ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

REVIEW ARTICLE

Artificial intelligence in skin cancer diagnosis and prognosis: A comprehensive narrative review of current applications and future perspectives

Harpreet Singh[1]*, Rajdeep Singh[2], Pinki Kushwaha[3], Jatin Agarwal[4], Sagar Varshney[1], Arvind Kumar[1], Arun Kumar Mishra[5], Hitesh Chopra[6], Shivani Chopra[7], and Tabarak Malik[8,9]*

[1]School of Pharmaceutical Sciences, Faculty of Pharmacy, IFTM University, Moradabad, Uttar Pradesh, India

[2]Department of Computer Applications, School of Computer Science & Applications, IFTM University, Moradabad, Uttar Pradesh, India

[3]Department of Pharmaceutical Chemistry, Moradabad Educational Trust, Group of Institutions, Faculty of Pharmacy, Moradabad, Uttar Pradesh, India

[4]Department of Pharmaceutics, Moradabad Educational Trust, Group of Institutions, Faculty of Pharmacy, Moradabad, Uttar Pradesh, India

[5]Sahu Onkar Saran School of Pharmacy, Faculty of Pharmacy, IFTM University, Moradabad, Uttar Pradesh, India

[6]Centre for Research Impact & Outcome, Chitkara College of Pharmacy, Chitkara University, Rajpura, Punjab, India

[7]Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India

[8]Department of Biomedical Sciences, Institute of Health, Jimma University, Jimma, Oromia Region, Ethiopia

[9]Division of Research & Development, Lovely Professional University, Phagwara, Punjab, India

***Corresponding authors:**
Harpreet Singh
(harpreetpharmacy@iftmuniversity.ac.in)
Tabarak Malik
(tabarak.malik@ju.edu.et)

Citation: Singh H, Singh R, Kushwaha P, *et al*. Artificial intelligence in skin cancer diagnosis and prognosis: A comprehensive narrative review of current applications and future perspectives. *Artif Intell Health.* doi: 10.36922/AIH025470106

Received: November 21, 2025

Revised: February 20, 2026

Accepted: March 4, 2026

Published online: May 5, 2026

Abstract

Skin cancer is among the most prevalent malignancies worldwide, and early, accurate diagnosis is crucial for improving patient outcomes. Recent advances in artificial intelligence (AI), particularly machine learning and deep learning, have shown substantial potential to enhance skin cancer detection, classification, and prognostic assessment. This review provides a comprehensive synthesis of current AI-based approaches for melanoma and non-melanoma skin cancers, highlighting methodological innovations and clinical applications. Literature from 2015 to 2025 was screened from PubMed/MEDLINE, Scopus, IEEE Xplore, and ScienceDirect, focusing on peer-reviewed studies reporting AI-driven diagnostic, classification, or prognostic outcomes. Deep learning models, especially convolutional neural networks, demonstrated high diagnostic performance in image-based skin cancer detection, often comparable to experienced dermatologists. AI has also shown promise in lesion segmentation, risk stratification, and prognostic modeling. However, challenges remain, including class imbalance, underrepresentation of darker skin tones, limited external validation, algorithmic opacity, and integration into clinical workflows. To enable broader clinical adoption, future research should prioritize diverse multicenter datasets, explainable AI systems, multimodal data integration, and prospective clinical validation studies. Overall, AI technologies offer significant potential to improve the accuracy and efficiency of skin cancer diagnosis and prognosis, but their translation into routine dermatological practice requires careful attention to reliability, equity, and interpretability.

Keywords:Artificial intelligence; Machine learning; Deep learning; Skin cancer; Melanoma; Diagnosis; Prognosis; Dermatology

1. Introduction

Skin cancer is among the most common malignancies worldwide, and its incidence continues to rise because of ultraviolet radiation exposure, environmental factors, and lifestyle-related influences. Generally, skin cancers can be divided into two groups: non-melanoma skin cancers, consisting of basal cell carcinoma and squamous cell carcinoma, and melanoma, which is less frequent but with a disproportionately high mortality rate.[1] Recent epidemiological studies estimate that millions of new non-melanoma skin cancer cases are diagnosed annually, whereas most skin cancer-related deaths are attributable to melanoma, which has aggressive biological behavior and high metastatic potential. Early detection and accurate prognosis assessment are of great importance in improving the survival and treatment outcomes.[2] For example, if melanoma is diagnosed early in the disease process, then the five-year survival rate is greater than 95%, whereas late-stage melanoma often becomes resistant to conventional therapies, and the prognosis remains poor.

Despite representing a smaller proportion of total skin cancer cases, melanoma accounts for most of the skin cancer–related deaths worldwide. In contrast, non-melanoma skin cancers, primarily basal cell carcinoma and squamous cell carcinoma, are markedly more prevalent but are associated with substantially lower mortality.[3] This disparity highlights the aggressive biological behavior and metastatic potential of melanoma, underscoring the critical importance of early detection and accurate risk stratification. Table 1 summarizes global incidence and mortality patterns of melanoma and non-melanoma skin cancers based on publicly available epidemiological data.[4]

The traditional diagnostic modalities include clinical examination, dermoscopy, histopathology, and, lately, molecular profiling. Although these modalities remain the clinical standard, they are limited by inter-observer variability, subjective image interpretation, and resource constraints.[5]

The increasing incidence of skin cancer worldwide requires urgent attention to develop new methods of diagnosis and prognosis that can complement the conventional methods. Advances in medical imaging, digital pathology, and computational analytics have enabled the development of AI algorithms that may improve diagnostic accuracy, reduce human error, and support personalized risk assessment. As AI continues to evolve, its application in dermatology and oncology has the potential to bring about a paradigm shift in the management of skin cancer.[8]

2. Importance of accurate diagnosis and prognosis

Early detection of melanoma, the most aggressive form of skin cancer, is critical for improving prognosis. Timely diagnosis increases the likelihood of successful treatment and may reduce the need for extensive intervention. Early detection is associated with less invasive surgery, shorter hospital stays, and improved treatment outcomes. Routine skin examination and prompt diagnosis are therefore key to improving patient outcomes. AI has paved the way for the integration of advanced computational methods into medical practice and provides additional support for diagnosis, prognosis, and treatment planning. In recent years, machine learning (ML) and deep learning (DL) approaches have been increasingly explored to support skin cancer diagnosis, prognostic assessment, and treatment planning.[9] This growing interest is largely attributable to the capacity of AI systems to analyze complex datasets and generate clinically meaningful predictions. In this context, such systems can support disease detection, prognostic assessment, and therapeutic decision-making. Diagnosis refers to identifying the presence and type of disease, whereas prognosis refers to predicting the likely course and outcome of the disease based on available clinical and diagnostic information. In skin cancer, ML and DL models are of particular interest because early detection is strongly associated with improved outcomes and more effective treatment.[10]

Table 1. Global incidence and mortality of melanoma vs. non-melanoma skin cancers

Skin cancer type	Estimated annual global incidence	Estimated annual global mortality	Proportion of total skin cancer deaths	Key clinical implication
Melanoma[6]	~325,000 cases	~57,000 deaths	~75–80%	High metastatic potential; early detection is critical
Non-melanoma skin cancers (BCC + SCC) [7]	>5 million cases	~20,000–25,000 deaths	~20–25%	High prevalence but generally low lethality

Data sources: Global Cancer Observatory, World Health Organization cancer statistics, and recent epidemiological reviews.
Abbreviations: BCC: Basal cell carcinoma; SCC: Squamous cell carcinoma.

3. Methods

A structured literature search was conducted to inform this narrative review, with the study selection process guided by principles from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. Relevant studies published between January 2015 and December 2025 were identified from PubMed/MEDLINE, Scopus, IEEE Xplore, and ScienceDirect, and the findings were synthesized narratively. After the removal of duplicate records, titles and abstracts were screened for relevance. Full-text articles were subsequently assessed against predefined eligibility criteria. Studies meeting the inclusion criteria were retained for qualitative synthesis. As shown in Figure 1, 1,248 records were identified across the databases searched, and 67 studies were ultimately included in the review.

3.1. Inclusion and exclusion criteria

The current review included studies published within the period of 2015–2025, in the English language, on the application of AI/ML techniques in the diagnosis or classification of skin cancer using either clinical images or dermoscopic data. Other inclusion criteria were studies that presented experimental results, comparative analyses, or clinical validation of AI models. The exclusion criteria involved non-peer-reviewed articles, studies unrelated to dermatology/skin cancer, papers lacking sufficient methodological detail, and those whose focus was solely on non-human subjects or theoretical models that are not empirically validated.

3.2. Article screening process

A systematic literature search was conducted across four major electronic databases, including PubMed/MEDLINE, Scopus, IEEE Xplore, and ScienceDirect, covering studies published between January 2015 and December 2025. A total of 1,248 records were initially identified, of which 276 duplicate entries were removed, leaving 972 unique records for screening. These records were evaluated based on titles and abstracts, leading to the exclusion of 781 articles that did not meet the predefined inclusion criteria. Following this, 191 full-text articles were retrieved and assessed for

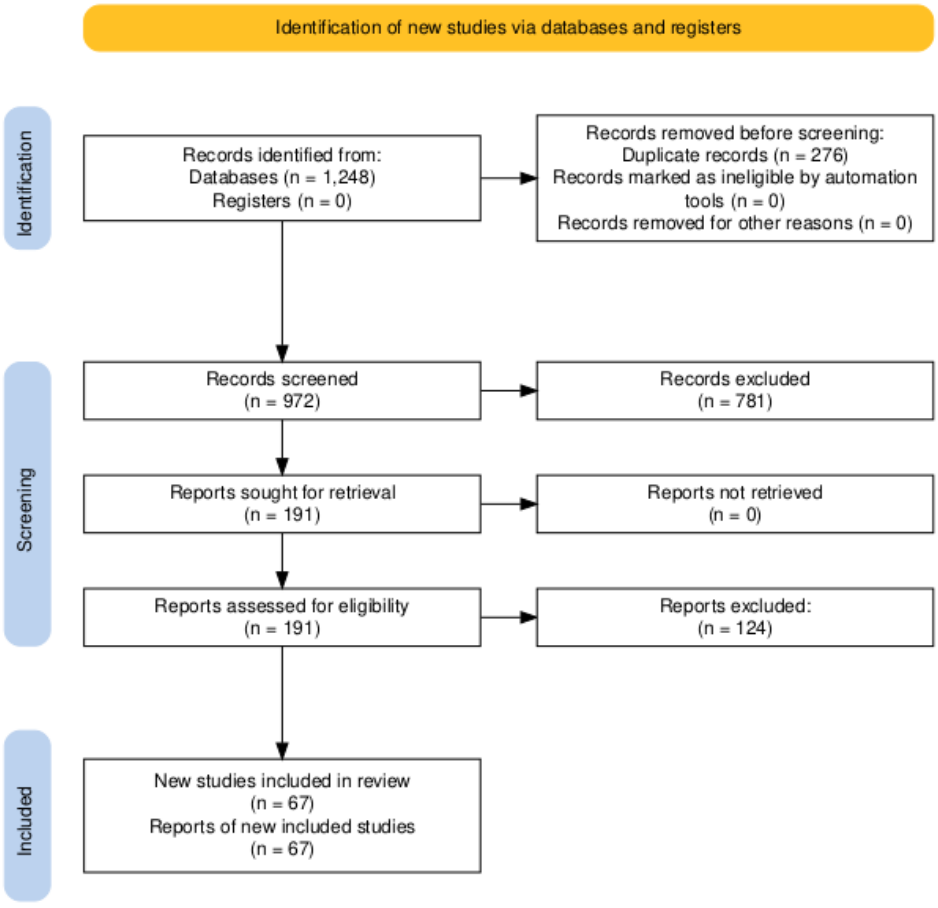

Figure 1. PRISMA flow diagram illustrating the study selection process

eligibility, with no reports missing or unretrieved. During the eligibility assessment, 124 articles were excluded due to reasons such as irrelevance to the study objectives, insufficient description of artificial intelligence (AI) methodologies, or lack of clinically relevant diagnostic or prognostic outcomes. Ultimately, 67 studies met all inclusion criteria and were included in the final qualitative synthesis. No additional studies were identified through other sources such as websites, organizational reports, or citation tracking. The entire study selection process was conducted in accordance with PRISMA guidelines to ensure transparency and reproducibility, and is illustrated in the PRISMA 2020 flow diagram (Figure 1).

