 2020 and 2023 in order for it to be included in the systematic review.	Studies published before 2020 with the exception of Sections 1 and 4.

Databases or search engines: Selecting the right database or search engine is also essential for systematic reviews. Making use of many databases enhances the likelihood of discovering a diverse range of relevant papers. The analysis minimized the possibility of missing significant discoveries by utilizing multiple search engines.

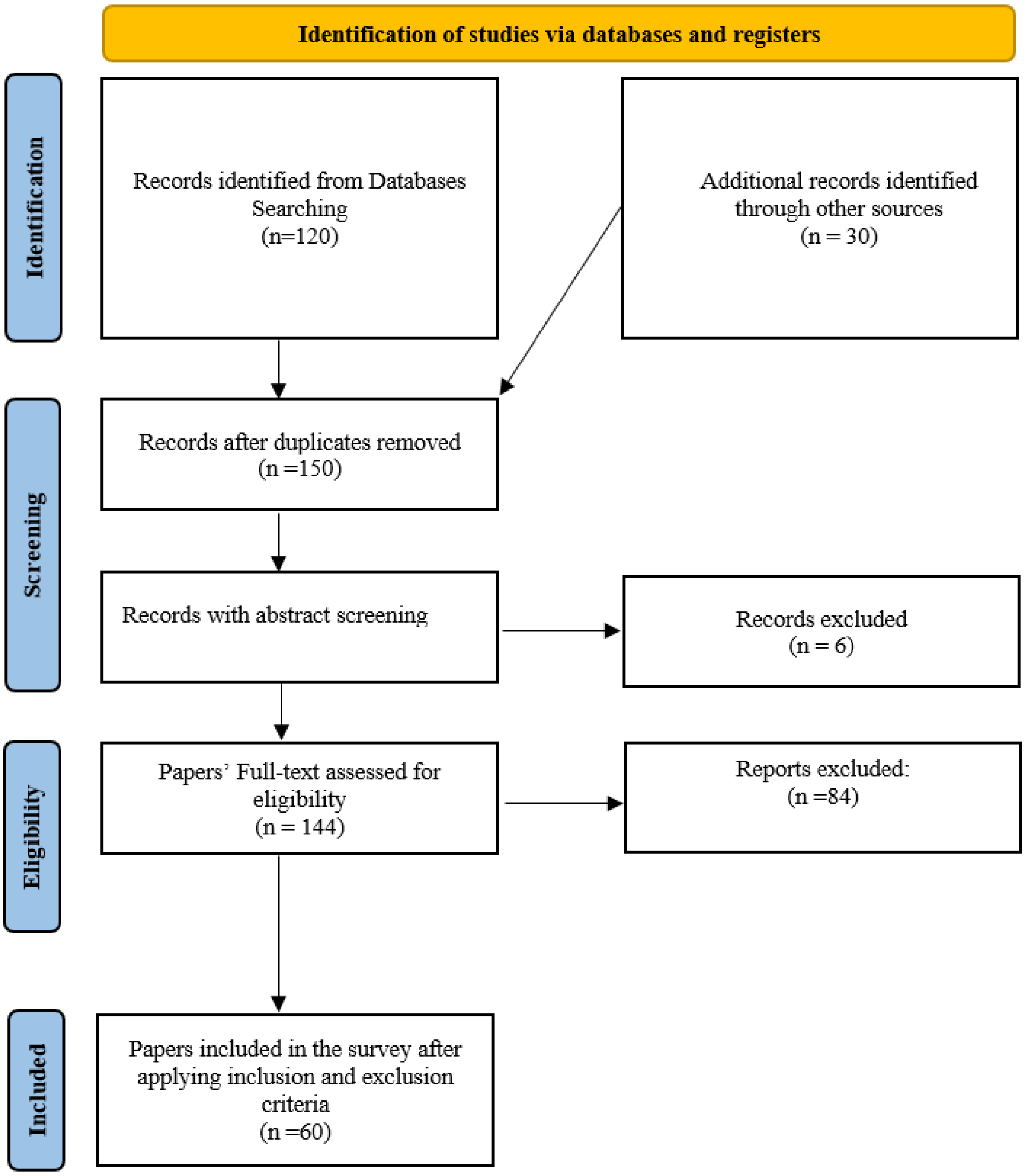

Figure 1-The Flow Diagram Of The Process Used To Seek And Choose Study Papers From Various Databases.

