ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

REVIEW ARTICLE

Artificial intelligence in diagnosis and monitoring of atopic dermatitis: From pixels to predictions

Pratheek Jain[1,2], Farhan Zameer[1*], Kounaina Khan[1], Vinay Alva[1,3], Ravish Huchegowda[4], Ali Jawad Akki[5], Raghu Anjanapura Venkataramanaiah[5], Muthuchelian Krishnasamy[6], Dilip Apturkar[3], and Raghavendra Hallur Laxmanashetty[2*]

[1]PathoGutOmics Laboratory, Alva's Traditional Medicinal Archive (ATMA), Department of Ayurveda Pharmacology, Alva's Ayurveda Medical College, Moodubidire, Karnataka, India

[2]Centre for Biotechnology, Pravara Institute of Medical Sciences, Ahmednagar, Maharashtra, India

[3]Department of Surgery, Dr. Balasaheb Vikhe Patil Rural Medical College, Pravara Institute of Medical Sciences (Deemed to be University), Ahmednagar, Maharashtra, India

[4]Department of Neurochemistry, National Institute of Mental Health and Neuro Sciences (NIMHANS), Bengaluru, Karnataka, India

[5]Department of Chemistry, Faculty of Science and Technology, Bijapur Lingayat District Educational (Deemed to be University), Vijayapura, Karnataka, India

[6]Department of Biotechnology, M.G.R. College, Dr. MGR Nagar, Hosur, Tamil Nadu, India

Abstract

In any ailment, the identification of the symptoms, detection, and diagnosis plays a pivotal role in treatment and therapy. However, certain diseases share similar symptoms, lacking signature key indicators, which can lead to fallacious or incorrect inferences. Skin disorders, such as pruritus, dermatitis, eczema, psoriasis, and ichthyosis, all present similar symptoms, which confound clinicians. One such commonly misunderstood condition is atopic dermatitis (AD), a chronic inflammatory skin condition characterized by its relapsing nature, which heightens the importance of diagnosis and disease monitoring for effective management. Recent strides in artificial intelligence (AI) have opened avenues for precise diagnosis and continuous monitoring of AD. This review explores and evaluates current applications of AI in the diagnosis and monitoring of individuals with AD emphasizing the need to address challenges and collaborate across intra-, inter-, trans-, and multi-disciplinary domains to maximize the benefits of AI in enhancing the precision of AD diagnosis, ultimately leading to improved patient care and satisfaction through technologically-driven biomedical tools in customized healthcare.

Keywords: Deep learning; Machine learning; Convolutional neural networks; Artificial neural network; Pruritus

***Corresponding authors:**
Raghavendra Hallur Laxmanashetty
(director.biotech@pmtpims.org)
Farhan Zameer
(farhanzameeruom@gmail.com)

Citation: Jain P, Zameer F, Khan K, *et al.* Artificial intelligence in diagnosis and monitoring of atopic dermatitis: From pixels to predictions. *Artif Intell Health.* 2024;1(2): 48-65. doi: 10.36922/aih.2775

Received: January 18, 2024

Accepted: March 1, 2024

Published Online: April 18, 2024

1. Introduction

Artificial intelligence (AI) has greatly improved our quality of life and is now widely used in many areas, including healthcare. Examples of its applications include autonomous driving, industrial automation, and the widespread use of cell phones. Recent developments have notably bolstered the efficiency, accuracy, and productivity of AI-optimized workflows in the health-care industry. AI involves the application of sophisticated computational algorithms to simulate complex cognitive functions of humans, accomplished through learning and adaptation to gathered data. Over the past decades, there has been a marked increase in both research and application of AI in healthcare, with the potential to completely transform the sector.[1-4]

Similar to AI, the field of dermatology has experienced rapid growth owing to advancements in technology and inventions. This evolution has brought dramatic changes in the diagnosis and treatment of dermatological illnesses. Computer algorithms have proven to be invaluable tools for dermatologists, particularly in diagnosing diseases such as malignant melanoma.[5] Dermatology boasts a vast archive of clinical, dermatoscopic, and dermatopathological images, positioning it as a leader in the application of AI in medicine. Therefore, having a basic understanding of AI becomes essential for designing and evaluating medical research in this area. Consequently, investigating the potential uses of AI in dermatological practice becomes imperative.

Atopic dermatitis (AD) is a common inflammatory skin disease that affects a significant proportion of dermatology patients worldwide, estimated to affect 1 – 2% of the global population.[6] Notably, between the 1980s and the early 2000s, there was a discernible global surge in the prevalence of AD, particularly pronounced among children under the age of five, with rates ranging from 10% to 16.5%.[7,8] Individuals with AD exhibit a wide range of clinical symptoms, categorized into six distinct subtypes based on their origin. Among these, the most prevalent subtype is early-onset, early-resolving; nonetheless, recurrence is frequently observed, with less severe symptoms than the original episode.[9,10] The three main symptoms of AD – skin inflammation, compromised skin barrier function, and persistent itching-exerta negative influence on the lives of those afflicted, significantly impacting their quality of life and level of satisfaction with therapy.[11-13] Moreover, patients' adherence to treatment procedures is severely hindered by these symptoms.[14-16] *Staphylococcus aureus* colonization often arises as a result of skin barrier dysfunction, exacerbating the degradation of the barrier function.[17,18] Notably, AD is linked to systemic inflammatory conditions, such as metabolic syndrome and cardiovascular disease, despite not being as visually conspicuous as psoriasis.[19]

Given its broad spectrum of clinical symptoms, AD presents as an unexpected illness that can prove challenging to diagnose and monitor effectively. Prompt and accurate identification is indispensable for the effective management of AD. However, conventional diagnostic techniques tend to heighten variability in diagnosis and may cause delays in the initiation of therapy, relying heavily on clinical judgment. A promising avenue toward achieving objective, accurate, and rapid diagnostic procedures lies in the integration of AI-based technology, which holds the potential to revolutionize AD management. This review highlights the rationale behind the potential of AI in completely transforming AD monitoring and diagnostic procedures. While acknowledging the efficiency of AI, it also emphasizes the importance of problem-solving and fostering teamwork. These initiatives are essential in maximizing the benefit of AI in improving the precision and effectiveness of AD monitoring and diagnosis procedures.

2. Methodology

In conducting a systematic literature review on the contemporary utilization of AI in diagnosing and monitoring individuals with AD, we implemented a rigorous methodology for gathering pertinent articles from prominent databases, including PubMed, Springer, and Elsevier. The search utilized specific keywords, such as "Artificial intelligence," "Machine learning," "Deep learning," and "Atopic dermatitis." Article selection prioritized peer-reviewed studies in dermatology that specifically examined the integration of AI in the diagnosis and monitoring of AD. Articles that failed to meet these criteria or published in languages other than English were systematically excluded from the study. Following data extraction, we summarized the key findings, and the synthesized information was then carefully compiled into a comprehensive literature review, providing valuable insights into the current state of knowledge, addressing challenges, and advocating for collaboration across intra-, inter-, trans-, and multi-disciplinary domains to optimize the benefits of AI in improving the accuracy of AD diagnosis.

3. Principle of AI

AI encompasses various computational subfields, including machine learning (ML) and natural language processing (NLP), enabling computer systems to mimic human cognitive functions (Figure 1). At present, ML, where computers anticipate data without explicit programming, stands as the frontier of AI advancement. Essentially, computers "learn" from data, offering analyses without explicit guidance on trait prioritization. Dermatologists offer compelling examples, such as identifying melanomas from clinical images,[5] predicting the effectiveness of

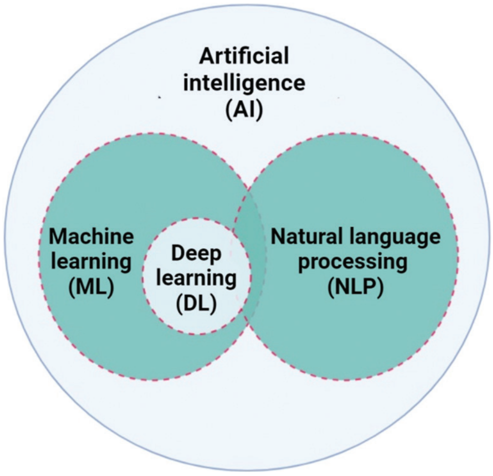

Figure 1. An overview of the principles of artificial intelligence. Artificial intelligence (AI) is a broad category of algorithms that includes subcategories such as machine learning (ML), natural language processing (NLP), and deep learning (DL).

biologic therapies for psoriasis,[20] and analyzing doctor notes in electronic health records to discern clinic visit purposes for AD.[21]

Deep learning (DL), a subset of ML, uses algorithms modeled after human neurons to discern complex patterns and relationships in data. DL permits the direct entry of raw data, unlike older ML methods that require domain expertise and human engineering to translate raw data into intelligible algorithm features.[22] For pattern recognition, the machine autonomously creates its own representations, which are arranged in a series of layers that build on one another to gradually abstract the data. Neural networks are represented by this layer architecture.[23] DL includes diverse methods, such as transformers,[24] which are adept at identifying sequential data relationships and extracting meaning, and convolutional neural networks (CNNs),[25] which are frequently used in imaging tasks. Due to its versatility and flexibility, DL is an effective tool for a wide range of applications.

