



## OPEN ACCESS

## EDITED BY

Mini Han Wang,  
Zhuhai People's Hospital, China

## REVIEWED BY

Ruoyu Zhou,  
City University of Macau, Macao SAR, China  
Shiva Marri,  
Sree Krishna Hospital, India

## \*CORRESPONDENCE

Zisis Kozlakidis  
✉ kozlakidis@who.int

RECEIVED 13 December 2024

ACCEPTED 21 July 2025

PUBLISHED 13 August 2025

## CITATION

Yu J, Cheong IH, Kozlakidis Z and Wang H  
(2025) Advancements and challenges of  
artificial intelligence in dermatology: a review  
of applications and perspectives in China.  
Front. Digit. Health 7:1544520.  
doi: 10.3389/fdgth.2025.1544520

## COPYRIGHT

© 2025 Yu, Cheong, Kozlakidis and Wang. This  
is an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with  
these terms.

# Advancements and challenges of artificial intelligence in dermatology: a review of applications and perspectives in China

Jiaao Yu<sup>1</sup>, Io Hong Cheong<sup>2</sup>, Zisis Kozlakidis<sup>3\*</sup> and Hui Wang<sup>2</sup>

<sup>1</sup>School of Medicine, Shanghai Jiao Tong University, Shanghai, China, <sup>2</sup>School of Public Health, Shanghai Jiao Tong University, Shanghai, China, <sup>3</sup>International Agency for Research on Cancer, World Health Organization, Lyon, France

The diagnosis of skin diseases can be challenging due to their diverse manifestations, while early detection of malignant skin cancers greatly improves the prognosis, highlighting the pressing need for efficient screening methods. In recent years, advancements in AI have paved the way for AI-aided diagnosis of skin lesions. Furthermore, the COVID-19 pandemic has spurred the demand of telemedicine, accelerating the integration of AI into medical domains, particularly in China. This article aims to provide an overview of the progress of AI-aided diagnosis in Chinese dermatology. Given the widespread use of public datasets in the reviewed studies, we compared the performance of AI models in segmentation and classification on public datasets. Despite the promising results of AI in experimental settings, we recognize the limitations of these public datasets in representing clinical scenarios in China. To address this gap, we reviewed the studies that used clinical datasets and conducted comparative analyses between AI and dermatologists. Although AI demonstrated comparable results to human experts, AI still cannot replace dermatologists due to limitations in generalizability and interpretability. We attempt to provide insights into improving the performance of AI through advancements in dataset quality, image pre-processing techniques, and integration of medical data. Finally, the role that AI will play in the medical practice and the relationship between AI and dermatologists are discussed. This systematic review addresses the gap in evaluating AI applications in Chinese dermatology, with a focus on dermatological datasets and real-world application.

## KEYWORDS

artificial intelligence, dermatology, China, AI-aided diagnosis, machine learning

## Introduction

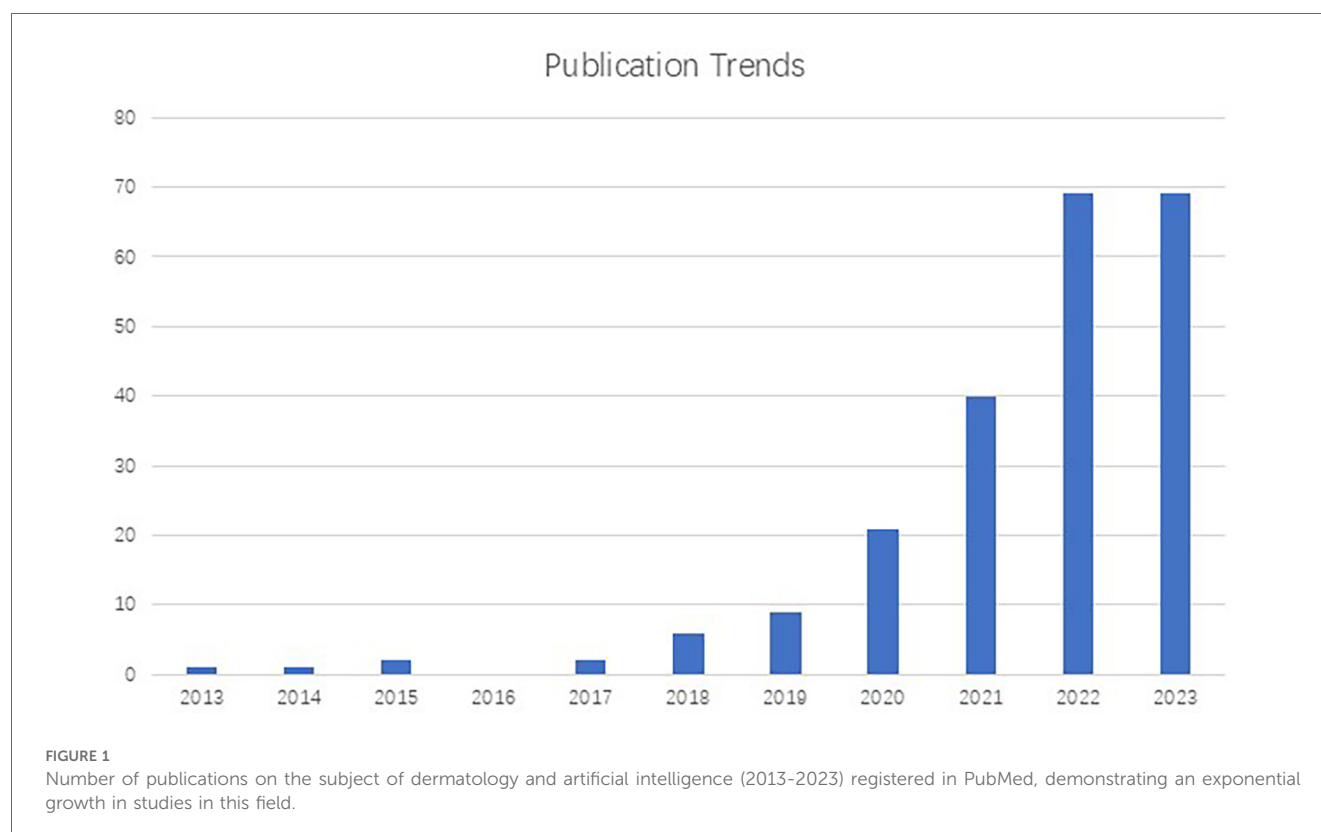
Dermatologists rely heavily on visual features of skin lesions in their diagnostic process, which can sometimes lead to errors, especially for inexperienced clinicians (1). Accurate dermatological diagnosis builds on years of clinical practice. Studies comparing the performance of different levels of Chinese dermatologists showed that the accuracy of dermatologists with high levels of expertise was between 85% to 95%, while the accuracy of dermatology residents ranged from 60% to 70% (2, 3). Because of the rarity of some unique cutaneous diseases and the similarity between different skin lesions, the misdiagnosis happened even in the tertiary care centers where expertise is anticipated to

accrue quicker due to the high volume of cases presented (4–6). What's more, skin diseases were the seventh causes of global disease burden in 2019, especially in the middle Sustainable Development Index (SDI) level countries such as China (7, 8). According to the Global Burden of Disease (GBD), skin and subcutaneous diseases ranked fourth leading cause of disability burden worldwide (9). Analysis based on GBD showed that the burden of disability in China had been growing continuously in the period of 1990 to 2019, with the prevalence of skin disease in the general population increasing by 5.5% in the same period, affecting many millions (8). Therefore, an objective, standardized and efficient assistant tool for dermatologic diagnosis is in need.

Artificial Intelligence (AI) refers to a machine learning-based system that simulates human cognition and can perform tasks typically requiring human intelligence. These tasks include learning from datasets, recognizing patterns, making decisions, understanding natural language, and solving complex problems (10). The application of AI in medical fields is rapidly growing, including but not limited to prediction, diagnosis, treatment, and long-term healthcare (11). Among all, the image diagnosis of AI seemed to be one of the most promising applications, attracting a growing interest from computer scientists and healthcare providers (12). In recent years, an increasing number of studies revealed the potential of AI in promoting the accuracy of skin lesion classification, ranging from malignant skin cancers to inflammatory skin diseases (13). However, although the future looks promising, the use of AI in dermatological diagnosis still faces challenges in everyday clinical practice, including dataset biases, poor interpretability, and ethical issues (14, 15).

Telemedicine, as defined by the World Health Organization (WHO), involves the delivery of health-care services over distance (16). While it cannot fully replace conventional medical care, telemedicine has been proved to be capable of improving medical services in terms of diagnosis and treatment, which can lead to reduced hospitalization rates and duration (17, 18). Besides, telemedicine has the potential of expanding the medical service to remote areas with poor access to high-quality medical facilities (19). Complementing the provision of healthcare in China, telemedicine is progressively emerging as an alternative of conventional medicine in rural areas, being impactful with limited resources (20). Lastly, during the COVID-19 pandemic, conventional face-to-face care was severely affected, which triggered the urgent need of developing and implementing telemedicine (21).

Owing to the visual nature of dermatology, dermatological telemedicine has embraced AI-based image identification as a diagnosis support and screening tool (19). In addition, in the pursuit of improving the classification of skin lesions, AI applications have extended to dermatopathology, medical metadata analysis, prognostic prediction, and biomarker screening (22). However, despite the increasing interest in this domain (Figure 1), few articles focus on the advancement of AI integration in dermatology in China, where the prevalence of skin diseases was 26.0% in 2019 (8). In this paper, we provide an overview of Chinese research on the use of AI in dermatology, focusing on machine learning models, datasets, and clinical applications. In addition, the paper addresses the concerns and challenges about the involvement of AI into dermatological



clinical practice in China. Furthermore, the relationships between AI and dermatologist are explored, with the insights into the future of the role that AI may play in dermatological practice.

## Methodology

This systematic literature review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (23). Databases in English (Pubmed and Web of Science) and databases in Chinese (Wanfang and CNKI) were searched using the keywords (“artificial intelligence” OR “machine learning”) AND (“dermatology” OR “skin lesion”) AND (“image diagnosis” OR “clinical application” OR “segmentation”) AND (“China” OR “Chinese”) for papers from January 2013 to April 2025. This starting point was chosen because the first research in China using AI in dermatology was published in 2013 (24). Exclusion criteria were as follows: (1) conference articles, reviews, or editorials; (2) irrelevant to China; (3) irrelevant to dermatology; (4) without use of AI; (5) no access to full text.

In total, 1394 publications were extracted from these four databases to the EndNote citation manager. 146 duplicates were excluded, leaving 1248 articles. According to the title and abstract screening, 885 articles were removed as non-relevant. Furthermore, 363 articles were reviewed by their full text, of which 136 were excluded because of the following reasons: (1) Review; (2) Topics irrelevant to China; (3) No AI used in the methodology. Finally, a total of 227 articles were selected for the present systematic review (Figure 2). The review of the articles took place by two independent researchers (IOC and YJ) and any contradicting opinions were resolved by a third independent reviewer (ZK and HW). The selected articles were then divided into two groups based on the datasets (public datasets or clinical datasets) used to train the AI model.