3.3. Data extraction

Data extraction from the included studies was performed in a structured and systematic manner. For each eligible article, information was collected on the authors and year of publication, the primary study objectives, and the specific AI task addressed, such as diagnosis, classification, segmentation, or prognosis. Additional extracted variables included dataset characteristics (data source, sample size, and image modality), the type of AI model or algorithm employed, and the validation strategy used, including train–test splits, cross-validation, or external validation approaches. Reported performance metrics—such as accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (AUC), and Dice coefficient—were recorded, along with any evidence of clinical evaluation or external validation against expert or histopathological standards.

3.4. Quality and risk-of-bias assessment

The methodological quality and risk of bias of the included studies were assessed using adapted criteria derived from the Prediction Model Risk of Bias Assessment Tool (PROBAST) and QUADAS-2, specifically tailored for AI-based diagnostic and prognostic research. The assessment focused on four key domains: data source and participant selection, model development and validation procedures, outcome definitions and reference standards, and the transparency and completeness of statistical analysis and reporting. Based on these domains, studies were categorized as having low, moderate, or high risk of bias. The most frequently identified sources of bias included class imbalance within datasets, absence of external or multicenter validation, and insufficient reporting of model calibration, robustness, or interpretability.

3.5. Data synthesis

Due to methodological heterogeneity across studies in terms of datasets, AI architectures, validation strategies, and outcome measures, a quantitative meta-analysis was not performed. Instead, findings were synthesized narratively and comparatively, supported by structured tables summarizing key methodological features and performance outcomes.

3.6. Artificial intelligence models in skin cancer diagnosis and prognosis: A focused synthesis

Rather than providing generic descriptions of ML paradigms, recent research has converged on task-specific AI implementations for skin cancer, primarily leveraging DL-based image analysis. The dominant application area is image-based diagnosis, particularly the classification of melanoma versus benign lesions using dermoscopic and clinical photographs. Across studies, convolutional neural networks (CNNs) remain the most widely adopted architecture due to their ability to capture hierarchical visual features such as asymmetry, border irregularity, color variation, and textural heterogeneity—hallmarks of malignant skin lesions. Architectures such as Inception-v3, ResNet-50/101, DenseNet-121, EfficientNet, and hybrid CNN–UNet models are frequently used, often with transfer learning from ImageNet to mitigate limited dataset sizes. These models are applied not only to binary melanoma detection but also to multi-class classification encompassing basal cell carcinoma, squamous cell carcinoma, benign keratoses, dermatofibroma, and melanocytic nevi. Beyond classification, AI models have been increasingly applied to lesion segmentation, a clinically important preprocessing step that improves diagnostic accuracy and enables objective assessment of lesion borders and size. Segmentation tasks commonly report Dice coefficients between 0.80 and 0.90, particularly when CNN-based architectures such as UNet or DeepLab variants are employed. In contrast, prognostic applications of AI in skin cancer remain comparatively limited. Existing studies primarily explore risk stratification or recurrence prediction using imaging features combined with clinical metadata. Few models incorporate longitudinal outcomes such as progression-free survival, metastatic risk, or response to immunotherapy, highlighting a major gap between diagnostic performance and meaningful clinical endpoints.

This review focuses on clinical applicability, validation, and translational relevance of AI systems in skin cancers, rather than reintroducing basic theoretical descriptions of established ML classifiers. Emphasis is placed on clinical implementation, comparative performance of these models, and major challenges regarding generalizability, interpretability, fairness, and clinical applicability of these models.

Recent state-of-the-art (SOTA) literature on skin cancer reveals an overwhelming predominance of CNN architecture, as expected from its superior capability in automated feature extraction and high-performance image classification. Early works were conducted mostly under the umbrella of conventional ML approaches, such as support vector machines (SVMs), random forests, and k-nearest neighbors (kNNs), highly dependent on handcrafted feature engineering. As larger annotated datasets became available and computational power increased, the field shifted toward DL approaches, especially CNN-driven architectures, often achieving higher accuracy and AUC values in benchmark datasets. More recently, prognostic modeling has come into focus, as well as multimodal AI systems incorporating dermoscopic images with clinical metadata, histopathology, and longitudinal data. This represents the transition of research interest from proof-of-concept diagnostic tools to clinically meaningful, explainable, and translationally relevant AI frameworks in dermatologic oncology.

3.7. Datasets and generalizability

Most AI studies in dermatology rely on a small number of publicly available datasets, notably ISIC, HAM10000, PH2, and Derm7pt. While these datasets have enabled benchmarking and rapid algorithm development, they suffer from class imbalance, overrepresentation of lighter skin tones, and limited geographic diversity. As a result, many high-performing models demonstrate strong internal validation yet lack robustness when applied to external cohorts.

Emerging datasets derived from three-dimensional (3D) total body photography and longitudinal imaging offer promising avenues for monitoring lesion evolution and predicting disease progression; however, their use remains limited to a small number of exploratory studies.

3.8. Model performance and clinical validation

Reported diagnostic performance is consistently high, with many CNN-based systems achieving accuracies above 90% and AUC values between 0.91 and 0.96. Several landmark studies have demonstrated performance comparable to, or exceeding, that of experienced dermatologists under controlled conditions. However, only a minority of studies include external validation, prospective testing, or head-to-head clinical evaluation, limiting conclusions about real-world effectiveness.[11]

Importantly, high accuracy alone is insufficient for clinical adoption. Sensitivity for melanoma detection, robustness across skin types, interpretability of predictions, and impact on clinical decision-making (e.g., biopsy reduction, earlier diagnosis) are increasingly recognized as critical endpoints.[12]

Artificial intelligence has emerged as a disruptive power in almost every field of human activity, and medical diagnosis is no exception. In dermatology, AI applications have received growing attention, and developments have occurred in computerized management and skin condition diagnosis. Early diagnosis and detection of skin lesions, including melanoma and other skin cancers, are essential for improved patient outcomes and early action.[13]

In ML, a branch of AI, computer programs learn from data rather than being programmed with explicit decision rules. ML methods can be broadly categorized as supervised, unsupervised, and semi-supervised approaches. In supervised learning, models are trained on labeled datasets, where input data are paired with corresponding correct outputs, enabling the model to learn mappings through iterative optimization. In contrast, unsupervised learning involves analyzing unlabeled data, allowing models to identify underlying patterns, structures, or relationships without prior knowledge of the correct answers. Semi-supervised learning combines both labeled and unlabeled data, leveraging the strengths of each to improve learning efficiency and performance, particularly when labeled data are limited. DL is a subset of ML that uses multilayered deep neural networks (DNNs), each capable of identifying and learning different features specific to the dataset. An artificial neural network (ANN) is a computational model inspired by biological neural networks.[14]

A basic feed-forward neural network is the simplest form of ANN, consisting of input, hidden, and output layers through which information passes sequentially. Such networks may contain multiple hidden levels. CNNs are a type of deep feed-forward ANN commonly used for image analysis. CNNs consist of convolutional and pooling layers that enable hierarchical feature extraction from images.[15]

4. Artificial intelligence, machine learning, and deep learning: A high-level overview for non-experts

Artificial Intelligence is a broad term used to describe computer systems that are designed to perform tasks that typically require human intelligence, such as recognizing patterns, making decisions, or learning from experience. In the context of healthcare, AI aims to assist clinicians by analyzing large volumes of medical data and providing decision support, rather than replacing human expertise. For skin cancer, AI systems can help analyze images of skin lesions, identify suspicious patterns, and support early

diagnosis.[16]

Machine learning is a subset of AI that enables computers to learn from data without being explicitly programmed with fixed rules. Instead of following predefined instructions, ML algorithms identify patterns by being trained on examples. For instance, in skin cancer diagnosis, ML models learn from thousands of labeled skin images (benign or malignant) and gradually improve their ability to classify new, unseen lesions. This data-driven learning approach makes ML particularly useful in medical imaging, where visual patterns may be subtle and complex.[17]

Deep learning is a specialized subset of ML that uses multilayered ANNs inspired by the human brain. These models automatically learn increasingly complex features from raw data, such as edges, shapes, colors, and textures in skin lesion images.[18] CNNs, a commonly used DL architecture, are especially effective for image-based tasks and have shown diagnostic performance comparable to that of experienced dermatologists in several studies. Unlike traditional methods that rely on manually selected features, DL systems can independently discover clinically relevant patterns, making them highly suitable for skin cancer detection and prognosis.[19]

From a clinical perspective, the distinction between AI, ML, and DL can be viewed as a hierarchy: AI represents the overall goal of intelligent decision support, ML provides the learning mechanism from medical data, and DL enables highly accurate image interpretation. Together, these technologies offer powerful tools to enhance diagnostic accuracy, reduce inter-observer variability, and support timely clinical decision-making in dermatology.[20]

5. Objectives of the current review

This review aims to synthesize the current literature on AI-based skin cancer diagnosis and prognosis, compare commonly used datasets and model classes, and identify key gaps, particularly in prognostic modeling, fairness, external validation, teledermatology, and regulatory considerations. This review also aims to highlight the strengths and limitations of previous approaches and to identify priorities for future research in AI-based skin cancer diagnosis and prognosis. Unlike prior reviews, this study emphasizes prognostic AI models, teledermatology applications, fairness across skin types, and regulatory considerations, highlighting areas critical for clinical translation.

6. Research gap of the review article

Although AI research in dermatology has expanded substantially, there remain major limitations in the current body of research. One major limitation is that most studies on AI systems have focused on classification tasks, particularly melanoma-versus-benign lesion discrimination using dermoscopic images. While there have been major advancements using DL networks, there has been less focus on prognostic AI systems, including systems used as predictors for recurrence, metastasis, treatment response, or disease evolution.