Algorithms in the broad field of ML employ diverse techniques to acquire knowledge. As depicted in Figure 2, these techniques include reinforcement learning, unsupervised learning, and supervised learning. Supervised learning, the most popular ML technique, relies on labeled datasets to predict outcomes. Predictions based on unseen data are made possible by the algorithm's ability to map input data to the correct output. During the training phase, the algorithm receives both the data and the corresponding answers (ground truths) from a set of training instances, enabling it to modify its weights accordingly. Subsequently, the algorithm's performance is assessed against a different test set that it had not

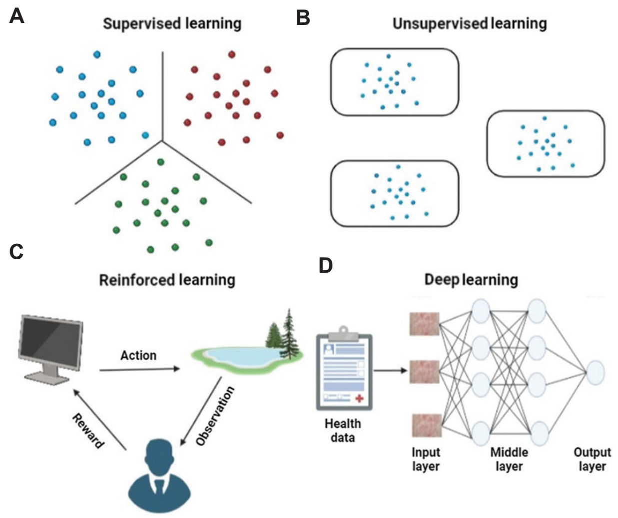

Figure 2. Several types of machine learning techniques, including supervised learning, unsupervised learning, and reinforcement learning. (A) Supervised learning involves using labeled datasets to categorize data, while (B) unsupervised learning does not use labeled datasets and instead finds patterns and relationships in the data to create categories. (C) Reinforcement learning uses iterative feedback loops to teach the algorithm. (D) Deep learning, a subset of machine learning, uses representation layers in a neural network to increase the abstraction of the data and employs techniques from all three types of machine learning.

encountered before post-training. Logistic regression and linear regression, two of the most frequently used methods in this domain, find common applications in image-based dermatology models. On the other hand, unsupervised learning involves training a model on datasets without labels, meaning that the input lacks a known right response. The algorithm's primary goal is to discover patterns and links in the data, such as clustering related data points. In the paradigm of reinforcement learning, an algorithm referred to as the agent interacts with the environment to accomplish predetermined objectives. Based on feedback, it receives from its actions in the form of rewards or penalties, the agent modifies its behavior to optimize rewards. Reinforcement learning learns through ongoing feedback loops, distinct from the predetermined data input found in supervised and unsupervised learning methods. This three-class categorization of ML techniques offers a thorough grasp of the various techniques algorithms use to learn from data.

Drawing on concepts from linguistics, statistics, ML, and DL, a discipline of AI known as NLP aims to interpret, analyze, and generate human language.[26] This complete technique enables the processing of human language in its entirety. Within NLP, two primary subfields exist: natural language generation (NLG) and natural language understanding (NLU). NLG focuses on generating new text, while NLU is committed to understanding textual material. NLG encapsulates recent advances in large language models, exemplified by OpenAI's freely available Chat Generative Pre-trained Transformer.[27] These advancements highlight the evolving landscape of NLP and its important role in the advancement of language-related applications.

In recent years, the adoption of multimodal techniques in algorithms has surged, driven by the utilization of diverse data sources for training. Given the inherently multifaceted nature of medicine, where doctors must interpret a wide range of data, including genetic information, laboratory results, clinical notes, and radiological images, these multimodal approaches have gained prominence. The latest strides in this discipline focus on building more reliable models and algorithms by leveraging the abundance of readily available data. Noteworthy examples of these multimodal technologies include Med-Flamingo,[28] LLaVa-Med,[29] Med-PaLM Multimodal (Med-PaLM M),[30] and MiniGPT-4.[31] At the core of these technologies, they lie foundation models (FMs), which undergo training on a variety of unlabeled datasets before being adjusted for certain downstream applications.[32] One particularly intriguing aspect is the ability of these models to absorb vast amounts of information from large datasets and subsequently apply this knowledge to specific applications, including those within the medical domain. This pattern represents a dynamic movement in the direction of using multimodal techniques to improve performance in medical applications.

4. AI for the diagnosis of AD

Accurate dermatological diagnosis and treatment of AD hinge on the quantitative evaluation of the disease, which emphasizes the molecular composition of the skin using non-invasive techniques. Confocal Raman micro-spectroscopy (CRM) serves as a tool for assessing the skin's biomolecular composition. Nevertheless, deciphering complex Raman spectroscopic signals requires multivariate analysis. Dev *et al.*[33] have presented a novel approach to classifying AD from healthy individuals by combining CRM with multivariate analysis, more precisely, partial least squares discriminant analysis (PLS-DA). While the current PLS-DA classification model is designed for binary classification, there is potential to explore its applicability for multiclass categorization based on the severity of eczema illness. The ML-aided PLS-DA classification approach used in the study simplifies dimensional reduction, variable selection, and classification for Raman micro-spectroscopy data. The cross-validated PLS-DA classification model exhibits remarkable sensitivity and specificity, scoring 0.94 and 0.85, respectively. Further enhancement of categorization accuracy is feasible by concentrating on wave number bands with a variable importance in projection (VIP) score of ≥1. In addition to bolstering the model's accuracy, the VIP score facilitates the identification of important Raman spectroscopic signatures associated with proteins, lipids, and nucleic acids, which can serve as biomarkers for therapeutic and clinical evaluation of AD patients' skin health. Using CRM and multivariate analysis, this quantitative method of assessing skin inflammatory disorders such as AD offers a viable path for next-generation diagnosis, departing from the subjective scoring systems currently used in clinical practice. The presented study describes a novel diagnostic method specific for AD using CRM and multivariate analysis. This non-invasive method will provide a new approach for molecular-based evaluation of skin conditions. Nevertheless, several challenges need to be addressed, such as sample size and diversity, independent dataset validation, clinical utility assessment, CRM standardization across different laboratories, patient data privacy and informed consent ethics issues, equipment accessibility, and cost. Regulatory approval for CRM technology's widespread use is also necessary. Overcoming these issues will improve the power and generalizability of this innovative diagnostic protocol for AD. Furthermore, the application assures the availability of significant datasets and ensures the repeatability and reliability of the model.

Multiphoton tomography (MPT) has previously demonstrated its utility as a diagnostic tool in dermatology. However, MPT data analysis has remained time-consuming and operator-dependent. In a study conducted by Guimarães *et al.*,[34] the potential of using AI for diagnosing AD from MPT images was substantiated. AD system was developed to discern images containing living cells and performs subsequent diagnostics accurately and reliably, thus eliminating the need for human operators. The study has demonstrated the potential of completely harnessing MPT through a CNN-based, fully automatic method. CNNs were trained and fine-tuned using 3663 MPT images, including morphological and metabolic information from both AD patients and healthy individuals. The primary objectives were to identify live cells and diagnose AD, irrespective of the imaging layer or location. Impressively, the suggested algorithm successfully diagnosed AD in 97.0±0.2% of the images containing live cells, with a sensitivity of 0.966±0.003, specificity of 0.977±0.003, and *F*-score of 0.964±0.002. The interpretability of the algorithm was enhanced using relevance propagation through deep Taylor decomposition, generating heat maps that highlighted important details for each classification. The study exemplifies the successful integration of MPT imaging and AI for AD diagnosis, marking a substantial improvement in the field. The proposed method establishes a framework for automating the identification of skin conditions using MPT.

Activation-regulated chemokine (TARC/CCL17) and immunoglobulin E (IgE) have served as biomarkers for AD in traditional approaches over the past few decades.[35-38] Common techniques used in these investigations include regression or correlation analyses between potential biomarkers and the intensity of AD symptoms, as well as univariate research comparing AD patients to healthy controls. However, accurate diagnosing and evaluating AD solely based on a single biomarker are considered extremely challenging.[39] Recent developments have ushered in the utilization of multivariate ML techniques to uncover hidden patterns between variables and develop more reliable predictive models in a variety of studies, including those in pain research.[40,41] The combination of multiple serum biomarkers, such as TARC, IL-22, and sIL-2R, has improved the model's capability to predict eczema area and severity index (EASI) scores compared to relying on a single biomarker.[39] In addition, the correlation coefficient between the combined biomarkers and the disease severity surpasses that of the individual biomarkers. This method emphasizes the potential of combining multiple biomarkers for a thorough comprehension and prediction of AD severity.

In a recent study, Lee *et al.*[42] investigated the potential of employing a multivariate ML technique to develop a diagnostic tool and severity prediction model for patients with AD. The authors conducted phase I ML analysis, wherein they collected multivariate data, divided it into training and test sets, trained the models, estimated prediction performance, and selected and estimated features. Clinical and serological indicators[43,44] from a prior clinical study were combined. The results indicate that the classification model significantly outperformed the random permutation model, boasting an area under the curve of 0.85±0.10 and a balanced accuracy of 0.81±0.15, compared to 0.50±0.15 for the latter. Correlation analysis unveiled a significant positive association between the objective SCORing AD score (SCORAD) (r=0.53), measured and projected total SCORAD (r=0.43), and eczema area and severity index scores (r=0.58, each p<0.001). Nevertheless, no discernible relationship was observed between the measured and anticipated itch scores (r=0.21, p=0.18). The research encompassed the creation and evaluation of multivariate prediction models, as well as the identification of critical characteristics using a range of serum biomarkers. These results underscore the potential of utilizing a multivariate ML approach to reveal complex connections between clinical and serum measures in patients with mild-to-moderate AD.