## Results

### General study characteristics

In total, 227 articles were reviewed, comprising 25 Chinese and 202 English articles. Among these, 81 focused on single-disease research, while 146 studied multiple skin diseases. Notably, melanoma emerged as the most extensively studied disease in single-disease research ( $n = 46$ ), followed by acne ( $n = 9$ ), psoriasis ( $n = 5$ ), vitiligo ( $n = 3$ ), monkeypox ( $n = 3$ ), actinic keratosis ( $n = 2$ ), and atopic dermatitis ( $n = 2$ ). Only a minority of researches investigating multiple skin diseases studied more than 8 diseases (11/146, 7.5%), with one study analyzing up to 14 diseases. Regarding the objectives of AI models, 132 were designed for classification, 66 for segmentation, and 10 for both segmentation and classification. Additionally, 12 articles aimed at predicting prognosis or treatment effects, 5 focused on screening for biomarkers or medication targets, and 2 were dedicated to image enhancement. According to the dataset used, the studies

were categorized into two groups. The group employing public datasets comprised 151 studies (66.5%), while the group employing clinical datasets included 76 studies (33.5%).

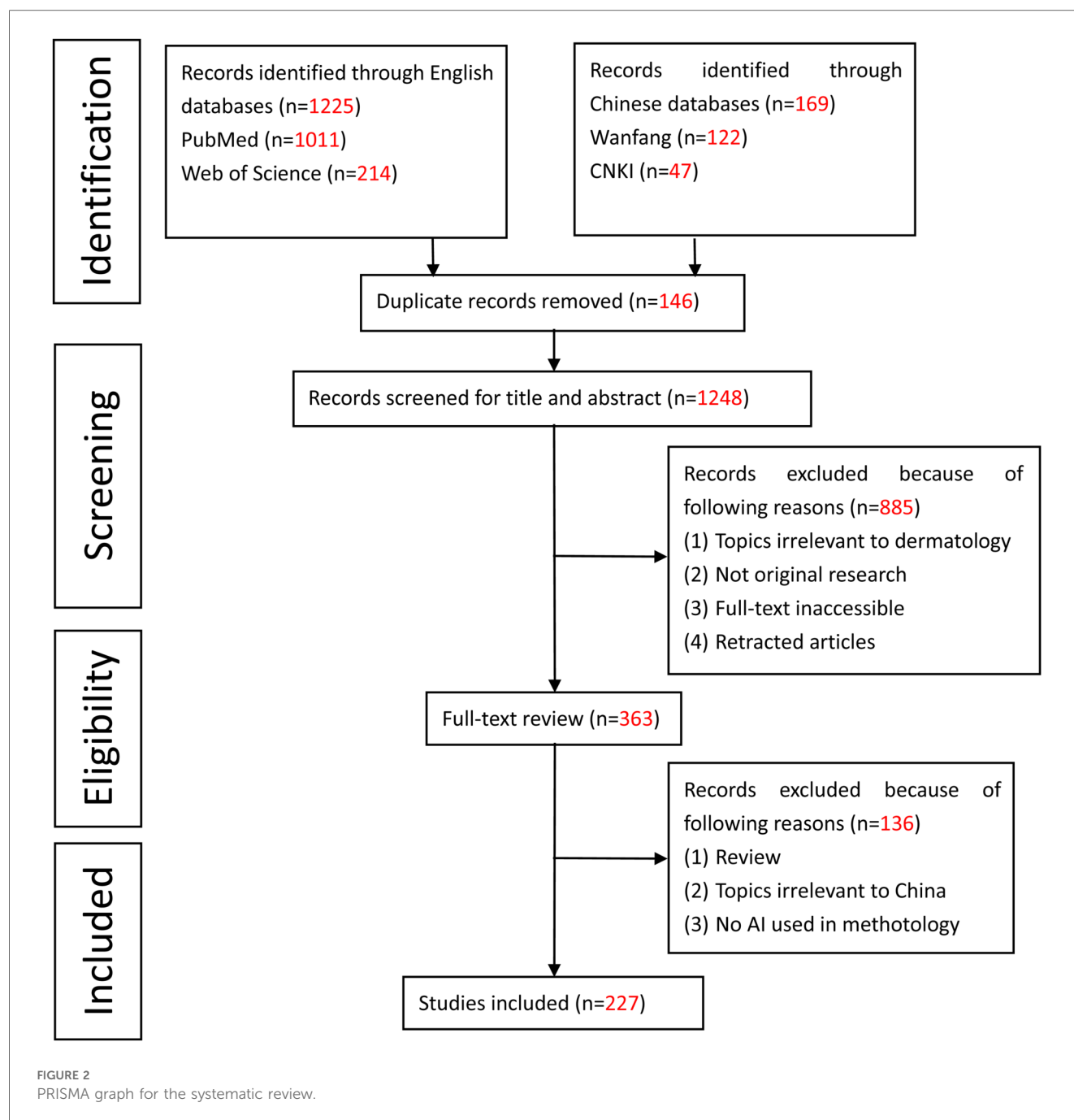
### Performance of AI models in skin lesion segmentation

Image segmentation serves as the foundational and pivotal step in medical image analysis (25). Accurate segmentation is vital for both training AI models and ensuring correct classification (26). According to a previous commentary on medical image segmentation, evaluation metrics employed in this sector include the Dice Similarity Coefficient (DSC), Intersection-over-Union (IoU, also known as Jaccard Index), sensitivity, specificity, and accuracy (27). Given the variance in AI performance across different datasets, a quantitative analysis of different AI models should be conducted on the same dataset (28, 29). ISIC 2018 dataset was chosen due to its more uniform sample distribution, making it suitable for both training and testing AI algorithms (30). Consequently, inclusion criteria for evaluating AI models in segmentation were listed as follows: trained and tested on the ISIC 2018 dataset, encompassing all the 5 evaluation metrics above, and addressing all the 7 diseases in ISIC 2018. Out of 66 articles focused on segmentation, a total of 17 studies met these inclusion criteria.

Among the studies included in this sector, CNN emerged as the most frequently utilized model (13/17, 76.5%), followed by Transformer (4/17, 23.5%) and Attention Network (2/17, 11.8%). Inspired by the structure of the animal visual cortex, CNNs are deep learning algorithms well-suited for processing image data, making them widely used in medical imaging fields such as dermatology, radiology, and pathology (48). A notable study introduced a CNN model with a multi-scale design, incorporating a Spatial Adaptation Module to minimize the loss of spatial location information and a Multi-scale Decoding Fusion Module to integrate information across different layers. Overall, this model design prevented the loss of crucial data during segmentation, resulting in more precise segmentation outcomes (43).

Despite the outstanding performance of CNN models in skin lesion segmentation, traditional CNN models often struggle to effectively capture contextual information (49). In light of this, Bahdanau et al. introduced a neural machine translation model with an attention mechanism, enabling the model to concentrate on relevant parts of the input sentence, thereby enhancing translation quality (50). Building on this concept, Attention Networks were subsequently used in biomedical image segmentation to gather contextual information across both long and short distances (51). Dong Y et al. proposed an Attention Network incorporating a combination of feedback fusion blocks and attention mechanism blocks to aggregate feature mapping. The model exhibited improvements compared to state-of-art models without any data augmentation (32).

Transformer, an algorithm that builds on attention mechanisms, was developed to address the limitations of CNN in modeling sequential data (49). Initially applied to Natural Language Processing (NLP) tasks, Transformer has since been utilized in multi-organ



segmentation, exhibiting competitive results to state-of-the-art methods (52, 53). It is worth noting that Transformer has a better parameter efficiency, which can be particularly advantageous on the devices with limited computational resources (54). He X et al. proposed a Fully Transformer Network, yielding improved sensitivity and specificity. With its linear computational complexity, the model required fewer computing and memory resources, surpassing CNN models in computational efficiency (35). Additionally, Transformer can also work together with CNN, Chen W et al. introduced a fusion module that combines the strengths of Transformer in extracting global information and those of CNN in extracting local information, resulting in superior segmentation outcomes (40). To overcome the limitations of traditional Transformer and CNNs in challenging

lesions, Fan C. et al. applied adaptive spatial-channel attention to reduce Transformer self-attention complexity and preserve spatial-channel interactions, while enhancing local feature extraction via cross-space multiscale attention. Ultimately, their model outperformed CNN/Transformer hybrids in datasets with irregular lesions and noisy backgrounds (47) (Table 1).

## Performance of AI models in multi-disease classification of skin lesions

Previous studies have demonstrated the potential of deep learning algorithms in binary classification of skin lesions,

TABLE 1 A summary table of the performance of AI models in skin lesion segmentation.

Study	Year	Model	DSC	IoU	SEN	SPE	ACC
Lei B et al. (31)	2020	GAN	0.885	0.824	0.953	0.911	0.933
Dong Y et al. (32)	2021	Attention Network	0.912	0.840	0.899	0.981	0.968
Bai R et al. (33)	2022	CNN	0.915	0.853	0.926	0.980	0.976
Dong Y et al. (34)	2022	CNN and Transformer	0.908	0.836	0.906	0.979	0.967
He X et al. (35)	2022	Transformer	0.898	0.828	<b>0.962</b>	0.975	0.966
Hu K et al. (36)	2022	Attention Network	0.896	0.831	0.931	0.947	0.957
Jiang Y et al. (37)	2022	CNN	0.934	0.879	<b>0.962</b>	0.950	0.956
Wang RX et al. (38)	2022	CNN	0.905	0.843	0.911	0.972	0.965
Zhang Z et al. (39)	2022	CNN	0.935	0.882	0.952	0.966	0.959
Chen W et al. (40)	2023	CNN and Transformer	0.908	0.843	0.919	0.973	0.969
Han Q et al. (41)	2023	CNN	0.914	0.859	0.930	0.977	0.963
Jiang Y et al. (42)	2023	CNN	0.935	0.881	0.950	0.968	0.959
Jiang Y et al. (43)	2023	CNN	0.936	0.885	0.950	0.966	0.957
Liu LZ et al. (44)	2023	CNN	0.902	0.832	0.909	0.977	0.944
Yu Z et al. (45)	2023	CNN	0.902	0.836	0.907	0.967	0.967
Zhang W et al. (46)	2023	CNN	0.908	0.835	0.886	0.983	0.961
Fan C. et al. (47)	2025	CNN and Transformer	<b>0.942</b>	<b>0.892</b>	0.953	<b>0.990</b>	<b>0.979</b>

Bold values indicate statistically significant values.  
SEN, sensitivity; SPE, specificity; ACC, accuracy; GAN, Generative Adversarial Network.

TABLE 2 A summary table of the performance of AI models in multi-disease classification of skin lesions.