Another important gap is the limited use of longitudinal outcome modeling. Most of the existing AI-based studies rely on cross-sectional image datasets and retrospective test–train splits, whereas time-dependent prediction, follow-up imaging, and survival analysis remain underexplored.

Furthermore, the development of fairness-aware AI systems is also an understudied area. In the existing datasets, lighter Fitzpatrick skin tones and lesion categories are over-represented, which can lead to algorithmic bias and decreased model performance in diverse populations. In addition, rare categories and darker skin tones are poorly represented. Despite a greater awareness of these problems, only a small number of studies examine model performance across different groups and incorporate bias mitigation. Similarly, other gaps range from inadequate external validation, especially across various cohorts, to limited prospective clinical trials, inadequate standardization of imaging, annotation, and explainability, among others. Furthermore, regulatory considerations and workflow harmonization have been insufficiently explored. While prior reviews have shown significant concentration on applications of AI, mainly with regard to the accuracy of various algorithms for diagnosing melanoma, the review is unique owing to the focus on the development of prognostic models, validation, fairness, and clinical integration, among other aspects.

7. Current status of diagnosis and prognosis for skin cancer

Skin cancer comprises a heterogeneous group of lesions, including melanoma and non-melanoma skin cancers, some of which can be life-threatening. Early diagnosis remains essential for improving patient outcomes. Current diagnostic approaches include clinical examination, dermoscopy, histopathology, and non-invasive imaging techniques. In addition, modern technologies such as deep learning, particularly convolutional neural networks, are increasingly being explored as adjunctive AI-based tools to support diagnosis.[21] These modalities are used to detect, characterize, and confirm skin cancer, which arises from the uncontrolled proliferation of abnormal skin cells, often associated with accumulated genetic damage. Skin lesions

may be broadly categorized as benign, pre-malignant, or malignant according to their biological behavior and histopathologic features.[22]

Benign lesions do not invade or metastasize, whereas malignant lesions can invade surrounding tissues and spread through lymphatic or hematogenous routes. Malignant skin lesions include melanoma and non-melanoma skin cancers. Because melanoma is associated with substantial mortality, numerous DL-based classification methods have been developed to support its early detection. Challenges such as image noise, irregular lesion boundaries, and variability in image size and quality have driven the development of ML-based approaches for lesion detection and classification.[23] A third category comprises pre-malignant lesions, which contain abnormal cells that are not yet malignant but have the potential to progress to cancer. These lesions can be difficult to classify because annotated image datasets are limited and because some lesions show overlapping clinical or histopathologic features. For example, actinic keratosis is generally regarded as a pre-malignant lesion with potential to progress to squamous cell carcinoma.[24]

8. Artificial intelligence algorithms used for skin cancer detection

Artificial intelligence has increasingly been integrated into skin cancer research to support diagnosis, prognosis, and treatment-related decision-making. Recent studies have applied both DL and ML approaches to improve early detection, lesion classification, and risk assessment. These systems can analyze complex clinical and imaging data to assist in disease detection, diagnostic assessment, and treatment planning.[25] While diagnosis involves assessing the patient's current condition, prognosis involves predicting the likely disease course and clinical outcome based on available clinical and diagnostic information. AI systems have shown potential to support earlier and more accurate detection in some controlled settings, but broader clinical validation remains necessary. In skin cancer research, such models are of growing importance because early detection is strongly associated with more effective treatment and better outcomes. As shown in **Figure 2**, several ML algorithms have been applied to the detection and classification of skin lesions.[26]

9. Artificial neural networks

Artificial neural networks are among the earliest ML models applied in medical diagnosis. In dermatology, ANNs have been utilized for skin-lesion classification and, less commonly, for outcome prediction for non-melanoma and melanoma-related conditions based on clinical, demographic, and imaging variables. Variables used for ANNs have ranged from characteristics of skin lesions (color, border irregularity, diameter) to demographic, comorbidity, and treatment-related factors.[27]

The preliminary ANN-based investigations revealed promising results for binary and multi-class classification of skin lesions, with sensitivity rates ranging from 86% to 88%, specificity rates ranging from 62% to 63%, and overall accuracy rates from 80% to 88%. Such results position the ANN models as powerful tools to aid clinicians with risk stratification of lesions, especially when employed alongside logistic regression or other feature-engineered models. Beyond image classification, ANNs have also been explored for risk stratification, outcome prediction, and structured-data analysis, because their architecture can integrate heterogeneous inputs such as clinical measurements, treatment variables, and healthcare-utilization data.[28]

Nevertheless, conventional feed-forward neural networks rely heavily on manually engineered features and generally scale less effectively than SOTA DL models. As annotated image databases grew, convolutional neural nets started to surpass conventional ANN models with respect to sensitivity and specificity, especially for melanoma image detection. Therefore, despite the core foundation for AI-based dermatological research emanating from ANN concepts, contemporary analysis on skin cancer research has shifted toward deep CNNs because of their superior feature-learning capacity.[29]

9.1. Conventional machine learning methods in early skin cancer classification

Early AI studies in skin-cancer diagnosis utilized conventional ML classifiers such as Naïve Bayes, decision trees, KNNs, k-means clustering, random forests, and SVMs. These approaches typically relied on handcrafted features derived from lesion asymmetry, border irregularity, colors, diameter, texture, and shape.[30]

These traditional ML models have shown reasonable diagnostic performance in small- to moderate-sized datasets. The accuracies reported were generally between 70% and 92%, depending on dataset size, feature quality, and validation strategy. Ensemble methods such as random forests may improve performance by reducing overfitting, whereas SVM-based systems have shown good discrimination in some melanoma-classification tasks. Naïve Bayes models had lower computational complexity but relatively low sensitivity for the detection of malignant lesions. Clustering techniques such as k-means and fuzzy c-means were used mostly during lesion segmentation and data preprocessing rather than the actual final diagnosis.[31]

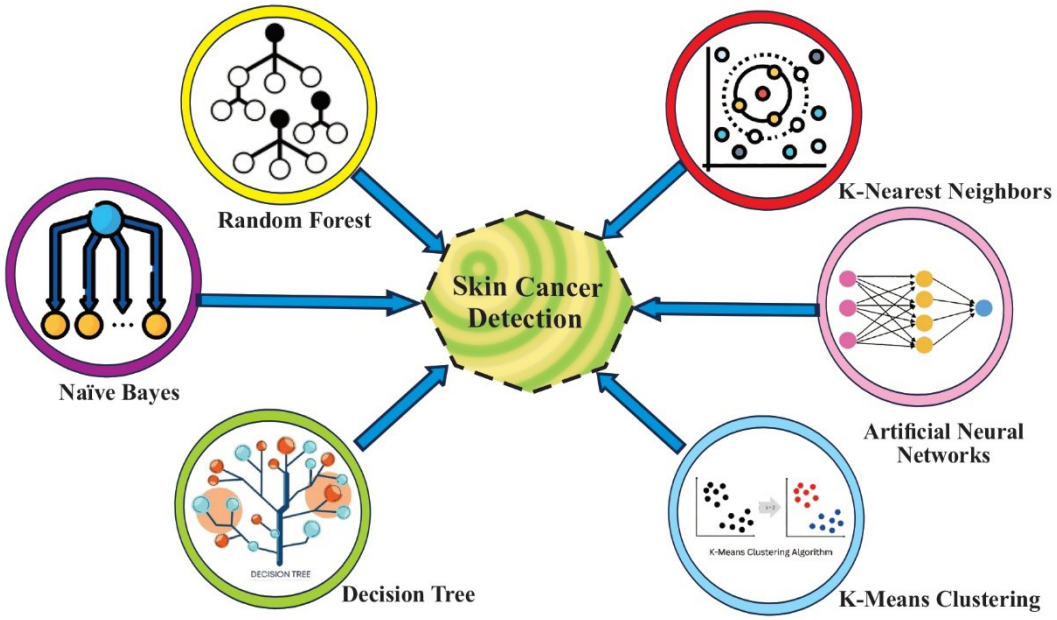

Figure 2. Machine learning algorithms for skin cancer detection. Image created by the authors using CorelDRAW software.

Despite the apparent success of the models, several shortcomings were observed. Model performance depended heavily on handcrafted feature engineering. The models' performance was further constrained in terms of generalization, considering the small number of samples within the dataset. Lastly, the models rarely involved any form of validation. Notably, however, when the volumes of the datasets and diversity of the images increased, CNNs generally outperformed these models since they learned features from the raw images.[32]

10. Deep learning techniques for the identification of skin cancer

Deep learning is a component of ML, a subfield of AI. DL models, made up of layers of ANNs that learn from enormous amounts of data, are based on the anatomy and physiology of the human brain. DL models are very advantageous in domains such as image identification, speech processing, and natural language processing, since they can automatically learn the extraction of patterns and features from raw data.[33]

In skin-lesion analysis, particularly melanoma detection, DL models are highly successful. They can be used, for instance, to scan medical images, including dermoscopic pictures of skin lesions. These models learn from large amounts of data, which are labeled images indicating whether lesions are benign or malignant. Upon training, the DL models can classify novel, unseen pictures with high precision, frequently reporting diagnostic accuracy akin to expert dermatologists. DL has improved skin cancer identification by allowing the models to extract intricate patterns within medical images for early and correct diagnosis. Due to their enormous image processing capabilities, CNNs are now the most advanced learning algorithms employed for skin cancer diagnosis.[34] Various DL approaches have been employed in skin cancer detection (Figure 3).