Colonocytes or colonic epithelial cells have recently garnered attention for their role in host-microbial interactions. During gut dysbiosis, which is linked to a number of chronic human disorders, colonocytes influence the composition and activity of the gut microbiota.[45] The diagnosis and prognosis of AD can now be achieved through the integration and correlation analyses of gut microbiota and host gene expression.[46,47] Notable correlations have been observed, including those between IL-17 and *Streptococcus* infection in AD, and between enzyme commission genes and microbiota in inflammatory bowel illnesses.[48,49] Despite these advancements, few researchers have investigated ML prediction analysis based on the gut transcriptome and microbiota in AD. In recent work, Jiang *et al.*[50] developed an ML classifier for precise and automated AD detection by utilizing the transcriptome of gut epithelial colonocytes and gut microbiota data (Figure 3). With an average F1-score of 0.84, the classifier demonstrated accurate discrimination and successfully predicted the risk of AD. It was trained on data from 161 participants, including both AD patients and healthy controls. The research identified three genes and three bacteria that are either directly or indirectly linked to AD, as well asa combination of 35 genes and 50 microbiome traits predictive for AD. These results suggest that the discovered genes and microbiota traits may provide fresh biological perspectives and serve as useful biomarkers for early detection of AD. However, replication studies with different populations are necessary to validate these findings. The study represents a major step toward the construction of an ML classifier for accurate and automated AD diagnosis, utilizing gut microbiota and transcriptome data from gut epithelial colonocytes. The robust ML pipeline used in the study, which comprises thorough procedures such as feature selection, model selection, cross-validation, classification, and follow-up statistical assessments, enables accurate distinction based on omics data (Figure 4).

Dautović *et al.*[51] developed an artificial neural network (ANN) specifically designed for the automated diagnosis of AD, aiming to facilitate the diagnosis process. The network uses a feed-forward ANN with nine input parameters and one output parameter for classification. After evaluating various configurations, the final design of the expert system chose a neural network with 15 neurons in a hidden layer based on training results. Demonstrating impressive sensitivity at 95.62% and accuracy at 94.44%, the ANN exhibits excellent performance in distinguishing AD from other skin disorders. However, it is imperative to recognize that despite its high sensitivity, the comparatively lower

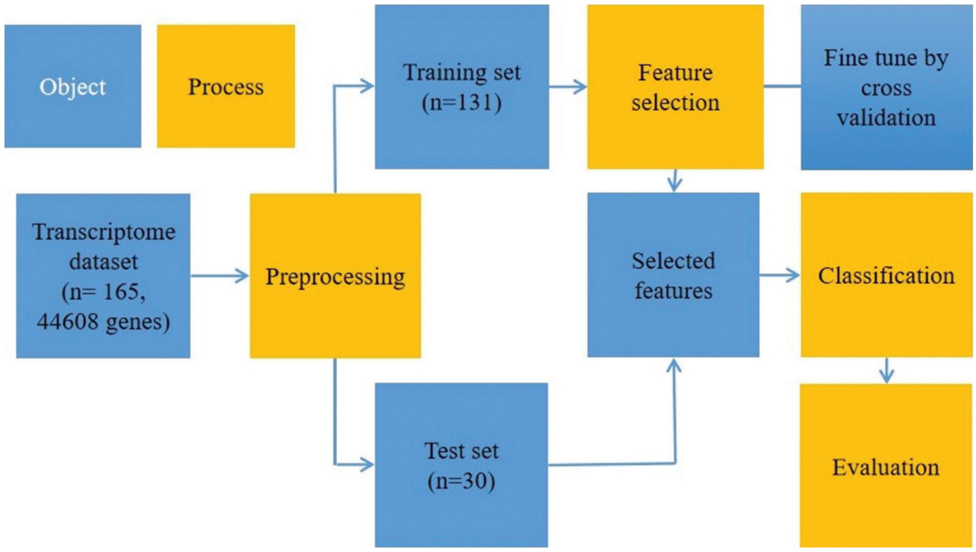

Figure 3. A comparison of two atopic dermatitis classification pipelines by Jiang *et al.*[50] (only the transcriptome dataset)

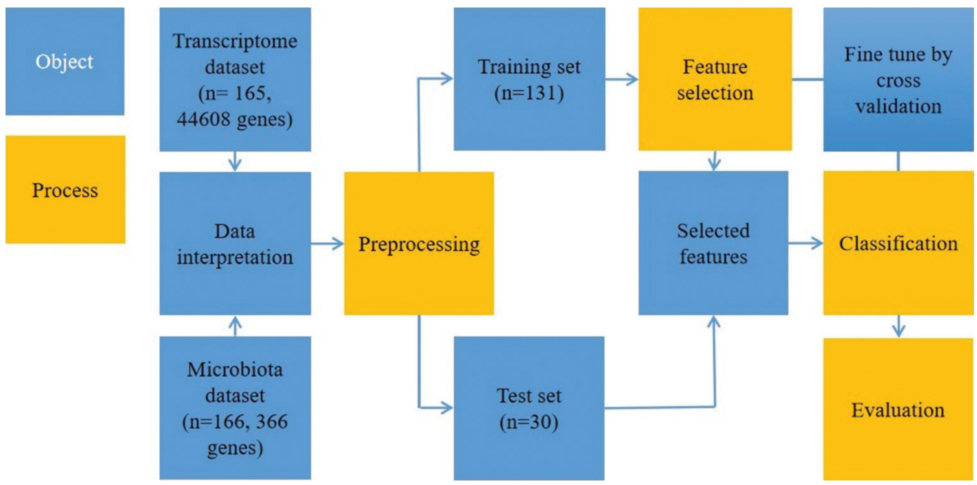

Figure 4. A comparison of two atopic dermatitis classification pipelines by Jiang *et al.*[50] (Both transcriptome and microbiota data)

specificity (85%) leads to a notable and unsatisfactory false positive rate of 15%. Furthermore, it is noteworthy that the ANN model shows diminishing resilience, and the model over fitting that is likely responsible for the provided accuracy is demonstrated by the F1 score of 0.964 and the Matthews correlation coefficient of 0.7454. A thorough summary of the performance results from the further validation of ANN is provided in Table 1, which also sheds light on the model's advantages and disadvantages.

AI-based techniques for diagnosing AD face various challenges stemming from diverse methodologies and data types. Although MPT coupled with AI has demonstrated success in automated diagnosis, it faces challenges such as robust generalization, real-time application feasibility, and ethical concerns. Similarly, serum biomarkers and multivariate ML techniques encounter issues, including selection and integration challenges, data standardization, and the necessity for extensive clinical validation. The complexity of AD, which is a multifactorial condition, renders accurate diagnosis and disease severity evaluation difficult using a single biomarker. Furthermore, integrating gut microbiota, host gene expression, and ML presents challenges concerning data reproducibility and validation across different populations. Replication studies are crucial for validating genes and microbiota traits as viable biomarkers. For instance, the ANN developed by Dautović *et al.*[51] exhibits sensitivity but grapples with specificity and resilience issues. Concerns such as overfitting and model validation necessitate thorough evaluation and validation across diverse datasets. Common challenges persist across these techniques, including the acquisition of diverse,

Table 1. Artificial neural network system performance

	Predicted positive output	Predicted negative output	Output class
Actual positive: 160	True positive (TP):153	False negative (FN): 7	1 – Subjects with disease Output class 0 – Healthy subjects
Actual negative: 20	False positive (FP):3	True negative (TN): 17	Accuracy: 94.44%
Σ = 180	Sensitivity: 95.62%	Specificity: 85%	F1 score: 0.9684
			MCC: 0.7454

Abbreviation: MCC: Matthews correlation coefficient.

representative datasets, ensuring the interpretability and explainability of AI models, addressing ethical considerations, and conducting rigorous clinical validation. The fulfillment of these requirements is essential to ensure the reliability and generalizability of proposed diagnostic tools for AD.

5. AI for monitoring AD

Clinical professionals' subjective visual inspections are frequently used to determine the severity of AD, which introduces significant inter- and intra-observer variability, especially in varied clinical study settings. In an attempt to standardize and automate the diagnosis of AD severity, Pan *et al.*[52] presented EczemaNet, a CNN computer vision pipeline. EczemaNet operates by initially identifying areas affected by AD in images, and subsequently generating probabilistic predictions regarding the severity of the condition. To generate its final predictions, EczemaNet uses ensemble approaches, including crops, ordinal categorization, and transfer and multitask learning. During evaluation in a published clinical trial, EczemaNet exhibited minimal root mean square error and well-calibrated prediction intervals. The research demonstrated the effectiveness of CNNs in treating non-neoplastic skin conditions, especially when dealing with medium-sized datasets. This finding highlights their potential for delivering an objective and more effective assessment of AD severity, which is a development with greater clinical significance compared to simple classification techniques.

Padilla *et al.*[53] used the MobileNet architecture, a CNN, to successfully distinguish between psoriasis and AD. They utilized publicly accessible dermatology to train the network. In a real-world experiment involving a Raspberry Pi camera and 30 subjects, the model successfully classified psoriasis with an impressive 90% accuracy rate and AD with an 88% accuracy rate. In a related study, Patella *et al.*[54] employed an ANN to examine the association between the severity of AD and exposure to air pollutants and environmental factors. Their results revealed a robust association, with the severity of AD lesions increasing by a considerable 200% in response to an increase in the diurnal temperature range, defined as the difference between the highest and lowest temperatures of the day. By predicting disease severity based on environmental parameters, the ANN exhibited promise in providing patients with early warnings to avoid potential irritants, in line with the overarching objective of predictive model-informed, tailored health-care actions.

Neural network algorithms provide a reliable and non-invasive method for classifying AD, frequently demonstrating efficacy in assessing lesion severity. This capability eliminates the need for direct clinical or specialized dermatologist intervention by allowing individuals to remotely monitor their condition using the cameras on their mobile phones. However, it is imperative to acknowledge that ANNs are not designed for constant observation, and existing models generally attain a maximum documented accuracy of 90%. Therefore, future research should prioritize enhancing accuracy and developing user-friendly mobile devices or applications to enable patients to confidently assess their condition using cutting-edge algorithms. In addition, additional scientific evidence is required to determine how the use of emollients or moisturizing creams affects the sensitivity of ANN. These studies could yield important information regarding whether these neural network systems can be used to measure the effectiveness of dermatitis treatment progress. Table 2 provides an overview of key information from the included studies, while Figure 5 illustrates the various data formats used in the previously described studies.