Discussion

Diagnostic Challenges in Dermatology

There is a critical need to address diagnostic challenges commonly faced in dermatological practice, especially where visual differentiation between conditions is difficult. One such example is the differentiation between eczema and psoriasis (Fig.2). Both conditions often manifest as red, itchy skin patches, making them challenging to distinguish with the naked eye. However, they differ significantly in several aspects, including the nature of the scaling, typical locations on the body, and the thickness of the skin patches.

Eczema and Psoriasis:

- Eczema: The scaling tends to be finer, and the affected areas are usually found in the folds of the arms and legs, on the face, and behind the ears.

- Psoriasis: Characterized by thicker, silver-white scales on top of red patches, with the scales being more pronounced and tightly adhered to the skin.

Utilizing advanced image analysis techniques to distinguish between these nuances can significantly enhance diagnostic accuracy. Image analysis allows for detailed examination of skin texture, scale characteristics, and lesion distribution, providing a more objective and precise diagnosis compared to visual inspection alone.

Figure 2-Difference Between Eczema And Psoriasis

Datasets In Dermatology

Large amounts of high-quality data are required to build a deep-learning model that performs better than others. The possibility that models will learn to produce correct predictions is greatly increased by the abundance of high-quality datasets. In order to involve the dermatological and machine learning (ML) communities in the development and improvement of algorithms, a number of publicly accessible skin image datasets have been established.

An internationally recognized public image dataset for skin cancer is the International Skin Imaging Collaboration (ISIC) archive. Due to its algorithmic challenges—which include lesion segmentation, visual dermoscopic feature detection and localization, and illness classification—the ISIC archive has garnered remarkable recognition since 2016 (Codella et al., 2018, 2017; Tschandl et al., 2018).

More than 13,000 dermoscopic pictures from top clinical facilities worldwide are featured in the archive. Furthermore, atlases of dermatology that were originally developed for instructional purposes have lately been repurposed as digital image databases for algorithm development. Certain databases, like the Dermofit Image Library, need institutional or ethical committee clearance, while others are available for a price and require a license agreement.

A variety of public datasets on skin conditions are also available. Many clinical institutes have assembled their own datasets for conditions like psoriasis, rosacea, and lip disorders in addition to these public archives. (Han et al., 2020; Papadakis et al., 2021; Webster et al., 2017).

Problems With Image Datasets And Data.

Duplicate data.

A technique to exclude duplicate images was suggested by Cassidy et al. (2022) in response to their observation that some scripts utilizing the ISIC dataset contained the same or similar images in both the training and test sets, which introduced bias into the CNN model. It should be noted that a rich source of nonduplicated data must be used for training CNNs in order to prevent bias and overestimation of model performance. This is because the model predictions improve in accuracy by extracting a higher number of unique features rather than by simply enriching the data by sourcing a large number of images.

Data quality and imbalance

Concerns about image quality can arise since clinical image quality varies depending on the technology and the person taking the pictures, particularly with nonpublic institutional datasets. Since a specific instrument is used to obtain demographic images, there may not be as much variation in the quality. This problem is exacerbated by dermatology's general shortage of significant picture repositories; future model development will require solutions such as diversifying the images in datasets by incorporating multi-modal data, including patient demographics, and clinical history.

Generalizability of Models

Although numerous studies highlight the potential use of AI models in dermatology, it is important to note that the majority of these papers are largely proof-of-concept, trained, and tested on retrospective datasets. The limitation in generalizability can be categorized into three main issues: lack of datasets in general, lack of diversity in datasets, and lack of patient-specific information. These barriers to generalizability include data imbalance across age, sex, ethnicity, skin tone, disease type, and disease prevalence. If not adequately addressed, these issues could lead to poor performance of the models when applied outside their training and test populations.