10.1. Recurrent neural networks

Recurrent neural network (RNN)-based models have had limited use in skin cancer diagnosis because dermoscopic analysis is primarily image-based rather than sequential. When used, RNNs are typically incorporated into hybrid architectures, often alongside CNN-derived features, to model dependencies among extracted representations. Bidirectional long short-term memory (LSTM) variants have also been explored in combined models to improve classification performance. Using hybrid models combined with optimization techniques, for example, has yielded a

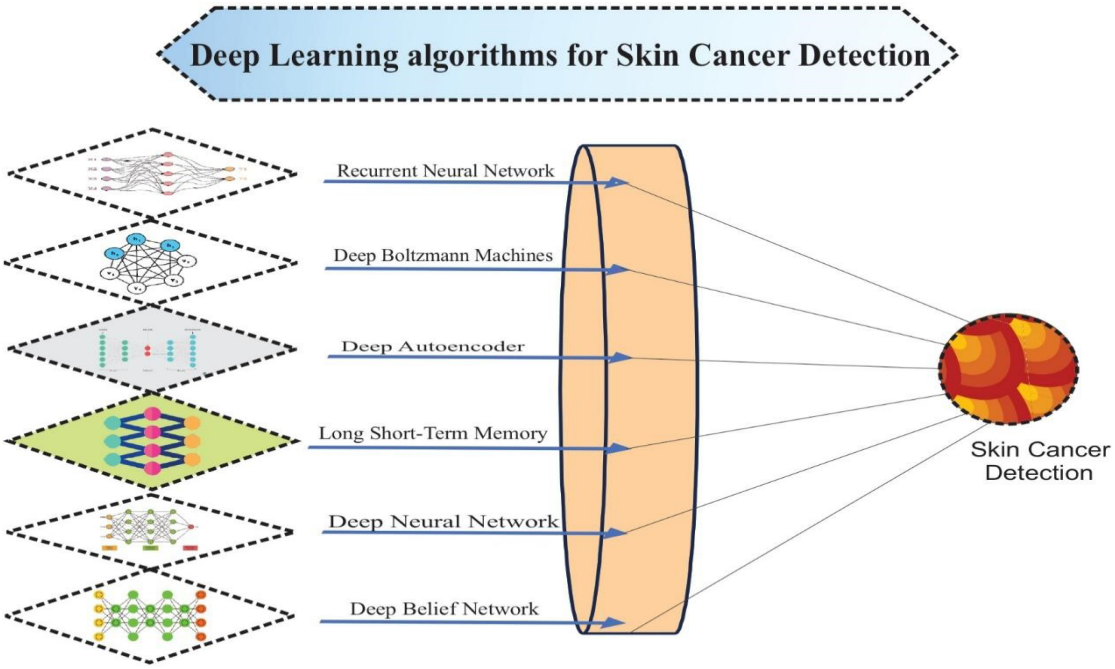

Figure 3. Deep learning algorithms for skin cancer detection. Image created by the authors using CorelDRAW software.

higher diagnostic accuracy than using a single model alone. In other studies, descriptions of texture, for instance, gray level co-occurrence matrices, have been combined with CNN outcomes and then filtered using RNN layer models, strengthening the classification. However, dataset size remains a limiting factor for RNN-based models, making them less common than CNN-based approaches.[35]

10.2. Deep Boltzmann machines

Deep Boltzmann machines have been explored for skin-lesion classification through probabilistic feature learning. Multimodal deep Boltzmann machine frameworks have also been investigated for integrating imaging or other data types in oncology. In dermoscopic applications, restricted Boltzmann machine-based pipelines have been proposed for segmentation and classification. However, these models are now used less frequently because they are computationally complex and have generally been surpassed by CNN-based systems in scalability and performance.[36]

10.3. Deep autoencoders

Deep autoencoders have mainly been used for feature reconstruction, dimensionality reduction, and mitigation of class imbalance in dermatologic datasets. Stacked autoencoder frameworks have been incorporated into melanoma detection pipelines, typically after preprocessing and alongside CNN-based feature extraction. Comparative studies have evaluated autoencoder-based approaches against architectures such as VGG19-UNet, ResNet50-UNet, MobileNetV2-UNet, DenseNet, and related systems. Their performance can be promising, but it remains highly dependent on careful parameter tuning and becomes increasingly difficult to optimize at scale.[37]

10.4. Long short-term memory

Long short-term memory models and related variants have been explored for sequential or temporal dermatologic data, including longitudinal risk modeling and disease progression analysis. Some studies have combined LSTM-based architectures with optimization algorithms, including hybrid squirrel butterfly search optimization, achieving greater precision and significance in various dermatologic classification tasks. Convolutional LSTM variants have been examined in broader spatiotemporal oncology settings. LSTM-based approaches remain much less common in skin cancer diagnosis than CNN-based models.[38]

10.5. Deep neural networks

General DNNs have also been applied to binary and multiclass classification of biopsy-confirmed skin lesions, including melanoma, seborrheic keratosis, and keratinocytic lesions. Several studies have reported performance comparable to that of expert dermatologists under controlled conditions.[39] Two-stream DNNs

that integrate complementary feature types have also improved discrimination between benign and malignant melanocytic lesions. Nonetheless, these models remain sensitive to dataset diversity and image quality, and their generalizability in clinical settings remains uncertain.[40]

10.6. Deep belief networks

Deep belief networks have been investigated for automated lesion detection and melanoma classification. A typical DBN-based dermoscopic pipeline includes preprocessing (using a filter such as a Gaussian filter), segmentation (using a technique such as k-means clustering), and classification stages, followed by fine-tuning of the network. Although DBNs have shown promising results over traditional techniques such as handcrafted features, their performance has generally lagged behind that of CNN-based methods.[41]

10.7. Convolutional neural networks

Convolutional neural networks remain the dominant architecture for skin cancer image analysis. They have been widely applied for the classification of melanoma, basal cell carcinoma, squamous cell carcinoma, and actinic keratosis. Reported performance on datasets such as ISIC and HAM10000 is often above 90% average accuracy, although it varies according to class balance, preprocessing, and validation design. Transfer learning, data augmentation, ensembling, and the incorporation of clinical metadata have further improved performance. Hybrid approaches, such as CNN+ECOC-SVM classifiers, and tuning the architecture, including learning rate, depth, and image resolution, allowed for further improvements in sensitivity and specificity. In several benchmark studies, CNN-based systems performed comparably to dermatologists.[42]

10.8. Deep reinforcement learning

Deep reinforcement learning has been explored mainly for lesion segmentation rather than primary diagnosis. Progressive refinement frameworks trained on ISIC and HAM10000 datasets have shown potential for improving boundary delineation in irregular lesions. Reinforcement learning-based systems may be useful for segmenting small or morphologically complex lesions; however, these approaches remain far less established than CNN-based segmentation models.[43]

10.9. Extreme learning machine

Extreme learning machines have been incorporated in hybrid DL frameworks for skin-lesion classification. In some studies, ELM-based classifiers can achieve an accuracy of greater than 93% in certain situations with the integration of transfer learning-based feature extraction methods, including SqueezeNet and DenseNet. ELM-based methods may converge faster than some traditional DNNs and may perform well in imbalanced datasets when combined with appropriate feature-extraction strategies. The integration of clinical data with images of skin lesions proves effective.[44]

11. Performance metrics in artificial intelligence-based skin cancer studies

Artificial intelligence studies in skin cancer commonly report multiple performance metrics, including accuracy, sensitivity, specificity, AUC, F1-score, and the Dice coefficient for segmentation tasks. Each metric provides distinct information: sensitivity reflects the algorithm's ability to correctly identify malignant lesions, specificity measures correct identification of benign lesions, and AUC summarizes overall discriminative performance. The F1-score balances precision and recall, particularly useful in datasets with class imbalance, while the Dice coefficient quantifies the overlap between predicted and ground-truth lesion boundaries. Understanding these metrics is crucial, as reliance on a single measure, such as accuracy, can be misleading, particularly in datasets where malignant cases are less frequent. Clinicians and researchers should interpret AI performance holistically, considering multiple metrics alongside clinical relevance.[45]

To better illustrate the range of AI methods used in skin cancer research, Table 2 summarizes representative studies using both traditional ML algorithms and SOTA models using DL architectures. Typically, traditional supervised ML methods, such as SVMs, random forests, kNN, Naïve Bayes, and decision trees, generally rely on handcrafted features (e.g., ABCD criteria of dermoscopic features, texture, and color). By contrast, SOTA DL methods, such as CNNs, learn directly from image pixels and therefore reduce reliance on manual feature extraction. Additional approaches, including RNNs, LSTM networks, DBNs, autoencoders, reinforcement learning, and ELM-based hybrids, have also been investigated for tasks such as segmentation, temporal modeling, and multimodal analysis. Emerging work on explainable AI aims to improve transparency, trust, and clinical usability.[46]

Overall, Table 2 shows that CNN-based DL systems dominate recent studies because of their strong representation learning for image data, although traditional ML models remain useful in smaller datasets. Ensemble methods may improve robustness, and segmentation-focused architectures can enhance downstream classification by improving lesion-boundary delineation. However, relatively few studies integrate explainability, external validation, or prognostic modeling, indicating an unaddressed research gap. Thus, Table 2 illustrates not only

Table 2. Summary of representative studies on AI applications in skin cancer diagnosis and image analysis

Author (year)	AI task	Dataset used (sample size)	Image modality	AI model/ technique	Validation strategy	Key numerical findings	Clinical validation	Major observations
Brinker *et al.* (2018)[47]	Melanoma vs. benign classification	ISIC + HAM10000 (~13,000 images)	Dermoscopic	CNN	10-fold cross-validation	Accuracy: 91.2%; AUC: 0.94; Sensitivity: 89.0%	Yes (vs. dermatologists)	CNN achieved dermatologist-level diagnostic accuracy
Esteva *et al.*(2017)[48]	Multi-class skin cancer classification	Private + public datasets (~129,450 images)	Clinical & dermoscopic	Deep CNN (Inception-v3)	Train/test split	AUC: 0.96 (melanoma); AUC: 0.95 (keratinocyte carcinoma)	Yes	Landmark study demonstrating AI parity with experts
Usama *et al.* (2022)[49]	Multi-class lesion classification	HAM10000 (10,015 images)	Dermoscopic	Deep feature extraction + ML	80/20 split	Accuracy: 92.4%; Sensitivity: 90.7%; Specificity: 93.6%	No	Strong performance; limited external generalization
Musthafa *et al.* (2024)[50]	Automated lesion classification	HAM10000 + PH2 (~10,200 images)	Dermoscopic	Optimized CNN	K-fold cross-validation	Accuracy: 94.8%; AUC: 0.96; F1-score: 0.95	Yes	Dataset fusion improved robustness
Gouda *et al.* (2022)[51]	Malignant vs. benign detection	DermNet + ISIC (~9,500 images)	Clinical & dermoscopic	Deep CNN	Hold-out validation	Accuracy: 93.1%; Sensitivity: 91.4%; Precision: 92.0%	No	Multi-source data improved feature diversity
Melbin & Raj (2021)[52]	Lesion type classification	ISIC subset (~2,000 images)	Dermoscopic	ABCD features + SVM	10-fold cross-validation	Accuracy: 90.8%; AUC: 0.91	Yes	Traditional ML remains competitive for small datasets
Divya & Ganeshbabu (2020)[53]	Melanoma detection	Private dataset (~1,200 images)	Clinical images	RNN + region growing	70/30 split	Accuracy: 89.5%; Sensitivity: 88.2%; Specificity: 90.3%	No	Early hybrid AI approach; dataset size limited
Nawaz *et al.* (2022)[54]	Lesion segmentation & detection	ISIC 2018 (2,594 images)	Dermoscopic	Fuzzy K-means + DL	Cross-validation	Dice: 0.87; Accuracy: 91.5%	No	Accurate segmentation improved classification

Abbreviations: AI: Artificial intelligence; AUC: Area under the receiver operating characteristic curve; CNN: Convolutional neural network; DL: Deep learning; ISIC: International Skin Imaging Collaboration; ML: Machine learning; RNN: Recurrent neural network; SVM: Support vector machine.

the reported performance metrics of these systems but also the methodological diversity and translational potential of current AI applications in dermatologic oncology.