6. Implications of AI in evaluating pruritus in AD

The sensation of itching, medically referred to as pruritus, is a complex problem that significantly affects one's overall quality of life. The persistent and intense prickly discomfort associated with AD can cause mental health issues. These problems may manifest as increased activity levels, generalized anxiety, and, in certain cases, major depressive disorders.[55-57] Despite the profound detrimental effects of pruritus in AD, the lack of standardized and established techniques for objectively evaluating it poses a challenge to physicians and researchers.[58] Pruritus is inherently subjective, with diagnosis primarily reliant on patient reporting. To address this issue, numerous metrics and surveys have been developed. One such commonly used tool is the Peak Pruritus Numerical Rating Scale (NRS),

Table 2. Overview of studies investigating diagnosis and monitoring of AD utilizing AI

Serial No.	Applications	Description	References
1	CRM-PLS-DA	Machine learning-assisted CRM along with PLS-DA aid in precise dermatological diagnosis of AD, distinguishing AD from healthy individuals and exploring multiclass categorization of eczema severity.	33
2	CNN-based MPT	An advanced learning system has been developed to diagnose AD using MPT images, eliminating the need for manual intervention.	34
3	Multivariate machine learning for AD severity prediction using combined biomarkers	The study utilized multivariate machine learning to create a diagnostic tool and severity prediction model for AD patients, revealing intricate connections between clinical and serum measures and enhancing disease understanding.	42
4.	Machine learning classifier for AD detection using gut microbiota and transcriptome data	The study uses a machine learning classifier to accurately detect AD using gut epithelial colonocytes and gut microbiota data. The robust pipeline includes techniques like feature selection, model selection, cross-validation, classification, and statistical assessments, enabling precise discrimination based on omics data.	50
5	ANN for automated AD diagnosis	The ANN has been developed for automated diagnosis of AD. It uses a feed-forward architecture with nine input and output parameters, aiming to improve accuracy and efficiency by distinguishing AD from other skin disorders and using its unique features for classification purposes.	51
6	EczemaNet	EczemaNet is a computer vision tool used to monitor AD severity. It uses CNN and transfer learning techniques to predict severity. Clinical trials confirm its efficacy, providing a standardized, effective monitoring approach.	52
7	Itch Tracker	The Itch Tracker is a software application for smartwatches that tracks nocturnal scratching and provides an objective assessment of itching. It uses an algorithm to analyze acceleration data, distinguishing scratching from other movements based on unique wrist motions. The device is effective in measuring pruritus severity in patients with AD.	71
8	Neurological imaging	The application uses neurological imaging, specifically positron emission tomography and functional magnetic resonance imaging, to objectively identify structural and functional changes in pruritus, detecting brain activity during itching episodes.AI techniques could be used in the analysis of neurological imaging data for pruritus research. These could include automated image analysis, pattern recognition, and predictive modeling.	75
9	Acoustic surveillance	Acoustic monitoring has been used to analyze scratching behavior in AD patients. Initially applied to transgenic mice, a software application automates data analysis. A sound detector integrated with wrist monitoring has accelerated data analysis, but more research is needed.	67,68

Abbreviations: AD: Atopic dermatitis; AI: Artificial intelligence; CNN: Convolutional neural network; MPT: Multiphoton tomography; CRM: Confocal Raman microspectroscopy; PLS-DA: Partial least squares discriminant analysis; ANN: Artificial neural network.

which asks patients a single question: “On a scale of 0 – 10, with 0 being ‘no itch’ and 10 being ‘worst itch imaginable,’ how would you rate your itch at the worst point over the past 24 h?” The peak pruritus NRS has proven to be a well-defined, reliable, sensitive, and accurate scale for determining the intensity of the most severe itching. Its clear and simple format makes it particularly appealing to busy clinicians.[59]

Although scales and questionnaires have proven quite useful in clinical settings, biases resulting from individual differences in perceiving and expressing pruritus diminish their utility in research contexts. For example, what one individual may rate as a “10 out of 10” or the “worst itch conceivable” might be noticeably less severe for another, rendering such measures subjective. In addition, not everyone experiences pruritus solely as an “itchy” sensation; some describe it as a “burning” or “tingling,” which could cause some clinical instruments to overlook genuine cases of pruritus. These cases demonstrate the limitations of relying solely on scales and questionnaires. While they may aid a doctor in managing an individual AD patient, more objective and supplementary tools are desperately needed.[60-64] More specifically, rigorous research in this area requires instruments capable of identifying minute differences and facilitating comparisons among cohorts before and after AD therapies or interventions.[65,66]

7. Acoustic surveillance

Initially, acoustic surveillance was used on transgenic mice specially engineered to mimic AD. The scratching behavior of these animals was recorded using a sound recording device, and an in-depth analysis of the recorded scratching noises, including the examination of frequency and wavelength data, was conducted. Subsequently, a software application was developed to identify and measure the scratching habits of the mice. This method offers an

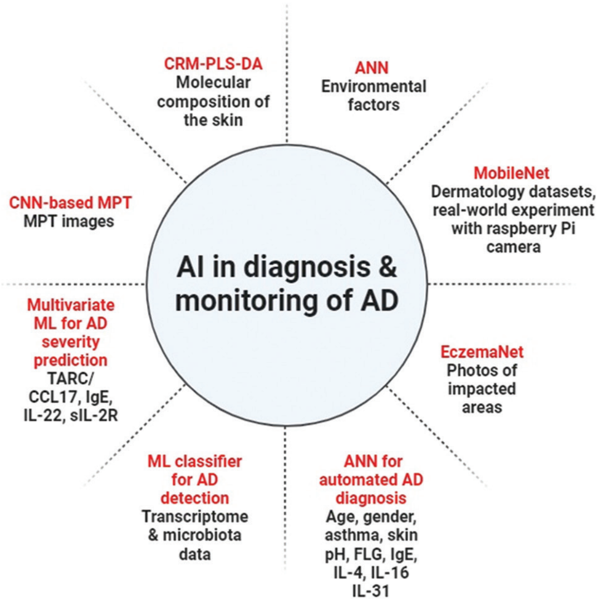

Figure 5. Diverse data modalities (such as images, genetic data, and biomarkers) employed across distinct AI studies.
Abbreviations: AD: Atopic dermatitis; AI: Artificial intelligence; ANN: Artificial neural network; ML: Machine learning; MPT: Multiphoton tomography.

automated and significantly more rapid analysis of scratch data, akin to video surveillance but without the need for direct observation. By contrasting the results of the acoustic counting method with those from video surveillance in the mouse model, the efficacy of the method was confirmed, revealing a close similarity between the two methods.[67]

In 2014, Noro *et al.*[68] expanded on the achievements of the mouse model by developing a sound detector integrated with a similar acoustic surveillance system, designed to be worn on the wrist to monitor scratching behavior. This groundbreaking technology distinguished a particular bone-conducted sound produced by motions through the finger and wrist bones, rather than relying on air-conducted noises. Over a 6-h sleep period, sound data were gathered from both AD patients and healthy controls, while infrared video captured individuals' scratching motions. The algorithm swiftly analyzed the audio data, taking only a few minutes compared to the several hours required for human observers to score video recordings. According to the study, the scratching time recorded by the sound detector was later discovered to be almost identical to the outcomes from video surveillance, which is considered the gold standard for comparison. As a result, the sound detector greatly accelerated and improved the objectivity of the study while also reducing the invasive need for human behavioral observation. Similar to advances in video surveillance, machine-learning algorithms are being actively researched to improve acoustic monitoring. These techniques analyze sound data to detect movement and quantify scratching activity. Despite the advantages of acoustic recording over visual surveillance, it has only been tested in a restricted capacity in the brief trial described earlier. Therefore, more extensive and comprehensive research is needed to confirm this strategy across a wider range of patient demographics.[68]

8. Smart devices

In the past 10 years, medical research has gained insights into a range of medical conditions, and novel therapies have been introduced subsequently, which include monitoring dietary intake and sending reminders for medication adherence, facilitated by the portability and processing power of smart devices.[69,70] Researchers investigating pruritus have modified wrist actigraphy for smartwatches while maintaining its core principles and keeping up with technological advancements. In a pilot study conducted by Lee *et al.*,[71] an accelerometer-equipped wristwatch was used with three subjects to identify scratching tendencies. Remarkably, when compared to infrared video surveillance, the wristwatch demonstrated remarkable accuracy, with detection rates ranging from 98.5 – 99.0% for right-hand scratching motions and 93.3 – 97.6% for left-hand scratching.

In 2017, the Itch Tracker device was developed through a joint effort between dermatologists, Nestle Skin Health, and Apple Inc. This invention consists of an application (software program) designed for well-known smartwatches, aimed at tracking nocturnal scratching and addressing the need for improved objective techniques in assessing itching. As stated in the study, the application incorporates an algorithm that evaluates acceleration data from smartwatches, using unique wrist motions to distinguish scratching from other types of movement. In addition, the application features a smartphone interface that enables users to respond to surveys linked to itching, thus combining subjective patient comments with objective data from wearables.[71] These results highlight the effectiveness of Itch Tracker as a smart gadget for tracking scratching, offering an objective and indirect measure of pruritus severity in patients with AD. With its subtle design and user-friendly interface, Itch Tracker is appropriate for general use. However, further studies are necessary to evaluate the application in various clinical settings and patient demographics.[72]

9. Neurological imaging

A unique method for objectively identifying structural and functional changes in both acute and chronic pruritus is neurological imaging. Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have been used to measure brain activity during experimentally induced itching episodes, offering insights into treatments for acute itching.[73] In a previous study, eight patients with AD and six healthy controls underwent PET scanning to examine acute histamine-induced itching. PET scans of AD patients revealed increased brain activity, especially in the basal ganglia, which are known to be a key regulator of the itch-scratch cycle.[74]

A study using arterial spin labeling fMRI, which involved histamine induction in seven healthy controls and eight AD patients, has yielded similar findings, demonstrating markedly enhanced cerebral perfusion after acute scratching in patients with AD.[75] In addition, two other studies investigating brain imaging in individuals with chronic scratching revealed increased activity in reward circuits and motor-related brain areas during scratching.[76,77] Moreover, patients with chronic itching exhibited reduced sensation of itch in response to scratching in these two studies that used arterial spin labeling fMRI to measure brain activity during active scratching. Concurrently, brain regions linked to the central reward system showed significant activation. Taken as a whole, these results emphasize the pleasurable aspect of scratching in relieving both acute and chronic pruritus, and they suggest the possibility that individuals with AD may develop an addiction to scratching.