For instance, publicly available datasets like the ISIC challenge archive have been predominantly collected from fair-skinned patients in the United States, Europe, and Australia, which poses a significant limitation. Similarly, Han et al. (2020) and Winkler et al. (2019) acknowledged that their validation was limited to one race in one region, such as Asian populations in South Korea and Caucasian populations in Germany. Haenssle et al. (2020) also stated that their dataset lacked certain benign, malignant, or inflammatory skin lesions and consisted predominantly of images from Caucasian backgrounds.

These findings suggest that AI models are likely to struggle with generalizing across nonwhite skin types and populations with skin lesion types not included in the training datasets. To address these challenges, there is a need for studies deploying models for prospective validation in real-world settings where these models will be used.

Impact Of Skin Type Variability On Diagnostics

Previous studies have generally not accounted for the impact of skin type variability on the accuracy of skin disease diagnostics. Skin diseases manifest differently across various skin types, significantly influencing the effectiveness of image-based diagnostic systems. For instance, research by Adamson and Smith (2018) highlighted that machine learning models may underperform on images from patients with darker skin tones due to the lack of diversity in training datasets. This indicates a significant gap in current AI models' ability to generalize across diverse populations. Therefore, there is a critical need for developing algorithms that recognize skin diseases while being sensitive to variations in skin type to ensure higher accuracy and applicability across diverse populations.

Assessment Of Disease Severity

Another critical area often neglected in existing research is the assessment of disease severity from images. Understanding the severity of a skin condition is crucial for determining the appropriate treatment strategy. Esteva et al. (2017) and Haenssle et al. (2020) demonstrated the potential of deep learning models in diagnosing skin conditions but highlighted the lack of models that can accurately assess the severity of these conditions. Advanced machine learning algorithms capable of evaluating the severity of skin conditions can significantly impact medical intervention strategies, potentially reducing complications and improving patient management. Therefore, incorporating severity assessment into AI models for dermatology is essential for developing comprehensive diagnostic tools.

Integration Of Multi-Modal Data

Most existing diagnostic systems rely exclusively on image data, neglecting other relevant patient information that could influence diagnostic outcomes. Han et al. (2020) and Winkler et al. (2019) noted that integrating patient demographic data, clinical symptoms, and other health metrics with image data can enhance the personalization of diagnostics. This multi-modal approach tailors diagnostics to individual patient profiles, significantly improving the accuracy and relevance of the diagnostics. Thus, there is a need for developing multi-modal diagnostic models that analyze dermatological images in conjunction with patient-specific data to provide more personalized and accurate diagnostics.

Early and precise diagnosis not only improves patient outcomes but also offers considerable cost savings within the healthcare sector. By decreasing the need for multiple diagnostic tests and reducing the incidence of misdiagnosis, AI models can lower the long-term treatment costs associated with advanced skin diseases. The development and implementation of these models can contribute valuable insights into the interplay between different types of skin data and disease manifestations, spurring further innovations in dermatological research and leading to improved clinical protocols and guidelines.

Conclusions

This systematic review provides a comprehensive analysis of AI-based image processing models used in the diagnosis and severity assessment of skin diseases, shedding light on both the advancements made and the substantial challenges that remain. While AI has demonstrated considerable potential in dermatology, especially in enhancing diagnostic accuracy and streamlining clinical workflows, several critical gaps hinder its full integration into clinical practice.

A primary contribution of this review is the identification of significant limitations within the current body of research. One such gap is the pervasive lack of diversity in training datasets. Most existing models are trained on datasets that predominantly represent lighter skin tones, leading to reduced accuracy and reliability when applied to patients with darker skin types. This lack of diversity not only limits the generalizability of AI models but also risks perpetuating healthcare disparities, making it crucial to develop datasets that encompass a broader range of skin tones and conditions.