12. Future perspectives and challenges

The use of AI in the diagnosis of skin cancer is an area that is growing rapidly, and although it is very promising, there are also enormous challenges that need to be met. The following are some of the key future trends and challenges in the application of AI in this field. AI algorithms, particularly DL models, have shown promise for image-based skin-cancer detection. Future systems will see even better algorithms that will be able to rival or even surpass the accuracy of human dermatologists. This will ensure that the detection of the disease in its early stages becomes more consistent, which is the only way the treatment of the disease will be improved. The widespread use of cameras in smartphones and wearables offers the opportunity for the real-time application of AI in the detection of skin cancer.[55] AI-powered applications may allow users to capture standardized skin images and receive preliminary risk assessments, which could support earlier referral. AI can also be applied to embed skin cancer detection in the routine healthcare setup so that its detection becomes a standard checkup. AI-powered automated systems can also be installed in clinics, dermatology clinics, or even general practitioners' offices to quickly check skin lesions for further testing or biopsy needs. AI can also contribute not only to detecting skin cancer but also to developing personalized treatment protocols. AI would identify the pattern and thus be able to tell what the most promising treatments of a patient would be according to individual patient characteristics and tumor type, and therefore maximize the effectiveness of the treatment in general.[56] AI may facilitate teledermatology triage, allowing dermatologists to remotely assess skin lesions received by way of photo or video transmission from patients. This would be especially helpful for underprivileged or rural areas where access to specialists may be limited, enhancing access to timely care. AI models rely substantially on large, high-quality training datasets. Most datasets, however, remain small and not very diverse, resulting in the risk of biased models that might not work as well across various populations, skin tones, or ethnicities. The absence of diverse data might result in reduced accuracy in minority populations, particularly among individuals with darker skin tones, where melanoma is more difficult to identify.[57]

Artificial intelligence systems, especially DL systems, are often referred to as "black boxes" because they give very little information on how they decide. Patients and medical professionals who must rely on the system's output suffer from this lack of interpretability. Explainable AI systems will be needed to improve transparency and clinician trust. As AI becomes more pervasive in the medical field, it will also have to comply with regulations and standards set by medical professionals. However, there is still a lot of work that needs to be done in order to ensure that these systems are secure and trustworthy, while also taking into consideration patient consent, privacy, and the possibility of error.[58] A model of AI that is trained on a single set, such as cases from a single hospital or region, will not generalize well to other regions. Skin cancer detection models will have to be robust enough to reach their intended level of performance without compromising in other geographic locations, healthcare systems, and populations. In order to effectively incorporate AI into the practice of healthcare, it must be incorporated seamlessly and with integrity into the practice of dermatologists and healthcare professionals. This remains challenging because clinical workflows prioritize efficiency and minimal disruption.[59]

Health AI needs large amounts of patient data for training and deployment. It is imperative to ensure that the data is stored and handled in a secure manner, as per the laws of the jurisdiction concerning privacy, such as the Health Insurance Portability and Accountability Act of 1996. Data breaches and misuse could be detrimental to the adoption of AI technologies by patients. Although AI has the ability to democratize healthcare, costs could be a factor to be taken into consideration, particularly in resource-poor settings. The training, deployment, and upkeep of AI systems are rather expensive; hence, the deployment of these systems to the majority of the population, especially in the healthcare system, will be an investment. AI has great potential in the future of skin cancer diagnosis, as it has the potential to revolutionize early diagnosis, treatment, and access to healthcare. There is a need to address the challenges of data diversity, explainability, regulatory issues, and clinical workflow integration. With technological advancements, these challenges could be addressed, and there would be greater adoption with improved patient outcomes.[60]

Artificial intelligence in skin cancer diagnosis is expected to progress through improved algorithms, data integration, and clinical validation. Future systems will increasingly adopt multimodal AI, combining dermoscopic and clinical images with histopathology and patient metadata to improve diagnostic and prognostic performance. The integration of AI with tele-dermatology may enable early screening and triage, particularly in resource-limited settings. Research is also shifting toward personalized risk stratification, allowing prediction of disease progression and treatment outcomes.[61] The development of explainable AI is essential to enhance clinician trust and clinical

adoption. Finally, large-scale multicenter prospective validation across diverse populations remains critical for ensuring generalizability, regulatory approval, and real-world implementation. Future research should focus on developing prognostic AI models, integrating AI with teledermatology for broader access, ensuring dataset diversity to mitigate bias, and aligning AI development with regulatory standards to facilitate clinical adoption.[62]

13. Recent developments in the diagnosis and prognosis of skin cancer

AI techniques that use image analysis can reliably classify skin cancer. These algorithms have the potential to be deployed either autonomously in controlled settings or as clinician-support tools, and skin cancer diagnosis can be aided by the use of DL models. These models can be employed in telemedicine screening devices or Smartphone apps. Skin cancer diagnosis accuracy can be improved using computer-aided diagnostic (CAD) algorithms. CAD algorithms incorporate ML models, algorithms, and data acquisition from automated equipment. Near-infrared light is used in optical coherence tomography (OCT) to create skin images of high quality. OCT is used to diagnose and evaluate the margins of skin cancers. Reflectance confocal microscopy produces high-resolution skin pictures by using near-infrared light. Immune checkpoint inhibitors have greatly influenced the survival of individuals with melanoma. Dermoscopy is an imaging method that observes skin lesions in real-time using a high-magnification lens. Dermoscopic images are usable in AI research.[63]

14. Recent advancements in artificial intelligence

Recent breakthroughs in AI, including DL, multimodal modeling, and explainable AI frameworks, have profoundly influenced various aspects of biomedical science. Apart from dermatology, AI possesses incredible transformative potential in areas like climate modeling, neurodegenerative disease prediction, precision medicine, and intelligent healthcare systems. The various breakthroughs in AI could be viewed as a shift from traditional rule-based computation towards adaptive intelligent systems. It is worth noting that breakthroughs in computational efficiencies, transfer learning, self-supervised learning, and large-scale dataset training allow AI systems to process medical information with even higher reliability.[64]

In dermatology, these technological advances have closely paralleled the development of AI-based skin-cancer diagnostic systems. In skin-cancer diagnostics systems, deep CNNs, which have proven to be effective in the recognition of natural images, have been further developed to become efficient in dermoscopy and clinical image analysis. Moreover, the development and availability of advanced AI architectures, such as vision transformers and other DNN models, have been incorporated, showing the integration of different features. In addition, multimodal systems that combine clinical images, dermoscopic images, and patient information reflect a shift from image-only classification toward broader risk assessment.[65]

A particularly significant advancement observed in the domain of AI for the field of dermatology is the emergence of so-called explainable AI, or XAI, in short. As the complexity of DL models rose, so did the concern regarding the so-called "black box" phenomenon of prediction modeling. Interpretability techniques such as gradient-weighted class activation mapping, saliency maps, Shapley Additive exPlanation values, and attention maps have thus far been introduced in the broader domain of AI for the field of dermatology. As a matter of fact, explainability is increasingly regarded as essential in the domain of AI. Additionally, the translational AI frameworks themselves are increasingly emphasized.[66] Thus, it is not just enough to aim for high internal accuracy metrics, but recent research has highlighted the need for external validation, multicentric dataset training, fairness to different skin types, and conformance to regulations. The AI algorithms are now designed to be integrable with the clinical workflows, allowing teledermatology, screening using smartphone devices, and incorporation within the electronic health record systems, which clearly indicate a paradigm shift in the field from experimental algorithm development approaches toward implementation science. These recent advances in AI collectively demonstrate the intersections between methodology, interpretability, and translation. In terms of skin cancer diagnosis and prognosis, what do the recent trends in AI suggest about the future of AI in this domain? They imply high performance, interpretability, fairness, and translatability to different clinical settings.[67]

15. Publicly available datasets and their challenges

A key requirement for AI-based skin-cancer systems for diagnosis and prognosis is access to large, diverse, and well-annotated image datasets. Publicly available data has considerably contributed to the training, validation, and benchmarking processes of the AI algorithms in dermatology. Prominent datasets in this area include the International Skin Imaging Collaboration (ISIC) archive, HAM10000, PH2, and Derm7pt, providing a wide variety of dermoscopic and clinical images of pigmented and non-pigmented skin lesions. These datasets form the foundation

for DL model development that can classify benign and malignant lesions with performance at the expert level.[68]

Recent efforts further extend the publicly available datasets. SLICE-3D is a dataset comprising more than 400,000 skin lesion image crops from 3D total body photography (TBP), which provides an important resource for lesion detection and longitudinal monitoring. Similarly, BCN20000 is a dataset of dermoscopic images captured "in the wild," reflecting natural variations due to acquisition conditions, illumination, and patient demographics. More recently, a dataset of skin region images extracted from 3D TBP was introduced to support lesion detection and segmentation, offering another large-scale benchmark for automated analysis.[69] Despite this importance, these datasets still face several persistent challenges that limit generalizability and clinical translation of AI models, such as class imbalance, where malignant lesions are highly under-represented compared to benign nevi, and biased model performance toward the majority class. Variations in image quality and resolution, and different acquisition devices, introduce noise and inconsistency, as mentioned in Table 3. A further lack of standardization in annotation protocols and metadata complicates cross-dataset training and evaluation. Furthermore, rare lesion types, such as Merkel cell carcinoma or amelanotic melanoma, are absent, impeding the establishment of robust AI tools for comprehensive skin cancer screening.[70]

Future dataset curation should therefore prioritize multi-institutional collaboration, inclusion of diverse skin tones, and standardized metadata to improve reproducibility and fairness in AI model development. Integration of 3D total body imaging and longitudinal follow-up data represents a promising direction for enabling early detection, personalized risk assessment, and disease progression modeling. Despite their widespread use, these datasets suffer from limitations such as class imbalance, underrepresentation of darker skin tones, and variability in acquisition protocols, underscoring the need for more diverse, standardized, and longitudinal datasets for clinically robust AI models.[78]

Table 3. Summary of major publicly available skin cancer datasets used in artificial intelligence research

Dataset name	Image modality	Sample size	Lesion types covered	Key limitations
ISIC Archive (International Skin Imaging Collaboration)[71]	Dermoscopic (some clinical images)	>70,000 images (ongoing expansion)	Melanoma, nevus, basal cell carcinoma, squamous cell carcinoma, benign keratoses, vascular lesions	Class imbalance with fewer malignant cases; limited metadata consistency across releases
HAM10000[72]	Dermoscopic	10,015 images	Melanoma, melanocytic nevi, basal cell carcinoma, actinic keratosis, benign keratosis, vascular lesions, dermatofibroma	Overrepresentation of lighter skin tones; single geographic source
PH2 Dataset[73]	Dermoscopic	200 images	Melanoma, atypical nevi, common nevi	Small sample size; limited lesion diversity
Derm7pt[74]	Dermoscopic + clinical metadata	1,011 images	Melanoma, nevi, basal cell carcinoma, other pigmented lesions	Limited dataset size; primarily focused on pigmented lesions
DermNet[75]	Clinical images	~23,000 images	Wide range of skin diseases including melanoma and non-melanoma cancers	Variable image quality; weak annotation standardization
BCN20000[76]	Dermoscopic	~19,000 images	Melanoma, nevi, keratoses, basal cell carcinoma	Images captured "in the wild"; annotation variability
SLICE-3D/TBP-derived datasets[69]	3D total body photography	>400,000 image crops	Longitudinal lesion monitoring (multiple lesion types)	Limited public availability; high storage and computational demands
ISIC 2018 Challenge Dataset[77]	Dermoscopic	2,594 images	Melanoma, nevus, seborrheic keratosis	Designed primarily for benchmarking; limited real-world diversity

16. Limitations and challenges

Artificial intelligence has shown remarkable potential for the diagnosis and prognosis of skin cancer, yet several limitations and challenges must be addressed to ensure clinical translation. A primary limitation is the lack of standardized and diverse datasets. Many AI models are trained on small, geographically limited datasets, often overrepresenting lighter skin types, while rare lesion types and darker skin tones remain underrepresented. This leads to bias and reduced generalizability across populations. Moreover, heterogeneity in image acquisition, annotation standards, and dermoscopic equipment further complicates reproducibility and benchmarking.