It is important to recognize three main limitations of neurological imaging in the context of AD-related acute and chronic pruritus.[73] First, many of these studies involve limited sample sizes, and the results are contingent on the techniques used to induce itching and perform imaging. Second, there is a scarcity of studies demonstrating differences in resting states between healthy controls and AD patients. Instead, these techniques need an itch stimulus other than AD-induced itch (e.g., histamine-induced itch) to identify changes that might not accurately mimic the natural state. Third, tests such as PET and fMRI are expensive for the health-care system and uncomfortable for patients, requiring extensive time for execution and analysis. These variables diminish the usefulness of brain imaging for routine clinical diagnosis and standard care. The different distribution of skin lesions with varying morphologies, intensities, and durations might complicate the diagnosis of AD further. However, in recent years, there has been a rapid development in AI-based techniques for image analysis. One method that represents complicated patterns is the use of CNNs, which rely on DL algorithms to identify correlations between neighboring images and integrate them into successive layers. CNNs are used in image analysis across a variety of medical specialties, including cardiology, neurology, and gastrointestinal research, in addition to dermatology.[34]

10. Limitations of AI in dermatology and possible solutions

At present, several noteworthy obstacles hinder the effective application of AI in the medical field, especially in dermatology. One of the primary challenges is the lack of appropriate quality manual annotation and limited sample numbers in existing dermatology training datasets for AI algorithms. This deficiency diminishes the accuracy and usefulness of AI algorithms, rendering them inadequate to meet the demands of routine clinical applications.[78] Furthermore, AI algorithms, usually developed using pre-existing samples, often fail to align with actual medical needs, resulting in a disconnect between them and practical clinical requirements. High-quality training sets are necessary for AI to learn from experience and evolve over time-a capability that human doctors possess naturally. Without such datasets, AI's capacity is hampered, making it challenging to satisfy the growing expectations of both clinical and scientific domains, especially in areas such as the hairy scalp, mucosal membranes, uncommon skin disorders, and the identification of picture artifacts such as colorful marks and tattoos on the skin.[79] Furthermore, training AI to recognize and diagnose a variety of skin problems is difficult due to the wide range of dermatological diseases and the lack of standard criteria for identification and diagnosis.[80] There remains a bottleneck in using AI for the automatic recognition and diagnosis of various dermatopathological images, with current AI applications more frequently used for differentiating between normal and abnormal instances.[81,82] In addition, another major barrier to the use of AI in dermatology is the presence of rare disorders, characterized by a low number of cases and insufficient specimens for appropriate ML training.[83]

At the center of these challenges, it lies the critical importance of guaranteeing the quality of data used in AI services. This intricacy is compounded by intertwined issues, such as potential inaccuracies in the accuracy and reliability of annotations, which can affect the accuracy of the model training assumptions. Furthermore, variability introduced in data collection procedures aggravates these problems, resulting in datasets that are highly inconsistent and pose significant challenges for applying general models to a multitude of clinical scenarios. The presence of such errors, artifacts, or the lack of proper data preprocessing algorithms adds another layer of complexity to raw dermatological data, which must be addressed through comprehensive data preprocessing strategies to enhance the reliability of the model outputs.[84]

One major challenge is addressing bias and limitations in dataset representation, which are evident in the diversity in skin types, conditions, and demographic factors. These aspects may not have been adequately emphasized in training datasets, leading to biases in model outcomes.[85] The complexity adds an extra layer to the dynamics related to clinical practices, impacting data relevance over time and asking for continuous adaptation of AI models. In addition, providing access to various datasets is crucial for adequate model training, but it poses challenges, and potential limitations may restrict effective model generalization. Moreover, ethical considerations and patient privacy issues further complicate matters, especially concerning sensitive dermatological information. Careful balancing of the use of patient information for AI research with individual privacy protection is essential. These factors make it challenging to predict the performance of AI models in actual clinical environments, which can significantly differ from controlled research environments due to diverse patient populations, variations in clinical workflows, and the dynamic nature of healthcare. Effectively recognizing and addressing these nuanced limitations in AI applications in dermatology are pivotal for developing models that not only demonstrate technical proficiency but also seamlessly align with the complex yet ever-changing realities of clinical practice.[86]

Overcoming the challenges of small datasets in AI applications in dermatology requires an appropriate and well-designed strategy. Data augmentation is the key approach that enhances dataset quality by transforming images into different forms, thus improving data diversity and allowing the model to learn more effective features.[87] Transfer learning proves advantageous when models initially trained on huge datasets are fine-tuned using particular dermatology datasets, allowing them to gain general knowledge from a wider context.[88] In addition, the incorporation of carefully developed artificial data, resembling the characteristics of dermatological conditions, serves as an effective pathway toward diversification, underscoring the importance of adequate representation.

Active learning adds a new iterative retraining paradigm where the model selectively prioritizes informative or challenging samples during every retraining cycle, thereby refining its performance.[89] Ensemble models, which employ different architectures and hyperparameters, help minimize the effect of limited data by mixing predictions. Collaboration with other institutions, clinics, or research groups, along with data pooling, ensures the development of a broader and more diverse dataset while adhering to established rules on privacy and ethics. Active data collection remains crucial, requiring regular acquisition through collaborations or utilizing telemedicine platforms to expand the spectrum of the dataset.

To reconcile the disparities between source and target datasets in AI scenarios for dermatology, domain adaptation techniques have been applied.[90] These techniques aim to align the distributions of data and increase adaptability without leading to overfitting incapacitation. By focusing on key features and applying an expert-driven targeted approach within the limitations posed by limited datasets, effective solutions can be developed to address limited data.[91] Overfitting can be mitigated by incorporating regularization techniques such as dropout or weight decay during training.[92] Ensuring the high quality of the limited dataset is crucial as it significantly influences the proper exploitation and effective performance of AI models. By refining these strategies through the iterative process that is in line with evolving research and clinical needs, practitioners could effectively overcome the hurdles imposed by limited datasets in AI applications to dermatology.

11. The acceptance of AI in dermatology: Attitude attribute

The application of AI to medical image recognition has garnered substantial attention recently, particularly in the fields of dermatopathology and dermatology. The growing advancements in AI technology make its use as a decision support tool for dermatologists – particularly in diagnosis support – increasingly relevant within the current legal and health-care frameworks. With the growing utilization of AI by both patients and medical professionals, numerous regional and international survey studies have been conducted to gauge perceptions and attitudes. Between January and June 2019, a comprehensive online survey was distributed to 1271 participants across 92 countries. The results revealed that respondents identified dermoscopic images as the most promising application of AI in dermatology. Significantly, 77.3% of participants expressed approval or strong approval of AI's potential in improving dermatology, with 79.8% incorporating AI into their medical education. However, only a minimal 5.5% (70 out of 1,271) agreed or strongly agreed with the notion that AI would replace dermatologists in the near future. A comparable international survey was conducted among dermatopathologists by the same research team, involving 718 respondents from 91 countries. The findings revealed that 84.1% of respondents thought AI should be included in medical education, and 72.3% of respondents agreed or strongly agreed that AI will improve dermatopathology. Only 6.0% of respondents thought AI would eventually replace human pathologists. Interestingly, 79.2% of respondents thought that automated suggestions for diagnosing skin tumors had strong or very strong potential in terms of diagnostic categorization, whereas 42.6% thought that automated detection of mitosis had the highest potential.[93,94]

Patients typically know little about AI than medical professionals. In a qualitative study conducted from May to July 2019, involving 48 patients and semi-structured interviews for analysis, around 60% of participants stated that shorter diagnosis times and easier access to healthcare were the two biggest advantages of AI for skin cancer surveillance. Nonetheless, 40% of participants expressed concerns about potential dangers, including a rise in patient anxiety. The patients identified the major benefits and drawbacks of AI as the ability to deliver more precise diagnoses (33 [69%]) and less precise diagnoses (41 [85%]). Notably, 35 out of 75 patients stated that they would recommend AI to friends and family.[95] In summary, pathologists and dermatologists generally hold an optimistic view of the prospective advantages and effects of AI in the field of dermatology. However, only a minority of respondents within the cohort exhibited a good or exceptional comprehension of AI. While most pathologists expect AI to be most useful in specific tasks rather than offering overall automated diagnostic advice, a majority of dermatologists believe that AI will improve diagnostic capabilities. Overall, only a small percentage (1 – 3%) of pathologists and dermatologists express concern that AI may soon replace them. As long as AI is used in a way that maintains the doctor-patient relationship, patients are amenable to using it to monitor skin conditions.