Study	Year	Model	ACC	SEN	SPE	F1	AUC	TL	EL
Gong A et al. (28)	2020	CNN	0.926	0.484	0.978	0.491	<b>0.918</b>	Y	Y
Qin Z et al. (60)	2020	CNN	<b>0.952</b>	0.743	0.966			Y	N
Iqbal I et al. (57)	2021	CNN	0.888	0.888	0.957			N	N
Shan P et al. (61)	2022	CNN	0.893					Y	N
Yao P et al. (29)	2022	CNN	0.875	0.819	<b>0.980</b>			Y	N
He X et al. (35)	2022	Transformer	0.927	0.857	0.936		0.897	Y	N
Liu Z et al. (62)	2023	CNN	0.873	0.873	<b>0.980</b>			N	N
Wang G et al. (63)	2023	CNN	0.912	<b>0.951</b>		<b>0.889</b>		Y	Y
Wang L et al. (64)	2023	CNN	0.841					Y	N
Yue G et al. (65)	2023	CNN	0.860			0.754	0.872	Y	N
Zhou S et al. (66)	2023	CNN	0.934	0.724	0.925	0.644		Y	N
Hu Z et al. (67)	2024	CNN	0.940	0.917	0.982	0.91	0.993	Y	N
Wan Z et al. (68)	2025	LLM	0.578			0.63		Y	N

Bold values indicate statistically significant values.  
ACC, accuracy; SEN, sensitivity; SPE, specificity; F1, F1-score; TL, transfer learning; EL, ensemble learning.

achieving performance comparable to or superior to that of dermatologists (55, 56). However, these methods are not directly translatable to the multi-class classification requirements of clinical settings (57). To extend the application of AI into real-world scenarios, the ability of AI models in multi-disease classification should be evaluated.

HAM10000 dataset encompasses 7 common skin diseases, including Melanoma, Melanocytic nevus (MN), Basal cell carcinoma (BCC), Actinic keratosis (AK), Benign keratosis (BKL), Dermatofibroma (DF), and Vascular lesion (VASC) (58). The Test 3 dataset of ISIC 2018 includes HAM10000, with only minor changes (59). To assess the efficacy of AI models in multi-disease classification, we reviewed and included 13 articles focusing on multi-classification trained and tested on ISIC 2018 or HAM10000. Classification evaluation metrics such as accuracy, sensitivity, specificity, F1-score, and Area Under the ROC Curve (AUC) were employed. Additionally, the learning strategies adopted in these articles were highlighted (Table 2).

The majority of studies reviewed in this sector adopted transfer learning strategy (11/13, 84.6%). Transfer learning is a machine learning approach that involves reusing an AI model pre-trained for a specific task on a new domain (69). With the advantages such as shorter training time, smaller dataset, and reduced inductive bias, transfer learning is getting increasingly popular in medical fields (70, 71). The application of transfer learning strategy has also extended to AI-CAD of skin lesions. Jain S et al. assessed 6 transfer learning networks on HAM10000 dataset and demonstrated that the Xception Net, pre-trained on the ImageNet dataset, achieved the highest accuracy of 0.905 (72). In China, Qin Z et al. developed a framework using a skin lesion style-based generative adversarial network (GAN) model to generate high quality images, which were then utilized to train another pretrained CNN model (ResNet 50) as the classifier. The image augmentation significantly promoted the performance of CNN model, achieving the highest accuracy of 0.952 among the articles reviewed (60). Instead of using CNN models, He X et al.



adopted the Transformer model. Following segmentation, they further applied their Transformer Network to classification. By pre-training their model on the segmentation dataset, they managed to enhance the classification accuracy from 89.6% to 92.7% (35). In light of the promising potential of large language models (LLMs) in medical fields, Wan Z et al. compared the performance of two LLMs (ChatGPT-4 and LLaVA-1.6) in skin disease identification. Although LLMs exhibited lower accuracy compared to state-of-the-art Transformers, they were less biased across different sex and age groups (68).

Ensemble learning, another machine learning approach, involves combining multiple learning algorithms to improve performance (73). With the potential to ensemble the strength of different AI models, the application of ensemble learning has extended to skin lesion diagnosis (74). For instance, Zillur et al. employed this strategy by combining five deep CNN models to classify skin lesions on ISIC 2019 and HAM10000. They proposed a weighted average ensemble learning model, achieving the recall outperforming the other existing systems (75). Among the 11 studies reviewed, 2 utilized ensemble learning methods. Gong A et al. employed the strategy of maximizing individual advantage and block-integrated voting. By integrating binary classification voting of CNN models, the ensemble model outperformed individual CNN in multi-disease classification tasks. Notably, they also used GAN to create a more balanced training dataset (28). In the other study, Wang G et al. combined two network models for feature fusion and incorporated a multi-receptive field module to capture pathological features. The ensemble network achieved classification performance comparable to that of state-of-art models (63). However, it is important to note that ensemble learning demands significant computing resources and can result in long response time, making it challenging to implement on portable devices (76). To address this issue, Yan P et al. designed a model combining a new loss function and cumulative learning strategy. This non-ensemble model demonstrated performance comparable to that of ensemble models with lower computing burden and less computing time (29).

## Public datasets used in AI training

The preparation of dataset stands as the primary and fundamental step of AI training (77). The quality of the dataset is important to the establishment of AI models, especially to segmentation and classification tasks (78). The appropriate selection of dataset can greatly enhance the ability and generalizability of AI models (79). Among the 144 studies trained on public datasets, ISIC 2018 is the most commonly used dataset (58/144, 40.3%), followed by ISIC 2017 (57/144, 39.6%), ISIC 2016 (35/144, 26.4%), PH2 (33/144, 22.9%), ISIC 2019 (13/144, 39.6%), and HAM 10000 (12/144, 8.3%). Some of the articles used more than one dataset. The ISIC datasets, provided by the International Skin Imaging Collaboration (ISIC), are widely used resources in dermatology imaging research (80). These datasets were established for the public benchmark challenges on dermoscopic image analysis held

TABLE 3 Table of public datasets used in AI training.

Dataset	Year	Data Size	Reference
PH2	2013	40 Melanoma, 160 Nevus	(83)
ISIC 2016	2016	248 Malignant; 1029 Benign	(84)
ISIC 2017	2017	521 Melanoma; 386 seborrheic keratosis; 1,843 benign nevi	(80)
ISIC 2018	2018	2594 for segmentation, 10003 for classification (1110 Melanoma, 6698 MN, 514 BCC, 130 AK, 1097 BKL, 115 DF, 142 VASC, and 197 SCC)	(85)
HAM 10000	2018	1113 Melanoma, 6705 MN, 514 BCC, 327 AK, 1099 BKL, 115 DF, and 142 VASC	(82)
ISIC 2019	2019	4522 Melanoma, 12875 MN, 3323 BCC, 867 AK, 2624 BKL, 239 DF, 253 VASC, 628 Squamous cell carcinoma (SCC)	(86)

annually from 2016 to 2020 (81). The HAM10000 dataset, also known as “Human Against Machine with 10,000 training images,” comprises a comprehensive collection of dermoscopic images, which are included in ISIC 2018 (82). The PH2 dataset, named after the Hospital Pedro Hispano (PH) where it was compiled, consists of dermoscopic primarily used in research related to CAD of melanoma (83) (Table 3). More features of the most frequently used public datasets are listed below:

The feature extraction ability of CNNs builds on large and balanced dataset. However, the challenge of insufficiency and uneven distribution poses a significant obstacle in AI training (28). The size of these public datasets varies greatly, from 200 images for PH2 to 33,569 images for ISIC 2019. Generally, AI models trained in small dataset are more easily challenged by radical bias (56, 87). Furthermore, the distribution of each skin disease within the dataset can impact the performance of AI models. Class imbalance occurs when one class is much more abundant than the other classes, leading to biased model training, misleading performance metrics, and reduced generalization (88). As shown in the table above, all the public datasets encounter the challenge of class imbalance, which may lead to diminished effectiveness (81). To address this issue, Li Z et al. used data augmentation to expand the images of minority classes and weighted random sampling method to avoid oversampling (89). Qian S et al. adopted an adaptive loss-weighted cross-entropy of specific categories to address significant imbalances across different categories of skin lesion images (90). Additionally, Qin Z adopted a skin lesion style-based GAN to augment the data, resulting in an improvement of accuracy from 0.944 to 0.952 (60).

The quality of dataset plays a crucial role in training AI models for skin lesion segmentation and classification (91). Inaccurate data can lead to impaired model performance and poor reproducibility (92). Biases present in datasets finally result in bias in AI models (93). Within public datasets, ground truth data are provided to identify the edges of skin lesions. However, in ISIC 2016, only binary classification of malignant or benign is provided, which offers limited practical value since not all malignant diseases are melanoma (81). Additionally, it has been reported that a number of images in ISIC 2017 were given incorrect ground truth information (94). The presence of pathologic verification in datasets significantly impacts the reliability of AI classification

(95). However, not all of the images in public datasets have corresponding pathologic diagnosis. The rate of pathologic verification of ISIC 2017 was only 26.3%, compared to 53.3% of HAM 10000 (82).

Thus, the generalization of AI models remains a significant concern in the application of AI in medical fields (96). Most of the studies reviewed in this article tested their models on the same dataset where they were trained, which may not reflect the performance in real-world contexts. Models trained in one dataset could show lower performance when tested on another dataset because of the differences in image settings (2). The performance of AI models relies on these details, and even minor deviations have the potential to cause considerable bias (97). Therefore, data pre-processing and image normalization are crucial to reduce the interference of deviations in datasets (28). For example, Tao S et al. converted the dermoscopic images provided by ISIC 2017 dataset from RGB format to HSV format, which is less influenced by the changes in external lighting. As a result, their model trained on ISIC 2017 exhibited great trans-dataset effectiveness when tested on PH2 (98). Despite the progress made by researchers, the application of these AI-CAD tools is still under question, and their performance should be evaluated extensively in real-world scenarios.

# Performance of AI-aided skin lesion classification on clinical datasets

It is worth noting that none of these public datasets mentioned above collect data from Asian countries (99). Previous researches have demonstrated that AI models may not achieve favorable results when transferred to a dataset collected from a population on which the AI models were not trained (100). Consequently, AI models trained on public AI datasets may not adequately address clinical needs in China. In addition to public datasets, many Chinese researchers opt to train and test their models on the dataset developed in local hospitals, collected via dermoscopy, camera, or smartphone. The comparison between AI and clinical physicians is widely discussed and has garnered intensive attention. In 2019, a comparative study was performed between 139 state-of-art machine learning algorithms and 511 dermatologists worldwide, demonstrating that the performance of AI classifiers surpassed that of experts with more than 10 years of experience (101). In this section, we include the articles that (1) were performed on self-developed clinical datasets; (2) focused on skin lesion classification; and those that (3) included a comparison between AI models and dermatologists. Further details are listed in Table 4 below.

Among the 15 studies included, 13 articles used a CNN model, while one article employed Support Vector Machines (SVM) and

TABLE 4 Summary table of the performance of AI-aided skin lesion classification on clinical datasets.