Another key challenge is interpretability.[79] DL models often act as "black boxes," providing predictions without clear explanations, which can undermine clinician trust and hinder adoption. In addition, regulatory barriers, ethical concerns, and data privacy issues pose significant obstacles. Questions regarding accountability for AI-assisted misdiagnosis, adherence to regional regulations, and integration with patient consent and privacy frameworks must be carefully managed. Clinical integration also remains a challenge. AI tools must seamlessly fit into existing workflows without disrupting clinicians' routines, while providing clear decision-support outputs. High development and deployment costs further limit access, particularly in resource-constrained settings. These issues collectively slow the adoption of AI in dermatology, despite its success in research environments.[80]

Key priorities include dataset diversification and standardization, incorporating multimodal data such as dermoscopic images, clinical photographs, 3D total body imaging, histopathology, and relevant clinical metadata (e.g., age, genetic predisposition, and medical history). Explainable AI is essential to enhance trust and interpretability, while large multicenter prospective studies are required to establish generalizability and support regulatory approval. Integration with teledermatology and personalized medicine may further expand the clinical utility of AI by enabling remote screening, risk stratification, and tailored treatment recommendations, particularly in under-resourced or rural settings. Another critical priority is ensuring robustness and fairness by developing models that perform accurately across diverse skin types, demographics, and healthcare systems, thereby reducing algorithmic bias.[81]

17. Challenges and barriers in clinical artificial intelligence for skin cancer

Notwithstanding the promising advances in AI for diagnosis and prognosis in skin cancer, several barriers remain to routine clinical implementation.

17.1. Data quality and standardization

Large, diverse, and high-quality datasets are needed to robustly train AI models; existing dermatological image databases are often biased toward particular skin types, lesion categories, or geographic locations, which limits model generalizability. Additionally, inconsistent imaging protocols, annotation standards, and variability in dermoscopic equipment further complicate reproducibility.[82]

17.2. Bias and underrepresentation of skin types

A major limitation is that darker skin tones and rare lesion types remain underrepresented in training datasets. Such bias may lead to reduced accuracy of diagnosis in populations that are not well represented, raising several concerns about fairness, equity, and inclusion within AI-driven dermatology.[83]

17.3. Interpretability and the "black box" problem

Most AI models, especially DL systems, provide limited transparency regarding how predictions are generated. Clinicians may remain skeptical about systems that do not offer justification for their results; hence, this lack of transparency can undermine clinical trust in algorithmic predictions.[84]

17.4. Clinical validation and regulatory approval

Few AI tools have received clinical validation, although many demonstrate exciting results in a research setting. Regulatory approval requires robust evidence of safety, efficacy, and reproducibility. The main challenges to overcome are related to compliance with the United States Food and Drug Administration, European Medicines Agency, or other regional regulatory frameworks.[85]

17.5. Integration into clinical workflow

Seamless integration into routine dermatologic practice will require interoperability with electronic health records, teledermatology platforms, and diagnostic devices. At present, a lack of standardization and disruption to established clinical workflows are common concerns among clinicians.[86]

17.6. Ethical, legal, and privacy concerns

Artificial intelligence training using patient images raises important issues related to informed consent, data security, and privacy. In addition, there is still confusion regarding accountability in cases of misdiagnosis—whether it lies with the clinician, the AI developer, or the healthcare institution.[87]

17.7. Cost, accessibility, and resource constraints

High development and deployment costs, combined with poor availability of sophisticated diagnostic imaging equipment in resource-poor settings, will further limit access to AI-based skin cancer diagnostics worldwide. This is a digital divide that must be overcome for equitable adoption to occur.[88]

17.8. Generic clinical workflow for artificial intelligence-assisted skin cancer diagnosis

Artificial intelligence-assisted skin cancer diagnosis follows a structured workflow integrated into routine dermatological practice. The process begins with patient presentation, either through in-person visits or teledermatology, where suspicious lesions are identified. This is followed by image acquisition using clinical or dermoscopic imaging under standardized conditions. Acquired images undergo preprocessing, including artifact removal, color normalization, and lesion segmentation, to improve analytical reliability.[89]

Subsequently, AI-based models, typically DL algorithms, analyze the images to classify lesions and estimate malignancy risk. The outputs are provided as probability scores or visual decision-support cues. During the clinical decision-support phase, AI results are interpreted alongside clinical examination and patient history, serving as an adjunct to clinician expertise. Lesions identified as high-risk undergo confirmatory biopsy and histopathological analysis, guiding treatment planning and follow-up. This workflow demonstrates practical AI integration while preserving clinician oversight.[90]

17.9. Artificial intelligence adoption in dermatology vs. radiology

While AI has made significant inroads in radiology, dermatology has lagged in clinical adoption. Several factors contribute to this disparity. Unlike radiology, dermatology lacks standardized imaging protocols, leading to high variability in image quality and acquisition conditions. Skin lesions themselves are highly heterogeneous in appearance across patients, body sites, and lighting conditions, which complicates algorithm training. Additionally, reimbursement pathways for AI-assisted dermatology tools remain limited, reducing incentives for clinical integration. Finally, compared to radiology, there are fewer large, well-annotated datasets publicly available for dermatology, constraining model development and validation. Together, these challenges slow the translation of AI from research to routine dermatologic practice.[91]

18. Future research directions

Current AI algorithms have largely been trained on datasets that are biased toward lighter skin types, common lesion types, and images captured under very controlled conditions. Hence, future work should aim to balance representation across the spectrum of Fitzpatrick skin types, ages, geographical locations, and even rare lesion categories. This kind of data curation will minimize algorithmic bias, ensure diagnostic performance for subpopulations of underrepresented patients, and foster fairness and inclusivity in clinical deployment. Variability in image capture, in the guidelines used for annotation, and in labeling data across datasets limits reproducibility and comparability between studies. Imaging protocols need to be harmonized, together with metadata documentation and standards for expert annotation. Future research should not be constrained to just dermoscopic images but include multimodal data like clinical photography, 3D total body scans, histopathological images, and other metadata of the patient, such as genetic predisposition and medical history.

19. Structured classification of referenced studies and research trends

The existing SOTA works mentioned and referenced throughout the review can be broadly grouped under four broad categories of methodology and/or approaches used. Firstly, the initial SOTA works followed more conventional approaches of traditional ML paradigms such as SVMs, random forests, kNN, Naïve Bayes, and decision trees, along with a collection of handcrafted features including the famous ABCD rule, textures, and morphology. Secondly, the overwhelming majority of SOTA works investigated and proposed DL approaches towards image classification, including CNNs, including the popular variants of Inception, ResNet, DenseNet, EfficientNet, and UNet networks. Thirdly, a considerable number of works highlighted the importance of lesion segmentation and preprocessing pipelines, including boundary detection. Lastly, the most recent and increasing numbers of SOTA works are focused on hybrid and multimodal approaches, including Explainable AI and/or limited but promising prognostic approaches.

20. Methodological evolution over time

Artificial intelligence methodologies in skin cancer research have evolved substantially over time. Early studies (2015–2017) mainly relied on conventional machine learning algorithms applied to relatively small datasets with handcrafted features. From 2017 to

2022, the field shifted toward deep learning methods, particularly convolutional neural networks (CNNs), driven in part by the increasing availability of labeled training data, including the ISIC archive and later the HAM10000 dataset. More recently, CNN-based models have shown strong diagnostic performance and, in some benchmark settings, performance comparable to that of dermatologists, as reflected in high AUC values. At the same time, methodological advances such as transfer learning, data augmentation, and cross-validation have become increasingly common to improve model training and evaluation robustness.

21. Timeline-based progression of research

Finally, the process of advancing the area of interest can be defined with a series of phases. The initial phase was associated with ML classification by making use of handcrafted features. The second phase was a breakthrough period, marked by the emergence of DL, particularly CNN-based architectures. The third phase was a period of improvement of segmentation, classification, and the validation of multiple datasets. The current period represents the translational period, with a focus on explainable AI, fairness in skin tones, multiple data integration, and clinical use.

22. Current translational focus

Recent studies are becoming more focused on clinical applicability and are moving on from simply assessing the diagnostic accuracy of AI tools. Current translational priorities in using AI tools in dermatology include teledermatology integration, remote triage in resource-limited settings, outcome prediction, explainability, and bias mitigation. In addition, regulatory compliance, multicenter technical and clinical validation, and workflow integration are increasingly recognized as essential for real-world implementation. This reflects a broader shift from proof-of-concept accuracy studies toward clinically deployable systems that are fair, reliable, and integrated into real-world workflows.

23. Conclusion

Skin cancer remains among the most common malignancies worldwide; its incidence is continuously increasing because of environmental, genetic, and lifestyle factors. Despite the significant advances in dermatological examination techniques and in therapeutic approaches, the burden of late diagnosis and recurrence makes the need to develop improved diagnostic and prognostic strategies very compelling. Melanoma is the most lethal major form of skin cancer and is associated with poor outcomes when diagnosed late. Although non-melanoma skin cancers generally have lower mortality, they impose a substantial clinical and socioeconomic burden. The heterogeneity in skin cancer types, along with their overlapping clinical and histopathological features, makes the accurate and timely diagnosis of skin cancer challenging. These challenges underscore the need for adjunctive technologies such as AI to complement current practice, improve diagnostic accuracy, and ultimately enhance patient outcomes.

Acknowledgments

All authors would like to express their sincere gratitude to their parent institutions for providing the necessary support, facilities, and encouragement to carry out this work successfully.