12. Perspectives and conclusion

The potential of AI in the field of AD presents an opportunity to significantly enhance diagnostic accuracy and provide personalized healthcare. However, several aspects must be addressed before this innovative approach can be seamlessly integrated into routine clinical practice. AI is gaining recognition at a pace in the field of dermatology, with researchers increasingly focusing on developing AI programs that require diverse data sources for training purposes. These data sources include clinical patient data, which encompasses various aspects such as demographics, comorbidities, characteristics of skin lesions, and relevant laboratory and imaging findings. Furthermore, molecular profiles obtained from biopsy data, such as proteomic analysis, provide valuable information.[5,96] Another avenue involves utilizing data from existing literature. Finally, images play a crucial role in the analysis and classification process. Notably, publicly available benchmarking image datasets such as the International Skin Imaging Collaboration and PH2 dermoscopic archives serve as instrumental training resources for AI models.[97-99] These datasets often consist of lesions that have been confirmed by pathology, follow-up examinations, expert consensus, or *in vivo* confocal microscopy, which enhances their reliability.

Prudent consideration is crucial for researchers involved in the development of AI programs, especially when facing challenges related to training datasets. Estimating the optimal number of training images can be challenging, as having an insufficient dataset may compromise the quality of the program, while an excessively large dataset runs the risk of overfitting the ML classifier to the data, limiting its applicability to external datasets. It is important to note that advanced mathematical techniques are available to address these challenges, such as dropout, data augmentation, batch normalization, and others.[92,100,101] These methods play a key role in preventing overfitting and ensuring the robustness and generalizability of the AI program, which holds significant clinical relevance from a scientific perspective. Efficient utilization of the dataset is crucial in achieving the desired accuracy for specific classifications. In addition, the dataset should include a diverse range of images from various demographics to ensure that resulting algorithms have external validity.[102,103] When acquiring images, it is important to consider potential systematic errors such as variations in lighting, tools, or processes, particularly in different clinical settings, to maintain the research's validity beyond its original context. Simplifying the program's classifications to those with significant prognostic implications can help reduce the size of the dataset and the complexity of algorithms.[102,104]

Randomized clinical trials must be carried out to evaluate the potential of new computer methods and DL in large-scale investigations. Given the limited research in this area, these studies are especially important for gathering data on therapeutic benefits and assisting with causal inference.[105] Moreover, addressing unmet demands such as cost-effectiveness and safety concerns is critical before transferring AI technology from research to clinical settings. Robust regulatory procedures are required to guarantee the safe handling and preservation of private information. Another important challenge is ensuring AI-based healthcare is equitable and inclusive. Healthcare AI should be trained and validated using population-representative data to achieve generalizable performance levels.[106] It is crucial to take into account social and health inequalities that can exclude kids from particular groups who typically have limited access to care. Relying mostly on data from majority ethnic groups or patients with high socioeconomic status could introduce bias into AI performance, as the system may pick up diagnostic tendencies from these over-represented groups.[107]

Developing instructional programs to help doctors successfully use and interpret AI products and services through continuous education is another important factor to take into account. Although AI holds great potential to improve diagnosis, there are several possible drawbacks, including the risk of false diagnoses and erroneous risk assessments. Establishing formal collaboration between industry, the research community, and health-care systems is critical. Such collaboration is necessary to handle every important facet of AI and guarantee a smooth transfer from academic research to practical implementation.

In the near future, physicians and AI-guided computers will probably work closely together. AI has the ability to change clinical management by assisting physicians in analyzing patient data on an individual basis, identifying trends in different diagnostic test results, and more.[108] In this setting, the role of medical experts in interdisciplinary teams becomes essential. The implementation of AI algorithms requires the collection and analysis of large amounts of data, which calls for interpretation within a particular healthcare setting. The domain of "augmented intelligence," where people and machines cooperate to enhance the diagnostic workflow, share judgments, and work in concert, is where clinicians and computers can effectively collaborate. This collaboration has the potential to effectively minimize emotional and economic burdens, leading to more efficient patient care in the health sector.

Acknowledgments

Pratheek Jain and Vinay Alva express gratitude towards Pavara Institute of Medical Sciences (PIMS), Loni, Maharastra, for research support. All authors thank the Management of Alva's Education Foundation (AEF), Moodubidire, ATMA Research Centre, Vidyagiri, and BLDE University, Karnataka, for the constant research motivation.

Funding

None.

Conflict of interest

The authors declare that they have no conflicts of interest.

Author contributions

Conceptualization: Farhan Zameer, Raghavendra Hallur Laxmanashetty
Data curation: Ravish Huchegowda, Ali Jawad Akki
Writing – original draft: Pratheek Jain, Kounaina Khan, Vinay Alva
Writing – review & editing: Raghu Anjanapura Venkataramanaiah, Muthuchelian Krishnasamy, Dilip Apturkar

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Huchegowda R, Huchegowda S, Pramer J, *et al*. Review on artificial intelligence and applications in healthcare. *Neuro Open J*. 2019;6(1):e1-e4.

   doi: 10.17140/noj-6-e010

2. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, *et al*. Artificial intelligence in cardiology: Present and future. *Mayo Clin Proc*. 2020;95(5):1015-1039.

   doi: 10.1016/j.mayocp.2020.01.038

3. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31-38.

   doi: 10.1038/s41591-021-01614-0

4. Cichocki A. Chief Editor's foreword to the inaugural issue of artificial intelligence in health. *Artif Intell Health*. 2024;1(1):2463.

   doi: 10.36922/aih.2463

5. Esteva A, Kuprel B, Novoa RA, *et al*. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.

   doi: 10.1038/nature21056

6. DaVeiga SP. Epidemiology of atopic dermatitis: A review. *Allergy Asthma Proc*. 2012;33(3):227-234.

   doi: 10.2500/aap.2012.33.3569

7. Williams H, Stewart A, Von Mutius E, Cookson W, Anderson HR, International Study of Asthma and Allergies in Childhood (ISAAC) Phase One and Three Study Groups. Is eczema really on the increase worldwide? *J Allergy Clin Immunol*. 2008;121(4):947-954.e15.

   doi: 10.1016/j.jaci.2007.11.004

8. Furue M, Chiba T, Takeuchi S. Current status of atopic dermatitis in Japan. *Asia Pac Allergy*. 2011;1(2):64-72.

   doi: 10.5415/apallergy.2011.1.2.64

9. Williams HC, Strachan DP. The natural history of childhood eczema: Observations from the British 1958 birth cohort study. *Br J Dermatol*. 1998;139(5):834-839.

   doi: 10.1046/j.1365-2133.1998.02509.x

10. Paternoster L, Savenije OE, Heron J, *et al*. Identification of atopic dermatitis subgroups in children from 2 longitudinal

birth cohorts. *J Allergy Clin Immunol*. 2018;141(3):964-971.

doi: 10.1016/j.jaci.2017.09.044

11. Kiebert G, Sorensen SV, Revicki D, *et al*. Atopic dermatitis is associated with a decrement in health-related quality of life. *Int J Dermatol*. 2002;41(3):151-158.

doi: 10.1046/j.1365-4362.2002.01436.x

12. Furue M, Chiba T, Tsuji G, *et al*. Atopic dermatitis: Immune deviation, barrier dysfunction, IgE autoreactivity and new therapies. *Allergol Int*. 2017;66(3):398-403.

doi: 10.1016/j.alit.2016.12.002

13. Nakahara T, Fujita H, Arima K, Taguchi Y, Motoyama S, Furue M. Treatment satisfaction in atopic dermatitis relates to patient-reported severity: A cross-sectional study. *Allergy*. 2019;74(6):1179-1181.

doi: 10.1111/all.13712

14. Furue M, Onozuka D, Takeuchi S, *et al*. Poor adherence to oral and topical medication in 3096 dermatological patients as assessed by the Morisky Medication Adherence Scale-8. *Br J Dermatol*. 2014;172(1):272-275.

doi: 10.1111/bjd.13377

15. Murota H, Takeuchi S, Sugaya M, *et al*. Characterization of socioeconomic status of Japanese patients with atopic dermatitis showing poor medical adherence and reasons for drug discontinuation. *J Dermatol Sci*. 2015;79(3):279-287.

doi: 10.1016/j.jdermsci.2015.05.010

16. Takeuchi S, Oba J, Esaki H, Furue, M. Non-corticosteroid adherence and itch severity influence perception of itch in atopic dermatitis. *J Dermatol*. 2017;45(2):158-164.

doi: 10.1111/1346-8138.14124

17. Furue M, Iida K, Imaji M, Nakahara T. Microbiome analysis of forehead skin in patients with atopic dermatitis and healthy subjects: Implication of *Staphylococcus* and *Corynebacterium*. *J Dermatol*. 2018;45(7):876-877.

doi: 10.1111/1346-8138.14486

18. Iwamoto K, Moriwaki M, Miyake R, Hide M. *Staphylococcus aureus* in atopic dermatitis: Strain-specific cell wall proteins and skin immunity. *Allergol Int*. 2019;68(3):309-315.

doi: 10.1016/j.alit.2019.02.006

19. Furue M, Kadono T. "Inflammatory skin march" in atopic dermatitis and psoriasis. *Inflamm Res*. 2017;66(10):833-842.

doi: 10.1007/s00011-017-1065-z

20. Emam S, Du AX, Surmanowicz P, Thomsen SF, Greiner R, Gniadecki R. Predicting the long-term outcomes of biologics in patients with psoriasis using machine learning. *Br J Dermatol*. 2020;182(5):1305-1307.

doi: 10.1111/bjd.18741

21. Pierce EJ, Boytsov NN, Vasey JJ, *et al*. A qualitative analysis of provider notes of atopic dermatitis-related visits using natural language processing methods. *Dermatol Ther (Heidelb)*. 2021;11(4):1305-1318.

doi: 10.1007/s13555-021-00553-5

22. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

doi: 10.1038/nature14539

23. Esteva A, Robicquet A, Ramsundar B, *et al*. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24-29.

doi: 10.1038/s41591-018-0316-z

24. Wolf T, Debut L, Sanh V, *et al*. Transformers: State-of-the-Art Natural Language Processing. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics; 2020:38-45.

doi: 10.18653/v1/2020.emnlp-demos.6

25. Gu J, Wang Z, Kuen J, *et al*. Recent advances in convolutional neural networks. *Pattern Recognit*. 2018;77:354-377.

doi: 10.1016/j.patcog.2017.10.013

26. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: An introduction. *J Am Med Inform Assoc*. 2011;18(5):544-551.

doi: 10.1136/amiajnl-2011-000464

27. Introducing ChatGPT. *Released on November by 2022 OpenAI*. Available from: https://chat.openai.com [Last accessed on 2024 Jan 05].