Study	Year	Disease	Algorithm	Type of image	Performance of AI	Performance of dermatologists
Chang W et al. (24)	2013	Melanoma	SVM	Conventional Photographs	ACC = 0.906 SEN = 0.856, SPE = 0.876	ACC = 0.833, SEN = 0.859, SPE = 0.853 (NS)
Xie B et al. (102)	2019	BCC and MN	CNN	Conventional Photographs	ACC = 0.92	ACC = 0.895 (NS)
Huang K et al. (103)	2020	BCC and Seborrheic Keratosis	CNN	Conventional Photographs	ACC = 0.856 SEN = 0.857 SPE = 0.857	TPR = 0.845, FPR = 0.114 (SP)
Li C et al. (104)	2020	11 Skin Diseases	CNN	Dermoscopic and Conventional Photographs	ACC = 0.764	ACC = 0.634 (NS)
Wang S et al. (105)	2020	4 Skin Diseases	CNN	Dermoscopic Image	SEN = 1.000, SPE = 0.605	SEN = 0.872, SPE = 0.838 (NS)
Yang Y et al. (106)	2021	6 PSLs	CNN	Conventional Photographs	SEN = 0.932, SPE = 0.989	SEN = 0.908, SPE = 0.982 (NS)
Yang Y et al. (107)	2021	Psoriasis and other papulosquamous skin diseases	CNN	Dermoscopic Image	SEN = 0.869, SPE = 0.956	SEN = 0.732, SPE = 0.912 (SP)
Zhang L et al. (2)	2021	Vitiligo	CNN	Conventional Photographs	F1 = 0.968, SEN = 0.972, SPE = 0.962	F1 = 0.893, SEN = 0.811, SPE = 0.999 (SP)
Zhao Z et al. (3)	2021	Rosacea and Acne	CNN	Conventional Photographs	ACC = 0.890, PRE = 0.867	ACC = 0.913, PRE = 0.881 (SP)
Zhu C et al. (108)	2021	14 Skin Diseases	CNN	Conventional Photographs	ACC = 0.928, SEN = 0.835, SPE = 0.941	ACC = 0.921, SEN = 0.685, SPE = 0.955 (NS)
Guo L et al. (109)	2021	Vitiligo	CNN	Conventional Photographs	ACC = 88.66%, SEN = 88.00%, SPE = 89.36%	ACC = 92.78%, SEN = 97.33%, SPE = 87.94% (NS)
Ding H et al. (110)	2022	6 PSLs	CNN	Conventional Photographs	PRE = 0.956, SEN = 0.962, SPE = 0.952	PRE = 0.994, SEN = 0.987, SPE = 0.981 (NS)
Ge L et al. (111)	2022	Acne, Rosacea, and Dermatitis	GBM	Dermoscopic Image	ACC = 84.4%	ACC = 35.5% (RD)
Yu Z et al. (112)	2022	Scalp Psoriasis and Seborrheic Dermatitis	CNN	Dermoscopic Image	SEN = 0.961, SPE = 0.882	SEN = 0.745, SPE = 0.882 (SP)
Zhu X et al. (113)	2022	7 Nail Diseases	CNN	Dermoscopic Image	SEN = 0.930, SPE = 0.785	SEN = 0.678, SPE = 0.766 (SP)

ACC, accuracy; SEN, sensitivity; SPE, specificity; F1, F1-score; PRE, precision; TPR, true positive rate; PR, faux positive rate; NS, not specified; SP, senior physician; RD, resident doctor; SVM, support vector machines; GBM, gradient boosting machine.

another used Gradient Boosting Machine (GBM). The skin diseases included in these studies ranged from single-disease classifications such as Melanoma to multiple-disease classifications of up to 14 diseases. 5 studies trained their model only with dermatologic images, while 9 studies trained their models with conventional images. Additionally, one study trained their model with both dermatologic images and conventional images. 6 studies claimed that their models achieved results comparable to those of experienced senior dermatologists. 8 studies claimed that the performance of their models was comparable to those of dermatologists, although the raters were not detailed. One study claimed their models achieved results significantly superior to those of resident doctors.

In 2013, Chang W et al. conducted the first AI-CAD research in China, employing SVM, a machine learning method used for classification and regression tasks, to eliminate the least informative features. The model achieved an accuracy comparable to that of clinicians. However, it was still limited to binary classification of benign or malignant lesions (24). Huang K et al. utilized a large-scale Xiangya Derm dataset established by Xiangya Hospital, consisting of 107,565 clinical images encompassing 541 types of skin diseases (114). Focusing on the binary classification of BCC and SK, they compared the performance of InceptionResNetV2 with that of dermatologists, demonstrating that the model was comparable to the average of 13 general dermatologists (103). Zhu C. et al. conducted a study using a self-developed dataset consisting of 13,603 dermoscopic images covering 14 diseases. They employed a fine-tuned EfficientNet-b4 model and received comparable results compared to 280 dermatologists of all levels on the classification task of 8 diseases (19).

In addition to comparing AI with dermatologists, some researchers are exploring deep learning models as assistants to dermatologists. Shen Y et al. developed a deep learning model to classify 22 common skin diseases, achieving a top-1 accuracy of 45.05%. They further developed an online dermatology diagnosis application based on the JAMA CLEAR dermatology guidelines. The AI diagnostic tool achieved a 63.04% acceptance rate among physicians across 18 tertiary hospitals (115).

In recent years, the application of Vision-Language Models (VLMs) has attracted growing attention in the field of medical diagnosis. The advantage of VLMs lies in their ability to integrate multimodal information from complex medical image data, making diagnoses more accurate and comprehensive (116). To explore the use of VLMs in dermatological diagnosis, Zhou J et al. developed SkinGPT by combining a pretrained vision transformer with an LLM through learning from over 50,000 skin disease images and clinical experiences of dermatologists. When working alongside dermatologists, 80.63% of the diagnoses of SkinGPT were considered useful, indicating that VLM can alleviate the burden on dermatologists by offering rapid suggestions (117).

## Use of AI in dermatology other than skin lesion diagnosis

Due to the shortage of experienced pathologists in fundamental hospitals in China, the pathological diagnosis of rare skin diseases can be challenging (118). AI-based diagnosis has emerged as a

potential solution to this issue (119). Zheng T et al. established a deep reinforcement learning model to detect melanoma cells from whole-slide histopathology images (WSI). Due to the lack of available annotations in WSI datasets, the researchers adopted weakly supervised learning, which does not require detailed annotation. The model finally achieved an accuracy, sensitivity and specificity of 0.966, 0.991 and 0.984, respectively (120). Wu H et al. applied pre-trained CNN models to the diagnosis of Extramammary Paget's disease (EMPD) using pathological images from a self-developed dataset. The model achieved an accuracy, sensitivity and specificity of 0.950, 0.923 and 0.9792, respectively (121). Jiang S et al. extended the AI-assisted pathological diagnosis to the classification of 11 types of skin diseases and achieved an accuracy of 86.8%, outperforming existing CNN models (122).

AI can be employed in predicting the prognosis of skin disease. Xue Y et al. developed a prediction model to estimate the probability of anti-melanoma differentiation associated gene 5 (MDA5) antibody, which is a biomarker associated with unfavorable outcomes in juvenile dermatomyositis (JDM) patients. The model was established using Stepwise logistic regression, least absolute shrinkage and selection operator (LASSO) regression, and random forest (RF) method based on clinical records and auxiliary examinations, achieving an AUC of 0.975 (123). Li W et al. established a SVM model to evaluate the risk types of melanoma using mRNA, miRNA, and DNA methylation data, indicating promising ability to identify different risk subgroups (124). Additionally, AI can be used in predicting the recurrence of skin disease. Cao C et al. introduced an RF model to estimate the recurrence rate of dermatofibrosarcoma protuberans (DFSP) based on MRI images. The model achieved the best concordance index score of 0.875, which is superior to Ki67 index (an independent predictor of recurrence) (125).

Finally, AI can serve as a valuable tool in screening for biomarker of skin diseases or new therapeutic targets. For instance, Liu J et al. developed a machine learning model using cuproptosis-related genes to elucidate their role in the metastasis of melanoma. This model successfully predicted the overall rates of melanoma patients and identified seven key genes, along with 98 potential drugs (126). Song J et al. screened ten pyroptosis-related genes using RF model, revealing the significance of pyroptosis in psoriasis and suggesting potential therapeutic targets. This model demonstrated promising results in external validation, with an AUC of 0.852, and unveiled the involvement of metabolic enhancement and the MAPK signaling pathway (127). Similarly, Xing L et al. introduced RF and LASSO methods to screen for psoriasis biomarkers among 33 differentially expressed methylated genes, identifying GJB2 as the potential target gene for the treatment of psoriasis (128).

## AI application in the market

In recent years, an increasing number of AI-aided medical applications have emerged in the market, particularly in China, and some have successfully made their way into clinical settings (129). Many of these applications are accessible via smartphones, facilitating widespread adoption of AI diagnosis (22). The five



most used, AI-aided skin lesion diagnosis tools in China are shown on the [Table 5](#) below.

AIDERMA, Youzhi AI, and VoxelCloud DermX are capable of detecting multiple skin lesions, covering most of common skin diseases. Among them, VoxelCloud DermX stands out for having the broadest range of skin disease classifications. Most the diagnosis tools above are accessible only to clinical doctors, which does not meet the demand of self-diagnosis by patients. VoxelCloud DermX is open to public and can serve as a screening tool for skin malignant cancer. However, no large-scale research has been performed to evaluate the performance of VoxelCloud DermX in real world scenarios. Following diagnosis, both AIDERMA and VoxelCloud DermX provide continuing education, treatment guidance, and auxiliary consultation ([129](#)). Notably, none of these AI applications have been widely adopted in clinical frontlines, indicating that the use of AI-CAD tools is still in its early stages in China. Further evaluation of reliability is necessary through large-scale clinical trials.

# Discussion

## Challenges of integration of AI into clinical settings

With the increasing interest in telemedicine following the COVID-19 pandemic, self-diagnosis on AI diagnosis platforms has garnered significant attention when conventional medical care was out of reach ([133](#)). However, it's worth noting that most of the frequently used pubic datasets and some of the clinical datasets consist only of dermoscopic images collected and processed exclusively by dermatologists, which may not reflect real-world scenarios. Dermoscopy, a widely used technique in dermatology, provides dermatologists with a detailed view of skin lesions in high resolution, significantly improving diagnostic accuracy ([134](#)). AI models also tend to perform better when trained on dermoscopic images compared to conventional photographs ([104](#)). However, for the dermatologists in restricted-resources areas and general practitioners who have limited access to dermoscopy, using AI-CAD with dermoscopy may not be realistic. If AI tools are only trained and tested with dermoscopic images, their effectiveness on conventional datasets may be questionable. Meanwhile, many clinical datasets use cameras, including smartphones, to capture clinical pictures ([135](#), [136](#)). For example, YOLOv5 model has

shown good performance on untrained datasets consisting of images taken by smartphone ([110](#)). These practices highlight the potential of AI in smart device-based self-diagnosis.