AI Declaration Statement

During the preparation of this work, the authors used "Grammarly" in order to "ensure grammatical accuracy and clarity". After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Funding

None.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: Harpreet Singh, Tabarak Malik, Rajdeep Singh, Pinki Kushwaha
Visualization: Arun Kumar Mishra, Hitesh Chopra
Writing-original draft: Harpreet Singh, Jatin Agarwal, Sagar Varshney, Arvind Kumar
Writing-review & editing: Shivani Chopra, Tabarak Malik

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Kristensen I, Iversen IB, Fenger-Grøn M, *et al*. Occupational solar ultraviolet radiation and incidence of non-melanoma skin cancer in a nationwide cohort. *Environ Pollut*. 2025;386:127275.

doi: 10.1016/J.ENVPOL.2025.127275

2. Roky AH, Islam MM, Ahasan AMF, *et al*. Overview of skin cancer types and prevalence rates across continents. *Cancer Pathog Ther.* 2025;3:89–100.

   doi: 10.1016/J.CPT.2024.08.002

3. Savage DJ, Switzer B, Parikh R, *et al*. Patterns in progression from early-stage melanoma to late-stage melanoma: implications for survivorship follow-up. *Melanoma Manag.* 2024;11:2424708.

   doi: 10.1080/20450885.2024.2424708

4. Kakish, D.R.K., Alsamhori, J.F., Qaqish, L.N, *et al*. Gender Disparities in Melanoma: Advances in Diagnosis, Treatment, and the Role of Artificial Intelligence. *Dermatol Rev.* 2025;6:e70022.

   doi: 10.1002/DER2.7002

5. Zhao J, Zhang X, Tang Q, *et al*. The correlation between dermoscopy and clinical and pathological tests in the evaluation of skin photoaging. *Ski Res Technol.* 2024;30:e13578.

   doi: 10.1111/SRT.13578

6. Skin cancer – IARC n.d. Available from: https://www.iarc.who.int/cancer-type/skin-cancer/ [Last accessed on April 1, 2026].

7. Non-melanoma skin cancer statistics|Cancer Research UK n.d. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-melanoma-skin-cancer [Last accessed on April 1, 2026].

8. Melarkode N, Srinivasan K, Qaisar SM, Plawiak P. AI-Powered Diagnosis of Skin Cancer: A Contemporary Review, Open Challenges and Future Research Directions. *Cancers*. 2023;15:1183.

   doi: 10.3390/CANCERS15041183

9. Garrison ZR, Hall CM, Fey RM, Clister T, Khan N, Nichols R, *et al*. Advances in Early Detection of Melanoma and the Future of At-Home Testing. *Life*. 2023;13:974.

   doi: 10.3390/LIFE13040974

10. Fahim YA, Hasani IW, Kabba S, Ragab WM. Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives. *Eur J Med Res.* 2025;30:848.

    doi: 10.1186/S40001-025-03196-W

11. Akram N, Mattoo S, Qaddoura B, Karimaghaei D, Hamarsheh I, Radha S. Computer Vision in Lower Limb Orthopaedics: A Scoping Review of Imaging-Based Artificial Intelligence Applications. *Cureus*. 2025;17:e98648.

    doi: 10.7759/cureus.98648

12. Górecki S, Tatka A, Brusey J. Artificial Intelligence and New Technologies in Melanoma Diagnosis: A Narrative Review. *Cancers*. 2025;17:3896.

    doi: 10.3390/cancers17243896

13. Sengupta D. Artificial Intelligence in Diagnostic Dermatology: Challenges and the Way Forward. *Indian Dermatol Online J.* 2023;14:782.

    doi: 10.4103/IDOJ.IDOJ_462_23

14. França RP, Borges Monteiro AC, Arthur R, Iano Y. An overview of deep learning in big data, image, and signal processing in the modern digital age. Trends in Deep Learning Methodologies: Algorithms, Applications, and Systems. 2020:63–87.

    doi: 10.1016/B978-0-12-822226-3.00003-9

15. Yu Z, Liu JK, Jia S, *et al*. Toward the Next Generation of Retinal Neuroprosthesis: Visual Computation with Spikes. *Engineering* 2020;6:449–61.

    doi: 10.1016/j.eng.2020.02.004

16. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthc J.* 2021;8:e188.

    doi: 10.7861/FHJ.2021-0095

17. Rashidi HH, Pantanowitz J, Hanna MG, *et al*. Introduction to Artificial Intelligence and Machine Learning in Pathology and Medicine: Generative and Nongenerative Artificial Intelligence Basics. *Mod Pathol.* 2025;38:100688.

    doi: 10.1016/J.MODPAT.2024.100688

18. Mienye ID, Swart TG. A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications. *Information*. 2024;15:755.

    doi: 10.3390/INFO15120755

19. Hermosilla P, Soto R, Vega E, Suazo C, Ponce J. Skin Cancer Detection and Classification Using Neural Network Algorithms: A Systematic Review. *Diagnostics*. 2024;14:454.

    doi: 10.3390/DIAGNOSTICS14040454

20. Sadr H, Nazari M, Khodaverdian Z, *et al*. Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. *Eur J Med Res.* 2025;30:418.

    doi: 10.1186/S40001-025-02680-7

21. Davis LE, Shalin SC, Tackett AJ. Current state of melanoma diagnosis and treatment. *Cancer Biol Ther.* 2019;20:1366.

    doi: 10.1080/15384047.2019.1640032

22. Das S, Dey MK, Devireddy R, Gartia MR. Biomarkers in Cancer Detection, Diagnosis, and Prognosis. *Sensors*. 2023;24:37.

    doi: 10.3390/S24010037

23. Marrapodi R, Bellei B. The Keratinocyte in the Picture Cutaneous Melanoma Microenvironment. *Cancers.* 2024;16:913.

    doi: 10.3390/CANCERS16050913

24. Lin T-L, Lee K-H, Karmakar R, *et al*. Artificial Intelligence-Assisted Dermatologic Screening: Epidemiology and Clinical Features of Basal Cell Carcinoma, Squamous Cell Carcinoma, Seborrheic Keratosis and Actinic Keratosis. *Bioengineering.* 2025;12:1258.

    doi: 10.3390/BIOENGINEERING12111258

25. Najar Najafi N, Hajihassani H, Azimzadeh Irani M. The Impact of Artificial Intelligence on Cancer Diagnosis and Treatment: A Review. *Cancer Inform* 2025;24:11769351251371272.

    doi: 10.1177/11769351251371273

26. Zhang B, Shi H, Wang H. Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment Selection: A Critical Approach. *J Multidiscip Healthc.* 2023;16:1779.

    doi: 10.2147/JMDH.S410301

27. Mao WB, Lyu JY, Vaishnani DK, *et al*. Application of artificial neural networks in detection and diagnosis of gastrointestinal and liver tumors. *World J Clin Cases.* 2020;8:3971.

    doi: 10.12998/wjcc.v8.i18.3971

28. Güler F, Ağraz M. Investigation of Binary and Multiclass Classification Performance of Skin Cancer Images Using Transfer Learning Methods. *J Clin Pract Res.* 2025;47:235.

    doi: 10.14744/cpr.2025.27623

29. Geradts Z, Filius N, Ruifrok A. Interpol review of imaging and video 2016–2019. *Forensic Sci Int*. 2020;2:540–62.

    doi: 10.1016/j.fsisyn.2020.01.017

30. Hussein AA, Abdulazeez AM. A Review on Utilizing Machine Learning Classification Algorithms for Skin Cancer. *J Appl Sci Technol Trends.* 2024;5:60–71.

    doi: 10.38094/jastt52191

31. Pennisi F, Borlini S, Harrison H, *et al*. Cancer Risk Prediction Using Machine Learning for Supporting Early Cancer Diagnosis in Symptomatic Patients: A Systematic Review of Model Types. *Cancer Med.* 2025;14:e71463.

    doi: 10.1002/cam4.71463

32. Ahmed SF, Alam MS Bin, Hassan M, *et al*. Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artif Intell Rev.* 2023;56:13521–617.

    doi: 10.1007/s10462-023-10466-8

33. Ali M, Dewan A, Sahu AK, Taye MM. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers.* 2023;12:91.

    doi: 10.3390/COMPUTERS12050091

34. Ahmad I, Alqurashi F. Early cancer detection using deep learning and medical imaging: A survey. *Crit Rev Oncol Hematol.* 2024;204:104528.

    doi: 10.1016/J.CRITREVONC.2024.104528

35. Hussein AA, Montaser AM, Elsayed HA. Skin cancer image classification using hybrid quantum deep learning model with BiLSTM and MobileNetV2. *Quantum Mach Intell.* 2025;7:66.

    doi: 10.1007/s42484-025-00288-y

36. Syafiandini AF, Wasito I, Yazid S, Fitriawan A, Amien M. Multimodal Deep Boltzmann Machines for feature selection on gene expression data. In: Proceedings of the 2016 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2016 2017:407–12.

    doi: 10.1109/ICACSIS.2016.7872733

37. Vinay S. Nalawade. Developing Improved Melanoma Detection Strategies Using Hybrid CNN and Autoencoder Models and Detailed Data Analysis. *J Inf Syst Eng Manag.* 2025;10:383–99.

    doi: 10.52783/jisem.v10i9s.1238

38. Kiser AC, Shi J, Bucher BT. An Explainable Long Short-Term Memory Network for Surgical Site Infection Identification. *Surgery.* 2024;176:24.

    doi: 10.1016/j.surg.2024.03.006

39. Han SS, Moon IJ, Kim SH, *et al*. Assessment of deep neural networks for the diagnosis of benign and malignant skin neoplasms in comparison with dermatologists: A retrospective validation study. *PLoS Med.* 2020;17:e1003381.

    doi: 10.1371/JOURNAL.PMED.1003381.

40. Puri P, Comfere N, Drage LA, *et al*. Deep Learning for Dermatologists: Part II. Current Applications. *J Am Acad Dermatol.* 2020;87:1352.

    doi: 10.1016/j.jaad.2020.05.053.

41. Farhi L, Kazmi SM, Imam H, Alqahtani M, Rehman FU. Dermoscopic Image Classification Using Deep Belief Learning Network Architecture. *Wirel Commun Mob Comput.* 2022;2022.

    doi: 10.1155/2022/2415726.

42. Qureshi AS, Roos T. Transfer Learning with Ensembles of Deep Neural Networks for Skin Cancer Detection in Imbalanced Data Sets. *Neural Process Lett.* 2023;55:4461–79.

    doi: 10.1007/s11063-022-11049-4.

43. Barata C, Rotemberg V, Codella NCF, *et al*. A reinforcement learning model for AI-based decision support in skin cancer. *Nat Med.* 2023;29:1941.

    doi: 10.1038/s41591-023-02475-5.

44. Afza F, Sharif M, Khan MA, Tariq U, Yong HS, Cha J. Multiclass Skin Lesion Classification Using Hybrid Deep Features Selection and Extreme Learning Machine. *Sensors.* 2022;22:799.

doi: 10.3390/s22030799.