28. Moor M, Huang Q, Wu S, *et al*. Med-Flamingo: A Multimodal Medical Few-Shot Learner. In: *Machine Learning for Health (ML4H)*; 2023:353-367. Available from: https://proceedings.mlr.press/v225/moor23a.html

29. Li C, Wong C, Zhang S, *et al*. LLaVA-Med: Training a Large Language-and-Vision Assistant for Biomedicine in One Day. In: *NIPS '23: Proceedings of the 37th International Conference on Neural Information Processing Systems*; 2023:28541-28564.

30. Tu T, Azizi S, Driess D, *et al*. Towards Generalist Biomedical AI. *arXiv*. Preprint posted online 2023.

doi: 10.48550/arXiv.2307.14334

31. Zhu D, Chen J, Shen X, Li X, Elhoseiny M. MiniGPT-4: Enhancing Vision-Language Understanding with Advanced Large Language Models. *arXiv*. Preprint posted online 2023.

doi: 10.48550/arXiv.2304.10592

32. Bommasani R, Hudson DA, Adeli E, *et al*. On the Opportunities and Risks of Foundation Models. *arXiv*. Preprint posted online 2021.

doi: 10.48550/arXiv.2108.07258

33. Dev K, Ho CJ, Bi R, *et al*. Machine learning assisted handheld

confocal raman micro-spectroscopy for identification of clinically relevant atopic eczema biomarkers. *Sensors (Basel)*. 2022;22(13):4674.

doi: 10.3390/s22134674

34. Guimarães P, Batista A, Zieger M, Kaatz M, Koenig K. Artificial intelligence in multiphoton tomography: Atopic dermatitis diagnosis. *Sci Rep*. 2020;10(1):7968.

doi: 10.1038/s41598-020-64937-x

35. Nomura I, Tanaka K, Tomita H, *et al*. Evaluation of the *Staphylococcal* exotoxins and their specific IgE in childhood atopic dermatitis. *J Allergy Clin Immunol*. 1999;104(2):441-446.

doi: 10.1016/S0091-6749(99)70390-8

36. Orfali RL, Sato MN, Santos VG, *et al*. *Staphylococcal* enterotoxin B induces specific IgG4 and IgE antibody serum levels in atopic dermatitis. *Int J Dermatol*. 2014;54(8):898-904.

doi: 10.1111/ijd.12533

37. Kakinuma T, Nakamura K, Wakugawa M, *et al*. Thymus and activation-regulated chemokine in atopic dermatitis: Serum thymus and activation-regulated chemokine level is closely related with disease activity. *J Allergy Clin Immunol*. 2001;107(3):535-541.

doi: 10.1067/mai.2001.113237

38. Hijnen D, De Bruin-Weller M, Oosting B, *et al*. Serum thymus and activation-regulated chemokine (TARC) and cutaneous T cell-attracting chemokine (CTACK) levels in allergic diseases TARC and CTACK are disease-specific markers for atopic dermatitis. *J Allergy Clin Immunol*. 2004;113(2):334-340.

doi: 10.1016/j.jaci.2003.12.007

39. Thijs JL, Nierkens S, Herath A, *et al*. A panel of biomarkers for disease severity in atopic dermatitis. *Clin Exp Allergy*. 2015;45(3):698-701.

doi: 10.1111/cea.12486

40. Krishnan A, Woo CW, Chang LJ, *et al*. Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *eLife*. 2016;5:e15166.

doi: 10.7554/elife.15166

41. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med*. 2013;368(15):1388-1397.

doi: 10.1056/NEJMoa1204471

42. Lee IS, Yeom M, Kim K, Hahm DH, Kang S, Park HJ. Prediction of disease severity using serum biomarkers in patients with mild-moderate atopic dermatitis: A pilot study. *PLoS One*. 2023;18(11):e0293332.

doi: 10.1371/journal.pone.0293332

43. Kang S, Kim YK, Yeom M, *et al*. Acupuncture improves symptoms in patients with mild-to-moderate atopic dermatitis: A randomized, sham-controlled preliminary trial. *Complement Ther Med*. 2018;41:90-98.

doi: 10.1016/j.ctim.2018.08.013

44. Kim J, Kwon SK, Lee IS, *et al.* Effect of acupuncture on gut-brain axis parameters in patients with atopic dermatitis: A study protocol for a randomized, participant-and assessor-blind, sham-controlled trial. *Evid Based Complement Alternat Med*. 2021;2021:5584247.

doi: 10.1155/2021/5584247

45. Litvak Y, Byndloss MX, Bäumler AJ. Colonocyte metabolism shapes the gut microbiota. *Science*. 2018;362(6418):eaat9076.

doi: 10.1126/science.aat9076

46. Ghosh D, Bernstein JA, Khurana Hershey GK, Rothenberg ME, Mersha TB. Leveraging multilayered "Omics" data for atopic dermatitis: A road map to precision medicine. *Front. Immunol*. 2018;9:2727.

doi: 10.3389/fimmu.2018.02727

47. Sacco KA, Milner JD. Gene-environment interactions in primary atopic disorders. *Curr Opin Immunol*. 2019;60:148-155.

doi: 10.1016/j.coi.2019.06.002

48. Lloyd-Price J, Arze C, Ananthakrishnan AN, *et al*. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*. 2019;569(7758):655-662.

doi: 10.1038/s41586-019-1237-9

49. Kang MJ, Lee SY, Park YM, *et al*. Interactions between IL-17 variants and *Streptococcus* in the gut contribute to the development of atopic dermatitis in infancy. *Allergy Asthma Immunol Res*. 2021;13(3):404-419.

doi: 10.4168/aair.2021.13.3.404

50. Jiang Z, Li J, Kong N, *et al*. Accurate diagnosis of atopic dermatitis by combining transcriptome and microbiota data with supervised machine learning. *Sci Rep*. 2022;12:290.

doi: 10.1038/s41598-021-04373-7

51. Dautović A, Dondras B, Dervisbegovic F, *et al*. Diagnosis of atopic dermatitis using artificial neural network. *IFAC Pap OnLine*. 2022;55(4):51-55.

doi: 10.1016/j.ifacol.2022.06.008

52. Pan K, Hurault G, Arulkumaran K, Williams HC, Tanaka RJ. EczemaNet: Automating Detection and Severity Assessment of Atopic Dermatitis. In: *Lecture Notes in Computer Science. Springer International Publishing*; 2020:220-230.

doi: 10.1007/978-3-030-59861-7_23

53. Padilla D, Yumang A, Diaz AL, Inlong G. Differentiating Atopic Dermatitis and Psoriasis Chronic Plaque using Convolutional Neural Network MobileNet Architecture. In: *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*. IEEE; 2019:1-6.

doi: 10.1109/hnicem48295.2019.9073482

54. Patella V, Florio G, Palmieri M, *et al*. Atopic dermatitis severity during exposure to air pollutants and weather changes with an artificial neural network (ANN) analysis. *Pediatr Allergy Immunol*. 2020;31(8):938-945.

doi: 10.1111/pai.13314

55. Slattery MJ, Essex MJ, Paletz EM, *et al*. Depression, anxiety, and dermatologic quality of life in adolescents with atopic dermatitis. *J Allergy Clin Immunol*. 2011;128(3):668-671.

doi: 10.1016/j.jaci.2011.05.003

56. Yaghmaie P, Koudelka CW, Simpson EL. Mental health comorbidity in patients with atopic dermatitis. *J Allergy Clin Immunol*. 2013;131(2):428-433.

doi: 10.1016/j.jaci.2012.10.041

57. Thyssen JP, Hamann CR, Linneberg A, *et al*. Atopic dermatitis is associated with anxiety, depression, and suicidal ideation, but not with psychiatric hospitalization or suicide. *Allergy*. 2017;73(1):214-220.

doi: 10.1111/all.13231

58. Silverberg JI. Practice gaps in pruritus. *Dermatol Clin*. 2016;34(3):257-261.

doi: 10.1016/j.det.2016.02.008

59. Yosipovitch G, Reaney M, Mastey V, *et al*. Peak pruritus numerical rating scale: Psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. *Br J Dermatol*. 2019;181(4):761-769.

doi: 10.1111/bjd.17744

60. Ikoma A, Rukwied R, Ständer S, Steinhoff M, Miyachi Y, Schmelz M. Neuronal sensitization for histamine-induced itch in lesional skin of patients with atopic dermatitis. *Arch Dermatol*. 2003;139(11):1455-1458.

doi: 10.1001/archderm.139.11.1455

61. Ikoma A, Steinhoff M, Ständer S, Yosipovitch G, Schmelz M. The neurobiology of itch. *Nat Rev Neurosci*. 2006;7(7):535-547.

doi: 10.1038/nrn1950

62. Paus R, Schmelz M, Biro T, Steinhoff M. Frontiers in pruritus research: Scratching the brain for more effective itch therapy. *J Clin Invest*. 2006;116(5):1174-1186.