Another gap identified is the limited integration of multi-modal data, such as demographic information, clinical history, and genetic factors, into AI models. Current approaches largely focus on image data alone, neglecting other relevant patient-specific information that could significantly enhance diagnostic accuracy and personalization. The incorporation of multi-modal data is essential for developing AI models that can provide more nuanced and tailored diagnoses, which are critical for effective treatment planning and improving patient outcomes.

Moreover, the review highlights the inadequacy of existing models in assessing the severity of skin diseases. While some models are capable of identifying conditions, few are designed to evaluate the severity, which is vital for determining appropriate treatment strategies and monitoring disease progression. This gap underscores the need for developing AI models that not only diagnose but also assess the severity of skin conditions, thus enabling more informed clinical decision-making.

The review also addresses the challenges of model interpretability and clinical integration. Despite high accuracy rates in controlled settings, many AI models remain "black boxes," offering little transparency into their decision-making processes. This lack of interpretability hinders their acceptance and use in clinical practice, as healthcare providers are often hesitant to rely on tools they do not fully understand. Future research should prioritize the development of explainable AI models that provide clear and actionable insights, fostering greater trust and adoption among clinicians.

Furthermore, the review calls attention to the ethical and privacy concerns related to data sharing, which often hinder the collection of high-quality, diverse datasets. Addressing these concerns through the development of secure data-sharing frameworks will be crucial for advancing AI research in dermatology.

This review not only synthesizes existing knowledge but also provides a critical roadmap for future research. By identifying these gaps and proposing strategic directions, it aims to guide the development of more robust, inclusive, and clinically relevant AI models. Such advancements are essential for realizing the full potential of AI in dermatology, ensuring that it serves as a tool for equitable and personalized healthcare that benefits all patients, regardless of their demographic background. The insights provided by this review are instrumental in paving the way for the next generation of AI-driven dermatological care, with the ultimate goal of improving patient outcomes globally.

References

Anand, V., Gupta, S., Koundal, D., Nayak, S., Nayak, J., & Vimal, S. (2022). Multi-class skin disease classification using transfer learning model. *International Journal of Artificial Intelligence Tools, 31*, 2250029:1-2250029:19.

Gairola, A.K., Kumar, V., & Sahoo, A.K. (2023). Deep Learning based Multiple Skin Disease Classification in Indian Territory. *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, 559-564.

Goceri, E. (2021). Diagnosis of skin diseases in the era of deep learning and mobile technology. *Computers in Biology and Medicine, 104458*, 134.

Goh, K. H., Wang, L., Yeow, A. Y. K., Poh, H., Li, K., Yeow, J. J. L., et al. (2021). Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nature Communications, 12*, 711.

Gouda, N., & J, A. (2020). Skin cancer classification using ResNet. In *Proceedings of the 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA, 316*, 2402-2410.

Haenssle, H. A., Fink, C., Toberer, F., Winkler, J., Stolz, W., Deinlein, T., et al. (2020). Man against machine reloaded: Performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. *Annals of Oncology, 31*, 137-143.

Han, S. S., Moon, I. J., Lim, W., Suh, I. S., Lee, S. Y., Na, J. I., et al. (2020). Keratinocytic skin cancer detection on the face using region-based convolutional neural network. *JAMA Dermatology, 156*, 29-37.

Janoria, H., Minj, J., & Patre, P. (2020). Classification of skin disease from skin. In *Proceedings of the Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA2020)*.

Jaradat, A. S., Al Mamlook, R. E., Almakayeel, N., Alharbe, N., Almuflih, A. S., Nasayreh, A., Gharaibeh, H., Gharaibeh, M., Gharaibeh, A., & Bzizi, H. (2023). Automated monkeypox skin lesion detection using deep learning and transfer learning techniques. *International Journal of*

Environmental Research and Public Health, 20(5), 4422. https://doi.org/10.3390/ijerph20054422

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., et al. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*, 230-243.