Due to the opacity of the process, deep learning models have long been regarded as a black box, casting doubt on its integration into real-world applications ([137](#)). To address this issue, future AI models should be more transparent and explainable. To improve the interpretability, Wang S et al. introduced the interpretability modules into their multimodal CNN model, providing diagnosis along with the visual analysis for explanations ([138](#)). Huang K et al. performed the evaluation methods to explore the interpretability of their model. They compared the Gradient-weighted Class Activation Mapping (Grad-CAM) and Local Interpretable Model-Agnostic Explanations (LIME), which are able to highlight the regions of image that lead to the decision of CNN models. Their findings demonstrated that their model was interpretable, with LIME showing more accurate results in localizing the responsible region in AI diagnosis compared to Grad-CAM ([103](#)).

Machine learning algorithms are built on data. Therefore, any biases present in the training dataset can result into biases in the model ([139](#)). These biases may stem from underrepresentation of certain patient subgroups within the dataset ([140](#)). For example, most of the AI skin lesion diagnosis models have been trained on datasets composed of individuals with light skin tones. Consequently, these models tend to perform significantly worse when tested by the skin lesion images from individuals with dark skin compared to those with light skin ([119](#)). However, the sharing of data across centers raises concerns regarding data privacy, as these datasets may contain sensitive patient information. One possible solution to this dilemma is federated learning, a machine learning strategy that allows AI models to be trained across different datasets without exchanging the data itself ([13](#)). So far, federated learning has shown promising results in various fields such as radiology and oncology, offering a potential option to address bias while protecting patient privacy ([141](#)).

## Methods to improve the performance of AI in clinical application

The quality of images greatly influences the training and performance of AI models ([142](#)). Datasets often contain images from diverse sources, indicating the necessity of image pre-

TABLE 5 Summary table of AI-aided medical applications in the market, in China.

Name	Year	Disease	Accessibility	Highlight	Reference
AIDERMA	2018	85 skin diseases	Only to doctors	The first comprehensive skin disease diagnosis platform in China, with an accuracy of 86% for 85 skin diseases and 95% for 34 common skin diseases	( <a href="#">129</a> )
Youzhi AI	2018	Multiple skin diseases	Only to doctors	Based on one of the biggest skin lesion datasets in China, with an accuracy of 0.912 in binary classification of benign and malignant skin tumors and 0.814 in multi-classification of skin diseases	( <a href="#">104</a> )
AIDDA	2020	Psoriasis, Eczema, and Atopic Dermatitis	Only to doctors	Trained on 4,740 clinical images, with an overall accuracy, sensitivity and specificity of 0.958, 0.944 and 0.972, respectively	( <a href="#">130</a> ).
VoxelCloud DermX	2020	143 skin diseases	Open to public	The top1, top3, and top5 diagnostic accuracies of the models are 71%, 89%, and 94%, respectively	( <a href="#">131</a> )
SkinTeller	2023	Psoriasis	Only to doctors	Designed to evaluate the Psoriasis Area and Severity Index (PASI), achieving results better than the average of dermatologists	( <a href="#">132</a> )

processing. Image pre-processing involves cropping, rescaling, contrast enhancement, noise reduction, and removal of artifacts. Cropping refers to the process of selecting a square region with the skin lesion in the center to minimize possible interference (143). To enhance the training of CNN models, manual cropping of clinical images was performed by Yang Y et al., resulting in a model with comparable sensitivity and specificity to expert dermatologists (106). Given the variation in image resolutions due to diverse data sources, Iqbal I et al. rescaled the images to 64\*64 pixels with inter nearest interpolation, which also alleviates the computational burden (57). Contrast stretch is employed to enhance contrast between diseased and normal skin regions (144). Due to the electronic noise brought by the digital camera, the interference of environmental factors, and the movement of patients, noise removal is necessary to image pre-processing. Hu L et al. adopted the Wang-Mendel algorithm, a fuzzy logic-based technique, to generate denoised images (145). Moreover, human hair may appear as an interference in skin lesion pictures (146). To remove the hair, Z. Al-Huda et al. used morphology top-hat to create the hair mask, followed by inpainting TELEA algorithm to fill in the masked area (144).

In clinical settings, dermatologists make the diagnosis based on not only the visual characteristics of skin lesions but also medical information such as age, gender, symptoms, and body site (103). However, public datasets often lack or have insufficient medical data, limiting the generalization and interpretability of AI models. To address this gap, metadata-rich clinical datasets should be developed (99). Cai G et al. proposed a model incorporating two encoders: a Transformer model to extract visual features and a Soft Label Encoder to embed medical data. By fusing these features using a Mutual Attention block, they demonstrated that adding metadata improved accuracy from 0.75 to 0.816 on ISIC 2018 compared to models using only images (135). Ou C et al. introduced a CNN model to extract visual features from smartphone images and a multi-layer perceptron model to extract metadata features. They found that including metadata significantly enhanced model performance, increasing accuracy from 0.616 to 0.768 (147). Similarly, Chen Q et al. developed a model combining clinical images and medical data using feature fusion and attention mechanisms. Integration of metadata improved accuracy from 0.716 to 0.804, highlighting the effectiveness of multimodal data fusion strategies in improving AI model performance (136).

Medical data are often high-dimensional and complex, making comprehensive feature extraction from clinical information challenging. The introduction of VLM into the medical field has enabled the integration of textual and visual data (148). For rare disease cases with limited labeled images, VLMs can be trained using few-shot or zero-shot learning approaches (116). In addition, VLMs generate human-readable reasoning for AI-generated diagnoses, making them suitable as health advisors for providing preliminary diagnoses (117). Although our review indicates that the application of VLMs in Chinese dermatology is still in its early stages, VLMs hold great potential as interpretable diagnostic tools in the medical imaging domain.

## AI assisted diagnosis

While there are concerns about AI potentially replacing human dermatologists, our research indicates that AI still cannot entirely replace them due to various reasons such as ethical concerns, limited disease spectrum coverage, technical limitations, insufficient interpretability, and unaddressed biases. The impressive performance of AI in experimental settings using specific high-quality datasets may not accurately reflect its real-world ability (79). However, AI can still serve as an assistant diagnosis tool to alleviate the burden of dermatologists or enhance the diagnostic accuracy of general practitioners who are not specialized in dermatology.

In general, the dermatologists regard the AI more as an assistant tool rather than a potential rival. According to an international survey conducted in 2019, the majority of the dermatologists (77.3%) agreed that AI would bring innovation to dermatology, while only a minority (5.5%) believed human dermatologist would be replaced by AI (149). Another survey including 1,228 dermatologists across China demonstrated that most of the participants (95.36%) believed that AI should serve as an assistant in diagnosis and treatment. Furthermore, a significant proportion (64.17%) of participants thought that AI should first be implemented in secondary hospitals in China (150).

Several studies have already demonstrated that AI could assist clinical doctors to make better decisions in the diagnosis of skin lesions. Instead of the competition between AI and human, Ba W et al. focused on the potential collaboration of dermatologists and CNN models in skin lesion diagnosis. They found that AI-assisted dermatologists achieved a higher accuracy in the classification of multiple skin diseases, with an improvement of accuracy from 0.628 to 0.766. The dermatologists with less than 10 years of experience benefited more from the assistance of AI (151). Another study using CNN models to improve the performance of three unexperienced doctors in the classification of psoriasis and seborrheic dermatitis, achieving an improvement of AUC from an average of 0.571 to 0.805 (112).

## Limitations

Meta-analysis couldn't be conducted due to diverse methodologies across studies and absence of specific evaluation metrics. As we wanted to focus on the application of AI in clinical setting, not all of the algorithms were introduced in detail. We extracted data from Web of Science, pubmed, CNKI, and Wanfang, but no gray literature was included. Clinical trials were not included as no AI-based skin lesion diagnosis tools were known to have undergone clinical trials in China. Certain skin conditions such as burn wounds and diabetic foot ulcers were not included because they were related to clinics and specialties other than dermatology. As this study concentrated on AI use in Chinese dermatology, some articles of international collaborations were excluded if neither the first author nor the corresponding author was Chinese and conducting work in China. The majority of studies reviewed trained their model on public dataset.

However, as most of the public datasets are not China-centric, these studies may not specifically reflect the application of AI in China. To address this discrepancy, we also included and evaluated diagnostic tools trained in Chinese clinical datasets, but these datasets often lack clarity and accessibility.

## Conclusion and prospective

Our study offers an overview of the ongoing progress in AI within the field of dermatology in China. While AI has demonstrated comparable performance to human experts in diagnosing multiple skin diseases, its clinical implementation remains at an early stage due to limitations in generalizability and interpretability. Large-scale clinical trials are warranted to further validate its efficacy in real world. However, the potential of AI in dermatology is promising. Given China's shortage of medical resources and the significant clinical burdens faced by physicians, AI technology is viewed as a solution to address healthcare disparities in remote areas. As a useful assistant to physicians, AI has the potential to enhance the quality of medical care, reduce costs, and alleviate the burden of dermatologists.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## Author contributions

JY: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. IC: Data curation, Formal analysis, Supervision, Visualization, Writing – original draft. ZK: Conceptualization, Project administration, Supervision, Validation, Writing – review & editing. HW: Funding acquisition, Methodology, Project administration, Resources, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. JY has been

supported in part of this work through an exchange program jointly funded by the Chinese Scholarship Council (CSC) and Cancéropôle Lyon Auvergne-Rhône-Alpes (CLARA), with the support of Shanghai Jiao Tong University.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Author disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer/WHO, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/WHO.