doi: 10.1172/JCI28553

63. Potenzieri C, Undem BJ. Basic mechanisms of itch. *Clin Exp Allergy*. 2012;42(1):8-19.

doi: 10.1111/j.1365-2222.2011.03791.x

64. Sanders KM, Nattkemper LA, Yosipovitch G. Advances in understanding itching and scratching: a new era of targeted treatments. *F1000Res.* 2016;5:2042.

doi: 10.12688/f1000research.8659.1

65. Weisshaar E, Gieler U, Kupfer J, *et al*. Questionnaires to assess chronic itch: A consensus paper of the special interest group of the international forum on the study of itch. *Acta Derm Venereol*. 2012;92(5):493-496.

doi: 10.2340/00015555-1402

66. Ständer S, Augustin M, Reich A, *et al*. Pruritus assessment in clinical trials: Consensus recommendations from the international forum for the study of itch (IFSI) special interest group scoring itch in clinical trials. *Acta Derm Venereol*. 2013;93(5):509-514.

doi: 10.2340/00015555-1620

67. Umeda K, Noro Y, Murakami T, *et al*. A novel acoustic evaluation system of scratching in mouse dermatitis: Rapid and specific detection of invisibly rapid scratch in an atopic dermatitis model mouse. *Life Sci*. 2006;79(22):2144-2150.

doi: 10.1016/j.lfs.2006.07.010

68. Noro Y, Omoto Y, Umeda K, *et al*. Novel acoustic evaluation system for scratching behavior in itching dermatitis: Rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients. *J Dermatol*. 2014;41(3):233-238.

doi: 10.1111/1346-8138.12405

69. Lu TC, Fu CM, Ma M, Fang CC, Turner A. Healthcare applications of smart watches. *Appl Clin Inform*. 2016;7(3):850-869.

doi: 10.4338/ACI-2016-03-R-0042

70. Reeder B, David A. Health at hand: A systematic review of smart watch uses for health and wellness. *J Biomed Inform*. 2016;63:269-276.

doi: 10.1016/j.jbi.2016.09.001

71. Lee J, Cho D, Song S, Kim S, Im E, Kim J. Mobile System Design for Scratch Recognition. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*; 2015:1567-1572.

doi: 10.1145/2702613.2732820

72. Ikoma A, Ebata T, Chantalat L, *et al*. Measurement of nocturnal scratching in patients with pruritus using a smartwatch: Initial clinical studies with the itch tracker app. *Acta Derm Venereol*. 2019;99(3):268-273.

doi: 10.2340/00015555-3105

73. Reich A, Szepietowski JC. Pruritus intensity assessment: Challenge for clinicians. *Expert Rev Dermatol*. 2013;8(3): 291-299.

doi: 10.1586/edm.13.25

74. Schneider G, Ständer S, Burgmer M, Driesch G, Heuft G, Weckesser M. Significant differences in central imaging of histamine-induced itch between atopic dermatitis and healthy subjects. *Eur J Pain*. 2008;12(7):834-841.

doi: 10.1016/j.ejpain.2007.12.003

75. Ishiuji Y, Coghill R, Patel T, Oshiro Y, Kraft R, Yosipovitch G. Distinct patterns of brain activity evoked by histamine-

induced itch reveal an association with itch intensity and disease severity in atopic dermatitis. *Br J Dermatol*. 2009;161(5):1072-1080.

doi: 10.1111/j.1365-2133.2009.09308.x

76. Papoiu AD, Nattkemper LA, Sanders KM, *et al*. Brain's reward circuits mediate itch relief. A functional MRI study of active scratching. *PLoS One*. 2013;8(12):e82389.

doi: 10.1371/journal.pone.0082389

77. Mochizuki H, Papoiu AD, Nattkemper LA, *et al*. Scratching induces overactivity in motor-related regions and reward system in chronic itch patients. *J Invest Dermatol*. 2015;135(11):2814-2823.

doi: 10.1038/jid.2015.223

78. Ching T, Himmelstein DS, Beaulieu-Jones BK, *et al*. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15(141):20170387.

doi: 10.1098/rsif.2017.0387

79. Winkler JK, Fink C, Toberer F, *et al*. Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA Dermatol*. 2019;155(10):1135-1141.

doi: 10.1001/jamadermatol.2019.1735

80. Haw WY, Al-Janabi A, Arents BW, *et al*. Global guidelines in dermatology mapping project (GUIDEMAP): A scoping review of dermatology clinical practice guidelines. *Br J Dermatol*. 2021;185(4):736-744.

doi: 10.1111/bjd.20428

81. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology-new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16(11):703-715.

doi: 10.1038/s41571-019-0252-y

82. Niazi MK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol*. 2019;20(5):e253-e261.

doi: 10.1016/S1470-2045(19)30154-8

83. Steele L, Velazquez-Pimentel D, Thomas B. Do AI models recognise rare, aggressive skin cancers? An assessment of a direct-to-consumer application in the diagnosis of Merkel cell carcinoma and amelanotic melanoma. *J Eur Acad Dermatol Venereol*. 2021;35(12):e877-e879.

doi: 10.1111/jdv.17517

84. Li Z, Koban KC, Schenck TL, Giunta RE, Li Q, Sun Y. Artificial intelligence in dermatology image analysis: Current developments and future trends. *J Clin Med*. 2022;11(22):6826.

doi: 10.3390/jcm11226826

85. Kleinberg G, Diaz MJ, Batchu S, Lucke-Wold B. Racial underrepresentation in dermatological datasets leads to biased machine learning models and inequitable healthcare. *J Biomed Res*. 2022;3(1):42-47.

doi: 10.46439/biomedres.3.025

86. Omiye JA, Gui H, Daneshjou R, Cai ZR, Muralidharan, V. Principles, applications, and future of artificial intelligence in dermatology. *Front Med* (*Lausanne*). 2023;10:1278232.

doi: 10.3389/fmed.2023.1278232

87. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):60.

doi: 10.1186/s40537-019-0197-0

88. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In: *NIPS'14: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*; 2014:3320-3328.

89. Settles B. *Active Learning Literature Survey. Computer Sciences Technical Report 1648, University of Wisconsin-Madison*; 2009. Available from: http://digital.library.wic.edu/1793/60660 [Last accessed on 2024 Jan 09].

90. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng*. 2010;22(10):1345-1359.

doi: 10.1109/TKDE.2009.191

91. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett*. 2006;27(8):861-874.

doi: 10.1016/j.patrec.2005.10.010

92. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15(1):1929-1958.

93. Polesie S, Gillstedt M, Kittler H, *et al*. Attitudes towards artificial intelligence within dermatology: An international online survey. *Br J Dermatol*. 2020;183(1):159-161.

doi: 10.1111/bjd.18875

94. Polesie S, McKee PH, Gardner JM, *et al*. Attitudes toward artificial intelligence within dermatopathology: An international online survey. *Front Med* (*Lausanne*). 2020;7:591952.

doi: 10.3389/fmed.2020.591952

95. Nelson CA, Pérez-Chada LM, Creadore A, *et al*. Patient perspectives on the use of artificial intelligence for skin cancer screening: A qualitative study. *JAMA Dermatol*. 2020;156(5):501-512.

doi: 10.1001/jamadermatol.2019.5014

96. Koguchi-Yoshioka H, Watanabe R, Fujisawa Y, *et al*. Skin resident memory T-cell population is not constructed effectively in systemic sclerosis. *Br J Dermatol*. 2019;180(1):219-220.

doi: 10.1111/bjd.17100

97. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset,

a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018;5:180161.

doi: 10.1038/sdata.2018.161

98. ISIC-International Skin Imaging Collaboration; 2018. Available from: https://www.isic-archive.com/#!/topwithheader/widecontenttop/main [Last accessed on 2024 Jan 09].

99. Mendonça T, Ferreira PM, Marques JS, Marcal AR, Rozeira J. PH2-A Dermoscopic Image Database for Research and Benchmarking. In*: 2013 35th Annual International Conference IEEE Engineering Medical Biology Society (EMBC)*. United States: IEEE; 2013:5437-5440.

doi: 10.1109/EMBC.2013.6610779

100. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv*. Preprint posted online 2015.

doi: 10.48550/arXiv.1502.03167

101. Lemley J, Bazrafkan S, Corcoran P. Smart augmentation-learning an optimal data augmentation strategy. *IEEE Access*. 2017;5:5858-5869.

doi: 10.1109/ACCESS.2017.2696121

102. Han SS, Lim W, Kim MS, Park I, Park GH, Chang SE. Interpretation of the outputs of a deep learning model trained with a skin cancer dataset. *J Invest Dermatol*. 2018;138(10):2275-2277.

doi: 10.1016/j.jid.2018.05.014

103. Navarrete-Dechent C, Dusza SW, Liopyris K, Marghoob AA, Halpern AC, Marchetti MA. Automated dermatological diagnosis: Hype or reality? *J Invest Dermatol*. 2018;138(10):2277-2279.

doi: 10.1016/j.jid.2018.04.040

104. Hogarty DT, Mackey DA, Hewitt AW. Current state and future prospects of artificial intelligence in ophthalmology: A review. *Clin Exp Ophthalmol*. 2019;47(1):128-139.

doi: 10.1111/ceo.13381

105. Angus DC. Randomized clinical trials of artificial intelligence. *JAMA*. 2020;323(11):1043.

doi: 10.1001/jama.2020.1039

106. Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care. A report from the national academy of medicine. *JAMA*. 2020;323(6):509-510.

doi: 10.1001/jama.2019.21579

107. Bergl PA, Wijesekera TP, Nassery N, Cosby KS. Controversies in diagnosis: Contemporary debates in the diagnostic safety literature. *Diagnosis (Berl)*. 2020;7(1):3-9.

doi: 10.1515/dx-2019-0016

108. Ferrante G, Licari A, Marseglia GL, La Grutta S. Artificial intelligence as an emerging diagnostic approach in paediatric pulmonology. *Respirology*. 2020;25(10):1029-1030.