45. Patel RH, Foltz EA, Witkowski A, Ludzik J. Analysis of Artificial Intelligence-Based Approaches Applied to Non-Invasive Imaging for Early Detection of Melanoma: A Systematic Review. *Cancers.* 2023;15:4694.

doi: 10.3390/CANCERS15194694

46. Bhatt H, Shah V, Shah K, Shah R, Shah M. State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review. *Intell Med.* 2023;3:180–90.

doi: 10.1016/j.imed.2022.08.004

47. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, *et al.* Deep neural networks are superior to dermatologists in melanoma image classification. *Eur J Cancer.* 2019;119:11–7.

doi: 10.1016/j.ejca.2019.05.023

48. 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;542:115–8.

doi: 10.1038/nature21056

49. Usama M, Naeem MA, Mirza F. Multi-Class Skin Lesions Classification Using Deep Features. *Sensors.* 2022;22:8311.

doi: 10.3390/s22218311

50. Musthafa MM, T R M, V VK, Guluwadi S. Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification. *BMC Med Imaging.* 2024 24:1 2024;24:201.

doi: 10.1186/s12880-024-01356-8

51. Gouda W, Sama NU, Al-Waakid G, Humayun M, Jhanjhi NZ. Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning. *Healthcare.* 2022;10.

doi: 10.3390/HEALTHCARE10071183

52. Melbin K, Raj YJV. Integration of modified ABCD features and support vector machine for skin lesion types classification. *Multimed Tools Appl.* 2020;80:6 2020;80:8909–29.

doi: 10.1007/s11042-020-10056-8

53. Divya D, Ganeshbabu TR. Fitness adaptive deer hunting-based region growing and recurrent neural network for melanoma skin cancer detection. *Int J Imaging Syst Technol.* 2020;30:731–52.

doi: 10.1002/ima.22414

54. Nawaz M, Mehmood Z, Nazir T, *et al.* Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microsc Res Tech.* 2022;85:339–51.

doi: 10.1002/jemt.23908

55. Rezk E, Haggag M, Eltorki M, El-Dakhakhni W. A comprehensive review of artificial intelligence methods and applications in skin cancer diagnosis and treatment: Emerging trends and challenges. *Healthcare Anal.* 2023;4:100259.

doi: 10.1016/J.HEALTH.2023.100259

56. Brancaccio G, Balato A, Malvehy J, Puig S, Argenziano G, Kittler H. Artificial Intelligence in Skin Cancer Diagnosis: A Reality Check. *J Investig Dermatol.* 2024;144:492–9.

doi: 10.1016/J.JID.2023.10.004

57. 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;11:6826.

doi: 10.3390/JCM11226826

58. Chan B. Black-box assisted medical decisions: AI power vs. ethical physician care. *Med Health Care Philos.* 2023;26:285.

doi: 10.1007/S11019-023-10153-Z

59. Combalia M, Codella N, Rotemberg V, *et al.* Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. *Lancet Digit Health.* 2022;4:e330.

doi: 10.1016/S2589-7500(22)00021-8

60. Yadav N, Pandey S, Gupta A, Dudani P, Gupta S, Rangarajan K. Data Privacy in Healthcare: In the Era of Artificial Intelligence. *Indian Dermatol Online J.* 2023;14:788.

doi: 10.4103/IDOJ.IDOJ_543_23

61. Wei ML, Tada M, So A, Torres R. Artificial intelligence and skin cancer. *Front Med.* 2024;11:1331895.

doi: 10.3389/FMED.2024.1331895/FULL

62. Ahmed F, Naz NS, Khan S, Rehman AU, Ismael WM, Khan MA. Explainable artificial intelligence (XAI) in medical imaging: a systematic review of techniques, applications, and challenges. *BMC Med Imaging.* 2026;26:37.

doi: 10.1186/S12880-025-02118-W

63. Al-Waisy AS, Al-Fahdawi S, Khalaf MI, Mohammed MA, Al-Attar B, Al-Andoli MN. A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images. *Sci Rep.* 2025 15:1 2025;15:31234-.

doi: 10.1038/s41598-025-15655-9

64. Parvin N, Joo SW, Jung JH, Mandal TK. Multimodal AI in Biomedicine: Pioneering the Future of Biomaterials, Diagnostics, and Personalized Healthcare. *Nanomaterials.* 2025;15:895.

doi: 10.3390/nano15120895

65. Okita AL, de Sousa RM, Rivero-Zavala EJ, *et al*. Development of an AI-Based Skin Cancer Recognition Model and Its Application in Enabling Patients to Self-Triage Their Lesions with Smartphone Pictures. *Dermato.* 2024;4:97–111.

doi: 10.3390/dermato4030011

66. Zhou J, Li H, Chen S, Chen Z, Han Z, Gao X. Large language models in biomedicine and healthcare. *Npj Artif Intell.* 2025;1:44.

doi: 10.1038/s44387-025-00047-1

67. Ribeiro Silva Fernandes T, Teles AS, Renan Neves Fernandes J, Daniel Batista Lima L, Lima Sousa D, de Castro Soares R, *et al*. External Validation of AI Models for Skin Diseases: A Systematic Review. *IEEE Access.* 2025;13:114411–27.

doi: 10.1109/ACCESS.2025.3584904

68. Yu Z, Xin C, Yu Y, Xia J, Han L. AI dermatology: Reviewing the frontiers of skin cancer detection technologies. *Intell Oncol.* 2025;1:89–104.

doi: 10.1016/J.INTONC.2025.03.002

69. Kurtansky NR, D'Alessandro BM, Gillis MC, *et al*. The SLICE-3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection. *Sci Data.* 2024;11.

doi: 10.1038/S41597-024-03743-W

70. Yu J, Cheong IH, Kozlakidis Z, Wang H. Advancements and challenges of artificial intelligence in dermatology: a review of applications and perspectives in China. *Front Digit Health.* 2025;7:1544520.

doi: 10.3389/FDGTH.2025.1544520

71. Ricci Lara MA, Rodríguez Kowalczuk MV, *et al*. A dataset of skin lesion images collected in Argentina for the evaluation of AI tools in this population. *Sci Data.* 2023;10.

doi: 10.1038/S41597-023-02630-0

72. 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

73. Accurate and Scalable System for Automatic Detection of Malignant Melanoma. *Dermoscopy Image Anal.* 2015:309–60.

doi: 10.1201/B19107-14

74. Somfai E, Baffy B, Fenech K, *et al*. Handling dataset dependence with model ensembles for skin lesion classification from dermoscopic and clinical images. *Int J Imaging Syst Technol.* 2023;33:556–71.

doi: 10.1002/IMA.22827

75. Cahyanto KA, Adi K, Widodo CE. The L2-EffCANet: A Novel Overfitting-Resistant EfficientNetV2S with Attention Mechanism and L2 Regularization for Skin Disease Classification. *Int J Online Biomed Eng.* 2025;21:113–129.

doi: 10.3991/IJOE.V21I13.57417

76. Pollastri F, Parreño M, Maroñas J, Bolelli F, Paredes R, Ramos D, *et al*. A deep analysis on high-resolution dermoscopic image classification. *IET Computer Vision.* 2021;15:514–526.

doi: 10.1049/CVI2.12048;PAGE:STRING:ARTICLE/CHAPTER

77. Jeong HK, Park C, Henao R, Kheterpal M. Deep Learning in Dermatology: A Systematic Review of Current Approaches, Outcomes, and Limitations. *JID Innov.* 2022;3:100150.

doi: 10.1016/J.XJIDI.2022.100150

78. Li Y, Taylor M, Chmielinski KS, Halpern AC, Daneshjou R, Lester JC, *et al*. Improving dataset transparency in dermatologic Artificial Intelligence using a dataset nutrition label. *NPJ Digit Med.* 2025;8:641.

doi: 10.1038/S41746-025-02125-9

79. Wang M, Chang W, Zhang Y. Artificial Intelligence for the Diagnosis and Management of Cancers: Potentials and Challenges. *MedComm.* 2025;6:e70460.

doi: 10.1002/MCO2.70460

80. Abd-Alrazaq A, Solaiman B, Mekki YM, Al-Thani D, Farooq F, Alkubeyyer M, *et al*. Hype vs Reality in the Integration of Artificial Intelligence in Clinical Workflows. *JMIR Form Res.* 2025;9:e70921.

doi: 10.2196/70921

81. Das A, Agarwal V, Shetty NP. Comparative analysis of multimodal architectures for effective skin lesion detection using clinical and image data. *Front Artif Intell.* 2025;8:1608837.

doi: 10.3389/FRAI.2025.1608837/BIBTEX

82. Behara K, Bhero E, Agee JT. AI in dermatology: a comprehensive review into skin cancer detection. *PeerJ Comput Sci.* 2024;10:1–42.

doi: 10.7717/PEERJ-CS.2530/SUPP-1

83. Alipour N, Burke T, Courtney J. Skin Type Diversity in Skin Lesion Datasets: A Review. *Curr Dermatol Rep.* 2024;13:198.

doi: 10.1007/S13671-024-00440-0

84. Fehr J, Citro B, Malpani R, Lippert C, Madai VI. A trustworthy AI reality-check: the lack of transparency of artificial intelligence products in healthcare. *Front Digit Health.* 2024;6:1267290.

doi: 10.3389/FDGTH.2024.1267290/FULL

85. Fu L, Jia G, Liu Z, Pang X, Cui Y. The applications and advances of artificial intelligence in drug regulation: A global perspective. *Acta Pharm Sin B.* 2025;15:1–14.

doi: 10.1016/J.APSB.2024.11.006

86. Park AJ, Weintraub GS, Asgari MM. Leveraging the

electronic health record to improve dermatologic care delivery: The importance of finding structure in data. *J Am Acad Dermatol.* 2019;82:773.

doi: 10.1016/J.JAAD.2019.10.064

87. Nisevic M, Milojevic D, Spajic D. Synthetic data in medicine: Legal and ethical considerations for patient profiling. *Comput Struct Biotechnol J.* 2025;28:190–8.

doi: 10.1016/J.CSBJ.2025.05.026

88. Singh S, Raucci A, Glovi A, Iula G, Mutti L, De Laurentiis M, *et al*. Socioeconomic impact of artificial intelligence–driven point-of-care testing devices for liquid biopsy in the OncoCheck system. *Cancer Metastasis Rev.* 2025;44:64.

doi: 10.1007/S10555-025-10281-3

89. Cugliari G, Bodini M, Zbrzezny AM, Krzywicki T. Artificial Intelligence in Dermatology: A Review of Methods, Clinical Applications, and Perspectives. *Appl Sci.* 2025;15:7856.

doi: 10.3390/APP15147856

90. Patel JC, Shukla M, Shukla M. From bench to bedside: translating mesenchymal stem cell therapies through preclinical and clinical evidence. *Front Bioeng Biotechnol.* 2025;13:1639439.

doi: 10.3389/FBIOE.2025.1639439

91. Eapen BR. Artificial Intelligence in Dermatology: A Practical Introduction to a Paradigm Shift. *Indian Dermatol Online J.* 2020;11:881.

doi: 10.4103/IDOJ.IDOJ_388_20