## References

- Ko CJ, Braverman I, Sidlow R, Lowenstein EJ. Visual perception, cognition, and error in dermatologic diagnosis: key cognitive principles. *J Am Acad Dermatol*. (2019) 81(6):1227–34. doi: 10.1016/j.jaad.2018.10.082
- Zhang L, Mishra S, Zhang T, Zhang Y, Zhang D, Lv Y, et al. Design and assessment of convolutional neural network based methods for vitiligo diagnosis. *Front Med (Lausanne)*. (2021) 8:754202. doi: 10.3389/fmed.2021.754202
- Zhao Z, Wu CM, Zhang S, He F, Liu F, Wang B, et al. A novel convolutional neural network for the diagnosis and classification of Rosacea: usability study. *JMIR Med Inform*. (2021) 9(3):e23415. doi: 10.2196/23415
- Onsoi W, Chaiyapit J, Techasatian L. Common misdiagnoses and prevalence of dermatological disorders at a pediatric tertiary care center. *J Int Med Res*. (2019) 48(2):0300060519873490. doi: 10.1177/0300060519873490
- Sikorska K, Gesing M, Olszański R, Roszko-Wysokińska A, Szostakowska B, Van Damme-Ostapowicz K. Misdiagnosis and inappropriate treatment of cutaneous leishmaniasis: a case report. *Trop Dis Travel Med Vaccines*. (2022) 8(1):18. doi: 10.1186/s40794-022-00175-5
- Urbancek S, Fedorcova P, Tomkova J, Sutka R. Misdiagnosis of melanoma: a 7 year single-center analysis. *Pigment Disord*. (2015) 2:208. doi: 10.4172/2376-0427.1000208

7. Yakupu A, Aimaier R, Yuan B, Chen B, Cheng J, Zhao Y, et al. The burden of skin and subcutaneous diseases: findings from the global burden of disease study 2019. *Front Public Health*. (2023) 11:1145513. doi: 10.3389/fpubh.2023.1145513
8. Peng D, Sun J, Wang J, Qi X, Li G. Burden of skin disease - China, 1990–2019. *China CDC wkly*. (2021) 3(22):472–5. doi: 10.46234/ccdcw2021.123
9. Chu S, Mehral S, Uppal P, Giesey RL, Delost ME, Delost GR. Burden of skin disease and associated socioeconomic status in Europe: an ecologic study from the global burden of disease study 2017. *JAAD Int*. (2020) 1(2):95–103. doi: 10.1016/j.jdin.2020.07.001
10. Collins C, Dennehy D, Conboy K, Mikalef P. Artificial intelligence in information systems research: a systematic literature review and research agenda. *Int J Inf Manage*. (2021) 60:102383. doi: 10.1016/j.jinfomgt.2021.102383
11. Haleem A, Javaid M, Khan IH. Current status and applications of artificial intelligence (AI) in medical field: an overview. *Curr Med Res Pract*. (2019) 9(6):231–7. doi: 10.1016/j.cmrp.2019.11.005
12. Pandey B, Kumar Pandey D, Pratap Mishra B, Rhmann W. A comprehensive survey of deep learning in the field of medical imaging and medical natural language processing: challenges and research directions. *J King Saud Univ Comput Inf Sci*. (2022) 34(8, Part A):5083–99. doi: 10.1016/j.jksuci.2021.01.007
13. 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
14. Liopyris K, Gregoriou S, Dias J, Stratigos AJ. Artificial intelligence in dermatology: challenges and perspectives. *Dermatol Ther (Heidelb)*. (2022) 12(12):2637–51. doi: 10.1007/s13555-022-00833-8
15. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov IV. Artificial intelligence applications in dermatology: where do we stand? *Front Med (Lausanne)*. (2020) 7:100. doi: 10.3389/fmed.2020.00100
16. *Consolidated Telemedicine Implementation Guide*. Geneva: World Health Organization (2022). Licence: CC BY-NC-SA 3.0 IGO.
17. Peters GM, Kooij L, Lenferink A, van Harten WH, Doggen CJM. The effect of telehealth on hospital services use: systematic review and meta-analysis. *J Med Internet Res*. (2021) 23(9):e25195. doi: 10.2196/25195
18. Lyth J, Lind L, Persson HL, Wihren AB. Can a telemonitoring system lead to decreased hospitalization in elderly patients? *J Telemed Telecare*. (2021) 27(1):46–53. doi: 10.1177/1357633X19858178
19. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. *Sens Int*. (2021) 2:100117. doi: 10.1016/j.sintl.2021.100117
20. Gao J, Fan C, Chen B, Fan Z, Li L, Wang L, et al. Telemedicine is becoming an increasingly popular way to resolve the unequal distribution of healthcare resources: evidence from China. *Front Public Health*. (2022) 10:916303. doi: 10.3389/fpubh.2022.916303
21. Luo JZ, Yang ZY, Xie Y, He Y, Wu MM, Fang X, et al. Emerging trends in teledermatology research: a scientometric analysis from 2002 to 2021. *Telemed E-Health*. (2023) 30:393–403. doi: 10.1089/tmj.2023.0101
22. 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
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J*. (2021) 372:n71. doi: 10.1136/bmj.n71
24. Chang WY, Huang A, Yang CY, Lee CH, Chen YC, Wu TY, et al. Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study. *PLoS One*. (2013) 8(11):e76212. doi: 10.1371/journal.pone.0076212
25. Vesal S, Ravikumar N, Maier A. Skinnet: a deep learning framework for skin lesion segmentation. 2018 *IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*; 2018 10–17 Nov. 2018
26. Burdick J, Marques O, Weinthal J, Furth B. Rethinking skin lesion segmentation in a convolutional classifier. *J Digit Imaging*. (2018) 31(4):435–40. doi: 10.1007/s10278-017-0026-y
27. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res Notes*. (2022) 15(1):210. doi: 10.1186/s13104-022-06096-y
28. Gong A, Yao XJ, Lin W. Classification for dermoscopy images using convolutional neural networks based on the ensemble of individual advantage and group decision. *IEEE Access*. (2020) 8:155337–51. doi: 10.1109/ACCESS.2020.3019210
29. Yao P, Shen SW, Xu MJ, Liu P, Zhang F, Xing JY, et al. Single model deep learning on imbalanced small datasets for skin lesion classification. *IEEE Trans Med Imaging*. (2022) 41(5):1242–54. doi: 10.1109/TMI.2021.3136682
30. Zhao C, Shuai RJ, Ma L, Liu WJ, Hu D, Wu ML. Dermoscopy image classification based on StyleGAN and DenseNet201. *IEEE Access*. (2021) 9:8659–79. doi: 10.1109/ACCESS.2021.3049600
31. Lei B, Xia Z, Jiang F, Jiang X, Ge Z, Xu Y, et al. Skin lesion segmentation via generative adversarial networks with dual discriminators. *Med Image Anal*. (2020) 64:101716. doi: 10.1016/j.media.2020.101716
32. Dong Y, Wang L, Cheng S, Li Y. FAC-Net: feedback attention network based on context encoder network for skin lesion segmentation. *Sensors (Basel)*. (2021) 21(15):5172. doi: 10.3390/s21155172
33. Bai R, Zhou M. SL-HarDNet: skin lesion segmentation with HarDNet. *Front Bioeng Biotechnol*. (2022) 10:1028690. doi: 10.3389/fbioe.2022.1028690
34. Dong Y, Wang L, Li Y. TC-Net: dual coding network of transformer and CNN for skin lesion segmentation. *PLoS One*. (2022) 17(11):e0277578. doi: 10.1371/journal.pone.0277578
35. He X, Tan EL, Bi H, Zhang X, Zhao S, Lei B. Fully transformer network for skin lesion analysis. *Med Image Anal*. (2022) 77:102357. doi: 10.1016/j.media.2022.102357
36. Hu K, Lu J, Lee DJ, Xiong DP, Chen ZE. AS-Net: attention synergy network for skin lesion segmentation. *Expert Syst Appl*. (2022) 201:117112. doi: 10.1016/j.eswa.2022.117112
37. Jiang Y, Cheng T, Dong J, Liang J, Zhang Y, Lin X, et al. Dermoscopic image segmentation based on pyramid residual attention module. *PLoS One*. (2022) 17(9):e0267380. doi: 10.1371/journal.pone.0267380
38. Wang RX, Chen SY, Ji CJ, Li Y. Cascaded context enhancement network for automatic skin lesion segmentation. *Expert Syst Appl*. (2022) 201:117069. doi: 10.1016/j.eswa.2022.117069
39. Zhang Z, Jiang Y, Qiao H, Wang M, Yan W, Chen J. SIL-Net: a semi-isotropic L-shaped network for dermoscopic image segmentation. *Comput Biol Med*. (2022) 150:106146. doi: 10.1016/j.combiomed.2022.106146
40. Chen W, Zhang R, Zhang Y, Bao F, Lv H, Li L, et al. Pact-Net: parallel CNNs and transformers for medical image segmentation. *Comput Methods Programs Biomed*. (2023) 242:107782. doi: 10.1016/j.cmpb.2023.107782
41. Han Q, Wang H, Hou M, Weng T, Pei Y, Li Z, et al. HWA-SegNet: multi-channel skin lesion image segmentation network with hierarchical analysis and weight adjustment. *Comput Biol Med*. (2023) 152:106343. doi: 10.1016/j.combiomed.2022.106343
42. Jiang Y, Dong J, Zhang Y, Cheng T, Lin X, Liang J. PCF-Net: position and context information fusion attention convolutional neural network for skin lesion segmentation. *Heliyon*. (2023) 9(3):e13942. doi: 10.1016/j.heliyon.2023.e13942
43. Jiang Y, Qiao H, Zhang ZQ, Wang MQ, Yan W, Chen J. MDSC-Net: a multi-scale depthwise separable convolutional neural network for skin lesion segmentation. *Let Image Processing*. (2023) 17:3713–27. doi: 10.1049/ipr.2.12892
44. Liu LZ, Zhang XX, Li YW, Xu ZA. An improved multi-scale feature fusion for skin lesion segmentation. *Appl Sci Basel*. (2023) 13(14):8512. doi: 10.3390/app13148512
45. Yu Z, Yu L, Zheng W, Wang S. EIU-Net: enhanced feature extraction and improved skip connections in U-net for skin lesion segmentation. *Comput Biol Med*. (2023) 162:107081. doi: 10.1016/j.combiomed.2023.107081
46. Zhang W, Lu F, Zhao W, Hu Y, Su H, Yuan M. ACCPG-Net: a skin lesion segmentation network with adaptive channel-context-aware pyramid attention and global feature fusion. *Comput Biol Med*. (2023) 154:106580. doi: 10.1016/j.combiomed.2023.106580
47. Fan C, Zhu Z, Peng B, Xuan Z, Zhu X. EAAC-Net: an efficient adaptive attention and convolution fusion network for skin lesion segmentation. *J Imaging Inform Med*. (2025) 38(2):1120–36. doi: 10.1007/s10278-024-01223-6
48. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med*. (2019) 25(1):24–9. doi: 10.1038/s41591-018-0316-z
49. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Adv Neural Inf Process Syst*. (2017) 30:1–11. doi: 10.48550/arXiv.1706.03762
50. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. *arXiv [Preprint]*. arXiv:14090473 (2014).
51. Cheng J, Tian S, Yu L, Lu H, Lv X. Fully convolutional attention network for biomedical image segmentation. *Artif Intell Med*. (2020) 107:101899. doi: 10.1016/j.artmed.2020.101899
52. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, et al. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv [Preprint]*. arXiv:210204306 (2021).
53. Strudel R, Garcia R, Laptev I, Schmid C. Segmenter: transformer for semantic segmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021).
54. Lin Y, Wang M, Zhang Z, Wang X, Xiao T, Zhu J. Understanding Parameter Sharing in Transformers. *arXiv [Preprint]*. arXiv:230609380 (2023).
55. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. (2017) 542(7639):115–8. doi: 10.1038/nature21056
56. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. *Diagnostics (Basel)*. (2021) 11(8):1390. doi: 10.3390/diagnostics11081390
57. Iqbal I, Younus M, Walayat K, Kakar MU, Ma J. Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. *Comput Med Imaging Graph*. (2021) 88:101843. doi: 10.1016/j.compmedimag.2020.101843