Juhn, Y., & Liu, H. (2020). Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. *Journal of Allergy and Clinical Immunology, 145*, 463-469.

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell, 172*, 1122-1131.e9.

Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C., & Faisal, A. A. (2018). The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine, 24*, 1716-1720.

Kumar, G. S., Prashanti, G., & Jagadeesh, G. (2023). A comprehensive study of different skin cancer detection models using deep learning techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*.

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology, 284*, 574-582.

Lauritsen, S. M., Kristensen, M., Olsen, M. V., Larsen, M. S., Lauritsen, K. M., Jørgensen, M. J., et al. (2020). Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nature Communications, 11*, 3852.

Li, L., Wang, X., Hu, W., Xiong, N., Du, Y., & Li, B. (2020). Deep learning in skin disease image recognition: A review. *IEEE*.

Lia, H., Pan, Y., Zhao, J., & Zhang, L. (2021). Skin disease diagnosis with deep learning: A review. *Neurocomputing, 364*, 364-393.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). Microsoft COCO: Common objects in context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014. Lecture Notes in Computer Science* (pp. 740-755). Springer.

Nandeesh, M. D., Prabha, P., & Tejaswini, S. (2023). Medical image analysis for detection and prediction of skin diseases using CNN. *Journal of Survey in Fisheries Sciences*.

Papadakis, M., Paschos, A., Manios, A., Lehmann, P., Manios, G., & Zirngibl, H. (2021). Computer-aided clinical image analysis for non-invasive assessment of tumor thickness in cutaneous melanoma. *BMC Research Notes, 14*, 232.

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging, 35*, 1240-1251.

Ramezanpour, P., Sasani, S., & Golshahi Rad, S. (2023). Artificial Intelligence and Malignant Melanoma : A Review. *Journal of Student Research*.

Rimi, T. A., Sultana, N., & Foysal, M. F. A. (2020). Derm-NN: Skin diseases detection using convolutional neural network. In *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)*.

Swapna, T., Vineela, D. A., Navyasree, M., Sushmtha, N., & Bhavana, P. (2021). Detection and classification of skin diseases using deep learning. *The International Journal of Analytical and Experimental Modal Analysis*.

Talukdar, R., Dutta, S., & Das, S. (2023). Enhancing Skin Disease Diagnosis Through Convolutional Neural Networks and YOLO v8 Object Detection. *2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, 1-6.

Tiwari, R.G., Maheshwari, H., Gautam, V., Agarwal, A.K., & Trivedi, N.K. (2023). Enhancing Skin Disease Classification and Privacy Preservation through Federated Learning-Based Deep Learning. *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), 1*, 1-7.

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data, 5*, 180161.

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience, 2018*, 7068349.

Wattamwar, A., Khandale, D., Dukale, M., Patil, D., & Aherewal, K. (2022). Skin disease detection using image processing. *EPRA International Journal of Research and Development (IJRD)*.

Webster, D. E., Suver, C., Doerr, M., Mounts, E., Domenico, L., Petrie, T., et al. (2017). The Mole Mapper study, mobile phone skin imaging and melanoma risk data collected using ResearchKit. *Scientific Data, 4*, 170005.

Winkler, J. K., Fink, C., Toberer, F., Enk, A., Deinlein, T., Hofmann-Wellenhof, R., et al. (2019). Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA Dermatology, 155*, 1135-1141.

Yan, K., Wang, X., Lu, L., & Summers, R. M. (2018). DeepLesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *Journal of Medical Imaging, 5*, 036501.

Yashu, Kukreja, V., & Garg, U. (2023). Advances in Skin Disease Recognition: Hybrid Deep Learning and Ensemble Models for Accurate Classification. *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*, 1-4.

Zhang, B., Zhou, X., Luo, Y., Zhang, H., Yang, H., Ma, J., & Ma, L. (2021). Opportunities and challenges: Classification of skin disease based on deep learning. *Chinese Journal of Mechanical Engineering*.