58. Yao P, Shen S, Xu M, Liu P, Zhang F, Xing J, et al. Single model deep learning on imbalanced small datasets for skin lesion classification. *IEEE Trans Med Imaging*. (2021) 41(5):1242–54. doi: 10.1109/TMI.2021.3136682
59. Lan Z, Cai S, He X, Wen X. Fixcaps: an improved capsules network for diagnosis of skin cancer. *IEEE Access*. (2022) 10:76261–7. doi: 10.1109/ACCESS.2022.3181225
60. Qin Z, Liu Z, Zhu P, Xue Y. A GAN-based image synthesis method for skin lesion classification. *Comput Methods Programs Biomed*. (2020) 195:105568. doi: 10.1016/j.cmpb.2020.105568
61. Shan P, Fu C, Dai L, Jia T, Tie M, Liu J. Automatic skin lesion classification using a new densely connected convolutional network with an SF module. *Med Biol Eng Comput*. (2022) 60(8):2173–88. doi: 10.1007/s11517-022-02583-3
62. Liu ZH, Xiong RQ, Jiang TT. CI-Net: clinical-inspired network for automated skin lesion recognition. *IEEE Trans Med Imaging*. (2023) 42(3):619–32. doi: 10.1109/TMI.2022.3215547
63. Wang G, Yan P, Tang Q, Yang L, Chen J. Multiscale feature fusion for skin lesion classification. *Biomed Res Int*. (2023) 2023:5146543. doi: 10.1155/2023/5146543
64. Wang L, Zhang L, Shu X, Yi Z. Intra-class consistency and inter-class discrimination feature learning for automatic skin lesion classification. *Med Image Anal*. (2023) 85:102746. doi: 10.1016/j.media.2023.102746
65. Yue GH, Wei PS, Zhou TW, Jiang QP, Yan WQ, Wang TF. Toward multicenter skin lesion classification using deep neural network with adaptively weighted balance loss. *IEEE Trans Med Imaging*. (2023) 42(1):119–31. doi: 10.1109/TMI.2022.3204646
66. Zhou S, Tian S, Yu L, Wu W, Zhang D, Peng Z, et al. ReFixMatch-LS: reusing pseudo-labels for semi-supervised skin lesion classification. *Med Biol Eng Comput*. (2023) 61(5):1033–45. doi: 10.1007/s11517-022-02743-5
67. Hu Z, Mei W, Chen H, Hou W. Multi-scale feature fusion and class weight loss for skin lesion classification. *Comput Biol Med*. (2024) 176:108594. doi: 10.1016/j.compbimed.2024.108594
68. Wan Z, Guo Y, Bao S, Wang Q, Malin BA. Evaluating sex and age biases in multimodal large language models for skin disease identification from dermatoscopic images. *Health Data Sci*. (2025) 5:0256. doi: 10.34133/hds.0256
69. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. *Proc IEEE*. (2021) 109(1):43–76. doi: 10.1109/JPROC.2020.3004555
70. Matsoukas C, Haslum JF, Sorkhei M, Söderberg M, Smith K. What makes transfer learning work for medical images: feature reuse & other factors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022).
71. Ayana G, Dese K, Choe S. Transfer learning in breast cancer diagnoses via ultrasound imaging. *Cancers (Basel)*. (2021) 13(4):738. doi: 10.3390/cancers13040738
72. Jain S, Singhania U, Tripathy B, Nasr EA, Aboudaif MK, Kamrani AK. Deep learning-based transfer learning for classification of skin cancer. *Sensors (Basel)*. (2021) 21(23):8142. doi: 10.3390/s21238142
73. Pintelas P, Livieris IE. Special issue on ensemble learning and applications. *Algorithms*. (2020) 13(6):140. doi: 10.3390/a13060140
74. Aboulmira A, Raouhi EM, Hrimch H, Lachgar M. Ensemble learning methods for deep learning: application to skin lesions classification. *2022 11th International Symposium on Signal, Image, Video and Communications (ISIVC)*; 2022 18–20 May 2022
75. Rahman Z, Hossain MS, Islam MR, Hasan MM, Hridhee RA. An approach for multiclass skin lesion classification based on ensemble learning. *Inform Med Unlocked*. (2021) 25:100659. doi: 10.1016/j.imu.2021.100659
76. Mahbod A, Schaefer G, Wang C, Dorffner G, Ecker R, Ellinger I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Comput Methods Programs Biomed*. (2020) 193:105475. doi: 10.1016/j.cmpb.2020.105475
77. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. *Radiology*. (2020) 295(1):4–15. doi: 10.1148/radiol.2020192224
78. Torras-Casas A, Paluzo-Hidalgo E, Gonzalez-Diaz R. A Topological Approach to Measuring Training Data Quality. *arXiv:230602411* (2023).
79. Liang W, Tadesse GA, Ho D, Fei-Fei L, Zaharia M, Zhang C, et al. Advances, challenges and opportunities in creating data for trustworthy AI. *Nat Mach Intell*. (2022) 4(8):669–77. doi: 10.1038/s42256-022-00516-1
80. Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*; 2018 4–7 April 2018
81. Cassidy B, Kendrick C, Brodzicki A, Jaworek-Korjakowska J, Yap MH. Analysis of the ISIC image datasets: usage, benchmarks and recommendations. *Med Image Anal*. (2022) 75:102305. doi: 10.1016/j.media.2021.102305
82. 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(1):180161. doi: 10.1038/sdata.2018.161
83. Mendonça T, Ferreira PM, Marques JS, Marcal AR, Rozeira J, editors. PH 2-A Dermoscopic image database for research and benchmarking. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; IEEE (2013).
84. Gutman D, Codella NC, Celebi E, Helba B, Marchetti M, Mishra N, et al. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). *arXiv [Preprint]*. *arXiv:160501397* (2016).
85. Codella N, Rotemberg V, Tschandl P, Celebi ME, Dusza S, Gutman D, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv [Preprint]*. *arXiv:190203368* (2019).
86. Gessert N, Nielsen M, Shaikh M, Werner R, Schlaefer A. Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. *MethodsX*. (2020) 7:100864. doi: 10.1016/j.mex.2020.100864
87. Sun C, Shrivastava A, Singh S, Gupta A. Revisiting unreasonable effectiveness of data in deep learning era. *Proceedings of the IEEE International Conference on Computer Vision* (2017).
88. Gao L, Zhang L, Liu C, Wu S. Handling imbalanced medical image data: a deep-learning-based one-class classification approach. *Artif Intell Med*. (2020) 108:101935. doi: 10.1016/j.artmed.2020.101935
89. Li ZY, Chen ZQ, Che XX, Wu YG, Huang D, Ma HY, et al. A classification method for multi-class skin damage images combining quantum computing and inception-ResNet-V1. *Front Phys*. (2022) 10:107166. doi: 10.3389/fphy.2022.1046314
90. Qian S, Ren K, Zhang W, Ning H. Skin lesion classification using CNNs with grouping of multi-scale attention and class-specific loss weighting. *Comput Methods Programs Biomed*. (2022) 226:107166. doi: 10.1016/j.cmpb.2022.107166
91. Bustillo A, Reis R, Machado AR, Pimenov DY. Improving the accuracy of machine-learning models with data from machine test repetitions. *J Intell Manuf*. (2022) 33(1):203–21. doi: 10.1007/s10845-020-01661-3
92. Alves VM, Auerbach SS, Kleinstreuer N, Rooney JP, Muratov EN, Rusyn I, et al. Curated data in - trustworthy in silico models out: the impact of data quality on the reliability of artificial intelligence models as alternatives to animal testing. *Altern Lab Anim*. (2021) 49(3):73–82. doi: 10.1177/02611929211029635
93. Bissoto A, Fornaciali M, Valle E, Avila S, editors. (De) constructing bias on skin lesion datasets. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; 2019 16–17 June 2019.
94. Shan P, Wang Y, Fu C, Song W, Chen J. Automatic skin lesion segmentation based on FC-DPN. *Comput Biol Med*. (2020) 123:103762. doi: 10.1016/j.compbimed.2020.103762
95. Schuster T, Shah DJ, Yeo YJS, Filizola D, Santus E, Barzilay R. Towards debiasing fact verification models. *arXiv [Preprint]*. *arXiv:190805267* (2019).
96. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*. (2021) 27(12):2176–82. doi: 10.1038/s41591-021-01595-0
97. Subbaswamy A, Saria S. From development to deployment: dataset shift, causality, and shift-stable models in health AI. *Biostatistics*. (2019) 21(2):345–52. doi: 10.1093/biostatistics/kxz041
98. Tao S, Jiang Y, Cao S, Wu C, Ma Z. Attention-guided network with densely connected convolution for skin lesion segmentation. *Sensors (Basel)*. (2021) 21(10):3462. doi: 10.3390/s21103462
99. Wen D, Khan SM, Ji Xu A, Ibrahim H, Smith L, Caballero J, et al. Characteristics of publicly available skin cancer image datasets: a systematic review. *Lancet Digital Health*. (2022) 4(1):e64–74. doi: 10.1016/S2589-7500(21)00252-1
100. Goyal M, Knackstedt T, Yan S, Hassanpour S. Artificial intelligence-based image classification methods for diagnosis of skin cancer: challenges and opportunities. *Comput Biol Med*. (2020) 127:104065. doi: 10.1016/j.compbimed.2020.104065
101. Tschandl P, Codella N, Akay BN, Argenziano G, Braun RP, Cabo H, et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *Lancet Oncol*. (2019) 20(7):938–47. doi: 10.1016/S1470-2045(19)30333-X
102. Xie B, He X, Huang W, Shen M, Li F, Zhao S. Clinical image identification of basal cell carcinoma and pigmented nevi based on convolutional neural network. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*. (2019) 44(9):1063–70. doi: 10.11817/j.issn.1672-7347.2019.190205
103. Huang K, He X, Jin Z, Wu L, Zhao X, Wu Z, et al. Assistant diagnosis of basal cell carcinoma and seborrheic keratosis in Chinese population using convolutional neural network. *J Healthc Eng*. (2020) 2020:1713904. doi: 10.1155/2020/1713904
104. Li CX, Fei WM, Shen CB, Wang ZY, Jing Y, Meng RS, et al. Diagnostic capacity of skin tumor artificial intelligence-assisted decision-making software in real-world clinical settings. *Chin Med J (Engl)*. (2020) 133(17):2020–6. doi: 10.1097/CM9.0000000000001002
105. Wang SQ, Zhang XY, Liu J, Tao C, Zhu CY, Shu C, et al. Deep learning-based, computer-aided classifier developed with dermoscopic images shows comparable performance to 164 dermatologists in cutaneous disease diagnosis in the Chinese



- population. *Chin Med J.* (2020) 133(17):2027–36. doi: 10.1097/CM9.0000000000001023
106. Yang Y, Ge Y, Guo L, Wu Q, Peng L, Zhang E, et al. Development and validation of two artificial intelligence models for diagnosing benign, pigmented facial skin lesions. *Skin Res Technol.* (2021) 27(1):74–9. doi: 10.1111/srt.12911
107. Yang Y, Wang J, Xie F, Liu J, Shu C, Wang Y, et al. A convolutional neural network trained with dermoscopic images of psoriasis performed on par with 230 dermatologists. *Comput Biol Med.* (2021) 139:104924. doi: 10.1016/j.compbiomed.2021.104924
108. Zhu CY, Wang YK, Chen HP, Gao KL, Shu C, Wang JC, et al. A deep learning based framework for diagnosing multiple skin diseases in a clinical environment. *Front Med (Lausanne).* (2021) 8:626369. doi: 10.3389/fmed.2021.626369
109. Guo LF, Ge YP, Yang Y, Lin T. Establishment and evaluation of an artificial intelligence model for the diagnosis of facial vitiligo. *Chin J Dermatol.* (2021) 54(7):586–9.
110. Ding H, Zhang E, Fang F, Liu X, Zheng H, Yang H, et al. Automatic identification of benign pigmented skin lesions from clinical images using deep convolutional neural network. *BMC Biotechnol.* (2022) 22(1):28. doi: 10.1186/s12896-022-00755-5
111. Ge L, Li Y, Wu Y, Fan Z, Song Z. Differential diagnosis of Rosacea using machine learning and dermoscopy. *Clin Cosmet Investig Dermatol.* (2022) 15:1465–73. doi: 10.2147/CCID.S373534
112. Yu Z, Kaizhi S, Jianwen H, Guanyu Y, Yonggang W. A deep learning-based approach toward differentiating scalp psoriasis and seborrheic dermatitis from dermoscopic images. *Front Med (Lausanne).* (2022) 9:965423. doi: 10.3389/fmed.2022.965423
113. Zhu X, Zheng B, Cai W, Zhang J, Lu S, Li X, et al. Deep learning-based diagnosis models for onychomycosis in dermoscopy. *Mycoses.* (2022) 65(4):466–72. doi: 10.1111/myc.13427
114. Xie B, He X, Zhao S, Li Y, Su J, Zhao X, et al. Xiangyaderm: a clinical image dataset of Asian race for skin disease aided diagnosis. *Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention: International Workshops, LABELS 2019, HAL-MICCAI 2019, and CuRIOUS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings*; Shenzhen, China: Springer-Verlag (2019). p. 22–31
115. Shen Y, Li H, Sun C, Ji H, Zhang D, Hu K, et al. Optimizing skin disease diagnosis: harnessing online community data with contrastive learning and clustering techniques. *NPJ Digit Med.* (2024) 7(1):28. doi: 10.1038/s41746-024-01014-x
116. Li X, Li L, Jiang Y, Wang H, Qiao X, Feng T, et al. Vision-language models in medical image analysis: from simple fusion to general large models. *Information Fusion.* (2025) 118:102995. doi: 10.1016/j.inffus.2025.102995
117. Zhou J, He X, Sun L, Xu J, Chen X, Chu Y, et al. Pre-trained multimodal large language model enhances dermatological diagnosis using SkinGPT-4. *Nat Commun.* (2024) 15(1):5649. doi: 10.1038/s41467-024-50043-3
118. Xu C, Li Y, Chen P, Pan M, Bu X. A survey on the attitudes of Chinese medical students towards current pathology education. *BMC Med Educ.* (2020) 20(1):259. doi: 10.1186/s12909-020-02167-5
119. Daneshjouri R, Vodrahalli K, Novoa RA, Jenkins M, Liang W, Rotemberg V, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. *Sci Adv.* (2022) 8(32):eabq6147. doi: 10.1126/sciadv.abq6147
120. Zheng T, Chen W, Li S, Quan H, Zou M, Zheng S, et al. Learning how to detect: a deep reinforcement learning method for whole-slide melanoma histopathology images. *Comput Med Imaging Graph.* (2023) 108:102275. doi: 10.1016/j.compmedimag.2023.102275
121. Wu H, Chen HY, Wang XC, Yu LH, Yu ZK, Shi ZJ, et al. Development and validation of an artificial intelligence-based image classification method for pathological diagnosis in patients with extramammary paget's disease. *Front Oncol.* (2022) 11:810909. doi: 10.3389/fonc.2021.810909
122. Jiang SC, Li HC, Jin Z. A visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis. *IEEE J Biomed Health Inform.* (2021) 25(5):1483–94. doi: 10.1109/JBHI.2021.3052044
123. Xue Y, Zhang J, Li C, Liu X, Kuang W, Deng J, et al. Machine learning for screening and predicting the risk of anti-MDA5 antibody in juvenile dermatomyositis children. *Front Immunol.* (2022) 13:940802. doi: 10.3389/fimmu.2022.940802
124. Li W, Huang Q, Peng Y, Pan S, Hu M, Wang P, et al. A deep learning approach based on multi-omics data integration to construct a risk stratification prediction model for skin cutaneous melanoma. *J Cancer Res Clin Oncol.* (2023) 17:15923–38. doi: 10.1007/s00432-023-05358-x
125. Cao C, Yi Z, Xie M, Xie Y, Tang X, Tu B, et al. Machine learning-based radiomics analysis for predicting local recurrence of primary dermatofibrosarcoma protuberans after surgical treatment. *Radiother Oncol.* (2023) 186:109737. doi: 10.1016/j.radonc.2023.109737
126. Liu JY, Liu LP, Li Z, Luo YW, Liang F. The role of cuproptosis-related gene in the classification and prognosis of melanoma. *Front Immunol.* (2022) 13:986214. doi: 10.3389/fimmu.2022.986214
127. Song JK, Zhang Y, Fei XY, Chen YR, Luo Y, Jiang JS, et al. Classification and biomarker gene selection of pyroptosis-related gene expression in psoriasis using a random forest algorithm. *Front Genet.* (2022) 13:850108. doi: 10.3389/fgene.2022.850108
128. Xing L, Wu T, Yu L, Zhou N, Zhang Z, Pu Y, et al. Exploration of biomarkers of psoriasis through combined multiomics analysis. *Mediators Inflamm.* (2022) 2022:7731082. doi: 10.1155/2022/7731082
129. Li CX, Shen CB, Xue K, Shen X, Jing Y, Wang ZY, et al. Artificial intelligence in dermatology: past, present, and future. *Chin Med J (Engl).* (2019) 132(17):2017–20. doi: 10.1097/CM9.0000000000000372
130. Wu H, Yin H, Chen H, Sun M, Liu X, Yu Y, et al. A deep learning, image based approach for automated diagnosis for inflammatory skin diseases. *Ann Transl Med.* (2020) 8(9):581. doi: 10.21037/atm.2020.04.39
131. VoxelCloud. VoxelCloud DermX (2017). Available online at: <https://www.voxelcloud.ai/solution/dermx> (Accessed August 03, 2025).
132. Huang K, Wu X, Li YX, Lv CZ, Yan YT, Wu Z, et al. Artificial intelligence-based psoriasis severity assessment: real-world study and application. *J Med Internet Res.* (2023) 25:e44932. doi: 10.2196/44932
133. Portnoy J, Waller M, Elliott T. Telemedicine in the era of COVID-19. *J Allergy Clin Immunol Pract.* (2020) 8(5):1489–91. doi: 10.1016/j.jaip.2020.03.008
134. Kato J, Horimoto K, Sato S, Minowa T, Uhara H. Dermoscopy of melanoma and non-melanoma skin cancers. *Front Med (Lausanne).* (2019) 6:180. doi: 10.3389/fmed.2019.00180
135. Cai G, Zhu Y, Wu Y, Jiang X, Ye J, Yang D. A multimodal transformer to fuse images and metadata for skin disease classification. *Vis Comput.* (2022) 39:1–13. doi: 10.1007/s00371-022-02492-4
136. Chen Q, Li M, Chen C, Zhou P, Lv X, Chen C. MDFNet: application of multimodal fusion method based on skin image and clinical data to skin cancer classification. *J Cancer Res Clin Oncol.* (2023) 149(7):3287–99. doi: 10.1007/s00432-022-04180-1
137. von Eschenbach WJ. Transparency and the black box problem: why we do not trust AI. *Philos Technol.* (2021) 34(4):1607–22. doi: 10.1007/s13347-021-00477-0
138. Wang ST, Yin YQ, Wang DJ, Wang YZ, Jin YC. Interpretability-based multimodal convolutional neural networks for skin lesion diagnosis. *IEEE Trans Cybern.* (2022) 52(12):12623–37. doi: 10.1109/TCYB.2021.3069920
139. Grzybowski A, Jin K, Wu H. Challenges of artificial intelligence in medicine and dermatology. *Clin Dermatol.* (2024) 42:210–5. doi: 10.1016/j.clindermatol.2023.12.013
140. Hacker P. A legal framework for AI training data—from first principles to the artificial intelligence act. *Innov Technol.* (2021) 13(2):257–301. doi: 10.1080/17579961.2021.1977219
141. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med.* (2020) 3(1):119. doi: 10.1038/s41746-020-00323-1
142. Nagaraju M, Chawla P, Kumar N. Performance improvement of deep learning models using image augmentation techniques. *Multimed Tools Appl.* (2022) 81(7):9177–200. doi: 10.1007/s11042-021-11869-x
143. Lu P, Zhang H, Peng X, Jin X. An end-to-end neural network for image cropping by learning composition from aesthetic photos. *arXiv [Preprint]. arXiv:1907.01432* (2019).
144. Al-Huda Y, Yao J, Peng B, Raza A. Weakly supervised skin lesion segmentation based on spot-seeds guided optimal regions. *Iet Image Processing.* (2023) 17(1):239–55. doi: 10.1049/ipr2.12631
145. Hu L, Zhang Y, Chen K, Mobayen S. A computer-aided melanoma detection using deep learning and an improved African vulture optimization algorithm. *Int J Imaging Syst Technol.* (2022) 32(6):2002–16. doi: 10.1002/ima.22738
146. Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors (Basel).* (2018) 18(2):556. doi: 10.3390/s18020556
147. Ou C, Zhou S, Yang R, Jiang W, He H, Gan W, et al. A deep learning based multimodal fusion model for skin lesion diagnosis using smartphone collected clinical images and metadata. *Front Surg.* (2022) 9:1029991. doi: 10.3389/fsurg.2022.1029991
148. Yang X, Li T, Su Q, Liu Y, Kang C, Lyu Y, et al. Application of large language models in disease diagnosis and treatment. *Chin Med J (Engl).* (2025) 138(2):130–42. doi: 10.1097/CM9.00000000000003456
149. Polesie S, Gillstedt M, Kittler H, Lallas A, Tschandl P, Zalaudek I, et al. Attitudes towards artificial intelligence within dermatology: an international online survey. *Br J Dermatol.* (2020) 183(1):159–61. doi: 10.1111/bjd.18875
150. Shen C, Li C, Xu F, Wang Z, Shen X, Gao J, et al. Web-based study on Chinese dermatologists' attitudes towards artificial intelligence. *Ann Transl Med.* (2020) 8(11):698. doi: 10.21037/atm.2019.12.102
151. Ba W, Wu H, Chen WW, Wang SH, Zhang ZY, Wei XJ, et al. Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images. *Eur J Cancer.* (2022) 169:156–65. doi: 10.1016/j.ejca.2022.04.015